On the additivity conjecture for channels with arbitrary constraints

M.E.Shirokov

1   Introduction

In [15] Shor proved equivalence of several open (sub)additivity problems related to the Holevo capacity and the entanglement of formation (EoF). In [6] we showed equivalence of these to the additivity of the Holevo capacity for channels with arbitrary linear constraints.

This note is the development of [6] in the direction of channels with general constraints. Introducing input constraints provides greater flexibility in the treatment of the additivity conjecture. On the other hand, while [15], [6] deal with the "global" additivity conjecture, i.e. properties valid for all possible channels, in this paper we make emphasis on results valid for individual channels. The Holevo capacity of the arbitrarily constrained channel is considered and the characteristic property of an optimal ensemble for such channel is derived (proposition 1), generalizing the maximal distance property of Schumacher and Westmoreland [13] . This property provides a useful estimate for the Holevo capacity of the constrained channel (proposition 2). Our main point of interest is the additivity conjecture for two arbitrarily constrained channels and its relations to Shor's channel extension [15]. An attempt to prove converse of the theorem 1 in [6], in which sufficient condition for "constrained additivity" was formulated, leads to the notion of asymptotic additivity for sequences of channels. This notion applied to the sequences of Shor's channel extensions makes it possible to formulate the main result of [6] in the form of necessary and sufficient condition. This implies that the additivity conjecture for two channels with *single* linear constraints is equivalent to the similar conjecture for two *arbitrarily* constrained channels and, hence, to an interesting subadditivity property of the χ-function for the tensor product of these channels.

This subadditivity property of the χ-function seems to be very appealing. It implies the unconstrained additivity and can be used for proving the latter in some cases. This property can be established for several types of channels (proposition 4). The characteristic property of the optimal ensembles for constrained channels provides necessary and sufficient condition for additivity of the Holevo capacity for two channels, constrained by fixing partial states of average state of input ensemble (theorem 2). This leads to an interesting characterization of the channels for which subadditivity of the χ-function holds (corollary 5).

Global subadditivity of the χ-function follows obviously from the strong superadditivity of the entanglement of formation and from the MSW channel representation [8]. But it turns out that the converse is also true (corollary 6). With the corollary 3 this gives another (in comparison to [15]) way of proving that the global additivity conjecture for unconstrained channels implies strong superadditivity of the entanglement of formation.

The arguments from the convex analysis raised in [2] provide another characterization of channels for which subadditivity of the χ-function holds (theorem 3). This characterization and some modification of Shor's channel extension provides a simple way of proving that global additivity of the minimum output entropy for unconstrained channels implies global subadditivity of the χ-function and strong superadditivity of the entanglement of formation.

2 Channels with arbitrary constraints

Let $\Phi : \mathfrak{S}(\mathcal{H}) \mapsto \mathfrak{S}(\mathcal{H}')$ be an arbitrary channel (here $\mathcal{H}, \mathcal{H}'$ are finite dimensional Hilbert spaces). Let $\{\rho_i\}$ be an arbitrary ensemble of states in $\mathfrak{S}(\mathcal{H})$ with the probability distribution $\{\pi_i\}$, which will be denoted as $\{\pi_i, \rho_i\}$.

The Holevo quantity [3] for this ensemble is defined by

$$\chi_\Phi \left(\{\pi_i, \rho_i\} \right) = H \left(\sum_i \pi_i \Phi \left(\rho_i \right) \right) - \sum_i \pi_i H \left(\Phi \left(\rho_i \right) \right).$$

For an arbitrary state ρ we denote

$$\chi_\Phi(\rho) = \max \chi_\Phi(\{\pi_i, \rho_i\}),$$

2

where the maximum is over all ensembles $\{\pi_i, \rho_i\}$ with $\sum_i \pi_i \rho_i = \rho$. The function $\chi_\Phi(\rho)$ (briefly χ-function) is continuous and concave on $\mathfrak{S}(\mathcal{H})$. The concavity of $\chi_\Phi(\rho)$ easily follows from the definition. The continuity of $\chi_\Phi(\rho)$ can be derived from the MSW-correspondence [8] and the continuity of the entanglement of formation [9].

Fixing a closed subset $\mathcal{A}$ of $\mathfrak{S}(\mathcal{H})$ we can consider the constraints on the input ensemble $\{\pi_i, \rho_i\}$ of the channel Φ defined by the requirement $\rho_{\mathrm{av}} \in \mathcal{A}$, where $\rho_{\mathrm{av}} = \sum_i \pi_i \rho_i$ is the average state of the ensemble. This type of constraints is a natural generalization of linear constraints considered in [4], where the subset $\mathcal{A}$ is defined by the inequality $\mathrm{Tr} A \rho_{\mathrm{av}} \leq \alpha$ for some positive operator A and positive number α.

Define the Holevo capacity of the $\mathcal{A}$-constrained channel Φ by

$$\bar{C}(\Phi; \mathcal{A}) = \max_{\rho \in \mathcal{A}} \chi_\Phi(\rho) = \max_{\sum \pi_i \rho_i \in \mathcal{A}} \chi_\Phi(\{\pi_i, \rho_i\}). \qquad (1)$$

Note that the unconstrained capacity $\bar{C}(\Phi) = \bar{C}(\Phi; \mathfrak{S}(\mathcal{H}))$.

Any ensemble $\{\pi_i, \rho_i\}$ on which the maximum in (1) is achieved is called an *optimal ensemble* for the $\mathcal{A}$-constrained channel Φ. In [13] it was shown that an optimal ensemble $\{\pi_i, \rho_i\}$ (with the average state ρ_{av}) for the unconstrained channel Φ is characterized the by the *maximal distance property*:

$$S(\Phi(\omega) \| \Phi(\rho_{\mathrm{av}})) \leq \chi_\Phi(\{\pi_i, \rho_i\}), \quad \forall \omega \in \mathfrak{S}(\mathcal{H}).$$

The generalization of the above property for the constrained channels is given by the following proposition.

Proposition 1. The ensemble $\{\pi_i, \rho_i\}$ with the average state $\rho_{\mathrm{av}} \in \mathcal{A}$ is optimal for the $\mathcal{A}$-constrained channel Φ with a *convex* set $\mathcal{A}$ if and only if

$$\sum_j \mu_j S(\Phi(\omega_j) \| \Phi(\rho_{\mathrm{av}})) \leq \chi_\Phi(\{\pi_i, \rho_i\})$$

for any ensemble $\{\mu_j, \omega_j\}$ of states in $\mathfrak{S}(\mathcal{H})$ with the average state $\omega_{\mathrm{av}} \in \mathcal{A}$.

Proof. The proof is a simple generalization of the arguments in [13]. The specific feature of the "constrained" situation consists in the necessity to consider variations of the initial ensemble by mixing not only one state as in [13], but a genuine ensemble.

Let $\{\pi_i, \rho_i\}_{i=1}^n$ and $\{\mu_j, \omega_j\}_{j=1}^m$ be two ensembles with average states ρ_{av} and ω_{av}, contained in $\mathcal{A}$. Consider the modification of the first ensemble by adding the second one with the weight coefficient η. The modified ensemble

$$\{(1 - \eta)\pi_1 \rho_1, ..., (1 - \eta)\pi_n \rho_n, \eta \mu_1 \omega_1, ..., \eta \mu_m \omega_m\}$$

3

has the average state $\rho_{\mathrm{av}}^{\eta} = (1-\eta)\rho_{\mathrm{av}} + \eta\omega_{\mathrm{av}}$, contained in $\mathcal{A}$ (by convexity). Let χ_{Φ} and χ_{Φ}^{η} be the Holevo quantities for the original and the modified ensembles correspondingly. Using the relative entropy expression for the Holevo quantity [14] we have

$$\chi_{\Phi}^{\eta} = (1-\eta)\sum_{i=1}^{n}\pi_i S(\Phi(\rho_i)\|\Phi(\rho_{\mathrm{av}}^{\eta})) + \eta\sum_{j=1}^{m}\mu_j S(\Phi(\omega_j)\|\Phi(\rho_{\mathrm{av}}^{\eta})). \qquad (2)$$

Applying Donald's identity [13],[14] to the original ensemble we obtain

$$\sum_{i=1}^{n}\pi_i S(\Phi(\rho_i)\|\Phi(\rho_{\mathrm{av}}^{\eta})) = \chi_{\Phi} + S(\Phi(\rho_{\mathrm{av}})\|\Phi(\rho_{\mathrm{av}}^{\eta})).$$

Substitution of the above expression into (2) gives

$$\chi_{\Phi}^{\eta} = \chi_{\Phi} + \eta\left[\sum_{j=1}^{m}\mu_j S(\Phi(\omega_j)\|\Phi(\rho_{\mathrm{av}}^{\eta})) - \chi_{\Phi}\right] + (1-\eta)S(\Phi(\rho_{\mathrm{av}})\|\Phi(\rho_{\mathrm{av}}^{\eta})). \qquad (3)$$

Applying Donald's identity to the modified ensemble we obtain

$$(1-\eta)\sum_{i=1}^{n}\pi_i S(\Phi(\rho_i)\|\Phi(\rho_{\mathrm{av}})) + \eta\sum_{j=1}^{m}\mu_j S(\Phi(\omega_j)\|\Phi(\rho_{\mathrm{av}}))$$
$$= \chi_{\Phi}^{\eta} + S(\Phi(\rho_{\mathrm{av}}^{\eta})\|\Phi(\rho_{\mathrm{av}}))$$

and hence

$$\chi_{\Phi}^{\eta} = \chi_{\Phi} + \eta\left[\sum_{j=1}^{m}\mu_j S(\Phi(\omega_j)\|\Phi(\rho_{\mathrm{av}})) - \chi_{\Phi}\right] - S(\Phi(\rho_{\mathrm{av}}^{\eta})\|\Phi(\rho_{\mathrm{av}})). \qquad (4)$$

Since the relative entropy is nonnegative the expressions (3) and (4) imply the following double inequality for the value $\Delta\chi_{\Phi} = \chi_{\Phi}^{\eta} - \chi_{\Phi}$:

$$\eta\left[\sum_{j=1}^{m}\mu_j S(\Phi(\omega_j)\|\Phi(\rho_{\mathrm{av}}^{\eta})) - \chi_{\Phi}\right]$$
$$\leq \Delta\chi_{\Phi} \leq \qquad (5)$$
$$\eta\left[\sum_{j=1}^{m}\mu_j S(\Phi(\omega_j)\|\Phi(\rho_{\mathrm{av}})) - \chi_{\Phi}\right].$$

4

Now the proof of the proposition is straightforward. If

$$\sum_j \mu_j S(\Phi(\omega_j)\|\Phi(\rho_{\mathrm{av}})) \leq \chi_\Phi$$

for any ensemble $\{\mu_j, \omega_j\}$ of states in $\mathfrak{S}(\mathcal{H})$ with the average state $\omega_{\mathrm{av}} \in \mathcal{A}$, then by (5) with $\eta = 1$ we have

$$\chi_\Phi(\{\mu_j, \omega_j\}) = \chi_\Phi^1 \leq \chi_\Phi = \chi_\Phi(\{\pi_i, \rho_i\}),$$

which means optimality of the ensemble $\{\pi_i, \rho_i\}$.

To prove the converse, suppose $\{\pi_i, \rho_i\}$ is an optimal ensemble and there exists an ensemble $\{\mu_j, \omega_j\}$ such that

$$\sum_j \mu_j S(\Phi(\omega_j)\|\Phi(\rho_{\mathrm{av}})) > \chi_\Phi.$$

By continuity of the relative entropy there is $\eta > 0$ such that

$$\sum_j \mu_j S(\Phi(\omega_j)\|\Phi(\rho_{\mathrm{av}}^\eta)) > \chi_\Phi,$$

because ρ_{av}^η tends to ρ_{av} when η tends to zero. By the first inequality in (5), the last inequality means that $\chi_\Phi^\eta > \chi_\Phi$ in contradiction with the optimality of the ensemble $\{\pi_i, \rho_i\}$.$\triangle$

The above characteristic of an optimal ensemble provides the following estimate for the Holevo capacity of the constrained channel.

Proposition 2. Let ρ_{av} be the average state of any optimal ensemble for the $\mathcal{A}$-constrained channel Φ with a closed convex subset $\mathcal{A}$ of $\mathfrak{S}(\mathcal{H})$. Then

$$\bar{C}(\Phi; \mathcal{A}) = \chi_\Phi(\rho_{\mathrm{av}}) \geq \chi_\Phi(\rho) + S(\Phi(\rho)\|\Phi(\rho_{\mathrm{av}})), \quad \forall \rho \in \mathcal{A}.$$

Proof. Let $\{\pi_i, \rho_i\}$ be an arbitrary ensemble such that $\sum_i \pi_i \rho_i = \rho \in \mathcal{A}$. By proposition 1

$$\sum_i \pi_i S(\Phi(\rho_i)\|\Phi(\rho_{\mathrm{av}})) \leq \chi_\Phi(\rho_{\mathrm{av}}).$$

This inequality and Donald's identity

$$\sum_i \pi_i S(\Phi(\rho_i)\|\Phi(\rho_{\mathrm{av}})) = \chi_\Phi(\{\pi_i, \rho_i\}) + S(\Phi(\rho)\|\Phi(\rho_{\mathrm{av}})).$$

complete the proof.$\triangle$

3 Shor's channel extension and constrained channels

Let $\Psi : \mathfrak{S}(\mathcal{K}) \mapsto \mathfrak{S}(\mathcal{K}')$ be another channel with the constraint, defined by a closed subset $\mathcal{B}$ of $\mathfrak{S}(\mathcal{K})$. For the channel $\Phi \otimes \Psi$ it is natural to consider the constraint defined by the requirements $\sigma_{\mathrm{av}}^{\Phi} = \mathrm{Tr}_{\mathcal{K}}\sigma_{\mathrm{av}} \in \mathcal{A}$ and $\sigma_{\mathrm{av}}^{\Psi} = \mathrm{Tr}_{\mathcal{H}}\sigma_{\mathrm{av}} \in \mathcal{B}$ on the average state σ_{av} of the input ensemble $\{\pi_i, \sigma_i\}$. The closed subset of $\mathfrak{S}(\mathcal{H} \otimes \mathcal{K})$ defined by the above requirements will be denoted as $\mathcal{A} \otimes \mathcal{B}$.

We conjecture the following additivity property for constrained channels

$$\bar{C}\left(\Phi \otimes \Psi; \mathcal{A} \otimes \mathcal{B}\right) \stackrel{?}{=} \bar{C}(\Phi; \mathcal{A}) + \bar{C}(\Psi; \mathcal{B}). \tag{6}$$

Note that additivity of the Holevo capacity is a partcular case of (6) with $\mathcal{A} = \mathfrak{S}(\mathcal{H})$ and $\mathcal{B} = \mathfrak{S}(\mathcal{K})$.

In [6] the role of Shor's channel extension [15] in connection with the constrained additivity problem was demonstrated. For reader's convenience we give here a brief description of this notion.

Let E be an operator in $\mathfrak{B}(\mathcal{H}), 0 \leq E \leq I_{\mathcal{H}}$ (the identity operator in the space $\mathcal{H}$), let $q \in [0;1]$ and $d \in \mathbb{N} = \{1, 2, \dots\}$. Consider the channel $\widehat{\Phi}(E, q, d)$, which maps $\mathfrak{B}(\mathcal{H}) \otimes \mathbf{C}^d$ into $\mathfrak{B}(\mathcal{H}') \oplus \mathbf{C}^{d+1}$, where $\mathbf{C}^d$ is the commutative algebra of complex d-dimensional vectors describing a classical system. By using the isomorphism of $\mathfrak{B}(\mathcal{H}) \otimes \mathbf{C}^d$ with the direct sum of d copies of $\mathfrak{B}(\mathcal{H})$, any state in $\mathfrak{B}(\mathcal{H}) \otimes \mathbf{C}^d$ can be represented as an array $\{\rho_j\}_{j=1}^d$ of positive operators in $\mathfrak{B}(\mathcal{H})$ such that $\mathrm{Tr}\sum_{j=1}^d \rho_j = 1$. The action of the channel $\widehat{\Phi}(E, q, d)$ on the state $\widehat{\rho} = \{\rho_j\}_{j=1}^d$ with $\rho = \sum_{j=1}^d \rho_j$ is defined by

$$\widehat{\Phi}(E, q, d)(\widehat{\rho}) = (1-q)\Phi_0(\widehat{\rho}) \oplus q\Phi_1(\widehat{\rho}),$$

where $\Phi_0(\widehat{\rho}) = \Phi(\rho) \in \mathfrak{S}(\mathcal{H}')$ and $\Phi_1(\widehat{\rho}) = [\mathrm{Tr}\rho\bar{E}, \mathrm{Tr}\rho_1 E, ..., \mathrm{Tr}\rho_d E] \in \mathbf{C}^{d+1}$ (throughout this paper we use the notation $\bar{A} = I - A$ for operators). Note that Φ_0 and Φ_1 are channels from $\mathfrak{B}(\mathcal{H}) \otimes \mathbf{C}^d$ to $\mathfrak{B}(\mathcal{H}')$ and to $\mathbf{C}^{d+1}$ correspondingly. The state space of the channel $\widehat{\Phi}(E, q, d)$ will be denoted as $\mathfrak{S}_{\widehat{\Phi}}$.

We will need some generalization of the proposition 1 in [6].

Proposition 3. Let $\Psi : \mathfrak{S}(\mathcal{K}) \mapsto \mathfrak{S}(\mathcal{K}')$ be an arbitrary $\mathcal{B}$ -constrained channel. Consider the channel $\widehat{\Phi}(E, q, d) \otimes \Psi$. Then

$$\left| \bar{C}\left(\widehat{\Phi}(E, q, d) \otimes \Psi, \mathfrak{S}_{\widehat{\Phi}} \otimes \mathcal{B} \right) - \max_{\sigma : \mathrm{Tr}_{\mathcal{H}} \sigma \in \mathcal{B}} \left[(1-q)\chi_{\Phi \otimes \Psi}(\sigma) + q \log d \mathrm{Tr}(E \otimes I_{\mathcal{K}}) \sigma \right] \right|$$

$$\leq q(\log \dim \mathcal{K}' + 1).$$

Proof. Due to the representation

$$\widehat{\Phi}(E, q, d) \otimes \Psi = ((1 - q)\Phi_0 \otimes \Psi) \oplus (q\Phi_1 \otimes \Psi), \qquad (7)$$

the lemma 1 in [6] reduces the calculation of the Holevo quantity $\chi_{\widehat{\Phi}(E,q,d)\otimes\Psi}$ for any ensemble of input states to the calculation of the Holevo quantities $\chi_{\Phi_k \otimes \Psi}$, $k = 0, 1$ for this ensemble:

$$\chi_{\widehat{\Phi}(E,q,d)\otimes\Psi} = (1 - q)\chi_{\Phi_0 \otimes \Psi} + q\chi_{\Phi_1 \otimes \Psi}. \qquad (8)$$

Note that any state $\widehat{\sigma}$ in $\mathfrak{B}(\mathcal{H}) \otimes \mathbf{C}^d \otimes \mathfrak{B}(\mathcal{K})$ can be represented as an array $\{\sigma_j\}_{j=1}^d$ of positive operators in $\mathfrak{B}(\mathcal{H} \otimes \mathcal{K})$ such that $\mathrm{Tr}\sum_{j=1}^d \sigma_j = 1$. Denote by $\delta_j(\sigma)$ the array $\widehat{\sigma}$ with the state σ in the j-th position and with zeroes in other places.

It is known [13] that for any channel there exists a pure state optimal ensemble. This fact and symmetry arguments imply the existence of an optimal ensemble for the channel $\widehat{\Phi}(E, q, d) \otimes \Psi$ consisting of the states $\widehat{\sigma}_{i,j} = \delta_j(\sigma_i)$ with the probabilities $\widehat{\mu}_{i,j} = d^{-1}\mu_i$, where $\{\mu_i, \sigma_i\}$ is a particular ensemble of states in $\mathfrak{S}(\mathcal{H} \otimes \mathcal{K})$ (cf. [15]). Let $\widehat{\sigma}_{av} = \sum_{i,j} \widehat{\mu}_{i,j}\widehat{\sigma}_{i,j}$ and $\sigma_{av} = \sum_i \mu_i\sigma_i$ be the average states of these ensembles. Note that $\widehat{\sigma}_{av} = [d^{-1}\sigma_{av}, ..., d^{-1}\sigma_{av}]$.

The action of the channel $\Phi_0 \otimes \Psi$ on the state $\widehat{\sigma} = \{\sigma_i\}_{i=1}^d$ with $\sigma = \sum_{i=1}^d \sigma_i$ is defined by

$$\Phi_0 \otimes \Psi(\widehat{\sigma}) = \Phi \otimes \Psi(\sigma).$$

Hence $\Phi_0 \otimes \Psi(\widehat{\sigma}_{i,j}) = \Phi \otimes \Psi(\sigma_i)$ and

$$\chi_{\Phi_0 \otimes \Psi}(\{\widehat{\mu}_{i,j}, \widehat{\sigma}_{i,j}\}) = \chi_{\Phi \otimes \Psi}(\{\mu_i, \sigma_i\}). \qquad (9)$$

Let us prove that

$$\chi_{\Phi_1 \otimes \Psi}(\{\widehat{\mu}_{i,j}, \widehat{\sigma}_{i,j}\}) = \log d \mathrm{Tr}(E \otimes I_{\mathcal{K}})\sigma_{av} + f_{\Psi}^E(\{\mu_i, \sigma_i\}), \qquad (10)$$

7

where $f_\Psi^E(\{\mu_i, \sigma_i\}) \le \log \dim \mathcal{K}' + 1$. It is easy to see that the action of the channel $\Phi_1 \otimes \Psi$ on the state $\widehat{\sigma} = \{\sigma_i\}_{i=1}^d$ with $\sigma = \sum_{i=1}^d \sigma_i$ is defined by

$$\Phi_1 \otimes \Psi(\widehat{\sigma}) = [\Psi_{\bar{E}}(\sigma), \Psi_E(\sigma_1), ..., \Psi_E(\sigma_d)],$$

where $\Psi_A(\cdot) = \mathrm{Tr}_\mathcal{H}\,(A \otimes I_\mathcal{K} \cdot (\mathrm{Id} \otimes \Psi)(\cdot))$ is a completely positive map from $\mathfrak{S}(\mathcal{H} \otimes \mathcal{K})$ into $\mathfrak{B}_+(\,\mathcal{K}')$ ($A = E, \bar{E}$ and Id is an identity map on $\mathfrak{S}(\mathcal{H})$).

Therefore,

$$H(\Phi_1 \otimes \Psi(\widehat{\sigma}_{i,j})) = H(\Psi_{\bar{E}}(\sigma_i)) + H(\Psi_E(\sigma_i)), \tag{11}$$

and

$$\Phi_1 \otimes \Psi(\widehat{\sigma}_{\mathrm{av}}) = \sum_{i,j} \widehat{\mu}_{i,j} \Phi_1 \otimes \Psi(\widehat{\sigma}_{i,j})$$
$$= [\Psi_{\bar{E}}(\sigma_{\mathrm{av}}), d^{-1}\Psi_E(\sigma_{\mathrm{av}}), ..., d^{-1}\Psi_E(\sigma_{\mathrm{av}})],$$

Due to this we can conclude that

$$H(\Phi_1 \otimes \Psi(\widehat{\sigma}_{\mathrm{av}})) = \log d\ \mathrm{Tr}\Psi_E(\sigma_{\mathrm{av}}) + H(\Psi_E(\sigma_{\mathrm{av}})) + H(\Psi_{\bar{E}}(\sigma_{\mathrm{av}})). \tag{12}$$

It is easy to see that

$$\mathrm{Tr}\Psi_E(\sigma) = \mathrm{Tr}(E \otimes I_\mathcal{K})\,\sigma \tag{13}$$

Using (11), (12) and (13), we obtain

$$\chi_{\Phi_1 \otimes \Psi}(\{\widehat{\mu}_{i,j}, \widehat{\sigma}_{i,j}\}) = \log d\, \mathrm{Tr}(E \otimes I_\mathcal{K})\sigma_{\mathrm{av}} +$$

$$H(\Psi_E(\sigma_{\mathrm{av}})) + H(\Psi_{\bar{E}}(\sigma_{\mathrm{av}})) - \sum_i \mu_i(H(\Psi_E(\sigma_i)) + H(\Psi_{\bar{E}}(\sigma_i))) \tag{14}$$

$$= \log d\mathrm{Tr}(E \otimes I_\mathcal{K})\sigma_{\mathrm{av}} + \chi_{\Psi_E}(\{\mu_i, \sigma_i\}) + \chi_{\Psi_{\bar{E}}}(\{\mu_i, \sigma_i\}).$$

Using the inequalities $0 \le H(S) \le \mathrm{Tr}S(\log \dim \mathcal{H} - \log \mathrm{Tr}S)$ for any positive operator $S \in \mathcal{B}(\mathcal{H})$ and $h_2(x) = x \log x + (1-x)\log(1-x) \le 1$ it is possible to show that

$$f_\Psi^E(\{\mu_i, \sigma_i\}) = \chi_{\Psi_E}(\{\mu_i, \sigma_i\}) + \chi_{\Psi_{\bar{E}}}(\{\mu_i, \sigma_i\}) \le \log \dim \mathcal{K}' + 1. \tag{15}$$

Applying this estimation to (14) we obtain (10).

The expression (8) with (9) and (10) imply

$$\chi_{\widehat{\Phi}(E,q,d) \otimes \Psi}(\{\widehat{\mu}_{i,j}, \widehat{\sigma}_{i,j}\}) = (1-q)\chi_{\Phi_0 \otimes \Psi}(\{\widehat{\mu}_{i,j}, \widehat{\sigma}_{i,j}\}) + q\chi_{\Phi_1 \otimes \Psi}(\{\widehat{\mu}_{i,j}, \widehat{\sigma}_{i,j}\})$$

$$= (1-q)\chi_{\Phi \otimes \Psi}(\{\mu_i, \sigma_i\}) + q \log d\mathrm{Tr}\sigma_{\mathrm{av}}(E \otimes I_\mathcal{K}) + qf_\Psi^E(\{\mu_i, \sigma_i\}).$$

The last equality with (15) completes the proof.△

When dealing with Shor's channel extension in connection with the constrained additivity problem it is convenient to slightly change notation. For nonnegative number p and operator $0 \leq A \leq I$ we will denote by $\widehat{\Phi}_d(A, p)$ the Shor's extension $\widehat{\Phi}(A, p/\log d, d)$ of the channel Φ.

Definition 1. We say that the additivity conjecture holds asymptotically for the sequence of unconstrained channels $\{\Phi_n : \mathfrak{S}(\mathcal{H}_n) \mapsto \mathfrak{S}(\mathcal{H}'_n)\}$ and the channel $\Psi : \mathfrak{S}(\mathcal{K}) \mapsto \mathfrak{S}(\mathcal{K}')$ with the constraint defined by the set $\mathcal{B}$ if

$$\lim_{n \to +\infty} \bar{C}(\Phi_n \otimes \Psi, \mathfrak{S}(\mathcal{H}_n) \otimes \mathcal{B}) = \lim_{n \to +\infty} \bar{C}(\Phi_n) + \bar{C}(\Psi, \mathcal{B}),$$

assuming that the limits exist and are finite.

Note that if the additivity conjecture holds for the channels Φ_n and Ψ for all sufficiently large n it obviously holds asymptotically for the sequence $\{\Phi_n\}$ and the channel Ψ.

Theorem 1. Let $\Phi : \mathfrak{S}(\mathcal{H}) \mapsto \mathfrak{S}(\mathcal{H}')$ and $\Psi : \mathfrak{S}(\mathcal{K}) \mapsto \mathfrak{S}(\mathcal{K}')$ be arbitrary channels with the fixed constraint on the second one defined by the set $\mathcal{B}$. The following statements are equivalent:

(i) The additivity conjecture is true for the channel Φ with the constraint $\mathrm{Tr}A\rho \leq \alpha$ for arbitrary (A, α) and the $\mathcal{B}$-constrained channel Ψ;

(ii) The additivity conjecture holds asymptotically for the sequence of the channels $\{\widehat{\Phi}_d(A, p)\}_{d \in \mathbb{N}}$ (without constraints) and the $\mathcal{B}$-constrained channel Ψ for arbitrary nonnegative number p and arbitrary operator $0 \leq A \leq I$;

(iii) The additivity conjecture is true for the $\mathcal{A}$-constrained channel Φ with arbitrary closed subset $\mathcal{A}$ in $\mathfrak{S}(\mathcal{H})$ and the $\mathcal{B}$-constrained channel Ψ .

Proof. The implication (iii) $\Rightarrow$ (i) is obvious. So it is sufficient to prove (i) $\Rightarrow$ (ii) and (ii) $\Rightarrow$ (iii).

Note, first of all, that for an operator $0 \leq A \leq I$ and a number $p \geq 0$ proposition 3 implies

$$\lim_{d \to +\infty} \bar{C}(\Phi_d(A, p)) = \max_{\rho} \left[\chi_\Phi(\rho) + p\mathrm{Tr}A\rho\right] \tag{16}$$

and

$$\lim_{d \to +\infty} \bar{C}(\Phi_d(A, p) \otimes \Psi) = \max_{\sigma : \mathrm{Tr}_\mathcal{H}\sigma \in \mathcal{B}} \left[\chi_{\Phi \otimes \Psi}(\sigma) + p\mathrm{Tr}(A \otimes I_\mathcal{K})\sigma\right] \tag{17}$$

correspondingly.

Begin with (i) $\Rightarrow$ (ii). Suppose that there exist an operator $0 \leq A \leq I$ and a number $p \geq 0$ such that the additivity conjecture is not asymptotically true for the sequence $\{\Phi_d(A, p)\}_{d \in \mathbb{N}}$ and the channel Ψ. Due to (16) and (17) it means that

$$\max_{\sigma: \mathrm{Tr}_{\mathcal{H}} \sigma \in \mathcal{B}} [\chi_{\Phi \otimes \Psi}(\sigma) + p\mathrm{Tr}(A \otimes I_{\mathcal{K}}) \sigma] > \max_{\rho} [\chi_{\Phi}(\rho) + p\mathrm{Tr}A\rho] + \bar{C}(\Psi, \mathcal{B}). \quad (18)$$

Let σ_* be a maximum point in the left side of the above inequality and $\alpha = \mathrm{Tr}(A \otimes I_{\mathcal{K}}) \sigma_*$. By the statement (i) the additivity conjecture is true for the channel Φ with the constraint $\mathrm{Tr}\bar{A}\rho \leq 1 - \alpha$ and the $\mathcal{B}$-constrained channel Ψ. So there exist such states σ^{Φ} and $\sigma^{\Psi} \in \mathcal{B}$ that $\mathrm{Tr}A\sigma^{\Phi} \geq \alpha$ and $\chi_{\Phi}(\sigma^{\Phi}) + \chi_{\Psi}(\sigma^{\Psi}) \geq \chi_{\Phi \otimes \Psi}(\sigma_*)$. Hence

$$\max_{\sigma} [\chi_{\Phi \otimes \Psi}(\sigma) + p\mathrm{Tr}(A \otimes I_{\mathcal{K}}) \sigma] = \chi_{\Phi \otimes \Psi}(\sigma_*) + p\mathrm{Tr}(A \otimes I_{\mathcal{K}}) \sigma_*$$
$$\leq \chi_{\Phi}(\sigma^{\Phi}) + \chi_{\Psi}(\sigma^{\Psi}) + p\mathrm{Tr}A\sigma^{\Phi} \leq \max_{\rho} [\chi_{\Phi}(\rho) + p\mathrm{Tr}A\rho] + \bar{C}(\Psi; \mathcal{B})$$

in contradiction with (18).

The proof of (ii) $\Rightarrow$ (iii) consists of two steps. First we will prove the statement (iii) with the set $\mathcal{A}$ defined by the system of n inequalities $\mathrm{Tr}A_k\rho \leq \alpha_k$, $k = \overline{1, n}$ with arbitrary arrays of operators $\{A_k\}_{k=1}^n$ and numbers $\{\alpha_k\}_{k=1}^n$ such that $0 \leq A_k \leq I$ and $0 \leq \alpha_k \leq 1$ for $k = \overline{1, n}$. Then we will pass to an arbitrary set $\mathcal{A}$.

Let $\mathcal{A}$ be a set of the above special type. Suppose that the interior of the set $\mathcal{A}$ is nonempty. It implies that there exist such a state ρ that $\mathrm{Tr}A_k\rho < \alpha_k, k = \overline{1, n}$. It is sufficient to show that

$$\bar{C} (\Phi \otimes \Psi; \mathcal{A} \otimes \mathcal{B}) \leq \bar{C}(\Phi; \mathcal{A}) + \bar{C}(\Psi; \mathcal{B}). \quad (19)$$

Suppose, " $>$ " takes place in (19). Then, there exists an ensemble $\{\mu_i, \sigma_i\}$ in $\mathfrak{S}(\mathcal{H} \otimes \mathcal{K})$ with the average state σ_{av}, such that $\mathrm{Tr}A_k\sigma_{\mathrm{av}}^{\Phi} \leq \alpha_k, k = \overline{1, n}$, $\sigma_{\mathrm{av}}^{\Psi} \in \mathcal{B}$ and

$$\chi_{\Phi \otimes \Psi}(\{\mu_i, \sigma_i\}) > \bar{C}(\Phi; \mathcal{A}) + \bar{C}(\Psi; \mathcal{B}). \quad (20)$$

Let ρ_{av} be the average of the optimal ensemble for the $\mathcal{A}$-constrained channel Φ so that $\bar{C}(\Phi; \mathcal{A}) = \chi_{\Phi}(\rho_{\mathrm{av}})$.

Note, that the state ρ_{av} is the point of maximum of the concave function $\chi_{\Phi}(\rho)$ with the constraints $\mathrm{Tr}A_k\rho \leq \alpha_k, k = \overline{1, n}$. By the Kuhn-Tucker

theorem [7][1] there exists a set of nonnegative numbers $\{p_k\}_{k=1}^n$, such that ρ_{av} is the point of the global maximum of the function $\chi_\Phi(\rho) - \sum_{k=1}^n p_k \mathrm{Tr} A_k \rho$ and the following conditions hold

$$p_k(\mathrm{Tr} A_k \rho_{\mathrm{av}} - \alpha_k) = 0; \quad k = \overline{1, n}. \tag{21}$$

It is clear that ρ_{av} is also the point of the global maximum of the concave function $\chi_\Phi(\rho) + \sum_{k=1}^n p_k \mathrm{Tr} \bar{A}_k \rho$, so that

$$\chi_\Phi(\rho) + \sum_{k=1}^n p_k \mathrm{Tr} \bar{A}_k \rho \leq \chi_\Phi(\rho_{\mathrm{av}}) + \sum_{k=1}^n p_k \mathrm{Tr} \bar{A}_k \rho_{\mathrm{av}}, \quad \forall \rho \in \mathfrak{S}(\mathcal{H}). \tag{22}$$

Let $p = \|\sum_{i=1}^n p_i A_i\|$. The case $p = 0$ means that ρ_{av} is the point of the global maximum of $\chi_\Phi(\rho)$. So in this case the channel Φ is in fact unconstrained. But the additivity for the unconstrained channel Φ and the $\mathcal{B}$ -constrained channel Ψ easily follows from the asymptotic additivity for the sequence $\widehat{\Phi}_d(A, 0)$ (with arbitrary operator A) and the $\mathcal{B}$-constrained channel Ψ, because each channel in this sequence is equivalent to Φ.

In the case $p > 0$ let $A = p^{-1} \sum_{i=1}^n p_i A_i$. Note that $0 \leq A \leq I$. Consider the sequence $\widehat{\Phi}_d(A, p)$ of Shor's extensions of the channel Φ. Assumed asymptotic additivity together with (16) and (17) implies

$$\max_\sigma \left[\chi_{\Phi \otimes \Psi}(\sigma) + p \mathrm{Tr}(A \otimes I_\mathcal{K}) \sigma \right] = \max_\rho \left[\chi_\Phi(\rho) + p \mathrm{Tr} A \rho \right] + \bar{C}(\Psi; \mathcal{B}). \tag{23}$$

Due to (21) and (22) we have

$$\max_\rho \left[\chi_\Phi(\rho) + p \mathrm{Tr} A \rho \right] = \max_\rho \left[\chi_\Phi(\rho) + \sum_{k=1}^n p_k \ \mathrm{Tr} \bar{A}_k \rho \right]$$
$$= \chi_\Phi(\rho_{\mathrm{av}}) + \sum_{k=1}^n p_k \mathrm{Tr} \bar{A}_k \rho_{\mathrm{av}} = \bar{C}(\Phi; \mathcal{A}) + \sum_{k=1}^n p_k(1 - \alpha_k). \tag{24}$$

Noting that

$$\mathrm{Tr}(A_k \otimes I_\mathcal{K}) \sigma_{\mathrm{av}} = \mathrm{Tr} A_k \sigma_{\mathrm{av}}^\Phi \leq \alpha_k, \quad k = \overline{1, n},$$

we have by (20)

$$\max_\sigma \left[\chi_{\Phi \otimes \Psi}(\sigma) + p \mathrm{Tr}(A \otimes I_\mathcal{K}) \sigma \right] \geq \chi_{\Phi \otimes \Psi}(\sigma_{\mathrm{av}}) + p \mathrm{Tr}(A \otimes I_\mathcal{K}) \sigma_{\mathrm{av}}$$
$$= \chi_{\Phi \otimes \Psi}(\sigma_{\mathrm{av}}) + \sum_{k=1}^n p_k \ \mathrm{Tr}(\bar{A}_k \otimes I_\mathcal{K}) \sigma_{\mathrm{av}} > \bar{C}(\Phi; \mathcal{A}) + \bar{C}(\Psi; \mathcal{B}) + \sum_{k=1}^n p_k(1 - \alpha_k).$$

[1]We use the strong version of this theorem with the Slater condition, which follows from the assumption that the interior of the set $\mathcal{A}$ is nonempty.

The contradiction of the last inequality with (23) and (24) completes the first step of the proof of (ii) $\Rightarrow$ (iii) under the condition that the interior of the set $\mathcal{A}$ is nonempty.

The case where the set $\mathcal{A}$ has no inner point may be reduced to the previous one with the help of the following lemma.

Lemma 1. Let $\{\mathcal{A}_n\}_{n \in \mathbb{N}}$ be a decreasing sequence of closed subsets of $\mathfrak{S}(\mathcal{H})$ and $\mathcal{A} = \bigcap_{n \in \mathbb{N}} \mathcal{A}_n$. If the additivity conjecture is true for the $\mathcal{A}_n$-constrained channel Φ and the $\mathcal{B}$-constrained channel Ψ for all $n \in \mathbb{N}$ then this conjecture is true for the $\mathcal{A}$-constrained channel Φ and the $\mathcal{B}$-constrained channel Ψ as well.

Proof. It is sufficient to prove that

$$\bar{C}(\Phi, \mathcal{A}) = \lim_{n \to +\infty} \bar{C}(\Phi, \mathcal{A}_n), \quad \bar{C}(\Phi \otimes \Psi, \mathcal{A} \otimes \mathcal{B}) = \lim_{n \to +\infty} \bar{C}(\Phi \otimes \Psi, \mathcal{A}_n \otimes \mathcal{B}).$$

Due to $\mathcal{A} \otimes \mathcal{B} = \bigcap_{n \in \mathbb{N}} \mathcal{A}_n \otimes \mathcal{B}$ the second equality is equivalent to the first.

The nonnegative sequence $\bar{C}(\Phi, \mathcal{A}_n)$ is decreasing, so the first of the above limits exists and " $\leq$ " is obvious. For each $n \in \mathbb{N}$ let ρ_n be the maximum point of the function $\chi_\Phi(\rho)$ with the constraint $\rho \in \mathcal{A}_n$. By compactness argument we can assume that there exists $\lim_{n \to +\infty} \rho_n = \rho_* \in \mathcal{A}$. The continuity property of the function $\chi_\Phi(\rho)$ gives $\lim_{n \to +\infty} \chi_\Phi(\rho_n) = \chi_\Phi(\rho_*)$, which proves the " $=$ ".$\triangle$

Consider the sequence $\{\mathcal{A}_m\}$ of the subsets in $\mathfrak{S}(\mathcal{H})$, in which subset $\mathcal{A}_m$ is defined by the system of inequalities $\mathrm{Tr} A_k \rho \leq \alpha_k + 1/m$, $k = \overline{1, n}$. The interior of each $\mathcal{A}_m$ is nonempty. By the previous consideration the additivity conjecture is true for the $\mathcal{A}_m$ -constrained channel Φ and the $\mathcal{B}$-constrained channel Ψ . The lemma 1 gives desired additivity for the $\mathcal{A}$-constrained channel Φ and the $\mathcal{B}$-constrained channel Ψ.

Now the proof of the additivity for arbitrary set $\mathcal{A}$ is very simple. Note that any state ρ in $\mathfrak{S}(\mathcal{H})$ may be considered as a set $\mathcal{A}_\rho = \{\rho\}$, defined by a finite number of linear inequalities. The result of the first step implies

$$\bar{C}\left(\Phi \otimes \Psi; \mathcal{A}_\rho \otimes \mathcal{B}\right) = \bar{C}(\Phi; \mathcal{A}_\rho) + \bar{C}(\Psi; \mathcal{B}) = \chi_\Phi(\rho) + \bar{C}(\Psi; \mathcal{B}). \qquad (25)$$

Suppose that the additivity conjecture is not true for the $\mathcal{A}$ -constrained channel Φ and the $\mathcal{B}$-constrained channel Ψ. Then there exists such a state $\sigma \in \mathfrak{S}(\mathcal{H} \otimes \mathcal{K})$ that $\sigma^\Phi \in \mathcal{A}$, $\sigma^\Psi \in \mathcal{B}$ and

$$\chi_{\Phi \otimes \Psi}(\sigma) > \bar{C}(\Phi; \mathcal{A}) + \bar{C}(\Psi; \mathcal{B}) \geq \chi_\Phi(\sigma^\Phi) + \bar{C}(\Psi; \mathcal{B}).$$

This inequality contradicts to (25) with $\rho = \sigma^\Phi$. The proof of (ii) $\Rightarrow$ (iii) is complete. $\triangle$

The above theorem implies two sorts of results. The equivalence (ii) $\Leftrightarrow$ (iii) implies the corollaries 1-3. The equivalence (i) $\Leftrightarrow$ (iii) gives the corollary 4.

For the verification of the constrained additivity the following sufficient condition will be convenient (see the proof of the proposition 4(B)).

Corollary 1. If the additivity conjecture holds true for the unconstrained channels $\widehat{\Phi}_d(A, p)$ with arbitrary pair (A, p) and the $\mathcal{B}$-constrained channel Ψ for all sufficiently large d, then the additivity conjecture is true for the $\mathcal{A}$ -constrained channel Φ and the $\mathcal{B}$-constrained channel Ψ with arbitrary $\mathcal{A} \subset \mathfrak{S}(\mathcal{H})$.

Corollary 2. The additivity for the Shor's channel extensions $\widehat{\Phi}_d(A, p)$ and $\widehat{\Psi}_e(B, r)$ with arbitrary pairs (A, p) and (B, r) for all sufficiently large d and e implies the additivity for the $\mathcal{A}$-constrained channel Φ and the $\mathcal{B}$-constrained channel Ψ with arbitrary $\mathcal{A} \subset \mathfrak{S}(\mathcal{H})$ and $\mathcal{B} \subset \mathfrak{S}(\mathcal{K})$.

Proof. This is obtained by double application of the corollary 1. $\triangle$

Corollary 3. If the additivity conjecture for *any* two unconstrained channels holds true then it holds for any two channels with arbitrary constraints.

Proof. This follows from corollary 2. $\triangle$

Corollary 4. If the additivity conjecture holds true for the channels Φ and Ψ with the *single* linear constraints $\mathrm{Tr}\rho A \leq \alpha$ and $\mathrm{Tr}\varrho B \leq \beta$ correspondingly then the additivity conjecture holds true for the $\mathcal{A}$-constrained channel Φ and the $\mathcal{B}$-constrained channel Ψ with *arbitrary* $\mathcal{A} \subset \mathfrak{S}(\mathcal{H})$ and $\mathcal{B} \subset \mathfrak{S}(\mathcal{K})$.

Proof. Double application of the equivalence (i) $\Leftrightarrow$ (iii) in the above theorem. $\triangle$

4 Subadditivity property of the χ-function

The additivity of the Holevo capacity for the channels Φ and Ψ with arbitrary constraints implies the following subadditivity property of the χ-function:

$$\chi_{\Phi \otimes \Psi}(\sigma) \le \chi_{\Phi}(\sigma^{\Phi}) + \chi_{\Psi}(\sigma^{\Psi}), \quad \forall \sigma \in \mathfrak{S}(\mathcal{H} \otimes \mathcal{K}). \qquad (26)$$

To see this it is sufficient to take $\mathcal{A} = \{\sigma^{\Phi}\}$, $\mathcal{B} = \{\sigma^{\Psi}\}$ and note that

$$\bar{C}(\Phi \otimes \Psi; \{\sigma^{\Phi}\} \otimes \{\sigma^{\Psi}\}) \ge \chi_{\Phi \otimes \Psi}(\sigma),$$

$$\bar{C}(\Phi; \{\sigma^{\Phi}\}) = \chi_{\Phi}(\sigma^{\Phi}), \quad \bar{C}(\Psi; \{\sigma^{\Psi}\}) = \chi_{\Psi}(\sigma^{\Psi}).$$

The subadditivity of the function $\chi_{\Phi \otimes \Psi}(\sigma)$ implies its additivity:

$$\chi_{\Phi \otimes \Psi}(\sigma^{\Phi} \otimes \sigma^{\Psi}) = \chi_{\Phi}(\sigma^{\Phi}) + \chi_{\Psi}(\sigma^{\Psi}), \quad \forall \sigma^{\Phi} \in \mathfrak{S}(\mathcal{H}), \; \forall \sigma^{\Psi} \in \mathfrak{S}(\mathcal{K}),$$

which is not obvious as well.

Subadditivity of the function $\chi_{\Phi \otimes \Psi}(\sigma)$ obviously implies the additivity of the Holevo capacity for these channels. On the other hand, it is equivalent to the additivity conjecture for the channels Φ and Ψ with arbitrary constraints, and, hence, by corollary 4, with arbitrary single linear constraints. Corollary 3 shows that global subadditivity of the χ-function is equivalent to the global additivity of the Holevo capacity for unconstrained channels. However, it is not clear whether (26) is implied by the additivity of the Holevo capacity for the channels Φ and Ψ.

There are several cases, where (26) can be indeed established.

Proposition 4. Let Ψ be arbitrary channel. The inequality (26) holds in each of the following cases:

(A) Φ is an entanglement breaking channel;

(B) Φ is a noiseless channel;

(C) Φ is a direct sum mixture of a noiseless channel Id and a channel Φ_0 such that the function $\chi_{\Phi_0 \otimes \Psi}$ is subadditive.

Proof. (A) Shor in [16] proved the additivity conjecture for two unconstrained channel if one of them is an entanglement breaking. But in the

proof of this theorem the subadditivity property of the χ -function was in fact established. It is interesting that in this case we can directly deduce the subadditivity of the χ-function from unconstrained additivity with the help of corollary 2. One should only verify that entanglement breaking property of any channel implies similar property of Shor's extension for this channel.

(B) The proof of this statement consists of two steps. First we will prove the additivity conjecture for two channels if one of them is noiseless and $\{\rho\}$-constrained while the other channel is arbitrary and unconstrained. Then we will apply the Shor's channel extension to pass to the $\{\varrho\}$-constrained second channel.

The proof of the first step is the modification of the proof in [5] of the "unconstrained" additivity for two channels if one of them is noiseless, which is based on the Groenevold-Lindblad-Ozawa inequality [11]

$$H(\sigma) \leq \sum_j p_j H(\sigma_j), \tag{27}$$

where σ is a state of a quantum system before measurement, σ_j is the state of this system after measurement with yield j and p_j is the probability of this yield.

Let $\Phi = \mathrm{Id}$ and ρ be an arbitrary state in $\mathfrak{S}(\mathcal{H})$. We want to prove that

$$\bar{C}(\mathrm{Id} \otimes \Psi, \{\rho\} \otimes \mathfrak{S}(\mathcal{K})) = \bar{C}(\mathrm{Id}, \{\rho\}) + \bar{C}(\Psi, \mathfrak{S}(\mathcal{K})) = H(\rho) + \bar{C}(\Psi) \tag{28}$$

Let $\{\mu_i, \sigma_i\}$ be an ensemble of states in $\mathfrak{S}(\mathcal{H} \otimes \mathcal{K})$ with $\sum_i \mu_i \sigma_i^\Phi = \rho$. By subadditivity of quantum entropy we obtain

$$\chi_{\mathrm{Id} \otimes \Psi}(\{\mu_i, \sigma_i\}) = H(\mathrm{Id} \otimes \Psi(\sum_i \mu_i \sigma_i)) - \sum_i \mu_i H(\mathrm{Id} \otimes \Psi(\sigma_i))$$
$$\leq H(\rho) + H(\Psi(\sum_i \mu_i \sigma_i^\Psi)) - \sum_i \mu_i H(\mathrm{Id} \otimes \Psi(\sigma_i)). \tag{29}$$

Consider the measurement, defined by the observable $\{|e_j\rangle\langle e_j| \otimes I_\mathcal{K}\}$, where $\{|e_j\rangle\}$ is an orthonormal basis in $\mathcal{H}$. By (27) we obtain

$$H(\mathrm{Id} \otimes \Psi(\sigma_i)) \leq \sum_j p_{ij} H(\Psi(\sigma_{ij}^\Psi)), \quad \forall i,$$

where $p_{ij} = \langle e_j|\sigma_i|e_j\rangle$ and $\sigma_{ij} = p_{ij}^{-1}|e_j\rangle\langle e_j| \otimes I_\mathcal{K} \cdot \sigma_i \cdot |e_j\rangle\langle e_j| \otimes I_\mathcal{K}$. Note that $\sum_j p_{ij}\sigma_{ij}^\Psi = \sigma_i^\Psi$. This and previous inequality show that two last terms in

15

(29) do not exceed $\chi_{\Psi}(\{\mu_i p_{ij}, \sigma_{ij}^{\Psi}\})$ and, hence, $\bar{C}(\Psi)$. With this observation (29) implies (28) and, hence, the first step of the proof is complete.

It follows that the additivity conjecture holds for the (unconstrained) Shor's extension of the channel Ψ (with arbitrary parameters) and the $\{\rho\}$-constrained channel $\Phi = \mathrm{Id}$. The corollary 1 gives the desired subadditivity property of the function $\chi_{\mathrm{Id} \otimes \Psi}$.

(C) Let $\Phi_q = q\mathrm{Id} \oplus (1-q)\Phi_0$. For an arbitrary channel Ψ we have $\Phi_q \otimes \Psi = q(\mathrm{Id} \otimes \Psi) \oplus (1-q)(\Phi_0 \otimes \Psi)$. By lemma 1 in [6] with the subadditivity of the functions $\chi_{\mathrm{Id} \otimes \Psi}$ and $\chi_{\Phi_0 \otimes \Psi}$ we obtain

$$\chi_{\Phi_q \otimes \Psi}(\sigma) \leq q\chi_{\mathrm{Id} \otimes \Psi}(\sigma) + (1-q)\chi_{\Phi_0 \otimes \Psi}(\sigma)$$

$$\leq q\chi_{\mathrm{Id}}(\sigma^{\Phi}) + q\chi_{\Psi}(\sigma^{\Psi}) + (1-q)\chi_{\Phi_0}(\sigma^{\Phi}) + (1-q)\chi_{\Psi}(\sigma^{\Psi})$$

$$= qH(\sigma^{\Phi}) + (1-q)\chi_{\Phi_0}(\sigma^{\Phi}) + \chi_{\Psi}(\sigma^{\Psi}) = \chi_{\Phi_q}(\sigma^{\Phi}) + \chi_{\Psi}(\sigma^{\Psi}),$$

where the last equality follows from the existence of a *pure* state ensemble on which the maximum in the definition of $\chi_{\Phi}(\sigma^{\Phi})$ is achieved.$\triangle$

Due to the MSW representation [8] and subadditivity of the quantum entropy the subadditivity of the χ-function easily follows from the strong superadditivity of the entanglement of formation. What is interesting that the converse is also true. The strong superadditivity of the entanglement of formation follows from the subadditivity of the χ -function for *any* two channels. Together with the corollary 3 it provides another way (as compared to [15],[12]) of proving that the global additivity conjecture for unconstrained channels (or additivity of the entanglement of formation) implies strong superadditivity of the entanglement of formation.

The above statement is a consequence of the following theorem.

Theorem 2. Let Φ and Ψ be fixed channels. For given arbitrary $\rho \in \mathfrak{S}(\mathcal{H})$ and $\varrho \in \mathfrak{S}(\mathcal{K})$ the additivity conjecture

$$\bar{C}(\Phi \otimes \Psi; \{\rho\} \otimes \{\varrho\}) = \bar{C}(\Phi; \{\rho\}) + \bar{C}(\Psi; \{\varrho\})$$

holds if and only if

$$\min \sum_{k} \mu_k H(\Phi \otimes \Psi(\sigma_k)) = \min_{\sum_i \pi_i \rho_i = \rho} \sum_i \pi_i H(\Phi(\rho_i)) + \min_{\sum_j \varpi_j \varrho_j = \varrho} \sum_j \varpi_j H(\Psi(\varrho_j)),$$

where the first minimum is over all ensembles $\{\mu_k, \sigma_k\}$ of states in $\mathfrak{S}(\mathcal{H} \otimes \mathcal{K})$ such that $\sum_k \mu_k \sigma_k^{\Phi} = \rho$ and $\sum_k \mu_k \sigma_k^{\Psi} = \varrho$.

Proof. The sufficiency of the above condition for the additivity of the Holevo capacity for the $\{\rho\}$-constrained channel Φ and the $\{\varrho\}$-constrained channel Ψ obviously follows from the subadditivity of quantum entropy.

Let us prove the necessity of this condition. The additivity conjecture for the $\{\rho\}$-constrained channel Φ and $\{\varrho\}$-constrained channel Ψ implies the existence of nonentangled ensemble with average state $\rho \otimes \varrho$, which is optimal for the $\{\rho\} \otimes \{\varrho\}$-constrained channel $\Phi \otimes \Psi$. By proposition 2 we have

$$\chi_{\Phi \otimes \Psi}(\rho \otimes \varrho) = \chi_{\Phi}(\rho) + \chi_{\Psi}(\varrho) \geq \chi_{\Phi \otimes \Psi}(\sigma) + S(\Phi \otimes \Psi(\sigma) \| \Phi(\rho) \otimes \Psi(\varrho)) \quad (30)$$

for any state σ in $\mathfrak{S}(\mathcal{H}) \otimes \mathfrak{S}(\mathcal{K})$ such that $\sigma^{\Phi} = \rho$ and $\sigma^{\Psi} = \varrho$. Note that

$$S(\Phi \otimes \Psi(\sigma) \| \Phi(\rho) \otimes \Psi(\varrho)) = H(\Phi(\rho)) + H(\Psi(\varrho)) - H(\Phi \otimes \Psi(\sigma)). \quad (31)$$

The inequality (30) with (31) and the definition of the χ-function provides " $\geq$ " in the condition of the theorem. Since " $\leq$ " in this condition is obvious, the proof is complete.$\triangle$

Corollary 5. The subadditivity (26) of the function $\chi_{\Phi \otimes \Psi}$ is equivalent to the following property:

$$\chi_{\Phi}(\sigma^{\Phi}) + \chi_{\Psi}(\sigma^{\Psi}) - \chi_{\Phi \otimes \Psi}(\sigma) \geq H(\Phi(\sigma^{\Phi})) + H(\Psi(\sigma^{\Psi})) - H(\Phi \otimes \Psi(\sigma)).$$

for all $\sigma \in \mathfrak{S}(\mathcal{H} \otimes \mathcal{K})$.

This means that the gap between $\chi_{\Phi}(\sigma^{\Phi}) + \chi_{\Psi}(\sigma^{\Psi})$ and $\chi_{\Phi \otimes \Psi}(\sigma)$ is no less than the gap between $H(\Phi(\sigma^{\Phi})) + H(\Psi(\sigma^{\Psi}))$ and $H(\Phi \otimes \Psi(\sigma))$.

Proof. The sufficiency of the above property for inequality (26) is clear.

To prove its necessity note that subadditivity of the function $\chi_{\Phi \otimes \Psi}$ implies additivity for the $\{\rho\}$-constrained channel Φ and $\{\varrho\}$-constrained channel Ψ with any $\rho \in \mathfrak{S}(\mathcal{H})$ and $\varrho \in \mathfrak{S}(\mathcal{K})$. By the definition of the χ-function the above property can be rewritten as

$$\min_{\sum_k \mu_k \sigma_k = \sigma} \sum_k \mu_k H(\Phi \otimes \Psi(\sigma_k)) \geq \min_{\sum_i \pi_i \rho_i = \rho} \sum_i \pi_i H(\Phi(\rho_i)) + \min_{\sum_j \varpi_j \varrho_j = \varrho} \sum_j \varpi_j H(\Psi(\varrho_j)).$$

with $\rho = \sigma^{\Phi}$ and $\varrho = \sigma^{\Psi}$. But this inequality follows from the theorem 2. $\triangle$

Corollary 6. The global subadditivity of the χ-function implies strong superadditivity of the entanglement of formation.

Proof. Let $\mathcal{H}_1, \mathcal{H}_2, \mathcal{K}_1, \mathcal{K}_2$ be finite dimensional Hilbert spaces and $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$, $\mathcal{K} = \mathcal{K}_1 \otimes \mathcal{K}_2$. Consider the channels $\Phi_i(\cdot) = \text{Tr}_{\mathcal{K}_i}(\cdot)$ from $\mathfrak{S}(\mathcal{H}_i \otimes \mathcal{K}_i)$ into $\mathfrak{S}(\mathcal{H}_i)$, $i = 1, 2$. It is clear that $\Phi_1 \otimes \Phi_2(\cdot) = \text{Tr}_{\mathcal{K}}(\cdot)$.

Let σ be an arbitrary state in $\mathfrak{S}(\mathcal{H}\otimes\mathcal{K})$, let σ_1 and σ_2 be the partial states of σ, corresponding to the decomposition $\mathcal{H}\otimes\mathcal{K} = (\mathcal{H}_1\otimes\mathcal{K}_1)\otimes(\mathcal{H}_2\otimes\mathcal{K}_2)$. By definition we have

$$\chi_{\Phi_1}(\sigma_1) = H(\mathrm{Tr}_{\mathcal{K}_1}(\sigma_1)) - E_F(\sigma_1), \quad \chi_{\Phi_2}(\sigma_2) = H(\mathrm{Tr}_{\mathcal{K}_2}(\sigma_2)) - E_F(\sigma_2),$$

$$\chi_{\Phi_1\otimes\Phi_2}(\sigma) = H(\mathrm{Tr}_{\mathcal{K}}(\sigma)) - E_F(\sigma).$$

By the corollary 5 with $\Phi = \Phi_1$ and $\Psi = \Phi_2$ we obtain the desired strong superadditivity of the entanglement of formation:

$$E_F(\sigma) \geq E_F(\sigma_1) + E_F(\sigma_2). \; \triangle$$

Note also the following lower bound for the Holevo capacity of the $\{\rho\}\otimes\{\varrho\}$-constrained channel $\Phi\otimes\Psi$.

Corollary 7. Let Φ and Ψ be an arbitrary channel. For given arbitrary states $\rho \in \mathfrak{S}(\mathcal{H})$ and $\varrho \in \mathfrak{S}(\mathcal{K})$

$$\bar{C}(\Phi\otimes\Psi; \{\rho\}\otimes\{\varrho\}) \geq \bar{C}(\Phi; \{\rho\}) + \bar{C}(\Psi; \{\varrho\}) + S(\Phi(\rho)\otimes\Psi(\varrho)\|\Phi\otimes\Psi(\sigma_{\mathrm{av}})),$$

where σ_{av} is the average state of *any* optimal ensemble for the $\{\rho\}\otimes\{\varrho\}$-constrained channel $\Phi\otimes\Psi$.

Proof. Direct application of the proposition 2.$\triangle$

5   Application of the convex duality

Corollary 5 of the theorem 2 and the arguments from the convex analysis raised in [2] provide the characterization of channels for which subadditivity of the χ-function holds. For a channel Φ and an operator $A \in \mathfrak{B}_+(\mathcal{H})$ we introduce the following modified output purity of the channel (cf. [1])

$$\nu_H\left(\Phi, A\right) = \min_{\rho \in \mathfrak{S}(\mathcal{H})} \left[H(\Phi(\rho)) + \mathrm{Tr} A\rho\right]. \tag{32}$$

We will show later that the additivity property for this value is equivalent to the subadditivity of the χ-function.

Let $\Phi : \mathfrak{S}(\mathcal{H}) \mapsto \mathfrak{S}(\mathcal{H}')$ be an arbitrary channel. For a given operator E in $\mathfrak{B}(\mathcal{H}), 0 \le E \le I_{\mathcal{H}}$, and numbers $q \in [0; 1]$, $d \in \mathbb{N} = \{1, 2, \dots\}$ consider the channel $\widetilde{\Phi}(E, q, d)$ from $\mathfrak{B}(\mathcal{H})$ into $\mathfrak{B}(\mathcal{H}') \oplus \mathbf{C}^{d+1}$, related to the Shor's channel extension $\widehat{\Phi}(E, q, d)$ by the following equality:

$$\widetilde{\Phi}(E, q, d)(\rho) = \widehat{\Phi}(E, q, d)(\rho \otimes \tau), \quad \forall \rho \in \mathfrak{S}(\mathcal{H}),$$

where τ is the chaotic classical state $[d^{-1}, ..., d^{-1}]$ in C^d . We have

$$\widetilde{\Phi}(E, q, d)(\rho) = (1 - q)\Phi(\rho) \oplus q\mathbf{E}(\rho),$$

where $\mathbf{E}(\rho) = [\mathrm{Tr}\rho\bar{E}, d^{-1}\mathrm{Tr}E\rho, ..., d^{-1}\mathrm{Tr}E\rho]$ is a channel from $\mathfrak{B}(\mathcal{H})$ into $\mathbf{C}^{d+1}$. The channel $\widetilde{\Phi}(E, q, d)$ was originally introduced by Shor in [15] with the aim to prove that additivity of the minimum output entropy for any pair of channels implies additivity of the entanglement of formation. We will show that subadditivity of the function $\chi_{\Phi \otimes \Psi}$ is closely connected to additivity of the minimum output entropy for the channels $\widetilde{\Phi}$ and $\widetilde{\Psi}$.

Definition 2. We say that additivity of the minimum output entropy holds asymptotically for the sequences of channels $\{\Phi_n : \mathfrak{S}(\mathcal{H}) \mapsto \mathfrak{S}(\mathcal{H}'_n)\}$ and $\{\Psi_m : \mathfrak{S}(\mathcal{K}) \mapsto \mathfrak{S}(\mathcal{K}'_m)\}$ if

$$\lim_{n,m \to +\infty} \min_{\sigma \in \mathfrak{S}(\mathcal{H} \otimes \mathcal{K})} H(\Phi_n \otimes \Psi_m(\sigma))$$

$$= \lim_{n \to +\infty} \min_{\rho \in \mathfrak{S}(\mathcal{H})} H(\Phi_n(\rho)) + \lim_{m \to +\infty} \min_{\varrho \in \mathfrak{S}(\mathcal{K})} H(\Psi_m(\varrho)),$$

assuming that the limits exist and are finite.

19

Note that if additivity of the minimum output entropy holds for the channels Φ_n and Ψ_m for all sufficiently large n and m it obviously holds asymptotically for the sequences $\{\Phi_n\}$ and $\{\Psi_m\}$.

As in the case of Shor's channel extension it is convenient to denote by $\widetilde{\Phi}_d(A, p)$ the channel $\widetilde{\Phi}(A, p/\log d, d)$.

Theorem 3. Let $\Phi : \mathfrak{S}(\mathcal{H}) \mapsto \mathfrak{S}(\mathcal{H}')$ and $\Psi : \mathfrak{S}(\mathcal{K}) \mapsto \mathfrak{S}(\mathcal{K}')$ be arbitrary fixed channels. The following statements are equivalent:

(i) The function $\chi_{\Phi \otimes \Psi}$ is subadditivite;

(ii) For all $A \in \mathfrak{B}_+(\mathcal{H})$ and $B \in \mathfrak{B}_+(\mathcal{K})$

$$\nu_H\left(\Phi \otimes \Psi, A \otimes I + I \otimes B\right) = \nu_H\left(\Phi, A\right) + \nu_H\left(\Psi, B\right);$$

(iii) The additivity of the minimum output entropy holds asymptotically for the sequences of channels $\{\widetilde{\Phi}_d(A, p)\}_{d \in \mathbb{N}}$ and $\{\widetilde{\Psi}_e(B, r)\}_{e \in \mathbb{N}}$ with arbitrary pairs (A, p) and (B, r).

Proof. (i) $\Leftrightarrow$ (ii) Note that the function

$$\hat{H}_\Phi(\rho) = \min_{\sum_i \pi_i \rho_i = \rho} \sum_i \pi_i H(\Phi(\rho_i))$$

can be considered as the convex closure of the function $H_\Phi(\rho) = H(\Phi(\rho))$, defined on the set of pure states [2], [7]. The conjugate function is defined on the set $\mathfrak{B}_h(\mathcal{H})$ of all hermitian operators by

$$H_\Phi^*(\rho)(X) = \max_\rho \left[\text{Tr} X\rho - H_\Phi(\rho)\right].$$

The inequality in the corollary 5 can be regarded as the strong superadditivity of the function $\hat{H}_\Phi(\rho)$. By lemma 1 in [2] this superadditivity is equivalent to the subadditivity of the conjugate function H_Φ^* with respect to the Kronecker sum:

$$H_\Phi^*(A \otimes I_\mathcal{K} + I_\mathcal{H} \otimes B) \leq H_\Phi^*(A) + H_\Phi^*(B), \quad \forall A \in \mathfrak{B}_h(\mathcal{H}), \; \forall B \in \mathfrak{B}_h(\mathcal{K}).$$

By the definition of H_Φ^* the last inequality is equivalent to

$$\max_{\sigma \in \mathfrak{S}(\mathcal{H} \otimes \mathcal{K})} \left[\text{Tr} A\sigma^\Phi + \text{Tr} B\sigma^\Psi - H(\Phi \otimes \Psi(\sigma))\right]$$

$$\leq \max_{\rho \in \mathfrak{S}(\mathcal{H})} \left[\text{Tr} A\rho - H(\Phi(\rho))\right] + \max_{\varrho \in \mathfrak{S}(\mathcal{K})} \left[\text{Tr} B\varrho - H(\Psi(\varrho))\right]$$

20

for all $A \in \mathfrak{B}_h(\mathcal{H})$ and $B \in \mathfrak{B}_h(\mathcal{K})$.

Noting that " $\geq$ " in the previous inequality is obvious and using the invariance of it after changing A and B on $A + \|A\|I_{\mathcal{H}}$ and $B + \|B\|I_{\mathcal{K}}$ correspondingly we obtain that (i) $\Leftrightarrow$ (ii).

(ii) $\Leftrightarrow$ (iii) It is necessary to obtain the expressions for $H(\widetilde{\Phi}_d(A, p)(\rho))$, $H(\widetilde{\Psi}_e(B, r)(\varrho))$ and $H(\widetilde{\Phi}_d(A, p) \otimes \widetilde{\Psi}_e(B, r)(\sigma))$ for arbitrary $\rho \in \mathfrak{S}(\mathcal{H})$, $\varrho \in \mathfrak{S}(\mathcal{K})$ and $\sigma \in \mathfrak{S}(\mathcal{H} \otimes \mathcal{K})$. For this aim we will use the following simple lemma.

Lemma 2. Let $\{\Phi_j\}_{j=1}^n$ be a collection of channels from $\mathfrak{S}(\mathcal{H})$ into $\mathfrak{S}(\mathcal{H}'_j)$, $\{q_j\}_{j=1}^n$ be a probability distribution. Then for the channel $\Phi = \bigoplus_{j=1}^n q_j \Phi_j$ from $\mathfrak{S}(\mathcal{H})$ into $\mathfrak{S}(\bigoplus_{j=1}^n \mathcal{H}'_j)$ one has

$$H(\Phi(\omega)) = H(\{q_j\}) + \sum_{j=1}^n q_j H(\Phi_j(\omega)), \quad \forall \omega \in \mathfrak{S}(\mathcal{H}).$$

Applying this lemma to the channels $\widetilde{\Phi}_d(A, p) = (1 - q')\Phi + q'\mathbf{A}$ and $\widetilde{\Psi}_e(B, r) = (1 - q'')\Psi + q''\mathbf{B}$ (where $q' = p/\log d$ and $q'' = r/\log e$) we obtain

$$
\begin{aligned}
H(\widetilde{\Phi}_d(A, p)(\rho)) &= h_2(q') + (1 - q')H(\Phi(\rho)) + p\mathrm{Tr}A\rho + q'h_2(\mathrm{Tr}A\rho), \\
H(\widetilde{\Psi}_e(B, r)(\varrho)) &= h_2(q'') + (1 - q'')H(\Psi(\varrho)) + r\mathrm{Tr}B\varrho + q''h_2(\mathrm{Tr}B\varrho).
\end{aligned}
$$

The above expressions imply existence of the limits:

$$\lim_{d \to +\infty} \min_{\rho \in \mathfrak{S}(\mathcal{H})} H(\widetilde{\Phi}_d(A, p)(\rho)) = \min_{\rho \in \mathfrak{S}(\mathcal{H})} [H(\Phi(\rho)) + p\mathrm{Tr}A\rho], \quad (33)$$

$$\lim_{e \to +\infty} \min_{\varrho \in \mathfrak{S}(\mathcal{K})} H(\widetilde{\Psi}_e(B, r)(\varrho)) = \min_{\varrho \in \mathfrak{S}(\mathcal{K})} [H(\Psi(\varrho)) + r\mathrm{Tr}B\varrho]. \quad (34)$$

Let us prove that

$$\lim_{d,e \to +\infty} \min_{\sigma \in \mathfrak{S}(\mathcal{H} \otimes \mathcal{K})} H(\widetilde{\Phi}_d(A, p) \otimes \widetilde{\Psi}_e(B, r)(\sigma))$$

$$= \min_{\sigma \in \mathfrak{S}(\mathcal{H} \otimes \mathcal{K})} \left[H(\Phi \otimes \Psi(\sigma)) + \mathrm{Tr}A\sigma^\Phi + \mathrm{Tr}B\sigma^\Psi \right]. \quad (35)$$

Due to the representation

$$\widetilde{\Phi}_d(A, p) \otimes \widetilde{\Psi}_e(B, r) = (1 - q')(1 - q'')\Phi \otimes \Psi$$

$$+ q'(1 - q'')\mathbf{A} \otimes \Psi + (1 - q')q''\Phi \otimes \mathbf{B} + q'q''\mathbf{A} \otimes \mathbf{B} \quad (36)$$

21

the lemma 2 reduces the calculation of the value $H(\widetilde{\Phi}_d(A,p) \otimes \widetilde{\Psi}_e(B,r)(\sigma))$ to the calculation of four entropies, the first of which is $H(\Phi \otimes \Psi(\sigma))$, while the fourth does not exceed $\log(d+1) + \log(e+1)$. To calculate the others note that

$$\mathbf{A} \otimes \Psi(\sigma) = [\Psi_{\bar{A}}(\sigma), d^{-1}\Psi_A(\sigma), ..., d^{-1}\Psi_A(\sigma)],$$

where $\Psi_X(\cdot) = \mathrm{Tr}_{\mathcal{H}}(X \otimes I_{\mathcal{K}} \cdot (\mathrm{Id} \otimes \Psi)(\cdot))$ is a completely positive map from $\mathfrak{S}(\mathcal{H} \otimes \mathcal{K})$ into $\mathfrak{B}_+(\mathcal{K}')$ ($X = A, \bar{A}$ and Id is an identity map on $\mathfrak{S}(\mathcal{H})$). Therefore,

$$H(\mathbf{A} \otimes \Psi(\sigma)) = H(\Psi_{\bar{A}}(\sigma)) + H(\Psi_A(\sigma)) + \log d \mathrm{Tr}\Psi_A(\sigma)$$

$$= F(\sigma|A, \Psi) + \log d \mathrm{Tr} A\sigma^{\Phi}, \tag{37}$$

where $F(\sigma|A, \Psi)$ does not depend on d and does not exceed $2 \log \dim \mathcal{K}'$. Similarly,

$$H(\Phi \otimes \mathbf{B}(\sigma)) = G(\sigma|B, \Phi) + \log e \mathrm{Tr} B\sigma^{\Psi}, \tag{38}$$

where $G(\sigma|B, \Phi)$ does not depend on e and does not exceed $2 \log \dim \mathcal{H}'$.

The lemma 2 and the representation (36) imply

$$H(\widetilde{\Phi}_d(A,p) \otimes \widetilde{\Psi}_e(B,r)(\sigma)) = H(\{q'q'', q'(1-q''), q''(1-q'), (1-q')(1-q'')\})$$

$$+(1-q')(1-q'')H(\Phi \otimes \Psi(\sigma)) + (1-q'')p\mathrm{Tr}A\sigma^{\Phi} + (1-q')r\mathrm{Tr}B\sigma^{\Psi}$$

$$+(1-q'')q'F(\sigma|A, \Psi) + (1-q')q''G(\sigma|B, \Phi) + q'q''H(\mathbf{A} \otimes \mathbf{B}(\sigma)).$$

The estimates for F, G and

$$q'q''H(\mathbf{A} \otimes \mathbf{B}(\sigma)) \leq \frac{pr(\log(d+1) + \log(e+1))}{\log d \log e}.$$

imply that the right side of the above expression tends to

$$H(\Phi \otimes \Psi(\sigma)) + \mathrm{Tr}A\sigma^{\Phi} + \mathrm{Tr}B\sigma^{\Psi}$$

uniformly in σ when d and e tend to infinity. This proves (35).

The expressions (33),(34) and (35) show that the statement (iii) is equivalent to the equality

$$\min_{\sigma \in \mathfrak{S}(\mathcal{H} \otimes \mathcal{K})} \left[H(\Phi \otimes \Psi(\sigma)) + p\mathrm{Tr}A\sigma^{\Phi} + r\mathrm{Tr}B\sigma^{\Psi} \right]$$

$$= \min_{\rho \in \mathfrak{S}(\mathcal{H})} \left[H(\Phi(\rho)) + p\mathrm{Tr}A\rho \right] + \min_{\varrho \in \mathfrak{S}(\mathcal{K})} \left[H(\Psi(\varrho)) + r\mathrm{Tr}B\varrho \right]$$

for arbitrary nonnegative numbers p, r and operators A, B such that $0 \le A \le I_{\mathcal{H}}$, $0 \le B \le I_{\mathcal{K}}$. But this is exactly the statement (ii). $\triangle$

The concavity argument shows that the minimum in (32) is achieved on a pure state. So, the statement (ii) of the above theorem characterizes the subadditivity of the χ-function for the particular channels in terms of action of these channels and its tensor product on pure states.

Note also that (ii) implies additivity of the minimum output entropy (the case $A = B = 0$).

Corollary 8. Additivity of the minimum output entropy for the channels $\widetilde{\Phi}_d(A, p)$ and $\widetilde{\Psi}_e(B, r)$ with arbitrary pairs (A, p) and (B, r) for all sufficiently large d and e implies subadditivity of the function $\chi_{\Phi \otimes \Psi}$.

Corollary 9. Global additivity of the minimum output entropy is equivalent to the global subadditivity of the χ-function.

Proof. By the corollary in previous section, subadditivity of the function $\chi_{\Phi \otimes \Psi}$ implies additivity of the minimum output entropy for the channels Φ and Ψ.

If additivity of the minimum output entropy holds for any pair of channels then, by corollary 8, the χ-function is globally subadditive. $\triangle$

Corollary 10. The asymptotic additivity of the minimum output entropy for the sequences of channels $\{\widetilde{\Phi}_d(A, p)\}_{d \in \mathbb{N}}$ and $\{\widetilde{\Psi}_e(B, r)\}_{e \in \mathbb{N}}$ with arbitrary pairs (A, p) and (B, r) is equivalent to the asymptotic additivity of the Holevo capacity for the sequences of channels $\{\widehat{\Phi}_d(A, p)\}_{d \in \mathbb{N}}$ and $\{\widehat{\Psi}_e(B, r)\}_{e \in \mathbb{N}}$ with arbitrary pairs (A, p) and (B, r).

Proof. By theorems 1 and 3 both properties are equivalent to subadditivity of the function $\chi_{\Phi \otimes \Psi}$. $\triangle$

Remark. The proof of theorem 3 can be modified to obtain a direct proof of the fact that additivity of the minimum output entropy for any pair of channels implies the strong subadditivity of the entanglement of formation (instead of simple additivity as in [15]). To see this consider channels Φ and Ψ of the form of the partial traces, as in the proof of the corollary 6. The arguments in the proof of theorem 3 shows that assumed additivity of the minimum output entropy for the channels $\widetilde{\Phi}_d(A, p)$ and $\widetilde{\Psi}_e(B, r)$ implies superadditivity of the function $H_{\Phi \otimes \Psi}(\sigma)$, which in this case coincides with the entanglement of formation $E_F(\sigma)$.

6   Some relations for the relative entropy

The proofs of the theorem 2 and corollary 7 were based on proposition 2. In this section we provide direct proofs of them considering some interesting relations for the relative entropy.

Let σ and ς be arbitrary states such that $\mathrm{supp}\,\sigma \subseteq \mathrm{supp}\,\varsigma$. The functions $f(x) = S(x\sigma + (1-x)\varsigma\|\varsigma)$ and $g(x) = S(\varsigma\|x\sigma + (1-x)\varsigma)$ are obviously continuous and convex on $[0;1]^2$.

Proposition 5. These functions are related by the following transformations:

$$g(x) = xf'(x) - f(x), \quad f(x) = x\int\limits_0^x \frac{g(t)}{t^2}dt, \qquad \forall x \in [0;1]$$

with the "initial conditions" $f(0) = g(0) = 0$ and

$$f'(0) = \frac{d}{dx}S(x\sigma + (1-x)\varsigma\|\varsigma)|_{x=0} = 0. \tag{39}$$

The proof of this proposition and some features of the above transformations can be found in the Appendix.

Note that in the case $\mathrm{supp}\,\sigma = \mathrm{supp}\,\varsigma$ proposition 5 implies

$$f'(1) = \frac{d}{dx}S(x\sigma + (1-x)\varsigma\|\varsigma)|_{x=1} = S(\sigma\|\varsigma) + S(\varsigma\|\sigma). \tag{40}$$

We will use proposition 5 with $\Phi \otimes \Psi(\sigma)$ and $\Phi(\sigma^\Phi) \otimes \Psi(\sigma^\Psi)$ in the role of σ and ς correspondingly. Taking into account the definition of the partial trace one can verify that $\mathrm{supp}\,\Phi \otimes \Psi(\sigma) \subseteq \mathrm{supp}\,\Phi(\sigma^\Phi) \otimes \Psi(\sigma^\Psi)$ $(\forall \sigma)$. In this case

$$f(x) = H(\Phi(\sigma^\Phi)) + H(\Psi(\sigma^\Psi)) - H(x\Phi \otimes \Psi(\sigma) + (1-x)\Phi(\sigma^\Phi) \otimes \Psi(\sigma^\Psi)).$$

The equality (39) may be used for proving theorem 2 while (40) provides the statement of the corollary 7.

Let $\{\pi_i, \rho_i\}$ and $\{\varpi_j, \varrho_j\}$ be the optimal ensembles for the $\{\rho\}$-constrained channel Φ and the $\{\varrho\}$-constrained channel Ψ (with the average states ρ and ϱ) correspondingly. Consider the mixture of an arbitrary ensemble $\{\mu_k, \sigma_k\}$

---

[2]The symbol } means) if $\mathrm{supp}\,\sigma \subset \mathrm{supp}\,\varsigma$ and] if $\mathrm{supp}\,\varsigma = \mathrm{supp}\,\sigma$

(with the average state σ_{av}) and the ensemble $\{\pi_i \varpi_j, \rho_i \otimes \varrho_j\}$ with the weights x and $1 - x$ correspondingly. This new ensemble has the average state $x\sigma_{\text{av}} + (1 - x)\rho \otimes \varrho \in \{\rho\} \otimes \{\varrho\}$ and the Holevo quantity

$$\chi^x_{\Phi \otimes \Psi} = H(\Phi \otimes \Psi(x\sigma_{\text{av}} + (1 - x)\rho \otimes \varrho)) - x \sum_k \mu_k H(\Phi \otimes \Psi(\sigma_k))$$

$$-(1 - x)\left(\sum_i \pi_i H(\Phi(\rho_i)) + \sum_j \varpi_j H(\Psi(\varrho_j)) \right)$$

$$= \chi_\Phi(\rho) + \chi_\Psi(\varrho) + H(\Phi \otimes \Psi(x\sigma_{\text{av}} + (1 - x)\rho \otimes \varrho)) - H(\Phi(\rho)) - H(\Psi(\varrho))$$

$$+x\left(\sum_i \pi_i H(\Phi(\rho_i)) + \sum_j \varpi_j H(\Psi(\varrho_j) - \sum_k \mu_k H(\Phi \otimes \Psi(\sigma_k)) \right)$$

Denoting the quantity in the last brackets by Δ we obtain

$$h(x) = \chi^x_{\Phi \otimes \Psi} - \chi^0_{\Phi \otimes \Psi} = x\Delta - f(x), \tag{41}$$

where $f(x) = H(\Phi(\rho)) + H(\Psi(\varrho)) - H(\Phi \otimes \Psi(x\sigma_{\text{av}} + (1-x)\rho \otimes \varrho))$ is concave function with $f'(0) = 0$ by proposition 5. If the value Δ is positive, then, due to $f'(0) = 0$, we will necessarily have $\chi^x_{\Phi \otimes \Psi} > \chi^0_{\Phi \otimes \Psi}$ for sufficiently small x . If the value Δ is not positive, then $\chi^x_{\Phi \otimes \Psi} < \chi^0_{\Phi \otimes \Psi}$ for all $x > 0$. This observation proves the theorem 2.

Consider the above construction in the case where $\{\mu_k, \sigma_k\}$ is an optimal ensemble for the $\{\rho\} \otimes \{\varrho\}$-constrained channel $\Phi \otimes \Psi$. In this case proposition 1 implies $\text{supp}\, \Phi \otimes \Psi(\sigma_{\text{av}}) = \text{supp}\, \Phi(\rho) \otimes \Phi(\varrho)$. Hence, by proposition 5, the function $f(x)$ is defined on $[0, 1]$ and

$$f'(1) = S(\Phi \otimes \Psi(\sigma_{\text{av}})\|\Phi(\rho) \otimes \Psi(\varrho)) + S(\Phi(\rho) \otimes \Psi(\varrho)\|\Phi \otimes \Psi(\sigma_{\text{av}})). \tag{42}$$

Due to optimality of the above ensemble the concave function $h(x)$, defined by (41), must be nondecreasing on $[0, 1]$ and hence $h'(x) = \Delta - f'(x) \geq 0$. By concavity, this implies $f'(1) \leq \Delta$. But in this case

$$\Delta = \sum_i \pi_i H(\Phi(\rho_i)) + \sum_j \varpi_j H(\Psi(\varrho_j) - \sum_k \mu_k H(\Phi \otimes \Psi(\sigma_k))$$

$$= \chi_{\Phi \otimes \Psi}(\sigma) - \chi_\Phi(\rho) - \chi_\Psi(\varrho) + S(\Phi \otimes \Psi(\sigma_{\text{av}})\|\Phi(\rho) \otimes \Psi(\varrho))$$

$$= \bar{C}(\Phi \otimes \Psi; \{\rho\} \otimes \{\varrho\}) - \bar{C}(\Phi; \{\rho\}) - \bar{C}(\Psi; \{\varrho\})$$

$$+ S(\Phi \otimes \Psi(\sigma_{\text{av}})\|\Phi(\rho) \otimes \Psi(\varrho)).$$

Taking into account (42) we obtain that the inequality $f'(1) \leq \Delta$ is equivalent to the inequality in the corollary 7.

7 Appendix

For proving the proposition 5 it is sufficient to show that $f'(0) = 0$ and $f'(x) = \langle f(x) + g(x) \rangle / x$ for $x > 0$. By definition

$$f'(x_0) = \frac{d}{dx} \left[\text{Tr}(x\sigma + (1-x)\varsigma) \log(x\sigma + (1-x)\varsigma) \right] |_{x=x_0}$$

$$- \frac{d}{dx} \left[\text{Tr}(x\sigma + (1-x)\varsigma) \log \varsigma \right] |_{x=x_0} = \text{Tr}(\sigma - \varsigma) \log(x\sigma + (1-x)\varsigma)$$

$$+ \frac{d}{dx} \left[\text{Tr}(x_0\sigma + (1-x_0)\varsigma) \log(x\sigma + (1-x)\varsigma) \right] |_{x=x_0} - \text{Tr}(\sigma - \varsigma) \log \varsigma \quad (43)$$

$$= \text{Tr}(\sigma - \varsigma)(\log(x\sigma + (1-x)\varsigma) - \log \varsigma),$$

where we use $\frac{d}{dx} \left[\text{Tr}(x_0\sigma + (1-x_0)\varsigma) \log(x\sigma + (1-x)\varsigma) \right] |_{x=x_0} = 0$ due to the fact that x_0 is the maximum point of the function

$$x \mapsto \text{Tr}(x_0\sigma + (1-x_0)\varsigma) \log(x\sigma + (1-x)\varsigma).$$

Expression (43) with $x_0 = 0$ gives $f'(0) = 0$. By the definition

$$f(x) + g(x) = x\text{Tr}(\sigma - \varsigma)(\log(x\sigma + (1-x)\varsigma) - \log \varsigma).$$

Comparing this with (43) completes the proof. $\triangle$

The transformation $f(x) \mapsto g(x) = xf'(x) - f(x)$ has simple geometric meaning and can be represented as $g(x) = f^*(f'(x))$, where f^* is the Legendre transform of f [7]. The eigenvectors of this transformation are powers x^α with corresponding eigenvalues $1 - \alpha$.

The inequality for relative entropy $S(\sigma' \| \sigma'') \geq \frac{1}{2} \| \sigma' - \sigma'' \|^2$ [10] shows that

$$f(x) \geq cx^2, \quad g(x) \geq cx^2, \quad c = \frac{1}{2} \| \sigma - \varsigma \|^2.$$

It is interesting to note that the "bound" cx^2 is (essentially unique) fixed point of the above transformations.

Proposition 5 implies that if $\text{supp}\sigma \subset \text{supp}\varsigma$ then

$$g(x) \to +\infty \quad \text{as} \quad x \to 1, \quad \text{but} \quad \int_0^1 g(x)dx < +\infty.$$

Note also the following relations for the derivatives

$$\frac{d^n}{dx^n}g(x)|_{x=0} = (n-1)\frac{d^n}{dx^n}f(x)|_{x=0}, \quad n \in \mathbb{N}.$$

It implies that $f(x) - g(x) = o(x^2)$ in a neighborhood of zero and its sign is defined by the sign of the third derivative at zero of $f(x)$.

References

[1] G.G.Amosov, A.S.Holevo, R.F.Werner, On some additivity problems in quantum information theory, Probl. Inf. Transm. 36, no.4, 25-34, 2000.

[2] K.M.R. Audenaert, S.L.Braunstein, On strong subadditivity of the entanglement of formation, e-print quant-ph/0303045, 2003.

[3] A. S. Holevo, Some estimates for information quantity transmitted by quantum communication channel, Probl. Peredachi Inform., vol.9, no. 3, pp. 177-183, 1973.

[4] A. S. Holevo, On quantum communication channels with constrained inputs, e-print quant-ph/9705054, 1997; Entanglement-assisted capacity of constrained quantum channels, e-print quant-ph/0211170.

[5] A. S. Holevo, Introduction to quantum information theory, Moscow Independent University, 2002.(in Russian)

[6] A. S. Holevo, M.E. Shirokov, On Shor's channel extension and constrained channels, e-print quant-ph/0306196, 2003;

[7] A. D. Joffe, B. M. Tikhomirov, Theory of extremum problems, Moscow: Nauka, 1974. (in Russian)

[8] K. Matsumoto, T. Shimono, A. Winter, Remarks on additivity of the Holevo channel capacity and of the entanglement of formation, e-print quant-ph/0206148, 2002.

[9] M.A. Nielsen, Continuity bounds for entangelment, Phys. Rev. A 61, N6, 064301, 2000.

[10] M.Ohya, D.Pets, Quantum entropy and its use, Springer, Berlin, 1993.

[11] M.Ozawa, On information gain by quantum measurement of continious observable, J.Math.Phys. V.27, P.759-763, 1986.

[12] A.A.Pomeransky, Strong superadditivity of the entanglement of formation follows from its additivity, e-print quant-ph/0305056, 2003.

[13] B. Schumacher, M.Westmoreland, Optimal signal ensemble, Phys.Re. A51, 2738, 1997.

[14] B. Schumacher, M.Westmoreland, Relative entropy in quantum information theory, e-print quant-ph/0004045, 2000.

[15] P. W. Shor, Equivalence of additivity questions in quantum information theory, e-print quant-ph/0305035, 2003.

[16] P. W. Shor, Additivity of the classical capacity of entanglement-breaking quantum channel, e-print quant-ph/0201149, 2002.