On Shor's Channel Extension and Constrained Channels

A.S. Holevo[1], M.E. Shirokov[2]

[1] Steklov Mathematical Institute, 119991 Moscow, Russia
[2] Moscow Institute of Physics and Technology, 141700 Moscow, Russia

Abstract: Several equivalent formulations of the additivity conjecture for constrained channels, which formally is substantially stronger than the unconstrained additivity, are given. To this end a characteristic property of the optimal ensemble for such a channel is derived, generalizing the maximal distance property. It is shown that the additivity conjecture for constrained channels holds true for certain nontrivial classes of channels. After giving an algebraic formulation for Shor's channel extension, its main asymptotic property is proved. It is then used to show that additivity for two constrained channels can be reduced to the same problem for unconstrained channels, and hence, "global" additivity for channels with arbitrary constraints is equivalent to additivity without constraints.

1. Introduction

In the recent paper [14] Shor gave arguments which show that conjectured additivity properties for several quantum information quantities, such as the minimal output entropy, the Holevo capacity (in what follows χ-capacity) and the entanglement of formation are in fact equivalent. An important new tool in these arguments is the construction of a special extension $\widehat{\Phi}$ for an arbitrary channel Φ which has desired properties lacking for the initial channel. In this paper we show that this extension allows us to deal with the additivity conjecture for quantum channels with constrained inputs. Introducing input constraints provides greater flexibility in the treatment of the additivity conjecture. In a sense, Shor's channel extension plays a role of the Lagrange function in optimization for the additivity questions. On the other hand, while [14] deals with the "global" additivity, i.e. properties valid for all possible channels, in this paper we make emphasis on results valid for individual channels.

We start with giving several equivalent formulations of the additivity conjecture for constrained channels (Theorem 1), which formally is substantially stronger than the unconstrained additivity. To this end a characteristic property of the optimal ensemble

for such a channel is derived (Proposition 1), generalizing the maximal distance property [11]. It is shown that the additivity conjecture for constrained channels holds true for certain nontrivial classes of channels (Proposition 2). After giving an algebraic formulation for Shor's channel extension [14], its main property (Proposition 3) is proved. It is then used to show that additivity for two constrained channels can be reduced to the same problem for unconstrained channels, and hence, global additivity for channels with arbitrary constraints is equivalent to global additivity without constraints (Theorem 2 and corollaries).

Further results in this direction can be found in [4], on which the present paper is based.

2. Basic Quantities

Let $\mathcal{H}$, $\mathcal{H}'$ be finite dimensional Hilbert spaces and let $\Phi : \mathfrak{S}(\mathcal{H}) \mapsto \mathfrak{S}(\mathcal{H}')$ be a channel, where $\mathfrak{S}(\mathcal{H})$ denotes the set of states (density operators) in $\mathcal{H}$. Let $\{\pi_i\}$ be a finite probability distribution and $\{\rho_i\}$ a collection of states in $\mathfrak{S}(\mathcal{H})$, then the collection $\{\pi_i, \rho_i\}$ is called an *ensemble,* and $\rho_{\mathrm{av}} = \sum_i \pi_i \rho_i$ is its *average.*

An important entropic characteristic of the ensemble is defined by

$$\chi_\Phi\left(\{\pi_i, \rho_i\}\right) = H\left(\sum_i \pi_i \Phi\left(\rho_i\right)\right) - \sum_i \pi_i H\left(\Phi\left(\rho_i\right)\right), \tag{1}$$

where $H\left(\cdot\right)$ is the von Neumann entropy. Following [6], we denote

$$\chi_\Phi(\rho) = \max_{\rho_{\mathrm{av}}=\rho} \chi_\Phi(\{\pi_i, \rho_i\}).$$

Notice that

$$\chi_\Phi(\rho) = H\left(\Phi\left(\rho\right)\right) - \hat{H}_\Phi\left(\rho\right), \tag{2}$$

where

$$\hat{H}_\Phi\left(\rho\right) = \min_{\rho_{\mathrm{av}}=\rho} \sum_i \pi_i H\left(\Phi\left(\rho_i\right)\right).$$

The function $\hat{H}_\Phi\left(\rho\right)$ is the *convex closure* [5, 1] (or the convex roof, cf. [15]) of the *output entropy* $H\left(\Phi\left(\rho\right)\right)$, which is a continuous concave function. The function $\hat{H}_\Phi\left(\rho\right)$ is a natural generalization of the entanglement of formation and coincides with it when the channel Φ is a partial trace. The continuity of $\hat{H}_\Phi\left(\rho\right)$ follows from the MSW correspondence [6] and the continuity of the entanglement of formation [7]. Thus the function $\chi_\Phi(\rho)$ (briefly χ -function) is itself continuous and concave on $\mathfrak{S}(\mathcal{H})$.

Consider the constraint on the ensemble $\{\pi_i, \rho_i\}$ defined by the requirement $\rho_{\mathrm{av}} \in \mathcal{A}$, where $\mathcal{A}$ is a closed subset of $\mathfrak{S}(\mathcal{H})$. A particular case is linear constraint, where the subset $\mathcal{A}^l$ is defined by the inequality $\mathrm{Tr} A \rho_{\mathrm{av}} \leq \alpha$ for some positive operator A and a number $\alpha \geq 0$. Define the χ-*capacity* of the $\mathcal{A}$-constrained channel Φ by

$$\bar{C}(\Phi; \mathcal{A}) = \max_{\rho \in \mathcal{A}} \chi_\Phi(\rho) = \max_{\rho_{\mathrm{av}} \in \mathcal{A}} \chi_\Phi(\{\pi_i, \rho_i\}). \tag{3}$$

In case of the linear constraint $\mathcal{A}^l$ we also use the notation $\bar{C}(\Phi; A, \alpha)$. Note that the χ-capacity for the unconstrained channel is $\bar{C}(\Phi) = \bar{C}(\Phi; \mathfrak{S}(\mathcal{H}))$.

Lemma 1. *For arbitrary channel* $\Phi : \mathfrak{S}(\mathcal{H}) \mapsto \mathfrak{S}(\mathcal{H}')$ *and arbitrary density opera-
tor* ρ_0 *of full rank there exists a positive operator* $A \leq I_{\mathcal{H}}$ *in* $\mathfrak{B}(\mathcal{H})$ *such that* ρ_0 *is the
maximum point of the function* $\chi_\Phi(\rho)$ *under the condition* $\text{Tr} A\rho \leq \alpha$, *where* $\alpha = \text{Tr} A\rho_0$.

The statement of the lemma is intuitively clear, but its proof (see Appendix, I) requires
an argument from the convex analysis due to the fact that the function $\chi_\Phi(\rho)$ may not
be smooth.

3. Optimal Ensembles

An ensemble $\{\pi_i, \rho_i\}$ on which the maximum in (3) is achieved is called an *optimal
ensemble* for the $\mathcal{A}$ -constrained channel Φ. The following proposition generalizes the
maximal distance property of optimal ensembles for unconstrained channels [11].

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

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

for any ensemble $\{\mu_j, \omega_j\}$ *with the average* $\omega_{\text{av}} \in \mathcal{A}$, *where* $H(\cdot \| \cdot)$ *is the relative
entropy.*

Proof. The proof generalizes the argument in [11] by considering variations of the initial
ensemble involving not a single component but the whole ensemble.

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

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

has the average $\rho_{\text{av}}^\eta = (1-\eta)\rho_{\text{av}} + \eta\omega_{\text{av}} \in \mathcal{A}$ (by convexity). Using the relative entropy
expression for the quantity (1), we have

$$\chi_\Phi\left(\Sigma^\eta\right) = (1-\eta) \sum_{i=1}^n \pi_i H(\Phi(\rho_i) \| \Phi(\rho_{\text{av}}^\eta)) + \eta \sum_{j=1}^m \mu_j H(\Phi(\omega_j) \| \Phi(\rho_{\text{av}}^\eta)). \quad (4)$$

Applying Donald's identity [11, 12] to the original ensemble we obtain

$$\sum_{i=1}^n \pi_i H(\Phi(\rho_i) \| \Phi(\rho_{\text{av}}^\eta)) = \chi_\Phi(\Sigma^0) + H(\Phi(\rho_{\text{av}}) \| \Phi(\rho_{\text{av}}^\eta)).$$

Substitution of the above expression into (4) gives

$$\chi_\Phi\left(\Sigma^\eta\right) = \chi_\Phi(\Sigma^0) + (1-\eta)H(\Phi(\rho_{\text{av}}) \| \Phi(\rho_{\text{av}}^\eta))$$

$$+\eta \left[\sum_{j=1}^m \mu_j H(\Phi(\omega_j) \| \Phi(\rho_{\text{av}}^\eta)) - \chi_\Phi(\Sigma^0) \right]. \quad (5)$$

Applying Donald's identity to the modified ensemble we obtain

$$(1 - \eta) \sum_{i=1}^{n} \pi_i H(\Phi(\rho_i) \| \Phi(\rho_{\text{av}})) + \eta \sum_{j=1}^{m} \mu_j H(\Phi(\omega_j) \| \Phi(\rho_{\text{av}}))$$
$$= \chi_\Phi \left(\Sigma^\eta \right) + H(\Phi(\rho_{\text{av}}^\eta) \| \Phi(\rho_{\text{av}})),$$

and hence

$$\chi_\Phi \left(\Sigma^\eta \right) = \chi_\Phi \left(\Sigma^0 \right) - H(\Phi(\rho_{\text{av}}^\eta) \| \Phi(\rho_{\text{av}}))$$
$$+ \eta \left[\sum_{j=1}^{m} \mu_j H(\Phi(\omega_j) \| \Phi(\rho_{\text{av}})) - \chi_\Phi \left(\Sigma^0 \right) \right]. \tag{6}$$

Since the relative entropy is nonnegative, the expressions (5) and (6) imply the following inequalities for the quantity $\Delta \chi_\Phi = \chi_\Phi \left(\Sigma^\eta \right) - \chi_\Phi \left(\Sigma^0 \right)$:

$$\eta \left[\sum_{j=1}^{m} \mu_j H(\Phi(\omega_j) \| \Phi(\rho_{\text{av}}^\eta)) - \chi_\Phi \left(\Sigma^0 \right) \right]$$
$$\leq \Delta \chi_\Phi \leq \tag{7}$$
$$\eta \left[\sum_{j=1}^{m} \mu_j H(\Phi(\omega_j) \| \Phi(\rho_{\text{av}})) - \chi_\Phi \left(\Sigma^0 \right) \right].$$

Now the proof of the proposition is straightforward. If

$$\sum_j \mu_j H(\Phi(\omega_j) \| \Phi(\rho_{\text{av}})) \leq \chi_\Phi \left(\Sigma^0 \right)$$

for any ensemble $\{\mu_j, \omega_j\}$ of states in $\mathfrak{S}(\mathcal{H})$ with the average $\omega_{\text{av}} \in \mathcal{A}$, then by the second inequality in (7) with $\eta = 1$ we have

$$\chi_\Phi(\{\mu_j, \omega_j\}) = \chi_\Phi(\Sigma^1) \leq \chi_\Phi \left(\Sigma^0 \right) = \chi_\Phi(\{\pi_i, \rho_i\}),$$

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

$$\sum_j \mu_j H(\Phi(\omega_j) \| \Phi(\rho_{\text{av}})) > \chi_\Phi \left(\Sigma^0 \right).$$

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

$$\sum_j \mu_j H(\Phi(\omega_j) \| \Phi(\rho_{\text{av}}^\eta)) > \chi_\Phi \left(\Sigma^0 \right).$$

By the first inequality in (7), this means that $\chi_\Phi \left(\Sigma^\eta \right) > \chi_\Phi \left(\Sigma^0 \right)$ in contradiction with the optimality of the ensemble $\{\pi_i, \rho_i\}$ $\square$

Corollary 1. *Let ρ_{av} be the average of an optimal ensemble for the $\mathcal{A}$-constrained channel Φ, then*

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

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

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

This inequality and Donald's identity,

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

complete the proof.   □

4. Additivity for Constrained Channels

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

We conjecture the following additivity property for constrained channels

$$\bar{C}(\Phi \otimes \Psi; \mathcal{A} \otimes \mathcal{B}) = \bar{C}(\Phi; \mathcal{A}) + \bar{C}(\Psi; \mathcal{B}). \tag{8}$$

The usual additivity conjecture for unconstrained channels is obtained by setting $\mathcal{A} = \mathfrak{S}(\mathcal{H})$, $\mathcal{B} = \mathfrak{S}(\mathcal{K})$.

Theorem 1. *Let Φ and Ψ be fixed channels. The following properties are equivalent:*

(i) equality (8) holds for arbitrary closed $\mathcal{A}$ and $\mathcal{B}$;
(ii) equality (8) holds for arbitrary linear constraints $\mathcal{A}^l$ and $\mathcal{B}^l$;
(iii) for arbitrary $\sigma \in \mathfrak{S}(\mathcal{H} \otimes \mathcal{K})$

$$\chi_{\Phi \otimes \Psi}(\sigma) \leq \chi_\Phi(\sigma^\Phi) + \chi_\Psi(\sigma^\Psi); \tag{9}$$

(iv) for arbitrary $\sigma \in \mathfrak{S}(\mathcal{H} \otimes \mathcal{K})$

$$\hat{H}_{\Phi \otimes \Psi}(\sigma) \geq \hat{H}_\Phi(\sigma^\Phi) + \hat{H}_\Psi(\sigma^\Psi). \tag{10}$$

These are also equivalent to the corresponding additivity properties of χ_Φ and $\hat{H}_\Phi$ for tensor product states. By using the MSW correspondence, the case of $\hat{H}_\Phi$ can be reduced to entanglement of formation, for which this was established in [14, 10].

Proof. $(i) \Rightarrow (ii)$ is obvious. $(ii) \Rightarrow (i)$ can be proved by double application of the following lemma.

Lemma 2. *The equality (8) holds for fixed closed $\mathcal{B}$ and arbitrary closed $\mathcal{A}$ if it holds for the set $\mathcal{B}$ and an arbitrary linear constraint $\mathcal{A}^l$, defined by the inequality $\mathrm{Tr} A\rho \leq \alpha$ with a positive operator A and a number α such that there exists a state ρ' with $\mathrm{Tr} A\rho' < \alpha$.*

Proof. Assume that the equality (8) holds for the set $\mathcal{B}$ and arbitrary set $\mathcal{A}^l$, satisfying the above condition. It is sufficient to prove that

$$\chi_{\Phi \otimes \Psi}(\sigma) \leq \chi_{\Phi}(\sigma^{\Phi}) + \bar{C}(\Psi; \mathcal{B}) \tag{11}$$

for any $\sigma \in \mathfrak{S}(\mathcal{H} \otimes \mathcal{K})$ such that $\sigma^{\Psi} \in \mathcal{B}$. Due to continuity of the χ -function, it is sufficient to prove (11) for a state σ with partial trace σ^{Φ} of full rank. For the state σ^{Φ} we can choose a positive operator A in $\mathfrak{B}(\mathcal{H})$ in accordance with Lemma 1. Let $\mathcal{A}^l = \{\rho \in \mathfrak{S}(\mathcal{H}) \,|\, \mathrm{Tr} A\rho \leq \alpha = \mathrm{Tr} A\sigma^{\Phi}\}$. The full rank of σ^{Φ} guarantees the existence of a state ρ' such that $\mathrm{Tr} A\rho' < \alpha = \mathrm{Tr} A\sigma^{\Phi}$. Let ω be the average state of the optimal ensemble for the $\mathcal{B}$-constrained channel Ψ. Due to the above assumption the state $\sigma^{\Phi} \otimes \omega$ is the average state of the optimal ensemble for the $\mathcal{A}^l \otimes \mathcal{B}$-constrained channel $\Phi \otimes \Psi$. But it is clear that this ensemble will also be optimal for $\{\sigma^{\Phi}\} \otimes \mathcal{B}$-constrained channel $\Phi \otimes \Psi$ and, hence, (11) is true. □

$(i) \Rightarrow (iv)$. Fix the states ρ and ω and take $\mathcal{A} = \{\rho\}$, $\mathcal{B} = \{\omega\}$, then (8) becomes

$$\bar{C}(\Phi \otimes \Psi; \{\rho\} \otimes \{\omega\}) = \bar{C}(\Phi; \{\rho\}) + \bar{C}(\Psi; \{\omega\}). \tag{12}$$

This implies existence of unentangled ensemble with the average $\rho \otimes \omega$, which is optimal for the $\{\rho\} \otimes \{\omega\}$ -constrained channel $\Phi \otimes \Psi$. By Corollary 1 we have

$$\chi_{\Phi \otimes \Psi}(\rho \otimes \omega) = \chi_{\Phi}(\rho) + \chi_{\Psi}(\omega) \geq \chi_{\Phi \otimes \Psi}(\sigma) + H((\Phi \otimes \Psi)(\sigma) \| \Phi(\rho) \otimes \Psi(\omega)) \tag{13}$$

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

$$H((\Phi \otimes \Psi)(\sigma) \| \Phi(\rho) \otimes \Psi(\omega)) = H(\Phi(\rho)) + H(\Psi(\omega)) - H((\Phi \otimes \Psi)(\sigma)). \tag{14}$$

The inequality (13) together with (14) and (2) implies (10).

$(iv) \Rightarrow (iii)$ obviously follows from the definition of the χ -function and subadditivity of the (output) entropy.

$(iii) \Rightarrow (i)$. From the definition of the χ-capacity and (9),

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

Since the converse inequality is obvious, there is equality here.

Remark 1. The additivity of the $\chi-$capacity for arbitrarily constrained channels is formally substantially stronger than the usual unconstrained additivity. Indeed, the latter holds trivially for channels that are (unconstrained) partial traces, but the additivity for constrained partial traces, by the MSW correspondence, would imply validity of the global additivity conjecture.

The following proposition implies that the set of quantum channels satisfying the properties in Theorem 1 is nonempty. We shall use the following obvious statement

Lemma 3. *Let $\{\Phi_j\}_{j=1}^n$ be a collection of channels from $\mathfrak{S}(\mathcal{H})$ into $\mathfrak{S}(\mathcal{H}_j)$, and let $\{q_j\}_{j=1}^n$ 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*

$$\chi_{\Phi}(\{\rho_i, \pi_i\}) = \sum_{j=1}^n q_j \chi_{\Phi_j}(\{\rho_i, \pi_i\}).$$

We shall call Φ the *direct sum mixture* of the channels $\{\Phi_j\}_{j=1}^n$.

Proposition 2. *Let Ψ be an arbitrary channel. The inequality (9) holds in each of the following cases:*

(i) *Φ is a noiseless channel;*

(ii) *Φ is an entanglement breaking channel;*

(iii) *Φ is a direct sum mixture of a noiseless channel and a channel Φ_0 such that (9) holds for Φ_0 and Ψ (in particular, an entanglement breaking channel).*

An obvious example of a channel of the type (iii) is the erasure channel.

Proof. (i) The proof is a modification of the proof in [3] of the "unconstrained" additivity for two channels with one of them noiseless, based on the Groenevold-Lindblad-Ozawa inequality [9]

$$H(\sigma) \geq \sum_j p_j H(\sigma_j), \tag{15}$$

where σ is a state of a quantum system before von Neumann measurement, σ_j — the posterior state with the outcome j and p_j is the probability of this outcome.

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

$$\bar{C}(\mathrm{Id} \otimes \Psi, \{\rho\} \otimes \{\omega\}) = \bar{C}(\mathrm{Id}, \{\rho\}) + \bar{C}(\Psi, \{\omega\}) = H(\rho) + \chi_\Psi(\omega). \tag{16}$$

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

$$\chi_{\mathrm{Id}\otimes\Psi}(\{\mu_i, \sigma_i\}) = H(\mathrm{Id} \otimes \Psi(\sum_i \mu_i\sigma_i)) - \sum_i \mu_i H(\mathrm{Id} \otimes \Psi(\sigma_i))$$

$$\leq H(\rho) + H(\Psi(\omega)) - \sum_i \mu_i H(\mathrm{Id} \otimes \Psi(\sigma_i)). \tag{17}$$

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 (15) we obtain

$$H(\mathrm{Id} \otimes \Psi(\sigma_i)) \geq \sum_j p_{ij} H(\Psi(\sigma_{ij}^\Psi)), \quad \text{for all } i,$$

where $p_{ij} = \langle e_j|\sigma_i|e_j\rangle$ and $\sigma_{ij}^\Psi = p_{ij}^{-1}|e_j\rangle\langle e_j| \otimes I_\mathcal{K} \cdot \sigma_i \cdot |e_j\rangle\langle e_j| \otimes I_\mathcal{K}$. Note that $\sum_j p_{ij}\sigma_{ij}^\Psi = \sigma_i^\Psi$ and $\sum_{ij} \mu_i p_{ij}\sigma_{ij}^\Psi = \omega$. This and the previous inequality show that the last two terms in (17) do not exceed $\chi_\Psi(\{\mu_i p_{ij}, \sigma_{ij}^\Psi\})$ and, hence, $\chi_\Psi(\omega)$. With this observation (17) implies (16) and hence the proof is complete.

(ii) See [13] where the additivity conjecture for two unconstrained channels with one of them entanglement breaking was proved. In the proof of this theorem the subadditivity property of the χ-function was in fact established. We can also deduce the subadditivity of the χ-function from the unconstrained additivity with the help of Corollary 2 (see Sect. 5 below). One should only verify that entanglement breaking property of a channel implies a similar property of Shor's extension for that channel.

(*iii*) 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 using Lemma 3 and subadditivity of the functions $\chi_{\mathrm{Id}\otimes\Psi}$ and $\chi_{\Phi_0\otimes\Psi}$,

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

where the last equality follows from the existence of a *pure* state ensemble on which the maximum in the definition of $\chi_{\Phi_0}(\sigma^\Phi)$ is achieved.   □

5. Shor's Channel Extension

Let Φ be a channel from $\mathfrak{S}(\mathcal{H})$ to $\mathfrak{S}(\mathcal{H}')$, and let E be an operator in $\mathfrak{B}(\mathcal{H})$, $0 \leq E \leq I$. Let $q \in [0; 1]$ and $d \in \mathbb{N} = \{1, 2, \dots\}$. Shor's channel extension $\widehat{\Phi}$ with probability $1 - q$ acts as the channel Φ and with probability q makes a measurement in $\mathcal{H}$ with the outcomes $\{0, 1\}$ corresponding to the resolution of the identity $\{E^\perp, E\}$, where we denote $E^\perp = I - E$. If the outcome is 1, then $\log d$ classical bits are sent to the receiver, otherwise – a failure signal [14]. Later q will tend to zero while d – to infinity, such that $q\log d = \lambda$ will be constant. The channel $\widehat{\Phi}$ will then mostly act on input states ρ as Φ, at the same time rarely sending a lot of classical information at the rate proportional to the value $\mathrm{Tr}\rho E$, which to some extent explains its relation to the capacity of the channel Φ with constrained inputs to be explored in this section.

Translating the definition into algebraic language, consider the following channel $\widehat{\Phi}(E, q, d)$, which maps states on $\mathfrak{B}(\mathcal{H}) \otimes \mathbf{C}^d$ into states on $\mathfrak{B}(\mathcal{H}') \oplus \mathbf{C}^{d+1}$, 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 $\Phi_1(\widehat{\rho}) = [\mathrm{Tr}\rho E^\perp, \mathrm{Tr}\rho_1 E, \dots, \mathrm{Tr}\rho_d E] \in \mathbf{C}^{d+1}$. 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 input state space of the channel $\widehat{\Phi}(E, q, d)$ will be denoted $\mathfrak{S}_{\widehat{\Phi}}$.

Remark 2. More precisely, since in this paper channel means a map defined on the algebra of all operators in the input Hilbert space, the action of $\widehat{\Phi}(E, q, d)$ should be extended correspondingly. Then $\mathbf{C}^d$ is considered as the algebra of diagonal matrices acting in d-dimensional Hilbert space $\mathcal{H}_d$, and the input algebra of the channel $\mathfrak{B}(\mathcal{H}) \otimes \mathbf{C}^d \subset \mathfrak{B}(\mathcal{H} \otimes \mathcal{H}_d)$, while the output algebra $\mathfrak{B}(\mathcal{H}') \oplus \mathbf{C}^{d+1} \subset \mathfrak{B}(\mathcal{H}' \oplus \mathcal{H}_{d+1})$. The action of $\widehat{\Phi}(E, q, d)$ can then be naturally extended to the whole of $\mathfrak{B}(\mathcal{H} \otimes \mathcal{H}_d)$ by letting $\widehat{\Phi}$ vanish on the elements $A \otimes B$, where $A \in \mathfrak{B}(\mathcal{H})$ and B is any matrix with zeroes on the diagonal, acting in $\mathcal{H}_d$. This is described in [14] by saying that the first action of $\widehat{\Phi}(E, q, d)$ is to make a measurement in the canonical basis of $\mathcal{H}_d$.

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*

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

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

Proof. Due to the representation

$$\widehat{\Phi}(E, q, d) \otimes \Psi = (1-q)(\Phi_0 \otimes \Psi) \oplus q(\Phi_1 \otimes \Psi), \tag{18}$$

Lemma 3 reduces the calculation of the quantity $\chi_{\widehat{\Phi}(E,q,d) \otimes \Psi}$ for any ensemble of input states to the calculation of the quantities $\chi_{\Phi_0 \otimes \Psi}$ and $\chi_{\Phi_1 \otimes \Psi}$ for this ensemble.

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 jth position and with zeroes in other places.

It is known that for any channel there exists a pure state optimal ensemble [11] and that the image of the average state of any optimal ensemble is the same (this follows from Corollary 1). These facts and symmetry arguments imply 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 an ensemble of states in $\mathfrak{S}(\mathcal{H} \otimes \mathcal{K})$ (cf. [14]). Let $\widehat{\sigma}_{\mathrm{av}} = \sum_{i,j} \widehat{\mu}_{i,j}\widehat{\sigma}_{i,j}$ and $\sigma_{\mathrm{av}} = \sum_i \mu_i\sigma_i$ be the averages of these ensembles. Note that $\widehat{\sigma}_{\mathrm{av}} = [d^{-1}\sigma_{\mathrm{av}}, ..., d^{-1}\sigma_{\mathrm{av}}]$.

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

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

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

$$\chi_{\Phi_0 \otimes \Psi}(\{\widehat{\mu}_{i,j}, \widehat{\sigma}_{i,j}\}) = \chi_{\Phi \otimes \Psi}(\{\mu_i, \sigma_i\}). \tag{19}$$

Let us prove that

$$\chi_{\Phi_1 \otimes \Psi}(\{\widehat{\mu}_{i,j}, \widehat{\sigma}_{i,j}\}) = \log d\mathrm{Tr}\sigma_{\mathrm{av}}(E \otimes I_{\mathcal{K}}) + f_{\Psi}^E(\{\mu_i, \sigma_i\}), \tag{20}$$

where $0 \leq f_{\Psi}^E(\{\mu_i, \sigma_i\}) \leq \log \dim \mathcal{K}' + 1$. It is easy to see that the action of the channel $\Phi_1 \otimes \Psi$ on the state $\widehat{\sigma} = [\sigma_j]_{j=1}^d$ with $\sigma = \sum_{i=1}^d \sigma_i$ is

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

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

Therefore,

$$H(\Phi_1 \otimes \Psi(\widehat{\sigma}_{i,j})) = H(\Psi_{E^\perp}(\sigma_i)) + H(\Psi_E(\sigma_i)), \tag{21}$$

and

$$\Phi_1 \otimes \Psi(\widehat{\sigma}_{\mathrm{av}}) = \sum_{i,j} \widehat{\mu}_{i,j}\Phi_1 \otimes \Psi(\widehat{\sigma}_{i,j})$$

$$= [\Psi_{E^\perp}(\sigma_{\mathrm{av}}), d^{-1}\Psi_E(\sigma_{\mathrm{av}}), ..., d^{-1}\Psi_E(\sigma_{\mathrm{av}})].$$

Due to this

$$H(\Phi_1 \otimes \Psi(\widehat{\sigma}_{\mathrm{av}})) = \log d \ \mathrm{Tr}\Psi_E(\sigma_{\mathrm{av}}) + H(\Psi_E(\sigma_{\mathrm{av}})) + H(\Psi_{E^\perp}(\sigma_{\mathrm{av}})). \qquad (22)$$

Using (21), (22) and $\mathrm{Tr}\Psi_E(\sigma) = \mathrm{Tr}\sigma(E \otimes I_\mathcal{K})$, we obtain

$$\chi_{\Phi_1 \otimes \Psi}(\{\widehat{\mu}_{i,j}, \widehat{\sigma}_{i,j}\}) = \log d \, \mathrm{Tr}\sigma_{\mathrm{av}}(E \otimes I_\mathcal{K}) + H(\Psi_E(\sigma_{\mathrm{av}})) + H(\Psi_{E^\perp}(\sigma_{\mathrm{av}}))$$
$$- \sum_i \mu_i(H(\Psi_E(\sigma_i)) + H(\Psi_{E^\perp}(\sigma_i)))$$
$$= \log d \mathrm{Tr}\sigma_{\mathrm{av}}(E \otimes I_\mathcal{K}) + \chi_{\Psi_E}(\{\mu_i, \sigma_i\}) + \chi_{\Psi_{E^\perp}}(\{\mu_i, \sigma_i\}). \quad (23)$$

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

$$f_\Psi^E(\{\mu_i, \sigma_i\}) := \chi_{\Psi_E}(\{\mu_i, \sigma_i\}) + \chi_{\Psi_{E^\perp}}(\{\mu_i, \sigma_i\}) \le \log \dim \mathcal{K}' + 1, \qquad (24)$$

hence we obtain (20).

Lemma 3 with (19) and (20) 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 (24) completes the proof. □

Theorem 2. *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 a closed set* $\mathcal{B}$. *The following statements are equivalent:*

(i) The additivity (8) holds for the $\mathcal{A}$-*constrained channel* Φ *with arbitrary closed* $\mathcal{A} \in \mathfrak{S}(\mathcal{H})$ *and the* $\mathcal{B}$-*constrained channel* Ψ;

(ii) The additivity holds asymptotically for the sequence of the channels $\{\widehat{\Phi}(E, \lambda/\log d, d)\}_{d \in \mathbb{N}}$ *with arbitrary operator* $0 \le E \le I$ *and arbitrary nonnegative number* λ *(without constraints) and the* $\mathcal{B}$-*constrained channel* Ψ, *in the sense that*

$$\lim_{d \to +\infty} \bar{C}(\widehat{\Phi}(E, \lambda/\log d, d) \otimes \Psi, \mathfrak{S}_{\widehat{\Phi}} \otimes \mathcal{B})$$
$$= \lim_{d \to +\infty} \bar{C}(\widehat{\Phi}(E, \lambda/\log d, d)) + \bar{C}(\Psi, \mathcal{B}).$$

Proof. Note, first of all, that for an operator $0 \le E \le I$ and a number $\lambda \ge 0$ Proposition 3 implies

$$\lim_{d \to +\infty} \bar{C}(\widehat{\Phi}(E, \lambda/\log d, d)) = \max_\rho [\chi_\Phi(\rho) + \lambda \mathrm{Tr}\rho E] \qquad (25)$$

and

$$\lim_{d \to +\infty} \bar{C}(\widehat{\Phi}(E, \lambda/\log d, d) \otimes \Psi, \mathfrak{S}_{\widehat{\Phi}} \otimes \mathcal{B}) = \max_{\sigma:\mathrm{Tr}_\mathcal{H}\sigma \in \mathcal{B}} \left[\chi_{\Phi \otimes \Psi}(\sigma) + \lambda \mathrm{Tr}\,\sigma(E \otimes I_\mathcal{K})\right]$$

$$(26)$$

correspondingly.

Begin with $(i) \Rightarrow (ii)$. Suppose that there exist an operator $0 \le E \le I$ and a number $\lambda \ge 0$ such that (ii) does not hold for the sequence $\widehat{\Phi}(E, \lambda/\log d, d)$ and the channel Ψ. Due to (25) and (26) this means that

$$\max_{\sigma:\sigma^\Psi \in \mathcal{B}} \left[\chi_{\Phi \otimes \Psi}(\sigma) + \lambda \mathrm{Tr}\, \sigma(E \otimes I_\mathcal{K}) \right] > \max_\rho \left[\chi_\Phi(\rho) + \lambda \mathrm{Tr}\rho E \right] + \bar{C}(\Psi, \mathcal{B}). \quad (27)$$

Let σ_* be a maximum point in the left side of the above inequality and $\alpha = \mathrm{Tr}\sigma_*(E \otimes I_\mathcal{K})$. By the statement (i) the additivity holds for the channel Φ with the constraint $\mathrm{Tr}\rho E^\perp \le 1 - \alpha$ and the $\mathcal{B}$-constrained channel Ψ. So there exist such states ρ and $\omega \in \mathcal{B}$ that $\mathrm{Tr}\rho E \ge \alpha$ and $\chi_\Phi(\rho) + \chi_\Psi(\omega) \ge \chi_{\Phi \otimes \Psi}(\sigma_*)$. Hence

$$\max_{\sigma:\sigma^\Psi \in \mathcal{B}} \left[\chi_{\Phi \otimes \Psi}(\sigma) + \lambda \mathrm{Tr}\sigma(E \otimes I_\mathcal{K}) \right] = \chi_{\Phi \otimes \Psi}(\sigma_*) + \lambda \mathrm{Tr}\, \sigma_*(E \otimes I_\mathcal{K})$$

$$\le \chi_\Phi(\rho) + \chi_\Psi(\omega) + \lambda \mathrm{Tr}\rho E$$

$$\le \max_\rho \left[\chi_\Phi(\rho) + \lambda \mathrm{Tr}\rho E \right] + \bar{C}(\Psi; \mathcal{B})$$

in contradiction with (27).

The proof of $(ii) \Rightarrow (i)$ is based on Lemma 2. Let $\mathcal{A}^l$ be a set defined by the inequality $\mathrm{Tr}\rho A \le \alpha$ with an operator $0 \le A \le I$ and a positive number α such that there exists a state ρ' with $\mathrm{Tr}\rho'A < \alpha$. Due to Lemma 2 it is sufficient to show that

$$\bar{C}\left(\Phi \otimes \Psi; \mathcal{A}^l \otimes \mathcal{B} \right) \le \bar{C}(\Phi; \mathcal{A}^l) + \bar{C}(\Psi; \mathcal{B}). \quad (28)$$

Suppose, " $>$ " takes place in (28). Then there exists an ensemble $\{\mu_i, \sigma_i\}$ in $\mathfrak{S}(\mathcal{H} \otimes \mathcal{K})$ with the average σ_{av}, such that $\mathrm{Tr}\sigma_{\mathrm{av}}^\Phi A \le \alpha$, $\sigma_{\mathrm{av}}^\Psi \in \mathcal{B}$ and

$$\chi_{\Phi \otimes \Psi}(\{\mu_i, \sigma_i\}) > \bar{C}(\Phi; \mathcal{A}^l) + \bar{C}(\Psi; \mathcal{B}). \quad (29)$$

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

Note that the state ρ_{av} is the point of maximum of the concave function $\chi_\Phi(\rho)$ with the constraint $\mathrm{Tr}\rho A \le \alpha$. By the Kuhn-Tucker theorem (we use the strong version of this theorem with the Slater condition, which follows from the existence of a state ρ' such that $\mathrm{Tr}\rho'A < \alpha$) [5], there exists a nonnegative number λ, such that ρ_{av} is the point of the global maximum of the function $\chi_\Phi(\rho) - \lambda \mathrm{Tr}\rho A$ and the following condition holds:

$$\lambda(\mathrm{Tr}A\rho_{\mathrm{av}} - \alpha) = 0. \quad (30)$$

It is clear that ρ_{av} is also the point of the global maximum of the concave function $\chi_\Phi(\rho) + \lambda \mathrm{Tr}\rho E$, where $E = I - A$, so that

$$\chi_\Phi(\rho) + \lambda \mathrm{Tr}\rho E \le \chi_\Phi(\rho_{\mathrm{av}}) + \lambda \mathrm{Tr}\rho_{\mathrm{av}} E, \quad \forall \rho \in \mathfrak{S}(\mathcal{H}). \quad (31)$$

Consider the sequence $\widehat{\Phi}(E, \lambda/\log d, d)$. Assumed asymptotic additivity together with (25) and (26) implies

$$\max_\sigma \left[\chi_{\Phi \otimes \Psi}(\sigma) + \lambda \mathrm{Tr}\, \sigma(E \otimes I_\mathcal{K}) \right] = \max_\rho \left[\chi_\Phi(\rho) + \lambda \mathrm{Tr}\rho E \right] + \bar{C}(\Psi; \mathcal{B}). \quad (32)$$

Due to (30) and (31) we have

$$\max_\rho \left[\chi_\Phi(\rho) + \lambda \mathrm{Tr}\rho E \right] = \chi_\Phi(\rho_{\mathrm{av}}) + \lambda \mathrm{Tr}\rho_{\mathrm{av}}(I - A) = \bar{C}(\Phi; \mathcal{A}^l) + \lambda(1 - \alpha). \quad (33)$$

Noting that

$$\mathrm{Tr}\, \sigma_{\mathrm{av}}(A \otimes I_{\mathcal{K}}) = \mathrm{Tr}\sigma_{\mathrm{av}}^{\Phi} A \le \alpha,$$

we have by (29)

$$\max_{\sigma : \sigma^{\Psi} \in \mathcal{B}} \left[\chi_{\Phi \otimes \Psi}(\sigma) + \lambda \mathrm{Tr}\sigma(E \otimes I_{\mathcal{K}}) \right]$$

$$\ge \chi_{\Phi \otimes \Psi}(\sigma_{\mathrm{av}}) + \lambda \mathrm{Tr}\sigma_{\mathrm{av}}(E \otimes I_{\mathcal{K}}) > \bar{C}(\Phi; \mathcal{A}^l) + \bar{C}(\Psi; \mathcal{B}) + \lambda(1 - \alpha).$$

The contradiction of the last inequality with (32) and (33) completes the proof of (28), and hence $(ii) \Rightarrow (i)$. $\square$

Corollary 2. *The additivity of χ-capacity for Shor's channel extensions $\widehat{\Phi}(E, q, d)$ and $\widehat{\Psi}(F, r, e)$ with arbitrary pairs (E, q, d) and (F, r, e) implies its 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 theorem 2. $\square$

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

Remark 3. The statement of Corollary 3 could be also deduced by combining results of [14] and [6], but we gave a direct proof here.

6. Additive Constraints

Let A be a positive operator in $\mathcal{H}$, and let

$$A^{(n)} = A \otimes \cdots \otimes I_{\mathcal{H}} + \cdots + I_{\mathcal{H}} \otimes \cdots \otimes A$$

be the corresponding operator in $\mathcal{H}^{\otimes n}$. The classical capacity of the channel Φ with inputs subject to the additive constraint

$$\mathrm{Tr}\rho^{(n)} A^{(n)} \le n\alpha; \quad n = 1, 2, \ldots$$

is shown [2] to be equal to

$$C(\Phi; A, \alpha) = \lim_{n \to \infty} \bar{C}(\Phi^{\otimes n}; A^{(n)}, n\alpha)/n.$$

In [6] the following *weak* additivity property was considered:

$$\bar{C}(\Phi \otimes \Psi; A \otimes I_{\mathcal{K}} + I_{\mathcal{H}} \otimes B, \, \gamma) = \max_{\alpha + \beta = \gamma} \left[\bar{C}(\Phi; A, \alpha) + \bar{C}(\Psi; B, \beta) \right], \quad (34)$$

where Φ and Ψ are channels with the input spaces $\mathcal{H}$ and $\mathcal{K}$, and the corresponding linear constraints $\mathrm{Tr}\rho A \le \alpha$ and $\mathrm{Tr}\rho B \le \beta$. It is easy to see that the additivity for the two constrained channels in the sense (8) implies the weak additivity (34). The extension of the latter to n channels implies

$$\bar{C}(\Phi^{\otimes n}; A^{(n)}, n\alpha) = n\bar{C}(\Phi; A, \alpha),$$

and hence the equality $C(\Phi; A, \alpha) = \bar{C}(\Phi; A, \alpha)$. Indeed, the function $f(\alpha) = \bar{C}(\Phi; A, \alpha)$ defined by (3) is nondecreasing and concave (see Appendix, II), whence

$$\max_{\alpha_1 + \cdots + \alpha_n = n\alpha} [f(\alpha_1) + \cdots + f(\alpha_n)]$$

is achieved for $\alpha_1 = \cdots = \alpha_n = \alpha$.

The weak additivity conjecture for constrained channels becomes equivalent to the additivity conjecture in the sense of this paper when this weak additivity holds true for *any* two channels. Indeed, the latter implies global additivity for channels without constraints, from which global additivity for constrained channels follows by Corollary 3.

Needless to say, however, that in applications constraints usually arise when the channel space is infinite-dimensional and the constraint operators are unbounded. The finite dimensionality (implying boundedness of the constraint operators) is crucial in this paper, and relaxing this restriction is both interesting and nontrivial problem.

7. Appendix

I. The main property underlying the proof of Lemma 1 is the concavity of the function $\chi_\Phi(\rho)$ on $\mathfrak{S}(\mathcal{H})$. This function may not be smooth, therefore we will use non-smooth convex analysis arguments instead of derivatives calculations.

Consider the Banach space $\mathfrak{B}_h(\mathcal{H})$ of all Hermitian operators on $\mathcal{H}$ and the concave extension $\widehat{\chi}_\Phi$ of the function χ_Φ to $\mathfrak{B}_h(\mathcal{H})$, defined by:

$$\widehat{\chi}_\Phi(\rho) = \begin{cases} [\mathrm{Tr}\rho] \cdot \chi_\Phi([\mathrm{Tr}\rho]^{-1}\rho), & \rho \in \mathfrak{B}_+(\mathcal{H}); \\ -\infty, & \rho \in \mathfrak{B}_h(\mathcal{H}) \backslash \mathfrak{B}_+(\mathcal{H}), \end{cases}$$

where $\mathfrak{B}_+(\mathcal{H})$ is the convex cone of positive operators in $\mathcal{H}$. The function $\widehat{\chi}_\Phi$ is bounded in a neighborhood of any internal point of $\mathfrak{B}_+(\mathcal{H})$ (and, hence, by the concavity it is continuous at all internal points of $\mathfrak{B}_+(\mathcal{H})$, which are nondegenerate positive operators, see [5], 3.2.3).

By the assumption ρ_0 is an internal point of the cone $\mathfrak{B}_+(\mathcal{H})$. Hence, the convex function $-\widehat{\chi}_\Phi$ is continuous at ρ_0. Due to the continuity, the subdifferential of the convex function $-\widehat{\chi}_\Phi$ at the point ρ_0 is not empty (see [5], 4.2.1). This means that there exists a linear function $l(\rho)$ such that ρ_0 is the minimum point of the function $-\widehat{\chi}_\Phi(\rho) - l(\rho)$. Any linear function on $\mathfrak{B}_h(\mathcal{H})$ has the form $l(\rho) = \mathrm{Tr}A\rho$ for some $A \in \mathfrak{B}_h(\mathcal{H})$. Hence, ρ_0 is also the minimum point of the function $-\widehat{\chi}_\Phi(\rho)$ under the conditions $\mathrm{Tr}A\rho = \alpha = \mathrm{Tr}A\rho_0$ and $\mathrm{Tr}\rho = 1$. Introduce the operator $A' = \frac{1}{2}[\|A\|^{-1}A + I]$ and the number $\alpha' = \frac{1}{2}[\|A\|^{-1}\alpha + 1]$. The linear variety defined by the conditions $\mathrm{Tr}\rho A = \alpha$ and $\mathrm{Tr}\rho = 1$ coincides with that defined by the conditions $\mathrm{Tr}A'\rho = \alpha'$ and $\mathrm{Tr}\rho = 1$. Therefore, ρ_0 is the minimum point of the function $-\widehat{\chi}_\Phi(\rho)$ under the conditions $\mathrm{Tr}A'\rho = \alpha'$ and $\mathrm{Tr}\rho = 1$, and, hence, ρ_0 is the maximum point of the function $\chi_\Phi(\rho)$ under the condition $\mathrm{Tr}A'\rho = \alpha'$. By concavity of the function $\chi_\Phi(\rho)$ it implies that ρ_0 is the maximum point of the function $\chi_\Phi(\rho)$ under the condition either $\mathrm{Tr}A'\rho \leq \alpha'$ or $\mathrm{Tr}A'\rho \geq \alpha'$ (see II below). By noting that $0 \leq A' \leq I$ and setting A and α to be equal to A' and α' in the first case and to $I - A'$ and $1 - \alpha'$ in the second, we complete the proof of Lemma 1.

II. If $F(x)$ is a concave continuous function and $l(x)$ is a linear function on a compact convex subset of a finite dimensional vector space, then the function

$$f(\alpha) = \max_{x: l(x) = \alpha} F(x)$$

is concave. Indeed, assume $f(\alpha)$ is not, then there exist α_1, α_2 such that $f(\frac{\alpha_1 + \alpha_2}{2}) < \frac{1}{2}[f(\alpha_1) + f(\alpha_2)]$. Let x_i be points at which the maxima are achieved, i. e. $l(x_i) = \alpha_i$ and $f(\alpha_i) = F(x_i)$, then $l(\frac{x_1 + x_2}{2}) = \frac{\alpha_1 + \alpha_2}{2}$ and $F(\frac{x_1 + x_2}{2}) \leq f(\frac{\alpha_1 + \alpha_2}{2}) < \frac{1}{2}[F(x_1) + F(x_2)]$, which contradicts the concavity of F. A similar argument applies to the functions $f_+(\alpha) = \max_{x:l(x) \leq \alpha} F(x)$ and $f_-(\alpha) = \max_{x:l(x) \geq \alpha} F(x)$ which are thus also concave.

With the same definitions one has either $f(\alpha) = f_+(\alpha)$ or $f(\alpha) = f_-(\alpha)$, for otherwise there exist x_1, x_2 such that

$$l(x_1) < \alpha; \quad F(x_1) > f(\alpha); \quad l(x_2) > \alpha; \quad F(x_2) > f(\alpha).$$

Then taking $\lambda = \frac{l(x_2) - \alpha}{l(x_2) - l(x_1)}$ one has $0 < \lambda < 1$, $l(\lambda x_1 + (1 - \lambda)x_2) = \alpha$ and

$$F(\lambda x_1 + (1 - \lambda)x_2) \leq f(\alpha) < \lambda F(x_1) + (1 - \lambda)F(x_2),$$

contradicting the concavity of F.

References

1. Audenaert, K.M.R., Braunstein, S.L.: On strong superadditivity of the entanglement of formation. Commun. Math. Phys. **246**(3), 443–452 (2004)
2. Holevo, A.S.: On quantum communication channels with constrained inputs. http://arxiv.org/abs/quant-ph/9705054, 1997; Entanglement-assisted capacity of constrained quantum channels. http://arxiv.org/abs/quant-ph/0211170, 2002
3. Holevo, A.S.: *Introduction to quantum information theory*. Moscow Independent University, 2002 (in Russian)
4. Holevo, A.S., Shirokov, M.E.: On Shor's channel extension and constrained channels. http://arxiv/abs/quant-ph/0306196, 2003; Shirokov, M.E.: On the additivity conjecture for channels with arbitrary constrains. http://arxiv.org/abs/quant-ph/0308168, 2003
5. Joffe, A.D., Tikhomirov, B.M.:*Theory of extremum problems*. Moscow: Nauka, 1974 (in Russian)
6. Matsumoto, K., Shimono, T., Winter, A.: Remarks on additivity of the Holevo channel capacity and of the entanglement of formation. http://arxiv.org/abs/quant-ph/0206148, 2002
7. Nielsen, M.A.: Continuity bounds for entanglement, Phys. Rev. A **61**, 064301 (2000)
8. Ohya, M., Pets, D.: *Quantum entropy and its use*. Berlin: Springer, 1993
9. Ozawa, M.: On information gain by quantum measurement of continuous observable, J.Math.Phys. **27**, 759–763 (1986)
10. Pomeransky, A.A.: Strong superadditivity of the entanglement of formation follows from its additivity. Phys. Rev. A **68**, 032317 (2003)
11. Schumacher, B., Westmoreland, M.: Optimal signal ensemble. Phys. Rev. A **51**, 2738, (1997)
12. Schumacher, B., Westmoreland, M.: Relative entropy in quantum information theory. In: Quantum Computation and Information, Lomonaco, S.J., Baamdt, H.E.(eds), Contemporary Mathematics **305**, Providence, RI: AMS, 2002
13. Shor, P. W.: Additivity of the classical capacity of entanglement-breaking quantum channel. J. Math. Phys. **43**, 4334–4340 (2003)
14. Shor, P. W.: Equivalence of additivity questions in quantum information theory. Commun. Math. Phys. **246**(3), 453–472,473 (2004)
15. Uhlmann, A.: Entropy and optimal decomposition of states relative to a maximal commutative subalgebra. Open Systems and Inf. Dynamics **5**(3), 209–228 (1998)

Communicated by M.B. Ruskai