THE DECREASING PROPERTY OF RELATIVE ENTROPY
AND THE STRONG SUPERADDITIVITY OF QUANTUM CHANNELS

GRIGORI G. AMOSOV[1] and STEFANO MANCINI[2]

[1] *Department of Higher Mathematics, Moscow Institute of Physics and Technology,*
141700 Dolgoprudny, Russia

[2] *Dipartimento di Fisica, Università di Camerino,*
62032 Camerino, Italy

We argue that a fundamental (conjectured) property of memoryless quantum channels, namely the strong superadditivity, is intimately related to the decreasing property of the quantum relative entropy. Using the latter we first give, for a wide class of input states, an estimation of the output entropy for phase damping channels and some Weyl quantum channels. Then we prove, without any input restriction, the strong superadditivity for several quantum channels, including depolarizing quantum channels, quantum-classical channels and quantum erasure channels.

Keywords: Quantum channel maps, Relative entropy, Classical capacity of quantum channels.
Communicated by: D Wineland & B Terhal

1. Introduction

The apparently simple concept of distinguishability is at the root of information processing, even at the quantum level. For instance, it is rather intuitive that the amount of classical information (symbols encoded into quantum states) that can be reliably transmitted through a quantum channel will ultimately depend upon the ability of the receiver to distinguish different quantum states. Unlike classical states, two different quantum states are not necessarily fully distinguishable. In [1] it was argued that the quantum relative entropy is the most appropriate quantity to measure distinguishability between different quantum states. Hence it could be a powerful tool for investigating quantum channels' properties. The quantum relative entropy does not increase under physical processes (described by completely and trace preserving maps) [2]. Thus two states can only become less distinguishable as they undergo any kind of physical transformation. This result will be central to this paper.

There is a single quantity that completely characterizes a quantum channel for transmitting classical information: its classical capacity [3]. It represents the maximum rate at which classical symbols can be transmitted through the channel in a reliably way. It should thus come from the average over a large number (actually infinity) of channel uses. However, it

was conjectured that memoryless channels possess the *additivity property*, that is the classical capacity adds up with the number of channel uses [4, 5]. Hence, it can be simply evaluated by considering one use (one shot) of the channel, likewise in the classical case due to the Shannon coding theorem [6]. This has the profound implication that entangled inputs do not matter for the capacity of memoryless quantum channels. The additive property has been proved for a class of quantum channels [7, 8, 10, 9] and it was suspected that l_p-norms play a crucial role for the global proof. Recently it has been shown that this is not the case [11, 12]. Moreover, the additivity seems not to be a global property of all quantum channels [13] as conjectured in [14]. Thus the need to devise a powerful method to single out the widest class of quantum channels for which it holds.

In reality, the additivity property as discussed above, can be traced back to the additivity of the minimal output entropy of two channels. In contrast, when we consider the minimum of the average output entropies, we are led to the *superadditivity property*. That is, the minimum of the average output entropies for the tensor product of two quantum channels is greater than or equal to the sum of the minima corresponding to the single channels. This property was conjectured in [15] and it turns out to be stronger than the simple additive property.[a]In fact, if the strong superadditivity property holds, then the additivity property follows [15].

Thus, it is of uppermost importance to prove the strong superadditivity for memoryless quantum channels. Actually, it has only been proved for entanglement-breaking channels and noiseless channels [15] and for the quantum depolarizing channel [17] using different methods.

In the present paper we argue that the strong superadditivity is related to the decreasing property of the relative entropy. Hence we shall give a proof of the strong superadditivity based on the decreasing property of the relative entropy for a class of quantum channels. This class not only includes the above mentioned channels (noiseless, entanglement -breaking and depolarizing), for which the proof turns out to be alternative to those of [15] and [17], but it also includes the quantum erasure channel, thus resulting as an extension over the already know results. Remarkably, our proof could pave the way for a unified approach to the strong superadditivity property.

The layout of the paper is the following. In Section II we recall some basic notions about quantum relative entropy and classical capacity of quantum channels. Section III is devoted to formalizing the additivity and the strong superadditivity properties. We give some estimates of the output entropy for the phase damping channels and for a subclass of Weyl channels in Section IV and Section V respectively . Finally, in Section VII we prove the strong superadditivity for a class of quantum channels without any restriction on the input states. Section VII is for conclusions.

2. Basic Notions

The von Neumann entropy of a quantum system described by a density matrix ρ belonging to the set of states $\mathfrak{S}(\mathcal{H})$ (positive unit trace operators) of the Hilbert space $\mathcal{H}$ of dimension $d < +\infty$,

$$S(\rho) := -\mathrm{Tr}(\rho \log \rho),$$

---

[a]The strong superadditivity of entanglement of formation discussed in [16] also implies the additivity of Holevo capacity.

can be considered as the proper quantum analogue of the Shannon entropy [18].

Moving on from Shannon relative entropy one can consider the von Neumann relative entropy between two states $\sigma, \rho \in \mathfrak{S}(\mathcal{H})$ as[b]

$$S(\sigma||\rho) := \mathrm{Tr}\left[\sigma(\log \sigma - \log \rho)\right].$$

This measure has the same statistical interpretation as its classical analogue: it tells us how difficult it is to distinguish the state σ from the state ρ [20]. In particular the following theorem holds [2]:

Theorem 1 (Decreasing property of relative entropy) *For any completely positive, trace preserving map* $\Phi : \mathfrak{S}(\mathcal{H}) \to \mathfrak{S}(\mathcal{H})$ *given by* $\Phi(\sigma) = \sum_i A_i \sigma A_i^*$ *such that* $\sum A_i^* A_i = 1$, *we have*

$$S(\Phi(\sigma)||\Phi(\rho)) \leq S(\sigma||\rho),$$

with $\sigma, \rho \in \mathfrak{S}(\mathcal{H})$.

We simply present a physical argument as to why we should expect this theorem to hold. A completely positive map (CP-map) can be represented as a unitary transformation on an extended Hilbert space. Unitary transformations do not change the relative entropy between two states. However, after this, we have to perform a partial trace to go back to the original Hilbert space which decreases the relative entropy as some information is invariably lost during this operation. Hence the relative entropy decreases under any CP-map.

A simple consequence of the fact that the quantum relative entropy itself does not increase under CP-maps is that quantum distinguishability never increases. Another consequence is that correlations (as measured by the quantum mutual information) also cannot increase, but now under *local* CP-maps.

In classical information theory the capacity for communication is given by the mutual information between sent message and received message [6]. This is intuitively clear, since mutual information quantifies correlations between sent and received messages and it thus tells us how faithful the transmission is. If we use quantum states to encode symbols, then the capacity is not given by the quantum mutual information, but is given by the so called Holevo bound [3] being achievable due to the quantum coding theorem proved independently in [4, 5].

The linear map $\Phi : \mathfrak{S}(\mathcal{H}) \to \mathfrak{S}(\mathcal{H})$ is said to be a quantum channel if it is completely positive [3]. Moreover, the quantum channel Φ is called bistochastic (or unital) if $\Phi(\frac{1}{d}I_{\mathcal{H}}) = \frac{1}{d}I_{\mathcal{H}}$, where $I_{\mathcal{H}}$ is the identity operator in $\mathcal{H}$.

According to [4, 5] the one shot capacity $C_1(\Phi)$ of a quantum channel Φ is defined by the formula

$$C_1(\Phi) := \sup\left[S\left(\sum_{j=1}^{r} \pi_j \Phi(x_j)\right) - \sum_{j=1}^{r} \pi_j S\left(\Phi(x_j)\right)\right], \tag{1}$$

where the supremum is taken over all probability distributions $\{\pi_j\}_{j=1}^{r}$ and states $x_j \in \mathfrak{S}(\mathcal{H})$.

---

[b]Actually, this quantity was first considered by Umegaki [19] and it is often referred to it as the Umegaki entropy.

Notice that

$$
S\left(\sum_{j=1}^{r}\pi_j\Phi(x_j)\right)-\sum_{j=1}^{r}\pi_j S\left(\Phi(x_j)\right)=\sum_{j=1}^{r}\pi_j S\left(\Phi(x_j)\,\bigg\|\,\sum_{l=1}^{r}\pi_l\Phi(x_j)\right),
$$

so that we have a direct link to the relative entropy.

The additivity conjecture states that for any two channels Φ and Ω

$$
C_1(\Phi\otimes\Omega)=C_1(\Phi)+C_1(\Omega). \tag{2}
$$

If the additivity conjecture holds, one can easily find the capacity $C(\Phi)$ of the channel Φ by the formula (see [4])

$$
C(\Phi)=\lim_{n\to+\infty}\frac{C_1(\Phi^{\otimes n})}{n}=C_1(\Phi). \tag{3}
$$

3. The strong superadditivity

Given a quantum channel Φ in a Hilbert space $\mathcal{H}$ let us put [15]

$$
H_\Phi(\rho):=\min\sum_{j=1}^{k}\pi_j S(\Phi(\rho_j)), \tag{4}
$$

where $\rho=\sum\limits_{j=1}^{k}\pi_j\rho_j$ and the minimum is taken over all probability distributions $\{\pi_j\}_{j=1}^{k}$ and states $\rho_j\in\mathfrak{S}(\mathcal{H})$.

The strong superadditivity conjecture for the channel Φ states that

$$
H_{\Phi\otimes\Omega}(\rho)\geq H_\Phi(\mathrm{Tr}_\mathcal{K}(\rho))+H_\Omega(\mathrm{Tr}_\mathcal{H}(\rho)), \tag{5}
$$

with $\rho\in\mathfrak{S}(\mathcal{H}\otimes\mathcal{K})$, for an arbitrary quantum channel Ω in the Hilbert space $\mathcal{K}$.

The infimum of the output entropy of a quantum channel Φ is defined by

$$
S_{min}(\Phi):=\inf_{\rho\in\mathfrak{S}(\mathcal{H})}S(\Phi(\rho)). \tag{6}
$$

The additivity conjecture for the quantity $S_{min}(\Phi)$ states that [4]

$$
S_{min}(\Phi\otimes\Omega)=S_{min}(\Phi)+S_{min}(\Omega) \tag{7}
$$

for an arbitrary quantum channel Ω. It was shown in [15] that if the strong superadditivity holds, then the additivity follows. Hence, the conjecture (5) is stronger than (7).

The additivity property (7) was proved for quantum depolarizing channel [8]. The method was based upon the estimation of l_p-norms of the channel. However, l_p-norms cannot be a general tool as it has been recently shown [11]. Thus, the need to devise alternative methods.

4. Estimation of the output entropy for the phase damping channel

Let $\{|e_s\rangle\}_{s=0}^{d-1}$ and $\{\lambda_s\}_{s=0}^{d-1}$ be an orthonormal basis in the Hilbert space $\mathcal{H}$ of dimension d and a probability distribution, respectively. Then, one can introduce the unitary operator

$$V := \sum_{s=0}^{d-1} \exp\left(i\frac{2\pi s}{d}\right) |e_s\rangle\langle e_s|,$$

so to define the phase damping channel as

$$\Phi(\rho) := \sum_{j=0}^{d-1} \lambda_j V^j \rho V^{*j}, \tag{8}$$

where $\rho \in \mathfrak{S}(\mathcal{H})$. Furthermore, the completely positive map defined as

$$E(\rho) := \frac{1}{d}\sum_{j=0}^{d-1} V^j \rho V^{*j} = \sum_{s=0}^{d-1} |e_s\rangle\langle e_s|\rho|e_s\rangle\langle e_s|,$$

represents the conditional expectation on the algebra of fixed elements of Φ.

We shall call a pure state $\rho = |f\rangle\langle f|\rangle \in \mathfrak{S}(\mathcal{H})$ *unbiased* with respect to the basis $\{|e_s\rangle\}$ if

$$\mathrm{Tr}(\rho|e_s\rangle\langle e_s|) = \frac{1}{d}, \quad 0 \le s \le d-1. \tag{9}$$

The above condition is equivalent to the property

$$|\langle f|e_s\rangle| = \frac{1}{\sqrt{d}}, \quad 0 \le s \le d-1. \tag{10}$$

Notice that if (10) is satisfied for vectors $|f\rangle = |f_j\rangle$, $0 \le j \le d-1$ forming an orthonormal basis in $\mathcal{H}$, then the bases $\{|f_j\rangle\}$ and $\{|e_s\rangle\}$ are said to be *mutually unbiased* [21].

Let us denote by $\mathcal{A}$ a convex set of states which can be represented as a convex linear combination of pure states $\rho = |f\rangle\langle f|$ being unbiased with respect to the basis $\{|e_s\rangle\}$ (eigenvectors of the unitary operators introduced in the definition of the phase damping channel (8)). As a consequence $\mathcal{A}$ is a convex set. Moreover the following proposition holds.

Proposition 2 *Suppose that $\rho \in \mathcal{A}$, then for the phase damping channel (8) we get*

$$H_\Phi(\rho) \le -\sum_{j=0}^{d-1} \lambda_j \log \lambda_j.$$

Proof. Proposition 2 Given $\rho \in \mathcal{A}$ we can write it as the convex linear combination $\rho = \sum_k \pi_k \rho_k$, $\rho_k = |f_k\rangle\langle f_k| \in \mathcal{A}$ such that

$$S(\rho_k) = -\sum_{j=0}^{d-1} \lambda_j \log \lambda_j.$$

Thus, the result follows from the definition of $H_\Phi(\rho)$. $\square$

Proposition 3 *Suppose that for $\rho \in \mathfrak{S}(\mathcal{H} \otimes \mathcal{K})$ the following inclusion holds,*

$$\mathrm{Tr}_{\mathcal{K}}(\rho) \in \mathcal{A}.$$

Then,

$$S((\Phi \otimes Id)(\rho)) \geq -\sum_{j=0}^{d-1} \lambda_j \log \lambda_j + \frac{1}{d} \sum_{j=0}^{d-1} S(\rho_j), \tag{11}$$

where $\rho_j = d \,\mathrm{Tr}_{\mathcal{H}}((|e_j\rangle\langle e_j| \otimes I_{\mathcal{K}})\rho) \in \mathfrak{S}(\mathcal{K})$.

Proof. Proposition 3 The proof treads [9] steps. Let us take $\rho \in \mathfrak{S}(\mathcal{H} \otimes \mathcal{K})$ such that $\mathrm{Tr}_{\mathcal{K}}(\rho) \in \mathcal{A}$ and define a quantum channel $\Xi_\rho : \mathfrak{S}(\mathcal{H} \otimes \mathcal{K}) \to \mathfrak{S}(\mathcal{H} \otimes \mathcal{K})$ by the formula

$$\Xi_\rho(\sigma) := \sum_{j=0}^{d-1} \mathrm{Tr}((|e_j\rangle\langle e_j| \otimes I_{\mathcal{K}})\sigma)(V^j \otimes I_{\mathcal{K}})\rho(V^{*j} \otimes I_{\mathcal{K}}),$$

with $\sigma \in \mathfrak{S}(\mathcal{H} \otimes \mathcal{K})$. Then, let

$$\sigma = \sum_{j=0}^{d-1} \lambda_j |e_j\rangle\langle e_j| \otimes y,$$

$$\overline{\sigma} = \sum_{j=0}^{d-1} \frac{1}{d} |e_j\rangle\langle e_j| \otimes y \equiv \frac{1}{d} I_{\mathcal{H}} \otimes y,$$

with $y \in \mathfrak{S}(\mathcal{K})$ an arbitrary fixed state. It follows that

$$\Xi_\rho(\sigma) = (\Phi \otimes Id)(\rho),$$

$$\Xi_\rho(\overline{\sigma}) = \frac{1}{d} \sum_{j=0}^{d-1} (V^j \otimes I_{\mathcal{K}})\rho(V^{*j} \otimes I_{\mathcal{K}}) := \tilde{E}(\rho).$$

Here and throughout the paper Id denotes the identity map. Also notice that $\tilde{E} = (E \otimes Id)$ is the conditional expectation to algebra of the elements being fixed with respect to the action of the cyclic group $\{V^j \otimes I_{\mathcal{K}}, \ 0 \leq j \leq d-1\}$.

Now, on the one hand, Theorem 1 gives us

$$S\left(\Xi_\rho(\sigma)\middle\|\Xi_\rho(\overline{\sigma})\right) \leq S(\sigma\|\overline{\sigma}) = \sum_{j=0}^{d-1} \lambda_j \log \lambda_j + \log d. \tag{12}$$

On the other hand, it is

$$\begin{aligned}
S\left(\Xi_\rho(\sigma)\middle\|\Xi_\rho(\overline{\sigma})\right) &= \mathrm{Tr}((\Phi \otimes Id)(\rho) \log(\Phi \otimes Id)(\rho)) - \mathrm{Tr}((\Phi \otimes Id)(\rho) \log \tilde{E}(\rho)) \\
&= -S((\Phi \otimes Id)(\rho)) - \mathrm{Tr}(\tilde{E} \circ (\Phi \otimes Id)(\rho) \log \tilde{E}(\rho)) \\
&= -S((\Phi \otimes Id)(\rho)) + S(\tilde{E}(\rho)).
\end{aligned} \tag{13}$$

In the above equations, we have used the equality $\tilde{E} \circ (\Phi \otimes Id) = \tilde{E}$ which holds because $\tilde{E}$ is the conditional expectation to the algebra of elements being fixed with respect to the action of $\Phi \otimes Id$.

Since, using $E(\mathrm{Tr}_{\mathcal{K}}(\rho)) = \frac{1}{d}Id$,

$$\tilde{E}(\rho) = \frac{1}{d}\sum_{j=0}^{d-1}|e_j\rangle\langle e_j| \otimes \rho_j, \quad \rho_j \in \mathfrak{S}(\mathcal{K}),$$

it follows

$$S(\tilde{E}(\rho)) = \log d + \frac{1}{d}\sum_{j=0}^{d-1}S(\rho_j), \tag{14}$$

with $\rho_j = d\,\mathrm{Tr}_{\mathcal{H}}((|e_j\rangle\langle e_j| \otimes I_{\mathcal{K}})\rho)$, $0 \le j \le d-1$. Then, combining (12), (13) and (14) we get the result of the proposition 3. $\square$.

We can now single out a class of input states for which the phase damping channels respect a kind of superadditivity property.

Theorem 4 *Suppose that $\rho \in \mathfrak{S}(\mathcal{H} \otimes \mathcal{K})$ is such that*

$$\mathrm{Tr}_{\mathcal{K}}(\rho) \in \mathcal{A}.$$

Let Φ be the phase damping channel (8), then the inequality

$$\begin{aligned}
S((\Phi \otimes \Omega)(\rho)) &\ge -\sum_{j=0}^{d-1}\lambda_j \log \lambda_j + H_\Omega(\mathrm{Tr}_{\mathcal{H}}(\rho)) \\
&\ge H_\Phi(\mathrm{Tr}_{\mathcal{K}}(\rho)) + H_\Omega(\mathrm{Tr}_{\mathcal{H}}(\rho)),
\end{aligned}$$

holds for an arbitrary quantum channel $\Omega : \mathfrak{S}(\mathcal{K}) \to \mathfrak{S}(\mathcal{K})$.

Proof. Theorem 4 Defining $\tilde{\rho} := (Id \otimes \Omega)(\rho)$, we notice that $\mathrm{Tr}_{\mathcal{K}}(\tilde{\rho}) \in \mathcal{A}$ and

$$S((\Phi \otimes \Omega)(\rho)) = S((\Phi \otimes Id)(\tilde{\rho})). \tag{15}$$

Applying the Proposition 3 we obtain

$$S((\Phi \otimes \Omega)(\rho)) \ge -\sum_{j=0}^{d-1}\lambda_j \log \lambda_j + \frac{1}{d}\sum_{j=0}^{d-1}S(\rho_j), \tag{16}$$

where $\rho_j = d\,\mathrm{Tr}_{\mathcal{H}}((|e_j\rangle\langle e_j| \otimes I_{\mathcal{K}})(Id \otimes \Omega)(\rho)) \in \mathfrak{S}(\mathcal{K})$. Using Proposition 2 we can rewrite (16) as

$$S((\Phi \otimes \Omega)(\rho)) \ge H_\Phi(\mathrm{Tr}_{\mathcal{K}}(\rho)) + \frac{1}{d}\sum_{j=0}^{d-1}S(\rho_j).$$

Finally, taking into account that $\frac{1}{d}\sum_{j=0}^{d-1}\rho_j = \Omega(\mathrm{Tr}_{\mathcal{H}}(\rho))$, we obtain

$$\sum_{j=0}^{d-1}S(\rho_j) \ge H_\Omega(\mathrm{Tr}_{\mathcal{H}}(\rho)).$$

The result of the theorem 4 then follows. $\square$.

5. Estimation of the Output Entropy for the Weyl Channels

Let us consider an orthonormal basis $|k\rangle$, $k = 0, 1, \ldots, d-1$ of the Hilbert space $\mathcal{H}$ of dimension d and define the unitary operators

$$U_{m,n} := \sum_{k=0}^{d-1} e^{\frac{2\pi i}{d} kn} |k \oplus m\rangle\langle k|, \tag{17}$$

where $0 \le m, n \le d - 1$ and $\oplus$ denotes the sum modulus d. The operators (17) satisfy the Weyl commutation relations

$$U_{m,n} U_{m',n'} = e^{2\pi i(m'n - mn')/d} U_{m',n'} U_{m,n}, \tag{18}$$

hence, we shall call them Weyl operators. Notice that

$$U_{m,0}|k\rangle = |k \oplus m\rangle, \quad U_{0,n}|k\rangle = e^{\frac{2\pi i}{d} kn}|k\rangle. \tag{19}$$

We shall consider bistochastic quantum channels of the following form

$$\Phi(\rho) := \sum_{m,n=0}^{d-1} \pi_{m,n} U_{m,n}\, \rho\, U_{m,n}^*, \tag{20}$$

where $\{\pi_{m,n}\}_{m,n=0}^{d-1}$ are probability distributions and $\rho \in \mathfrak{S}(\mathcal{H})$ states. The channels (20) are called Weyl channels.

Now, let us fix positive numbers $0 \le p_n, r_m \le 1$, $1 \le n \le d - 1$, $0 \le m \le d - 1$ such that $d \sum_{n=1}^{d-1} p_n + \sum_{m=0}^{d-1} r_m = 1$ and let us consider the Weyl channel

$$\Phi(\rho) = \sum_{m=0}^{d-1} r_m U_{m,0}\, \rho\, U_{m,0}^* + \sum_{m=0}^{d-1}\sum_{n=1}^{d-1} p_n U_{m,n}\, \rho\, U_{m,n}^*, \tag{21}$$

$\rho \in \mathfrak{S}(\mathcal{H})$.

It is shown in [9] that the channels (21) is covariant with respect to the maximum commutative group of unitary operators. Moreover, if the dimension of the space d is a prime number, the following decomposition holds

$$\Phi(\rho) = \sum_{k=0}^{d-1}\sum_{m=0}^{d-1} c_m U_{m,0} \Psi_k(\rho) U_{m,0}^*, \tag{22}$$

where $\rho \in \mathfrak{S}(\mathcal{H})$ and

$$\Psi_k(\rho) = \sum_{n=0}^{d-1} \lambda_n U_{nk \bmod d, n}\, \rho\, U_{nk \bmod d, n}^*,$$

are phase damping channels. Furthermore, it is

$$\begin{aligned}
\lambda_0 &= 1 - d\sum_{n=1}^{d-1} p_n, \\
\lambda_n &= dp_n, \quad 1 \le n \le d-1, \\
c_m &= \frac{r_m}{d\left(1 - d\sum_{n=1}^{d-1} p_n\right)}, \quad 0 \le m \le d-1.
\end{aligned}$$

We can now single out a wide class (over the totality) of input states for which the Weyl channels (21) respect a kind of superadditivity property.

Let us denote by $\mathcal{A}$ the maximum commutative algebra generated by the projectors $|k\rangle\langle k|$, $0 \le k \le d - 1$. Notice that the states $\rho \in \mathcal{A}$ are mutually unbiased with respect to the eigenvectors of the unitary operators $U_{nk,n}$, $0 \le k, n \le d - 1$ [9]. Then, the following theorem holds.

Theorem 5 *Let the dimension d of the space $\mathcal{H}$ be a prime number. Suppose that $\rho \in \mathfrak{S}(\mathcal{H} \otimes \mathcal{K})$ is such that*

$$\mathrm{Tr}_{\mathcal{K}}(\rho) \in \mathcal{A}.$$

Let Φ be the Weyl channel (21), then the inequality

$$S((\Phi \otimes \Omega)(\rho)) \ge H_{\Phi}(\mathrm{Tr}_{\mathcal{K}}(\rho)) + H_{\Omega}(\mathrm{Tr}_{\mathcal{H}}(\rho)),$$

holds for an arbitrary quantum channel $\Omega : \mathfrak{S}(\mathcal{K}) \to \mathfrak{S}(\mathcal{K})$.

Proof. Theorem 5 Using the decomposition (22) we easily arrive at

$$S((\Phi \otimes \Omega)(\rho)) \ge \frac{1}{d} \sum_{k=0}^{d-1} S((\Psi_k \otimes \Omega)(\rho)). \tag{23}$$

Then, by applying Theorem 4 to each term of the right hand side of (23) we obtain the result of Theorem 5. $\square$

6. Quantum Channels Respecting the Strong Superadditivity

We shall provide hereafter a class of quantum channels that fully respect the strong superadditivity, i.e. without any restriction on the input states.

6.1. *The quantum noiseless channel*

The quantum noiseless channel in the Hilbert space $\mathcal{H}$ of the dimension d is simply defined as the identity operation

$$\Phi(\rho) := Id(\rho) = \rho, \tag{24}$$

with $\rho \in \mathfrak{S}(\mathcal{H})$.

Theorem 6 *Let Φ be the quantum noiseless channel of Eq.(24), then the inequality*

$$H_{\Phi \otimes \Omega}(\rho) \ge H_{\Omega}(\mathrm{Tr}_{\mathcal{H}}(\rho)), \tag{25}$$

holds for an arbitrary quantum channel $\Omega : \mathfrak{S}(\mathcal{K}) \to \mathfrak{S}(\mathcal{K})$.

Proof. Theorem 6 Actually this theorem was proved in [15]. Our prove is alternative and based upon the decreasing property of the relative entropy. Let us take the optimal ensemble $\{\rho_k\}$ such that

$$H_{\Phi \otimes \Omega}(\rho) = \sum_k \pi_k S((\Phi \otimes \Omega)(\rho_k)).$$

Given a state $\rho_k \in \mathfrak{S}(\mathcal{H} \otimes \mathcal{K})$, the identity channel can be considered as the phase damping channel Φ with $\lambda_0 = 1$, $\lambda_j = 0$, $1 \le j \le d - 1$ (see (8)), for which the state $\mathrm{Tr}_{\mathcal{K}}(\rho) \in \mathcal{A}$,

where $\mathcal{A}$ is the convex set generated by pure states unbiased with respect to the basis of eigenvectors of the unitary operator determining Φ. Hence, the result follows from Theorem 4. $\square$.

6.2. *The quantum-classical channel*

Let $\{M_j,\ 1 \leq j \leq d\}$ be a resolution of the identity in $\mathcal{H}$ consisting of positive operators $M_j > 0,\ \sum_{j=1}^{d} M_j = I_{\mathcal{H}}$. The quantum channel Φ is said to be a quantum-classical channel (shortly q-c channel) if there exists an orthogonal basis $\{|e_j\rangle\}$ in $\mathcal{H}$ such that [4]:

$$\Phi(\rho) = \sum_{j=1}^{d} \mathrm{Tr}(M_j \rho)|e_j\rangle\langle e_j|. \tag{26}$$

The additivity for quantum-classical channels was proved in [4] and the result was extended to the entanglement-breaking channels in [22].

Theorem 7 *Let Φ be the q-c channel (26), then the inequality*

$$H_{\Phi \otimes \Omega}(\rho) \geq H_{\Phi}(\mathrm{Tr}_{\mathcal{K}}(\rho)) + H_{\Omega}(\mathrm{Tr}_{\mathcal{H}}(\rho)), \tag{27}$$

holds for an arbitrary quantum channel $\Omega : \mathfrak{S}(\mathcal{K}) \to \mathfrak{S}(\mathcal{K})$.

To prove the theorem we need of the following lemma.

Lemma 1 *Let Φ be the q-c channel (26). Then, given a state $\rho \in \mathfrak{S}(\mathcal{H} \otimes \mathcal{K})$,*

$$S((\Phi \otimes Id)(\rho)) \geq S(\Phi(\mathrm{Tr}_{\mathcal{K}}(\rho))) + \sum_{j=1}^{d} \lambda_j S(\rho_j), \tag{28}$$

where $\lambda_j = \mathrm{Tr}(M_j \mathrm{Tr}_{\mathcal{K}}(\rho)),\ \rho_j = \frac{1}{\lambda_j}\mathrm{Tr}_{\mathcal{H}}((M_j \otimes I_{\mathcal{K}})\rho) \in \mathfrak{S}(\mathcal{K})$.

Proof. Lemma 1 Let us define a quantum channel $\Sigma_\rho : \mathfrak{S}(\mathcal{H}) \to \mathfrak{S}(\mathcal{H} \otimes \mathcal{K})$ by the formula

$$\Sigma_\rho(\sigma) := \sum_{j=1}^{d} \mathrm{Tr}(|e_j\rangle\langle e_j|\sigma)|e_j\rangle\langle e_j| \otimes \rho_j, \tag{29}$$

where the states $\rho_j \in \mathfrak{S}(\mathcal{K})$ are the same as in the formulation of the Lemma 1. One can see that

$$\Sigma_\rho(\Phi(\mathrm{Tr}_{\mathcal{K}}(\rho))) = (\Phi \otimes Id)(\rho), \tag{30}$$

$$\Sigma_\rho\left(\frac{1}{d}I_{\mathcal{H}}\right) = \frac{1}{d}\sum_{j=1}^{d}|e_j\rangle\langle e_j| \otimes \rho_j. \tag{31}$$

The decreasing property of the relative entropy, Theorem 1, gives us

$$S\left(\Sigma_\rho(\Phi(\mathrm{Tr}_{\mathcal{K}}(\rho)))\,\middle\|\,\Sigma_\rho(\tfrac{1}{d}I_{\mathcal{H}})\right) \leq S\left(\Phi(\mathrm{Tr}_{\mathcal{K}}(\rho))\,\middle\|\,\tfrac{1}{d}I_{\mathcal{H}}\right). \tag{32}$$

Taking into account Eq.(30) and (31) we get (using same technique as proof of Proposition 3)

$$S\left(\Phi(\mathrm{Tr}_{\mathcal{K}}(\rho))\,\middle\|\,\frac{1}{d}I_{\mathcal{H}}\right) = \log d - S\left(\Phi(\mathrm{Tr}_{\mathcal{K}}(\rho))\right), \tag{33}$$

and

$$S\left(\Sigma_\rho(\Phi(\mathrm{Tr}_\mathcal{K}(\rho)))\,\Big\|\,\Sigma_\rho(\tfrac{1}{d}I_\mathcal{H})\right) = \log d + \sum_{j=1}^d \lambda_j S(\rho_j) - S((\Phi \otimes Id)(\rho)),$$

from which the result of Lemma 1 follows. $\square$.

Proof. Theorem 7 Let Φ be the q-c channel (26). Suppose that Ω is an arbitrary channel and

$$\rho = \sum_{j=1}^k p_j \rho_j, \tag{34}$$

such that the states ρ_j, $1 \le j \le k$, form the optimal ensemble for the output entropy of $\Phi \otimes \Omega$, i.e.

$$H_{\Phi \otimes \Omega}(\rho) = \sum_j p_j S((\Phi \otimes \Omega)(\rho_j)). \tag{35}$$

Applying Lemma 1 to each term in the sum on the right hand side we get

$$H_{\Phi \otimes \Omega}(\rho) \ge \sum_j p_j S(\Phi(\mathrm{Tr}_\mathcal{K}(\rho_j))) + \sum_j p_j \sum_{k=1}^d \lambda_{jk} S(\Omega(\rho_{jk})),$$

where $\lambda_{jk} = \mathrm{Tr}(M_k \mathrm{Tr}_\mathcal{K}(\rho_j))$ and $\rho_{jk} = \frac{1}{\lambda_{jk}} \mathrm{Tr}_\mathcal{H}((M_k \otimes I_\mathcal{K})\rho_j) \in \mathfrak{S}(\mathcal{K})$. By the definitions (34) and (4) we obtain on the one hand

$$\sum_j p_j S(\Phi(\mathrm{Tr}_\mathcal{K}(\rho_j))) \ge H_\Phi(\mathrm{Tr}_\mathcal{K}(\rho)). \tag{36}$$

On the other hand,

$$\sum_j p_j \sum_{k=1}^d \lambda_{jk} \Omega(\rho_{jk}) = \Omega(\mathrm{Tr}_\mathcal{H}(\rho)). \tag{37}$$

The last formula implies that

$$\sum_j p_j \sum_{k=1}^d \lambda_{jk} S(\Omega(\rho_{jk})) \ge H_\Omega(\mathrm{Tr}_\mathcal{H}(\rho)). \tag{38}$$

Then the result of Theorem 7 follows. $\square$.

Notice that a q-c channel is a partial case of the entanglement-breaking channels considered in [15]. So our proof is alternative to the one given in [15] for entanglement-breaking channels.

6.3. *The quantum erasure channel*

Let $\mathcal{H}$ and $\mathcal{H}'$ be Hilbert spaces of dimension d and $d+1$ respectively. We claim that $\mathcal{H} \subset \mathcal{H}'$ which results in the inclusion $\mathfrak{S}(\mathcal{H}) \subset \mathfrak{S}(\mathcal{H}')$. Suppose that $|\omega\rangle \in \mathcal{K}$ is orthogonal to $\mathcal{H}$. Fix ϵ such that $0 \le \epsilon \le 1$, then we call quantum erasure channel the CP-map $\Phi : \mathfrak{S}(\mathcal{H}) \to \mathfrak{S}(\mathcal{H}')$ defined by

$$\Phi(\rho) := \epsilon |\omega\rangle\langle\omega| + (1 - \epsilon)\rho, \tag{39}$$

with $\rho \in \mathfrak{S}(\mathcal{H})$. Notice that this is a generalization to dimension d of the qubit erasure channel introduced in [23].

Theorem 8 *Let Φ be the erasure channel (39), then the inequality*

$$H_{\Phi \otimes \Omega}(\rho) \geq H_{\Phi}(\mathrm{Tr}_{\mathcal{K}}(\rho)) + H_{\Omega}(\mathrm{Tr}_{\mathcal{H}}(\rho)), \tag{40}$$

holds for an arbitrary quantum channel $\Omega : \mathfrak{S}(\mathcal{K}) \to \mathfrak{S}(\mathcal{K})$.

To prove the theorem we need of the following lemma.

Lemma 2 *Let Φ be the quantum erasure channel (39). Then, given a state $\rho \in \mathfrak{S}(\mathcal{H} \otimes \mathcal{K})$*

$$S((\Phi \otimes Id)(\rho)) \geq \epsilon S(\mathrm{Tr}_{\mathcal{H}}(\rho)) + (1 - \epsilon)S(\rho) + S(\Phi(\mathrm{Tr}_{\mathcal{K}}(\rho))).$$

Proof. Lemma 2 Denote by $P_{\mathcal{H}}$ the orthogonal projection in $\mathcal{H}'$ onto the subspace $\mathcal{H}$. Given $\rho \in \mathfrak{S}(\mathcal{H} \otimes \mathcal{K})$ let us define a quantum channel $\Sigma_{\rho} : \mathfrak{S}(\mathcal{H}') \to \mathfrak{S}(\mathcal{H}' \otimes \mathcal{K})$ by the formula

$$\Sigma_{\rho}(\sigma) := \mathrm{Tr}(|\omega\rangle\langle\omega|\sigma)|\omega\rangle\langle\omega| \otimes \mathrm{Tr}_{\mathcal{H}}(\rho) + \mathrm{Tr}(P_{\mathcal{H}}\sigma)\rho, \tag{41}$$

with $\sigma \in \mathfrak{S}(\mathcal{H}')$.

Pick up the orthogonal projection $|e\rangle\langle e|$ from the spectral decomposition of the state $\mathrm{Tr}_{\mathcal{K}}(\rho)$. One can see that

$$\Sigma_{\rho}(\Phi(\mathrm{Tr}_{\mathcal{K}}(\rho))) = (\Phi \otimes Id)(\rho), \tag{42}$$

$$\Sigma_{\rho}\left(\frac{1}{2}|\omega\rangle\langle\omega| + \frac{1}{2}|e\rangle\langle e|\right) = \frac{1}{2}|\omega\rangle\langle\omega| \otimes \mathrm{Tr}_{\mathcal{H}}(\rho) + \frac{1}{2}\rho. \tag{43}$$

The decreasing property of the relative entropy, Theorem 1, gives us

$$S\left(\Sigma_{\rho}\left(\Phi(\mathrm{Tr}_{\mathcal{K}}(\rho))\right) \middle\| \Sigma_{\rho}\left(\frac{1}{2}|\omega\rangle\langle\omega| + \frac{1}{2}|e\rangle\langle e|\right)\right) \leq S\left(\Phi(\mathrm{Tr}_{\mathcal{K}}(\rho)) \middle\| \frac{1}{2}|\omega\rangle\langle\omega| + \frac{1}{2}|e\rangle\langle e|\right).$$

Taking into account (42) and (43) we get

$$
\begin{aligned}
S\left(\Phi(\mathrm{Tr}_{\mathcal{K}}(\rho)) \middle\| \frac{1}{2}|\omega\rangle\langle\omega| + \frac{1}{2}|e\rangle\langle e|\right) &= (\epsilon + (1 - \epsilon)\langle e|\mathrm{Tr}_{\mathcal{K}}(\rho)|e\rangle)\log d - S(\Phi(\mathrm{Tr}_{\mathcal{K}}(\rho))) \\
&\leq \log d - S(\Phi(\mathrm{Tr}_{\mathcal{K}}(\rho))),
\end{aligned}
$$

and

$$S\left(\Sigma_{\rho}(\Phi(\rho)) \middle\| \Sigma_{\rho}(\frac{1}{2}|\omega\rangle\langle\omega| + \frac{1}{2}|e\rangle\langle e|)\right) = \log d + \epsilon S(\mathrm{Tr}_{\mathcal{H}}(\rho)) + (1 - \epsilon)S(\rho) - S((\Phi \otimes Id)(\rho)).$$

The result of Lemma 2 then follows. $\square$.

Proof. Theorem 8 Let Φ be the erasure channel (39). Suppose that Ω is an arbitrary channel and

$$\rho = \sum_{j=1}^{k} p_j \rho_j \tag{44}$$

is such that the states ρ_j, $1 \leq j \leq k$, form the optimal ensemble for for the output entropy of $\Phi \otimes \Omega$, i.e.

$$H_{\Phi \otimes \Omega}(\rho) = \sum_{j} p_j S((\Phi \otimes \Omega)(\rho_j)) = \sum_{j} p_j S((\Phi \otimes Id)(\tilde{\rho}_j)), \tag{45}$$

with $\tilde{\rho}_j = (Id \otimes \Omega)(\rho_j)$. Applying Lemma 2 to each term in the sum on the right hand side of the above equation we get

$$H_{\Phi \otimes \Omega}(\rho) \geq \sum_j \ p_j \ \left[\epsilon S(\Omega(\mathrm{Tr}_{\mathcal{H}}(\rho_j))) + (1 - \epsilon) S((Id \otimes \Omega)(\rho_j)) + S(\Phi(\mathrm{Tr}_{\mathcal{K}}(\rho_j))) \right].$$

Notice also that

$$\sum_j p_j S((Id \otimes \Omega)(\rho_j)) \geq H_{Id \otimes \Omega}(\rho) \geq H_{\Omega}(\mathrm{Tr}_{\mathcal{H}}(\rho))$$

because the strong superadditivity conjecture holds for the noiseless channel [15]. Then, the result of Theorem 8 follows.   □.

6.4. *The quantum depolarizing channel*

The quantum depolarizing channel in the Hilbert space $\mathcal{H}$ of dimension d is defined as [17]

$$\Phi(\rho) := (1 - p)\rho + \frac{p}{d} I_{\mathcal{H}}, \tag{46}$$

with $\rho \in \mathfrak{S}(\mathcal{H})$, $0 \leq p \leq d^2/(d^2 - 1)$.

Theorem 9 *Let Φ be the quantum depolarizing channel (46), then the inequality*

$$H_{\Phi \otimes \Omega}(\rho) \geq H_{\Phi}(\mathrm{Tr}_{\mathcal{K}}(\rho)) + H_{\Omega}(\mathrm{Tr}_{\mathcal{H}}(\rho)), \tag{47}$$

holds for an arbitrary quantum channel $\Omega : \mathfrak{S}(\mathcal{K}) \to \mathfrak{S}(\mathcal{K})$.

To prove Theorem 9 we need of some properties of the quantum depolarizing channel.

Following Ref.[8], by choosing an orthonormal basis $\{|f_j\rangle\}$ in $\mathcal{H}$, we can define a set of orthonormal bases $\{\{|e_j^k\rangle\}_{j=0}^{d-1}\}_{k=1}^{2d^2}$ as

$$|e_j^k\rangle := \sum_{s=0}^{d-1} \exp\left(i\frac{2\pi s^2 k}{2d^2}\right) \exp\left(i\frac{2\pi j}{d}\right) |f_s\rangle, \tag{48}$$

with $1 \leq k \leq 2d^2$. Moreover, let

$$\begin{aligned} U &:= \sum_{s=0}^{d-1} \exp\left(i\frac{2\pi s}{d}\right) |f_s\rangle\langle f_s|, \\ V_k &:= \sum_{s=0}^{d-1} \exp\left(i\frac{2\pi s}{d}\right) |e_s^k\rangle\langle e_s^k|, \end{aligned}$$

be unitary operators in $\mathcal{H}$. We introduce phase damping channels as follows

$$\Psi_k(\rho) = \left(1 - \frac{d-1}{d}p\right)\rho + \frac{p}{d}\sum_{s=1}^{d-1} V_k^s \rho V_k^s, \tag{49}$$

with $\rho \in \mathfrak{S}(\mathcal{H})$, $0 \leq p \leq 1$ and $1 \leq k \leq 2d^2$.

Then, the quantum depolarizing Φ can be expressed in terms of the above phase damping channels as

$$
\begin{aligned}
\Phi(\rho) \quad = \quad & \frac{1-p}{1+(d-1)(1-p)} \frac{1}{2d} \sum_{k=1}^{2d^2} \Psi_k(\rho) \\
+ \quad & \frac{p}{1+(d-1)(1-p)} \frac{1}{2d^3} \sum_{j=1}^{d-1} \sum_{k=1}^{2d^2} U^j \Psi_k(\rho) U^{*j},
\end{aligned}
$$

$$(50)$$

with $\rho \in \mathfrak{S}(\mathcal{H})$. By defining

$$
E_k(\rho) := \frac{1}{d} \sum_{s=0}^{d-1} V_k^s \rho U_k^{*s}, \tag{51}
$$

the conditional expectations on the algebras of fixed elements for the phase dampings Ψ_k, we have

$$
E_k(|f_j\rangle\langle f_j|) = \frac{1}{d} I_{\mathcal{H}}, \tag{52}
$$

for $1 \le k \le 2d^2$, $0 \le j \le d-1$. This property guarantees that the basis $\{|f_j\rangle\}$ is mutually unbiased with respect to all the bases $\{|e^k\rangle\}$ defined by (48).

Proof. Theorem 9 Let us take the optimal ensemble corresponding to the state ρ such that

$$
H_{\Phi\otimes\Omega}(\rho) = \sum_s \pi_s S((\Phi \otimes \Omega)(\rho_s)).
$$

In the following we shall estimate $S((\Phi \otimes \Omega)(\rho_s))$ for each fixed s.

Let us consider ϱ instead of a ρ_s. Let us choose a unitary operator T such that the state

$$
\tilde\varrho = (T \otimes I_{\mathcal{K}})(Id \otimes \Omega)(\varrho)(T^* \otimes I_{\mathcal{K}}), \tag{53}
$$

satisfies the property

$$
E_k(\mathrm{Tr}_{\mathcal{K}}(\tilde\varrho)) = \frac{1}{d} I_{\mathcal{H}}. \tag{54}
$$

Using the covariance property $\Phi(\sigma) = T^*\Phi(T\sigma T^*)T$, taking place for all states $\sigma \in \mathfrak{S}(\mathcal{H})$, we can rewrite the decomposition (50) as follows

$$
\begin{aligned}
\Phi(\sigma) \quad = \quad & \frac{1-p}{1+(d-1)(1-p)} \frac{1}{2d} \sum_{k=1}^{2d^2} \tilde\Psi_k(\sigma) \\
+ \quad & \frac{p}{1+(d-1)(1-p)} \frac{1}{2d^3} \sum_{j=1}^{d-1} \sum_{k=1}^{2d^2} T^*U^jT\tilde\Psi_k(\sigma)TU^{*j}T^*,
\end{aligned}
$$

$$(55)$$

where $\tilde\Psi_k(\sigma) = T^*\Psi_k(T\sigma T^*)T$ are the phase damping channels with the property

$$
\mathrm{Tr}(\tilde E_k(\mathrm{Tr}_{\mathcal{K}}(\varrho))) = \frac{1}{d} I_{\mathcal{H}}. \tag{56}
$$

Here $\tilde{E}_k(\sigma) = T^* E_k(T\sigma T^*)T$, $\varrho \in \mathfrak{S}(\mathcal{H} \otimes \mathcal{K})$ and $\sigma \in \mathfrak{S}(\mathcal{H})$. The above equality guarantees that the state ϱ is unbiased with respect to all the orthonormal bases which form the unitary operators determining the action of the phase damping channels Ψ_k, $0 \le k \le d-1$.

It follows from the decomposition (55) that

$$S((\Phi \otimes \Omega)(\varrho)) \ge \frac{1}{d} \sum_{k=0}^{d-1} S((\tilde{\Psi}_k \otimes \Omega)(\varrho)). \tag{57}$$

Applying Theorem 4 to each term of the sum in the right hand side and taking into account that $-\sum_{j=0}^{d-1} \lambda_j \log \lambda_j = H_\Phi(\mathrm{Tr}_\mathcal{K}(\varrho))$ for $\lambda_0 = 1 - \frac{d-1}{d}p$, $\lambda_j = \frac{p}{d}$, $1 \le j \le d-1$ due to (56), we get

$$S((\Phi \otimes \Omega)(\rho_s)) \ge H_\Phi(\mathrm{Tr}_\mathcal{K}(\rho_s)) + H_\Omega(\mathrm{Tr}_\mathcal{H}(\rho_s)), \tag{58}$$

hence the result of Theorem 9. $\square$.

7. Conclusion

By using the decreasing property of the relative entropy, we have proved the strong super-additivity for a class of quantum channels. This class includes the channels for which the property was already shown by using other methods (thus giving an alternative proof) as well as others channels (thus providing an extension of the class).

We guess that the decreasing property of the relative entropy could be a powerful tool for a further extension of such class of channels. More generally, it could constitute a universal method to investigate relevant properties of memoryless quantum channels. In fact, as a consequence of the strong superadditivity property we get the additivity property. Thus for our class of channels, the additivity results automatically proved.

The perspective of determining the broadest class of channels for which additivity holds through strong superadditivity seems fascinating and motivate further investigations, especially in consideration of the limits of other methods [11, 12].

Acknowledgments

References

1. V. Vedral (2002), *The role of relative entropy in quantum information theory*, Rev. Mod. Phys. vol. 74. pp. 197-234.
2. G. Lindblad (1975), *Completely positive maps and entropy inequalities*, Comm. Math. Phys. vol. 40, pp. 147-151.
3. A. S. Holevo (1972), *On the mathematical theory of quantum communication channels*, Probl. Inf. Transm. vol. 8, pp. 62-71.
4. A. S. Holevo (1998), *Quantum coding theorems*, Russ. Math. Surveys vol. 53, pp. 1295-1331.
5. B. Schumacher and M.D. Westmoreland (1997), *Sending classical information via noisy quantum channels*, Phys. Rev. A vol. 56, pp. 131-138.

6. T. M. Cover and J. A. Thomas (1991), *Elements of Information Theory*, Wiley-Interscience Publication (New York).

7. G. G. Amosov, A. S. Holevo and R. F. Werner (2000), *On some additivity problems in quantum information theory*, Probl. Inf. Transm. vol. 36, pp. 305-313.

8. C. King (2003), *The capacity of the quantum depolarizing channel*, IEEE Trans. Inf. Th. vol. 49, pp. 221-229.

9. G. G. Amosov (2007), *On the Weyl channels being covariant with respect to the maximum commutative group of unitaries*, J. Math. Phys. vol. 48, pp. 012104-01–012104-14.

10. G. G. Amosov (2006), *Remark on the additivity conjecture for the depolarizing quantum channel*, Probl. Inf. Transm. vol. 42, pp. 69-76.

11. P. Hayden (2007), *The maximal p-norm multiplicativity conjecture is false*, http://arxiv.org/abs/0707.3291.

12. P. Hayden, and A. Winter (2008), *Counterexamples to the maximal p-norm multiplicativity conjecture for all p¿1*, http//arxiv.org/abs/0807.4753.

13. M. B. Hastings (2008), *A counterexample to additivity of minimum output entropy*, http://arxiv.org/abs/0809.3972.

14. P. Shor (2004), *Equivalence of additivity questions in quantum information theory*, Commun. Math. Phys. vol. 246, pp. 453-472.

15. A. S. Holevo and M. E. Shirokov (2004), *On Shor's channel extension and constrained channels*, Commun. Math. Phys. vol. 249, pp. 417-430.

16. K. Matsumoto, T. Shimono and A. Winter (2004), *Remarks on additivity of the Holevo channel capacity and of the entanglement of formation*, Comm. Math. Phys. vol. 246, pp. 427-442.

17. G. G. Amosov (2007), *The strong superadditivity conjecture holds for a quantum depolarizing channel in any dimension*, Phys. Rev. A vol. 75, pp. 060304-1–060304-2.

18. M. Ohya and D. Petz (1993), *Quantum Entropy and Its Use, Texts and Monographs in Physics*, Springer-Verlag (Berlin).

19. H. Umegaki (1962), *Conditional expectation in an operator algebra. IV. Entropy and information*, Kodai Math. Sem. Rep. vol. 14, pp. 59-85.

20. F. Hiai and D. Petz (1991), *The proper formula for relative entropy and its asymptotics in quantum probability*, Comm. Math. Phys. vol. 143, pp. 99-114.

21. I. D. Ivanovich (1981), *Geometrical description of quantum state determination*, J. Phys. A vol. 14, pp. 3241-3245.

22. P. Shor (2002), *Additivity of the classical capacity of entanglement-breaking quantum channels*, J. Math. Phys. vol. 43, pp. 4334-4340.

23. C. H. Bennett, D. P. DiVincenzo and J. A. Smolin (1997), *Capacities of quantum erasure channels*, Phys. Rev. Lett. vol. 78, pp. 3217-3220.