Entanglement Assisted Classical Capacity for a Class of Quantum Channels with Long-Term Memory

Nilanjana Datta and Yurii Suhov
Statistical Laboratory
Centre for Mathematical Sciences
University of Cambridge
Wilberforce Road, Cambridge CB30WB
email: n.datta@statslab.cam.ac.uk

Tony C. Dorlas
Dublin Institute for Advanced Studies
School of Theoretical Physics
10 Burlington Road, Dublin 4, Ireland.
email: dorlas@stp.dias.ie

Keywords: quantum channels with long-term memory, entanglement assisted classical capacity

Abstract

In this paper we evaluate the entanglement assisted classical capacity of a class of quantum channels with long-term memory, which are convex combinations of memoryless channels. The memory of such channels can be considered to be given by a Markov chain which is aperiodic but not irreducible. This class of channels was introduced in [7] and its product state capacity was evaluated.

1 Introduction

The biggest hurdle in the path of efficient information transmission is the presence of noise, in both classical and quantum channels. This noise causes a distortion of the information sent through the channel. Error–correcting codes are used to overcome this problem. Instead of transmitting the original messages, they are encoded into codewords, which are then sent through the channel. Information transmission is said to be reliable if the probability of error, in decoding the output of the channel, vanishes asymptotically in the number of uses of the channel (see e.g. [4] and [21]). The aim is to achieve reliable transmission, whilst optimizing the rate, i.e., the ratio between the size of the message and its correponding codeword. The optimal rate of reliable transmission is referred to as the capacity of the the channel.

A classical communications channel has a unique capacity, the formula for which was obtained by Shannon in 1948. A quantum channel, in contrast, has various distinct capacities. This is because there is flexibility in the use of a quantum channel. The particular definition of the capacity which is applicable, depends on the following: (*i*) whether the information transmitted is classical or quantum; (*ii*) whether the sender[1], Alice, is allowed to use inputs *entangled* over various uses of the channel or whether she is only allowed to use product inputs; (*iii*) whether the receiver, Bob is allowed to make collective measurements over various outputs of the channel or whether he is only allowed to measure the output of each channel use separately; (*iv*) whether Alice and Bob have additional resources e.g. prior shared entanglement.

The different capacities resulting from the different choices mentioned above were evaluated intitially for memoryless[2] quantum channels. The capacity of a quantum memoryless channel for transmitting classical information, obtained under the restriction that the inputs are product states and collective measurements on the outputs, is referred to as the product state (classical) capacity of the channel. The formula for this capacity is given by the Holevo-Schumacher-Westmoreland (HSW) Theorem [15, 27]. The formula for the quantum capacity of a memoryless channel, i.e., its capacity for transmitting quantum information was established through a series of papers [24, 19, 26, 10, 12]. The maximum asymptotic rate of reliable transmission of classical information with the help of unlimited prior entanglement between the sender and the receiver is known as entanglement assisted capacity. The formula for this was first proved by Bennett, Shor, Smolin and Thapliyal

[1]We follow the normal conventiona and refer to the sender as Alice, and the receiver receiving it as Bob

[2]For such a channel, the noise affecting successive input states, is assumed to be perfectly uncorrelated.

[5, 6] and the proof was later simplified by Holevo [16].

For real world communication channels, the assumption that noise is uncorrelated between successive uses of a channel cannot be justified. Hence memory effects need to be taken into account. This leads us to the consideration of quantum channels with memory. The first model of such a channel was studied by Macchiavello and Palma [20]. They showed that the transmission of classical information through two successive uses of a quantum depolarising channel, with Markovian correlated noise, is enhanced by using inputs entangled over the two uses. An important class of quantum channels with memory consists of the so-called *forgetful channels*. The channel studied in [20] falls in this class. Roughly speaking, a forgetful channel is one for which the output after a large number of successive uses, does not depend on the initial input state. Forgetful channels have been studied by Bowen and Mancini [3] and more recently by Kretschmann and Werner [18]. In [18], coding theorems for arbitrary forgetful channels were proved. The proof of the direct channel coding theorem for a class of quantum channels with Markovian correlated noise, where the underlying Markov Chain was aperiodic and irreducible, was sketched out in [8]. Recently Bjelaković and Boche [2] have proved a coding theorem for causal ergodic classical-quantum channels with decaying input memory.

The capacities of channels with long-term memory (i.e., channels which are "not forgetful"), had remained an open problem until recently. In [7], the classical capacity of a class of quantum channels with long-term memory, which are given by convex combinations of memoryless channels, was evaluated. In this paper we evaluate the entanglement-assisted classical capacity of the same class of channels. For a channel Φ in this class, $\Phi^{(n)} : \mathcal{B}(\mathcal{H}^{\otimes n}) \rightarrow \mathcal{B}(\mathcal{K}^{\otimes n})$ and the action of $\Phi^{(n)}$ on any state $\rho^{(n)} \in \mathcal{B}(\mathcal{H}^{\otimes n})$ is given as follows:

$$\Phi^{(n)}(\rho^{(n)}) = \sum_{i=1}^{M} \gamma_i \Phi_i^{\otimes n}(\rho^{(n)}), \quad (1)$$

where $\Phi_i : \mathcal{B}(\mathcal{H}) \rightarrow \mathcal{B}(\mathcal{K})$, $(i = 1, \ldots, M)$ are completely positive, trace-preserving (CPT) maps and $\gamma_i > 0$, $\sum_{i=1}^{M} \gamma_i = 1$. Here $\mathcal{H}$ and $\mathcal{K}$ denote Hilbert spaces. On using the channel, an initial random choice is made as to which memoryless channel the successive input states are transmitted through. A classical version of such a channel was introduced by Jacobs [17] and studied further by Ahlswede [1], who obtained an expression for its capacity which is analogous to one obtained in [7].

Note that the memory of the class of channels that we study, can be considered to be given by a Markov chain which is aperiodic but not irreducible. This can be seen as follows. A quantum channel (of length n) with

Markovian correlated noise is a CPT map $\Phi^{(n)} : \mathcal{B}(\mathcal{H}^{\otimes n}) \to \mathcal{B}(\mathcal{K}^{\otimes n})$ defined as follows

$$\Phi^{(n)}(\rho^{(n)}) = \sum_{i_1,\ldots,i_n} q_{i_n|i_{n-1}} \ldots q_{i_2|i_1} \gamma_{i_1} (\Phi_{i_1} \otimes .. \otimes \Phi_{i_n})(\rho^{(n)}),$$

Here (i) $q_{j|i}$ denote the elements of the transition matrix of a discrete–time Markov chain with a finite state space I; (ii) $\{\gamma_i\}$, denotes the invariant distribution of the chain, and (iii) for each $i \in I$, $\Phi_i : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{K})$ is a CPT map. Casting the channel defined by (1). in this form yields $q_{j|i} = \delta_{ij}$. Hence the transition matrix of the Markov chain, in this case, is the identity matrix. In other words, once a particular branch $i = 1, \ldots, M$ has been chosen, the successive inputs are sent through this branch. The Markov chain is therefore aperiodic but not irreducible.

We start the main body of our paper with some preliminaries in Section 2. Our main result, giving the expression for the entanglement-assisted classical capacity of the channels in question, is stated as a theorem in Section 3. The proofs of the converse and direct parts of this theorem our given in Sections 3.1 and 3.2 respectively. In proving the direct part of the theorem we make use of the expression for the product state capacity of the channel, which was obtained in [7].

2 Preliminaries

Let $\mathcal{B}(\mathcal{H})$ denote the algebra of linear operators acting on a finite–dimensional Hilbert space $\mathcal{H}$. The von Neumann entropy of a state ρ, i.e., a positive operator of unit trace in $\mathcal{B}(\mathcal{H})$, is defined as $S(\rho) = -\mathrm{Tr}\,\rho \log \rho$, where the logarithm is taken to base 2. A quantum channel is given by a completely positive trace–preserving (CPT) map $\Phi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{K})$, where $\mathcal{H}$ and $\mathcal{K}$ are the input and output Hilbert spaces of the channel. Let $\dim \mathcal{H} = d$ and $\dim \mathcal{K} = d'$. For any ensemble $\{p_j, \rho_j\}$ of states ρ_j and probability distributions $\{p_j\}$, the Holevo χ quantity is defined as

$$\chi(\{p_j, \rho_j\}) := S\left(\sum_j p_j\, \rho_j\right) - \sum_j p_j\, S(\rho_j). \tag{2}$$

The Holevo capacity of a memoryless quantum channel Φ is given by

$$\chi^*(\Phi) := \max_{\{p_j,\rho_j\}} \chi\big(\{p_j, \Phi(\rho_j)\}\big), \tag{3}$$

where the maximum is taken over all ensembles $\{p_j, \rho_j\}$ of possible input states $\rho_j \in \mathcal{B}(\mathcal{H})$ occurring with probabilities p_j. It is known that the maximum in (3) can be achieved by using an ensemble of pure states, and that it suffices to restrict the maximum to ensembles of at most d^2 pure states.

3 Main Result

As mentioned in the Introduction, in this paper we evaluate the entanglement-assisted classical capacity of a class of channels with long-term memory defined by (1).

Consider the following protocol for the entanglement-assisted transmission of classical information through the quantum channel defined by (1). Suppose two parties, Alice and Bob, share indefinitely many copies of an entangled pure state $\rho^{AB} = |\psi^{AB}\rangle\langle\psi^{AB}| \in \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)$. Here the system $A\,(B)$, with Hilbert space $\mathcal{H}_A\,(\mathcal{H}_B)$ is in Alice's (Bob's) possession and $\dim \mathcal{H}_A = \dim \mathcal{H}_B$. Suppose Alice has a set of messages, labelled by the elements of the set $\mathcal{M}_n = \{1, 2, \ldots, M_n\}$, which she would like to communicate via the quantum channel (1) to Bob, exploiting this shared entanglement. For this purpose she has an ensemble $\{\pi_j, \mathcal{E}_j\}$ of completely positive trace-preserving (CPT) encoding maps $\mathcal{E}_j$ acting on $\mathcal{B}(\mathcal{H}_A)$ chosen with probabilities π_j. In order to transmit her classical messages through the quantum channel, Alice encodes each of her messages in a quantum state in $(\mathcal{H}_A \otimes \mathcal{H}_B)^{\otimes n}$ in the following manner. To each $\alpha \in \mathcal{M}_n$ she assigns a quantum state (or codeword)

$$\rho_\alpha^{AB;n} := \rho_{j_1(\alpha)}^{AB} \otimes \ldots \otimes \rho_{j_n(\alpha)}^{AB} \in \mathcal{B}((\mathcal{H}_A \otimes \mathcal{H}_B)^{\otimes n}). \tag{4}$$

Here $j_k(\alpha) = j$ with probability π_j independently for every $k = 1, 2, \ldots, n$, and

$$\rho_j^{AB} = \big(\mathcal{E}_j \otimes id_B\big)\rho^{AB}, \tag{5}$$

where id_B denotes the identity map in $\mathcal{B}(\mathcal{H}_B)$. Thus, the probability of assigning the codeword $\rho_\alpha^{AB;n}$ to the message α is $\pi_{j_1(\alpha)} \ldots \pi_{j_n(\alpha)}$, which generates an ensemble of quantum encodings. Note that the codewords are states shared between Alice and Bob. Alice then sends her part of these shared states to Bob through n subsequent uses of the quantum channel (1). Consequently, Alice's attempt to send the classical message α to Bob results in him having the state

$$\sigma_\alpha^{AB;n} := \big(\Phi^{(n)} \otimes id_B^{\otimes n}\big)\rho_\alpha^{AB;n}. \tag{6}$$

In order to infer the message that Alice communicated to him, Bob makes a measurement (described by a set of POVM elements) on the state $\sigma_\alpha^{AB;n}$.

The encoding and decoding operations, employed to achieve reliable transmission of information by means of this protocol, together define an ensemble of quantum codes. More precisely, in this case a sample code $\mathcal{C}^{(n)}$ of size M_n is given by a sequence $\{\rho_\alpha^{AB;n}, F_\alpha^{B;n}\}_{\alpha=1}^{M_n}$ where each $\rho_\alpha^{AB;n}$ is a state in $\mathcal{B}((\mathcal{H}_A \otimes \mathcal{H}_B)^{\otimes n})$ of the form (4) and $F_\alpha^{AB;n}$ is a positive operator acting in $\mathcal{H}_A \otimes \mathcal{H}_B)^{\otimes n}$, such that $\sum_{\alpha=1}^{M_n} F_\alpha^{AB;n} \leq I^{AB;n}$. Here $I^{AB;n}$ denotes the identity operator in $\mathcal{B}((\mathcal{H}_A \otimes \mathcal{H}_B)^{\otimes n})$. Defining $F_0^{AB;n} = I^{AB;n} - \sum_{\alpha=1}^{M_n} F_\alpha^{AB;n}$ yields a resolution of identity in $(\mathcal{H}_A \otimes \mathcal{H}_B)^{\otimes n}$. Hence, $\{F_\alpha^{AB;n}\}_{i=0}^{M_n}$ defines a POVM. An output $\beta \in \mathcal{M}_n$ would lead Bob to conclude that the state (or codeword) was $\rho_\beta^{AB;n}$, whereas the output 0 is interpreted as a failure of any inference. Assuming equidistribution of messages, the average probability of error for the sample code $\mathcal{C}^{(n)}$ is given by

$$P_e(\mathcal{C}^{(n)}) := \frac{1}{M_n} \sum_{\alpha=1}^{M_n} \Big(1 - \mathrm{Tr}\left((\Phi_A^{(n)} \otimes \mathrm{id}_B^{\otimes n})(\rho_\alpha^{AB;n}) F_\alpha^{AB;n}\right)\Big), \tag{7}$$

where $\mathrm{id}_B^{(n)}$ denotes the identity operator in $\mathcal{B}(\mathcal{H}_B^{\otimes n})$.

The expected error probability

$$\overline{P_e^{(n)}} = \mathbb{E} P_e(\mathcal{C}^{(n)}) = \frac{1}{M_n} \sum_{\alpha=1}^{M_n} \Big(1 - \mathbb{E}\mathrm{Tr}\left((\Phi_A^{(n)} \otimes \mathrm{id}_B^{\otimes n})(\rho_\alpha^{AB;n}) F_\alpha^{AB;n}\right)\Big), \tag{8}$$

is obtained by further averaging over the sample codes $\mathcal{C}^{(n)}$. If for a given $R > 0$ there exists a sequence of M_n's with

$$\lim_{n\to\infty} \frac{1}{n} \log M_n = R,$$

such that

$$\lim_{n\to\infty} \overline{P_e^{(n)}} = 0,$$

then R is said to be an *achievable* rate under the choice of the ensemble $\{\pi_j, \mathcal{E}_j\}$ and the initial shared state ρ^{AB}.

The *one-shot* entanglement-assisted classical capacity is then defined as

$$C_{ea}^{(1)}(\Phi) := \sup_{\{\pi_j, \mathcal{E}_j\}, \rho^{AB}} \sup\big[R : R \text{ achievable}\big], \tag{9}$$

where the internal supremum is over the rates achievable under the choice of the ensemble $\{\pi_j, \mathcal{E}_j\}$ and the initial shared state ρ^{AB}.

The same construction can be performed for m-shot ensembles $\{\pi_j^{(m)}, \mathcal{E}_j^{(m)}\}$ where $\mathcal{E}_j^{(m)}$ is a CPT encoding map in $\mathcal{B}(\mathcal{H}_A^{\otimes n})$ chosen with probabilities $\pi_j^{(m)}$.

In this case Alice uses block encoding, and a message $\alpha \in \mathcal{M}_n$ is encoded by the state

$$\rho_{\alpha|m}^{AB;n} := \rho_{j_1(\alpha)|m}^{AB} \otimes \ldots \otimes \rho_{j_n(\alpha)|m}^{AB} \in \mathcal{B}((\mathcal{H}_A \otimes \mathcal{H}_B)^{\otimes mn})$$

where $j_k(\alpha) = i$ with probability $\pi_j^{(m)}$ independently for every $k = 1, 2, \ldots, n$, and

$$\rho_{j|m}^{AB} = \big(\mathcal{E}_j^{(m)} \otimes id_B^{\otimes m}\big)(\rho^{AB})^{\otimes m}.$$

As before, Bob uses decoding POVM elements $F_{\alpha|m}^{AB;n}$.

The average probability of error for the resultant sample code (which we denote by $\mathcal{C}_m^{(n)}$) is given by

$$P_e(\mathcal{C}_m^{(n)}) := \frac{1}{M_n} \sum_{\alpha=1}^{M_n} \Big(1 - \mathrm{Tr}\left((\Phi_A^{(mn)} \otimes \mathrm{id}_B^{\otimes mn})(\rho_{\alpha|m}^{AB;n}) F_{\alpha|m}^{AB;n}\right)\Big), \tag{10}$$

which is then averaged over the sample codes generated by the ensemble $\{\pi_j^{(m)}, \mathcal{E}_j^{(m)}\}$ to yield

$$\overline{P_e^{(n)}} = \mathbb{E} P_e(\mathcal{C}_m^{(n)}) = \frac{1}{M_n} \sum_{\alpha=1}^{M_n} \Big(1 - \mathbb{E}\mathrm{Tr}\left((\Phi_A^{(mn)} \otimes \mathrm{id}_B^{\otimes mn})(\rho_{\alpha|m}^{AB;n}) F_\alpha^{AB|mn}\right)\Big), \tag{11}$$

This gives rise to the *m-shot* entanglement-assisted classical capacity of Φ :

$$C_{ea}^{(m)}(\Phi) := \sup_{\{\pi_j^{(m)}, \mathcal{E}_j^{(m)}\}, \rho^{AB}} \sup\big[R : R \text{ achievable}\big]. \tag{12}$$

Finally, the full entanglement-assisted classical capacity of Φ is given by

$$C_{ea}(\Phi) := \limsup_{m \to \infty} \frac{1}{m} C_{ea}^{(m)}(\Phi) \tag{13}$$

Our main result is given by the following theorem.

Theorem 3.1 *The entanglement assisted classical capacity of a channel Φ, with long-term memory, defined through (1), is given by*

$$C_{ea}(\Phi) = \max_\rho \left[\bigwedge_{i=1}^{M} I(\rho; \Phi_i)\right],$$

with $I(\rho; \Phi_i) := S(\rho) + S(\Phi_i(\rho)) - S(\rho; \Phi_i)$, where $S(\rho) = -\mathrm{Tr}\, \rho \log \rho$ is the von Neumann entropy of ρ and $S(\rho; \Phi_i)$ is the entropy exchange, defined as follows:

$$S(\rho; \Phi_i) := S\big((\Phi_i^A \otimes \mathrm{id}^R)\psi_\rho^{AR})\big), \tag{14}$$

with ψ_ρ^{AR} being a purification of ρ on R.

Here we use the standard notation $\bigwedge$ to denote the minimum.

3.1 Proof of the converse part of Theorem 3.1

In this section we prove that for any rate $R > C_{ea}(\Phi)$, reliable transmission of entanglement-assisted transmission of classical information from Alice to Bob via the quantum channel Φ (eq.(1)) is impossible, regardless of the encoding used.

Recall that under our protocol, Alice and Bob share multiple copies of an entangled bipartite pure state $\rho^{AB} = |\psi^{AB}\rangle\langle\psi^{AB}| \in \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)$. Then, given $m, n \in \mathbb{Z}$, Alice encodes her classical messages by applying chosen m-block encoding CPT maps, n times to her part of the shared state $\rho_{AB}^{\otimes mn}$. Here we show that the average error probability of the corresponding code, as defined in (7), does not tend to zero as $n \to \infty$, for any m and any choice of encoding maps. Henceforth, for notational simplicity, the index m will be omitted from the subscripts.

Let $\sigma_{\alpha,i}^{AB;n}$ denote Bob's final state, if Alice sends the message α, and her corresponding codeword

$$\rho_\alpha^{AB;n} = \rho_{\alpha,1}^{AB} \otimes \ldots \otimes \rho_{\alpha,n}^{AB}$$

is transmitted through the i-th branch of the channel. Also let

$$\sigma_\alpha^{AB;n} := \sum_{i=1}^{M} \gamma_i \sigma_\alpha^{AB;n}(i) \quad ; \quad \bar{\sigma}^{AB;n}(i) = \frac{1}{|\mathcal{M}_n|} \sum_{\alpha \in \mathcal{M}_n} \sigma_\alpha^{AB;n}(i)$$

$$\text{and} \quad \bar{\sigma}_k(i) = \frac{1}{|\mathcal{M}_n|} \sum_{\alpha \in \mathcal{M}_n} \sigma_{\alpha,k}(i) \ ,$$

where $\sigma_{\alpha,k}(i) = (\Phi_i \otimes id_B)\rho_{\alpha k}^{AB}$.

Then the average probability of error (7) equals

$$\bar{p}_e^{(n)} := 1 - \frac{1}{|\mathcal{M}_n|} \sum_{\alpha \in \mathcal{M}_n} \mathrm{Tr}\ \left[\sigma_\alpha^{AB;n} F_\alpha^{AB;n}\right]. \tag{15}$$

We also define the average probability of error corresponding to the i^{th} branch of the channel as

$$\bar{p}_{i,e}^{(n)} := 1 - \frac{1}{|\mathcal{M}_n|} \sum_{\alpha \in \mathcal{M}_n} \mathrm{Tr}\ \left[\sigma_\alpha^{AB;n}(i) F_\alpha^{AB;n}\right] \quad \text{so that} \ \ \bar{p}_e^{(n)} = \sum_{i=1}^{M} \gamma_i \bar{p}_{i,e}^{(n)} \tag{16}$$

Let $X^{(n)}$ be a random variable with a uniform distribution over the set $\mathcal{M}_n$, characterizing the classical message sent by Alice to Bob. Let $Y_i^{(n)}$ be the random variable corresponding to Bob's inference of Alice's message, when the codeword is transmitted through the i^{th} branch of the channel. It is defined by the conditional probabilities

$$\mathbb{P}\,[Y_i^{(n)} = \beta \,|\, X^{(n)} = \alpha] = \mathrm{Tr}\ [\big(\Phi_i^{\otimes n} \otimes id_B^{\otimes n}\big)(\rho_\alpha^{AB;n})F_\beta^{AB;n}]. \tag{17}$$

By Fano's inequality,

$$h(\bar{p}_{i,e}^{(n)}) + \bar{p}_{i,e}^{(n)} \log(|\mathcal{M}_n| - 1) \geq H(X^{(n)} \,|\, Y_i^{(n)}) = H(X^{(n)}) - H(X^{(n)} \,:\, Y_i^{(n)}). \tag{18}$$

Here $h(\cdot)$ denotes the binary entropy and $H(\cdot)$ denotes the Shannon entropy. Using the Holevo bound and the subadditivity of the von Neumann entropy we have

$$\begin{aligned}
H(X^{(n)} \,:\, Y_i^{(n)}) &\leq S\left(\frac{1}{|\mathcal{M}_n|}\sum_{\alpha\in\mathcal{M}_n} \sigma_\alpha^{AB;n}(i)\right) - \frac{1}{|\mathcal{M}_n|}\sum_{\alpha\in\mathcal{M}_n} S\left(\sigma_\alpha^{AB;n}(i)\right) \\
&\leq \sum_{k=1}^{n}\left[S\left(\bar{\sigma}_k^{AB;n}(i)\right) - \frac{1}{|\mathcal{M}_n|}\sum_{\alpha\in\mathcal{M}_n} S\left(\sigma_{\alpha,k}^{AB}(i)\right)\right] \\
&= \sum_{k=1}^{n}\chi\left(\left\{\frac{1}{|\mathcal{M}_n|}, \sigma_{\alpha,k}^{AB}(i)\right\}_{\alpha\in\mathcal{M}_n}\right) \\
&= \sum_{k=1}^{n}\frac{1}{|\mathcal{M}_n|}\sum_{\alpha\in\mathcal{M}_n} S\left(\sigma_{\alpha,k}^{AB}(i) \,||\, \bar{\sigma}_k^{AB}(i)\right) := \sum_{k=1}^{n} A_k.
\end{aligned} \tag{19}$$

The expression A_k can be rewritten using Donald's identity:

$$\sum_\alpha p_\alpha S(\omega_\alpha \,||\, \rho) = \sum_\alpha p_\alpha S(\omega_\alpha \,||\, \bar{\omega}) + S(\bar{\omega} \,||\, \rho), \tag{20}$$

where $\bar{\omega} = \sum_\alpha p_\alpha \omega_\alpha$. We apply this with ρ replaced by

$$\bar{\sigma}^{AB}(i) = \frac{1}{n|\mathcal{M}_n|}\sum_{k=1}^{n}\sum_{\alpha\in\mathcal{M}_n} \sigma_{\alpha,k}^{AB}(i) \tag{21}$$

ω_α replaced by $\sigma_{\alpha,k}^{AB}(i)$, p_α replaced by $1/|\mathcal{M}_n|$, and consequently $\bar{\omega}$ replaced by $\bar{\sigma}_k^{AB}(i)$. Hence,

$$\begin{aligned}
\frac{1}{|\mathcal{M}_n|}\sum_{\alpha\in\mathcal{M}_n} S\left(\sigma_{\alpha,k}^{AB}(i) \,||\, \bar{\sigma}_k^{AB}(i)\right) &\leq \frac{1}{|\mathcal{M}_n|}\sum_{\alpha\in\mathcal{M}_n} S(\sigma_{\alpha,k}^{AB}(i) \,||\, \bar{\sigma}_i^{AB}) - S(\bar{\sigma}_k^{AB}(i) \,||\, \bar{\sigma}^{AB}(i)) \\
&\leq \frac{1}{|\mathcal{M}_n|}\sum_{\alpha\in\mathcal{M}_n} S(\sigma_{\alpha,k}^{AB}(i) \,||\, \bar{\sigma}^{AB}(i)),
\end{aligned} \tag{22}$$

where we have used the non-negativity of the von Neumann entropy. Inserting into (19) we now have:

$$\begin{aligned}\frac{1}{n}H(X^{(n)}:Y_i^{(n)}) &\leq \frac{1}{n|\mathcal{M}_n|}\sum_{k=1}^{n}\sum_{\alpha\in\mathcal{M}_n} S(\sigma_{\alpha,k}^{AB}(i)\,||\,\bar{\sigma}^{AB}(i)) \\ &= \chi\left(\left\{\frac{1}{n|\mathcal{M}_n|},\sigma_{\alpha,k}^{AB}(i)\right\}_{(\alpha,k)}\right). \end{aligned} \tag{23}$$

The inequality (18) now yields (cf. eq.(17) of [16])

$$\begin{aligned} h(\bar{p}_{i,e}^{(n)})+\bar{p}_{i,e}^{(n)}\log|\mathcal{M}_n| &\geq \log|\mathcal{M}_n| - n\,\chi\left(\left\{\frac{1}{n|\mathcal{M}_n|},\sigma_{\alpha,k}^{AB}(i)\right\}_{(\alpha,k)}\right) \\ &\geq \log|\mathcal{M}_n| - n\,I(\rho_A,\Phi_i), \end{aligned} \tag{24}$$

where

$$\rho_A := \sum_{\alpha,k} p_{\alpha,k}\,\rho_{\alpha,k}^A,$$

with $p_{\alpha,k} := \dfrac{1}{n|\mathcal{M}_n|}$ for each α and k, and $\rho_{\alpha,k}^A = \mathrm{Tr}_B\big(\rho_{\alpha,k}^{AB}\big)$. However, since

$$C_{ea}(\Phi) \geq \bigwedge_{i=1}^{M}\chi_i\left(\left\{\frac{1}{n|\mathcal{M}_n|},\rho_{\alpha,k}\right\}_{(\alpha,k)}\right) \tag{25}$$

and $R = \frac{1}{n}\log|\mathcal{M}_n| > C_{ea}(\Phi)$, there must be at least one branch i such that

$$\bar{p}_{i,e}^{(n)} \geq 1 - \frac{C_{ea}(\Phi)+1/n}{R} > 0. \tag{26}$$

We conclude from (16) and (26) that

$$\bar{p}_e^{(n)} \geq \left(1-\frac{C_{ea}(\Phi)+1/n}{R}\right)\bigwedge_{i=1}^{M}\gamma_i. \tag{27}$$

Hence $\bar{p}_{i,e}^{(n)}$ does not tend to zero as $n\to\infty$. □

3.2 Proof of the direct part of Theorem 3.1

In this section we prove that $C_{ea}(\Phi)$, defined by (13) satisfies the lower bound

$$C_{ea}(\Phi) \geq \max_\rho\left[\bigwedge\nolimits_{i=1}^{M} I(\rho;\Phi_i)\right] \tag{28}$$

To prove this we employ the following result which we proved in [7]

Theorem 3.2 *The product state capacity of a channel Φ, with long-term memory, defined through (1), is given by*

$$C(\Phi) = \sup_{\{\pi_j,\rho_j\}} \Big[\bigwedge_{i=1}^{M} \chi_i(\{\pi_j, \rho_j\}) \Big] , \tag{29}$$

where $\chi_i(\{\pi_j, \rho_j\}) := \chi\left(\{p_j, \Phi_i(\rho_j)\}\right)$. The supremum is taken over all finite ensembles of states $\rho_j \in \mathcal{B}(\mathcal{H})$ with probabilities π_j.

Here the notation $\bigwedge$ denotes the minimum.

From the definition (9) of the one-shot entanglement assisted capacity and (29) it follows that

$$C_{ea}^{(1)}(\Phi) = \sup_{\{\pi_j,\mathcal{E}_j\},\rho^{AB}} \Big[\bigwedge_{i=1}^{M} \chi(\{\pi_j, \big(\Phi_i \otimes id_B\big)\rho_j^{AB}\}) \Big] , \tag{30}$$

where (i) ρ^{AB} is the bipartite entangled pure state, indefinitely many copies of which are shared by Alice and Bob and (ii) $\mathcal{E}_j$ are encoding maps acting on $\mathcal{B}(\mathcal{H}_A)$, as described in Section 3.

Moreover, from the definition (12) of the m-shot entanglement assisted capacity it follows that

$$C_{ea}^{(m)}(\Phi) = \sup_{\{\pi_j^{(m)},\mathcal{E}_j^{(m)}\},\rho_j^{AB|m}} \Big[\bigwedge_{i=1}^{M} \chi(\{\pi_j^{(m)}, \big(\Phi_i^{\otimes m} \otimes id_B^{\otimes m}\big)\rho_j^{AB|m}\}) \Big] . \tag{31}$$

Now, following [16] and [6], consider a specific encoding ensemble $\{\pi_{(a,b)}^{(m)}, \mathcal{E}_{(a,b)}^{(m)}\}$, where $a, b = 1, 2, \ldots, q$, and

$$\pi_{(a,b)}^{(m)} = \frac{1}{q^2} \quad ; \quad \mathcal{E}_{(a,b)}^{(m)} = W_{a,b}^{(m)},$$

the discrete Weyl-Segal operators (see e.g.[16]) for a q-dimensional subspace $\mathcal{Q}_m$ of $\mathcal{H}_A^{\otimes m}$. Further set

$$\varrho_{a,b}^{AB|m} = \big(W_{a,b}^{(m)} \otimes id_B^{\otimes m}\big)(|\psi_m^{AB}\rangle\langle\psi_m^{AB}|),$$

with $|\psi_m^{AB}\rangle$ being a maximally entangled state in $\mathcal{Q}_m$. Hence,

$$C_{ea}^{(m)}(\Phi) \geq \bigwedge_{i=1}^{M} \chi(\{\frac{1}{q^2}, \big(\Phi_i^{\otimes m} \otimes id_B^{\otimes m}\big)\varrho_{a,b}^{AB|m}\}) \tag{32}$$

From [16] it follows that

$$\chi(\{\frac{1}{q^2}, \big(\Phi_i^{\otimes m} \otimes id_B^{\otimes m}\big)\varrho_{a,b}^{AB|m}\}) = I\big(\frac{P^{(m)}}{\mathrm{Tr}\,(P^{(m)})}; \Phi_i^{\otimes m}\big), \tag{33}$$

where $P^{(m)}$ is the orthoprojection onto $\mathcal{Q}_m$. Further, it was proved in [16] that if $\mathcal{Q}_m$ is chosen to be the strongly δ-typical subspace for an arbitrary state $\rho^{\otimes n}$ in $\mathcal{B}(\mathcal{H}_A^{\otimes m})$, and $P^{m,\delta}$ is its orthoprojection, then

$$\lim_{\delta \to 0} \lim_{m \to \infty} \frac{1}{m} I\Big(\frac{P^{m,\delta}}{\mathrm{Tr}\,(P^{m,\delta})}; \Phi_i^{\otimes m}\Big), = I(\rho; \Phi_i) \tag{34}$$

From (32), (33), (34) and the definition (13) of the full entanglement-assisted capacity, it follows that

$$C_{ea}(\Phi) \geq \bigwedge\nolimits_{i=1}^{M} I(\rho; \Phi_i). \tag{35}$$

□

References

[1] R. Ahlswede, "The Weak Capacity of Averaged Channels", Z. Wahrscheinlichkeitstheorie verw. Geb. **11**, 61–73 (1968).

[2] I. Bjelaković and H. Boche, "Ergodic Classical-Quantum Channels: Structure and Coding Theorems", *quant-ph/0609229.*

[3] G. Bowen and S. Mancini, "Quantum channels with a finite memory", *Phys. Rev. A* **69**, 01236, 2004.

[4] T. M. Cover and J. A. Thomas, *Elements of Information Theory,* John Wiley & Sons, Inc.

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

[6] C.H.Bennett, P.W.Shor, J.A.Smolin, A.V.Thapliyal, "Entanglement-assisted capacity of a quantum channel and the reverse Shannon theorem", *quant-ph/0106052*

[7] N. Datta and T. C. Dorlas, "Coding Theorem for a class of quantum channels with long term memory", *quant-ph/0610049.* To appear in a special issue of the *Jour. of Phys.* **A** on Quantum Information Theory,

[8] N.Datta and T.Dorlas, "A Quantum Version of Feinstein's Lemma and its application to Channel Coding", *Proc. of Int. Symp. Inf. Th. ISIT 2006, Seattle,* 441-445 (2006).

[9] N. Datta and T. C. Dorlas, "Classical capacity of quantum channels with general Markovian correlated noise", in preparation.

[10] I. Devetak, "The private classical capacity and quantum capacity of a quantum channel", *IEEE Trans. Inform. Theory* **51**, 44–55, 2005.

[11] M.J.Donald, Math. Proc. Cam. Phil. Soc. **101**, 363, 1987.

[12] P.Hayden, M.Horodecki, J.Yard and A.Winter, "A decoupling approach to the quantum capacity", *quant-ph/0702005.*

[13] M.Hayashi and H. Nagaoka, "General formulas for capacity of classical-quantum channels," *IEEE Trans. Inform. Theory* **49**, 1753–1768, 2003.

[14] A.S. Holevo, "Capacity of a quantum communications channel", Prob. of Inf. Trasm¿, **5**, 247-253, 1979.

[15] A.S. Holevo, "The capacity of a quantum channel with general signal states," *IEEE Trans. Info. Theory*, **44**, 269-273, 1998.

[16] A.S. Holevo, "On entanglement-assisted classical capacity" *quant-ph/0106075.*

[17] K.Jacobs, "Almost periodic channels." Colloqium on Comb. Methods in Prob. Theory. Ahrhus 1962.

[18] D. Kretschmann and R.F.Werner, "Quantum channels with memory," *quant-ph/0502106.*

[19] S. LLoyd, "Capacity of the noisy quantum channel", *Phys. Rev. A* **55**, 1613, 1996.

[20] C. Macchiavello and G. M. Palma, "Entanglement–enhanced information transmission over a quantum channel with correlated noise", *Phys. Rev. A* **65**, 050301, 2002.

[21] M.A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information,* Cambridge University Press, Cambridge, 2000.

[22] J.R.Norris, *Markov Chains, Cambridge Series in Statistical and Probabilistic Mathematics* Cambridge University Press, Cambridge, 1997.

[23] M. Ohya and D. Petz, *Quantum Entropy and Its Use*, Springer-Verlag, 1993.

[24] B. Schumacher, "Sending entanglement through noisy quantum channels", *Phys. Rev. A* **54**, 2614-2628, 1996.

[25] B. Schumacher, "Quantum Coding", *Phys. Rev. A* **51**, 2738-2747, 1995.

[26] P.W. Shor, The quantum channel capacity and coherent information, MRSI seminar, November 2002.

[27] B. Schumacher and M. D. Westmoreland, "Sending classical information via noisy quantum channels," *Phys. Rev. A* **56**, 131-138, 1997.