The product state capacity of a periodic channel

Nilanjana Datta
Statistical Laboratory
Centre for Mathematical Sciences
University of Cambridge
Wilberforce Road, Cambridge CB3 0WB, U.K.
email: n.datta@statslab.cam.ac.uk.

Teunis C. Dorlas
Dublin Institute for Advanced Studies
School of Theoretical Physics
10 Burlington Road, Dublin 4, Ireland.

August 7, 2008

Abstract

In this paper we develop a quantum version of Feinstein's Lemma and use it to give a new proof of the direct channel coding theorem for transmission of classical information through a periodic quantum channel, when the inputs to multiple uses of the channel are restricted to product states. Moreover, we also prove the analogue of the converse channel coding theorem for this class of channels.

1 Preliminaries

Let $\mathcal{B}(\mathcal{H})$ denote the algebra of linear operators acting on a finite–dimensional Hilbert space $\mathcal{H}$, and $\mathcal{S}(\mathcal{H})$ denote the set of all positive operators of unit trace in $\mathcal{B}(\mathcal{H})$, i.e., states (or density matrices). The von Neumann entropy of a state ρ is defined as $S(\rho) = -\mathrm{Tr}\,\rho \log \rho$, where the logarithm is taken to base 2. A general quantum channel is given by completely positive trace–preserving (CPT) maps $\Phi^{(n)} : \mathcal{B}(\mathcal{H}^{\otimes n}) \to \mathcal{B}(\mathcal{K}^{\otimes n})$, where $\mathcal{H}$ and $\mathcal{K}$ are the input and output Hilbert spaces of the channel. Here we consider a periodic channel of the following form:

$$\Phi^{(n)}(\rho^{(n)}) = \frac{1}{L}\sum_{i=0}^{L-1}(\Phi_i \otimes \Phi_{i+1} \otimes \cdots \otimes \Phi_{i+n-1})(\rho^{(n)}), \tag{1}$$

where we assume that a set of CPT maps $\Phi_i : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{K})$ $(i = 0, \ldots, L-1)$ is given, and the index is cyclic modulo the period L.

If we denote the Holevo quantity for the i-th branch by χ_i, i.e.

$$\chi_i(\{p_j, \rho_j\}) = S\left(\sum_j p_j \Phi_i(\rho_j)\right) - \sum_j p_j S(\Phi_i(\rho_j)),$$

then we shall prove that the product capacity of the channel (1) is given by

$$C(\Phi) = \sup_{\{p_j, \rho_j\}} \frac{1}{L}\sum_{i=0}^{L-1} \chi_i(\{p_j, \rho_j\}). \tag{2}$$

2 The Quantum Feinstein Lemma

The direct part of the theorem follows from

Theorem 1 *Given $\epsilon > 0$, there exists $n_0 \in \mathbf{N}$ such that for all $n \geq n_0$ there exists $N_n \geq 2^{n(C(\Phi)-\epsilon)}$ and there exist product states $\tilde{\rho}_1^{(n)}, \ldots, \tilde{\rho}_{N_n}^{(n)} \in \mathcal{S}(\mathcal{H}^{\otimes n})$ and positive operators $E_1^{(n)}, \ldots, E_{N_n}^{(n)} \in \mathcal{B}(\mathcal{K}^{\otimes n})$ such that $\sum_{k=1}^{N_n} E_k^{(n)} \leq \mathbf{I}_n$ and*

$$\mathrm{Tr}\,\Phi^{(n)}\left(\tilde{\rho}_k^{(n)}\right) E_k^{(n)} > 1 - \epsilon, \tag{3}$$

for each k.

Proof. We first construct a preamble to the code which serves to identify the first branch i chosen. To distinguish the initial branch, notice first of all that the corresponding CPT maps Φ_i need not all be distinct! However, we may assume that there is no internal periodicity of these maps; otherwise the channel be contracted to a single such period. This means, that for any two states $i, i' \in \{0, \ldots, L-1\}$ $(i < i')$ there exists $k \leq L-1$ such that $\Phi_{i+k} \neq \Phi_{i'+k}$. Then choose $\omega = \omega_{i,i'}$ such that

$$f := F(\Phi_{i+k}(\omega), \Phi_{i'+k}(\omega)) < 1. \tag{4}$$

In the following we write $\Phi_i^{(n)}$ for the branch of the channel with initial state i, i.e.

$$\Phi_i^{(n)}(\rho^{(n)}) = (\Phi_i \otimes \Phi_{i+1} \otimes \cdots \otimes \Phi_{i+n-1})(\rho^{(n)}). \tag{5}$$

Lemma 1 *For any $0 \le i < i' \le L-1$, let ω be a state as above. Then*

$$F\left(\Phi_i^{(mL)}(\omega^{\otimes mL}), \Phi_{i'}^{(mL)}(\omega^{\otimes mL})\right) \to 0 \tag{6}$$

as $m \to \infty$.

Proof.

$$\begin{aligned} & F\left(\Phi_i^{(mL)}(\omega^{\otimes mL}), \Phi_{i'}^{(mL)}(\omega^{\otimes mL})\right) \\ = & \left[F\left(\Phi_i^{(L)}(\omega^{\otimes L}), \Phi_{i'}^{(L)}(\omega^{\otimes m})\right)\right]^m \\ \le & \left[F(\Phi_{i+k}(\omega), \Phi_{i'+k}(\omega))\right]^m = f^m \to 0. \end{aligned} \tag{7}$$

□

We now introduce, for any pair of states σ, σ' on $\mathcal{K}$, and $\gamma, \gamma' > 0$, the difference operators

$$A_{\sigma,\sigma'}^{(M)} = \gamma\sigma^{\otimes M} - \gamma'(\sigma')^{\otimes M}. \tag{8}$$

Let $\Pi^{\pm}$ be the orthogonal projections onto the eigenspaces of $A_{\sigma,\sigma'}^{(M)}$ corresponding to all non-negative, and all negative eigenvalues, respectively. In [DD] we proved:

Lemma 2 *Suppose that for a given $\delta > 0$,*

$$|\mathrm{Tr}\,[|A_{\sigma,\sigma'}^{(M)}|] - (\gamma + \gamma')| \le \delta. \tag{9}$$

Then

$$|\mathrm{Tr}\,[\Pi^+(\sigma)^{\otimes M}] - 1| \le \frac{\delta}{2\gamma} \tag{10}$$

and

$$|\mathrm{Tr}\,[\Pi^-(\sigma')^{\otimes M}] - 1| \le \frac{\delta}{2\gamma'}. \tag{11}$$

To compare the outputs of all the different branches of the channel, we define projections $\tilde{\Pi}_i$ on the tensor product space $\bigotimes_{0\le i<i'<L}\mathcal{K}^{\otimes M} = \mathcal{K}^{\otimes ML_2}$ with $L_2 = \binom{L}{2}$ as follows:

$$\tilde{\Pi}_i = \bigotimes_{0\le i_1<i_2<L} \Gamma_{i_1,i_2}^{(i)}, \text{ where } \Gamma_{i_1,i_2}^{(i)} = \begin{cases} \mathbf{I}_{nM} & \text{if} \quad i_1 \ne i \text{ and } i_2 \ne i \\ \Pi_{i_1,i}^- & \text{if} \quad i_2 = i \\ \Pi_{i,i_2}^+ & \text{if} \quad i_1 = i. \end{cases} \tag{12}$$

Notice that it follows from the fact that $\Pi_{i,i'}^+\Pi_{i,i'}^- = 0$, that the projections $\tilde{\Pi}_i$ are also disjoint:

$$\tilde{\Pi}_i\tilde{\Pi}_{i'} = 0 \quad \text{for } i \ne i'. \tag{13}$$

It now follows easily with the help of the previous lemma and the inequalities [11]

$$\mathrm{Tr}\,(A_1) + \mathrm{Tr}\,(A_2) - 2F(A_1, A_2) \leq ||A_1 - A_2||_1 \leq \mathrm{Tr}\,(A_1) + \mathrm{Tr}\,(A_2) \qquad (14)$$

for any two positive operators A_1 and A_2, that these projections distinguish the relevant initial branches. Indeed, if we introduce the corresponding preamble state

$$\omega^{(ML_2)} = \left(\bigotimes_{i_1<i_2} \omega_{i_1,i_2}^{\otimes M} \right), \qquad (15)$$

then we have

Lemma 3 *For all $i \in \{0, \ldots, L-1\}$,*

$$\lim_{M\to\infty} \mathrm{Tr}\, \left[\tilde{\Pi}_i\, \Phi_i^{(ML_2)} \left(\omega^{(ML_2)} \right) \right] = 1. \qquad (16)$$

In the following we fix M so large that

$$\mathrm{Tr}\, \left[\tilde{\Pi}_i\, \Phi_i^{\otimes ML_2} \left(\omega^{(ML_2)} \right) \right] > 1 - \delta \qquad (17)$$

for all $i \in \{0, \ldots, L-1\}$. We also assume that M is a multiple of L so that (1) applies. The product state $\omega^{(ML_2)}$, defined through (15) is used as a preamble to the input state encoding each message, and serves to distinguish between the different branches, $\Phi_i^{(n)}$, of the channel. If $\rho_k^{(n)} \in \mathcal{B}(\mathcal{H}^{\otimes n})$ is a state encoding the k^{th} classical message in the set $\mathcal{M}_n$, then the k^{th} codeword is given by the product state

$$\omega^{(ML_2)} \otimes \rho_k^{(n)}.$$

Note that, since M is a multiple of L, the index of the first channel branch applying to ρ_k is also i.

Continuing with the proof of Theorem 1, let the maximum of the mean Holevo quantity $\frac{1}{L}\sum_{i=0}^{L-1} \chi_i$ be attained for an ensemble $\{p_j, \rho_j\}_{j=1}^{J(}$. Denote $\sigma_{i,j} = \Phi_i(\rho_j)$, $\bar{\sigma}_i = \sum_{j=1}^{J} p_j \Phi_i(\rho_j)$.

Choose $\delta > 0$. We will relate δ to ϵ at a later stage. Consider the typical subspaces $\overline{\mathcal{T}}_{i,\epsilon}^{(n)}$ of $\mathcal{K}^{\otimes n}$, with projection $\bar{P}_{i,n}$ such that if $\bar{\sigma}_i$ has a spectral decomposition

$$\bar{\sigma}_i = \sum_k \bar{\lambda}_{i,k} |\psi_{i,k}\rangle\langle\psi_{i,k}| \qquad (18)$$

then if $\underline{k} = (k_1, \ldots, k_n)$, $|\psi_{i,k_1}\rangle \otimes \cdots \otimes |\psi_{i,k_n}\rangle \in \overline{\mathcal{T}}^{(n)}_{i,\epsilon}$ if and only if

$$\left| \frac{1}{n} \sum_{j=1}^{n} \log \bar{\lambda}_{i,k_j} + S(\bar{\sigma}_i) \right| < \frac{\epsilon}{4}. \tag{19}$$

Then, for n large enough,

$$\mathrm{Tr}\,(\bar{P}_{i,n} \bar{\sigma}_i^{\otimes n}) > 1 - \delta^2. \tag{20}$$

For any given initial index i, we let $\overline{\mathcal{V}}^{(n)}_{i,\epsilon}$ be the subspace of $\mathcal{K}^{\otimes n}$ spanned by the vectors $|\psi_{i,k_1}\rangle \otimes |\psi_{i+1,k_2}\rangle \otimes |\psi_{i+n-1,k_n}\rangle$, where $|\psi_{i,k_1}\rangle \otimes |\psi_{i,k_{L+1}}\rangle \otimes \cdots \otimes |\psi_{i,k_{[(n-1)/L]L+1}}\rangle \in \overline{\mathcal{T}}^{([(n-1)/L]+1)}_{i,\epsilon}$, etc. Clearly, if we denote $\bar{P}^{(n)}_i$ the projection onto $\overline{\mathcal{V}}^{(n)}_{i,\epsilon}$, then for n large enough,

$$\mathrm{Tr}\,(\bar{P}^{(n)}_i \bar{\sigma}_i \otimes \bar{\sigma}_{i+1} \otimes \cdots \otimes \bar{\sigma}_{i+n-1}) > 1 - \delta^2. \tag{21}$$

Moreover, if $|\psi_{i,k_1}\rangle \otimes |\psi_{i+1,k_2}\rangle \otimes |\psi_{i+n-1,k_n}\rangle \in \overline{\mathcal{V}}^{(n)}_{i,\epsilon}$ then

$$\left| \frac{1}{n} \sum_{j=1}^{n} \log \bar{\lambda}_{i+j-1,k_j} + \frac{1}{L} \sum_{i=0}^{L-1} S(\bar{\sigma}_i) \right| < \frac{\epsilon}{4}. \tag{22}$$

Let n_1 be so large that (21) and (22) hold for $n \geq n_1$.

We need a similar result for the average entropy

$$\bar{S} = \frac{1}{L} \sum_{i=0}^{L-1} \sum_{j=1}^{J} p_j\, S(\sigma_{i,j}). \tag{23}$$

Lemma 4 *Fix $i \in \{0, \ldots, L-1\}$. Given a sequence $\underline{j} = (j_1, \ldots, j_n)$ with $1 \leq j_r \leq J(i+r-1)$, let $P^{(n)}_{i,\underline{j}}$ be the projection onto the subspace of $\mathcal{K}^{\otimes n}$ spanned by the eigenvectors of $\sigma^{(n)}_{i,\underline{j}} = \sigma_{i,j_1} \otimes \cdots \otimes \sigma_{i+n-1,j_n}$ with eigenvalues $\lambda^{(n)}_{\underline{j},\underline{k}} = \prod_{r=1}^{n} \lambda_{i+r-1,j_r,k_r}$ such that*

$$\left| \frac{1}{n} \log \lambda^{(n)}_{\underline{j},\underline{k}} + \bar{S} \right| < \frac{\epsilon}{4}. \tag{24}$$

For any $\delta > 0$ there exists $n_2 \in \mathbf{N}$ such that for $n \geq n_2$,

$$\mathbf{E}\left(\mathrm{Tr}\,\left(\sigma^{(n)}_{i,\underline{j}} P^{(n)}_{i,\underline{j}}\right)\right) > 1 - \delta^2, \tag{25}$$

where $\mathbf{E}$ denotes the expectation with respect to the probability distribution $\{p^{(n)}_{\underline{j}}\}$ on the states $\rho^{(n)}_{\underline{j}}$.

Proof. Define i.i.d. random variables $X_1, \ldots, X_n$ with distribution given by

$$\text{Prob}\,(X_r = \lambda_{i+r-1,j,k}) = p_{i+r-1,j}\,\lambda_{i+r-1,j,k}. \tag{26}$$

By the Weak Law of Large Numbers,

$$\begin{aligned} \frac{1}{n}\sum_{r=1}^{n}\log X_r \quad &\to \quad \frac{1}{L}\sum_{i=0}^{L-1}\sum_{j=1}^{J}\sum_{k} p_j\,\lambda_{i,j,k}\log\lambda_{i,j,k} \\ &= \quad -\frac{1}{L}\sum_{i=1}^{L}\sum_{j=1}^{J(i)} p_j\,S(\sigma_{i,j}) = -\bar{S}. \end{aligned} \tag{27}$$

It follows that there exists n_2 such that for $n \geq n_2$, the typical set $T_{i,\epsilon}^{(n)}$ of sequences of pairs $((j_1,k_1),\ldots,(j_n,k_n))$ such that

$$\left|\frac{1}{n}\sum_{r=1}^{n}\log\lambda_{i+r-1,j_r,k_r} + \bar{S}\right| < \frac{\epsilon}{3} \tag{28}$$

satisfies

$$\mathbf{P}\left(T_{i,\epsilon}^{(n)}\right) = \sum_{((j_1,k_1),\ldots,(j_n,k_n))\in T_\epsilon^{(n)}}\prod_{r=1}^{n} p_{j_r}\lambda_{i+r-1j_r,k_r} > 1-\delta^2. \tag{29}$$

Obviously,

$$P_{i,\underline{j}}^{(n)} \geq \sum_{\substack{\underline{k}=(k_1,\ldots,k_n)\\ ((j_1,k_1),\ldots,(j_n,k_n))\in T_{i,\epsilon}^{(n)}}} |\psi_{\underline{j},\underline{k}}^{(n)}\rangle\langle\psi_{\underline{j},\underline{k}}^{(n)}| \tag{30}$$

and

$$\mathbf{E}\left(\text{Tr}\left(\sigma_{\underline{j}}^{(n)}P_{i,\underline{j}}^{(n)}\right)\right) \geq \mathbf{P}\left(T_{i,\epsilon}^{(n)}\right) > 1-\delta^2. \tag{31}$$

□

The remainder of the proof is essentially the same as that in [DD]. Let $N = \tilde{N}(n)$ be the maximal number of product states $\tilde{\rho}_1^{(n)}, \ldots, \tilde{\rho}_N^{(n)}$ on $\mathcal{H}^{\otimes n}$ (each of which is a tensor product of states in the maximising ensemble $\{p_j,\rho_j\}_{j=1}^J$) for which there exist positive operators $E_1^{(n)}, \ldots, E_N^{(n)}$ on $\mathcal{K}^{\otimes ML_2} \otimes \mathcal{K}^{\otimes n}$ such that

(i) $E_k^{(n)} = \sum_{i=1}^L \tilde{\Pi}_i \otimes E_{k,i}^{(n)}$ and $\sum_{k=1}^N E_{k,i}^{(n)} \leq \bar{P}_i^{(n)}$ and

(ii) $\displaystyle\frac{1}{L}\sum_{i=1}^{L}\text{Tr}\left[\left(\tilde{\Pi}_i\otimes E_{k,i}^{(n)}\right)\Phi_i^{(ML_2+n)}\left(\omega^{(ML_2)}\otimes\tilde{\rho}_k^{(n)}\right)\right] > 1-\epsilon$ and

(iii) $\dfrac{1}{L}\sum_{i=1}^{L} \mathrm{Tr}\,\left[\left(\tilde{\Pi}_i \otimes E_{k,i}^{(n)}\right)\Phi_i^{(ML_2+n)}\left(\omega^{(ML_2)} \otimes \bar{\rho}^{\otimes n}\right)\right] \le 2^{-n[C(\Phi)-\frac{1}{2}\epsilon]}$

where $\bar{\rho} = \sum_{j=1}^{J} p_j \rho_j$.

For each $i = 1, \ldots, M$ and $\underline{j} = (j_1, \ldots, j_n)$ such that $1 \le j_r \le J$, we define, as before,

$$V_{i,\underline{j}}^{(n)} = \left(\bar{P}_i^{(n)} - \sum_{k=1}^{N} E_{k,i}^{(n)}\right)^{1/2} \bar{P}_i^{(n)} P_{i,\underline{j}}^{(n)} \bar{P}_i^{(n)} \left(\bar{P}_i^{(n)} - \sum_{k=1}^{N} E_{k,i}^{(n)}\right)^{1/2}. \qquad (32)$$

and we put

$$V_{\underline{j}}^{(n)} := \sum_{i=1}^{M} \tilde{\Pi}_i \otimes V_{i,\underline{j}}^{(n)}. \qquad (33)$$

Clearly $V_{i,\underline{j}}^{(n)} \le \bar{P}_i^{(n)} - \sum_{k=1}^{N} E_{k,i}^{(n)}$.

$V_{\underline{j}}^{(n)}$ is a candidate for an additional measurement operator, $E_{N+1}^{(n)}$, for Bob with corresponding input state $\tilde{\rho}_{N+1}^{(n)} = \rho_{\underline{j}}^{(n)} = \rho_{j_1} \otimes \rho_{j_2} \ldots \otimes \rho_{j_n}$. Clearly, the condition (i), given above, is satisfied and we also have

Lemma 5

$$\frac{1}{L}\sum_{i=1}^{L} \mathrm{Tr}\,\left[\left(\tilde{\Pi}_i \otimes V_{i,\underline{j}}^{(n)}\right)\Phi_i^{(ML_2+n)}\left(\omega^{(ML_2)} \otimes \bar{\rho}^{\otimes n}\right)\right] \le 2^{-n[C(\Phi)-\frac{1}{2}\epsilon]}. \qquad (34)$$

Proof. Put $Q_{n,i} = \sum_{k=1}^{N} E_{k,i}^{(n)}$. Note that $Q_{n,i}$ commutes with $\bar{P}_i^{(n)}$. Using the fact that $\bar{P}_i^{(n)}\Phi_i^{(n)}(\bar{\rho}^{\otimes n})\bar{P}_i^{(n)} \le 2^{-n[\frac{1}{L}\sum_{i=1}^{L} S(\bar{\sigma}_i)-\frac{1}{4}\epsilon]}$ by (22), we have, denoting $\bar{\sigma}_i^{(n)} = \Phi_i^{(n)}(\bar{\rho}^{\otimes n})$,

$$\begin{aligned}
\mathrm{Tr}\,(\bar{\sigma}_i^{(n)} V_{i,\underline{j}}^{(n)}) &= \mathrm{Tr}\,\left[\bar{\sigma}_i^{(n)}(\bar{P}_i^{(n)} - Q_{n,i})^{1/2}\bar{P}_i^{(n)} P_{i,\underline{j}}^{(n)} \bar{P}_i^{(n)}(\bar{P}_i^{(n)} - Q_{n,i})^{1/2}\right] \\
&= \mathrm{Tr}\,\left[\bar{P}_i^{(n)}\bar{\sigma}_i^{(n)}\bar{P}_i^{(n)}(\bar{P}_i^{(n)} - Q_{n,i})^{1/2} P_{i,\underline{j}}^{(n)}(\bar{P}_i^{(n)} - Q_{n,i})^{1/2}\right] \\
&\le 2^{-n[\frac{1}{L}\sum_{i=1}^{L} S(\bar{\sigma}_i)-\frac{1}{4}\epsilon]}\mathrm{Tr}\,\left[(\bar{P}_i^{(n)} - Q_{n,i})^{1/2} P_{i,\underline{j}}^{(n)}(\bar{P}_i^{(n)} - Q_{n,i})^{1/2}\right] \\
&\le 2^{-n[\frac{1}{L}\sum_{i=1}^{L} S(\bar{\sigma}_i)-\frac{1}{4}\epsilon]}\mathrm{Tr}\,(P_{i,\underline{j}}^{(n)}) \le 2^{-n[\frac{1}{L}\sum_{i=1}^{L} S(\bar{\sigma}_i)-\bar{S}-\frac{1}{2}\epsilon]}, \qquad (35)
\end{aligned}$$

where, in the last inequality, we used the standard upper bound on the dimension of the typical subspace: $\mathrm{Tr}\,(P_{i,\underline{j}}^{(n)}) \le 2^{n[\bar{S}+\frac{1}{4}\epsilon]}$, which follows from Lemma 4. □

By maximality of N it now follows that the condition (ii) above cannot hold, that is,

$$\frac{1}{L}\sum_{i=1}^{L}\mathrm{Tr}\left[\left(\tilde{\Pi}_i\otimes V_{i,\underline{j}}^{(n)}\right)\,\Phi_i^{(ML_2+n)}\left(\omega^{(ML_2)}\otimes\rho_{\underline{j}}^{(n)}\right)\right]\leq 1-\epsilon \tag{36}$$

for every $\underline{j}$, and this yields the following:

Corollary 1

$$\frac{1}{L}\sum_{i=1}^{L}\mathbf{E}\left(\mathrm{Tr}\left[\left(\tilde{\Pi}_i\otimes V_{i,\underline{j}}^{(n)}\right)\,\Phi_i^{(ML_2+n)}\left(\omega^{(ML_2)}\otimes\rho_{\underline{j}}^{(n)}\right)\right]\right)\leq 1-\epsilon. \tag{37}$$

We also need the following lemma:

Lemma 6 *For all $\eta'>\delta^2+3\delta$,*

$$\frac{1}{L}\sum_{i=1}^{L}\mathrm{Tr}\left[\left(\tilde{\Pi}_i\otimes \bar{P}_i^{(n)}P_{i,\underline{j}}^{(n)}\bar{P}_i^{(n)}\right)\,\Phi_i^{(ML_2+n)}\left(\omega^{(ML_2)}\otimes\Phi_i^{(n)}(\rho_{\underline{j}}^{(n)})\right)\right]>1-\eta' \tag{38}$$

if n is large enough.

Proof. This is proved as in [DD]. □

Lemma 7 *Assume $\eta'<\frac{1}{3}\epsilon$ and write*

$$Q_{n,i}=\sum_{k=1}^{N}E_{k,i}^{(n)}. \tag{39}$$

Then for n large enough,

$$\frac{1}{L}\sum_{i=1}^{L}\mathbf{E}\left(\mathrm{Tr}\left[\left(\tilde{\Pi}_i\otimes Q_{n,i}\right)\,\Phi_i^{(ML_2+n)}\left(\omega^{(ML_2)}\otimes\rho_{\underline{j}}^{(n)}\right)\right]\right)\geq\eta'^2. \tag{40}$$

Proof. This follows as before from the previous lemma using the Cauchy-Schwarz inequality. □

It now follows that for n large enough, $\tilde{N}(n)\geq(\eta')^2\,2^{n[C(\Phi)-\frac{3}{4}\epsilon]}$. We take the following states as codewords:

$$\rho_k^{(ML_2+n)}=\omega^{(ML_2)}\otimes\tilde{\rho}_k^{(n)}. \tag{41}$$

For n sufficiently large we then have

$$N_{n+ML_2} = \tilde{N}(n) \geq (\eta')^2\, 2^{n[C(\Phi)-\frac{3}{4}\epsilon]} \geq 2^{(ML_2+n)[C(\Phi)-\epsilon]}. \qquad (42)$$

To complete the proof we need to show that the set $E_k^{(n)}$ satisfies (3). But this follows immediately from condition (ii):

$$\mathrm{Tr}\left[\Phi^{(ML_2+n)}\left(\rho_k^{(ML_2+n)}\right) E_k^{(n)}\right] =$$
$$= \frac{1}{L}\sum_{i=0}^{L-1} \mathrm{Tr}\left[\Phi_i^{\otimes(ML_2+n)}\left(\omega^{(ML_2)} \otimes \tilde{\rho}_k^{(n)}\right)\left(\tilde{\Pi}_i \otimes E_{i,k}^{(n)}\right)\right] > 1-\epsilon. \quad (43)$$

□

3 The Direct Channel Coding Theorem

Theorem 1 can now be used to prove that for any rate $R < \chi(\Phi)$, where $\chi(\Phi)$ is the Holevo capacity defined in (**??**), there exists a code using product state inputs, which enables reliable transmission of classical information through a quantum memoryless channel at a rate R.

Consider the sender Alice to have a classical ergodic information source with alphabet A and probability distribution μ. Let μ_n denote the restriction of μ to A^n. Alice would like to achieve reliable transmission of messages emitted from her source to Bob, through multiple uses of a memoryless quantum channel Φ, using product state inputs. For this purpose, Alice assigns a label k_x to each message x emitted by her source. By McMillan's theorem (see e.g. [3]), for n large enough, there is a set T_n of typical messages, which has a high probability of occurrence. Moreover, the number, $|T_n|$, of typical messages is $\simeq 2^{nH}$, where H is the Kolmogorov–Sinai entropy of the ergodic source (see e.g. [3]).

By Theorem 2, there exists a set of product states $\tilde{\rho}_k^{(n)}$ and a set of positive operators $E_k^{(n)}$ with $k = 1, \ldots, N_n$ and $N_n > 2^{n(\chi(\Phi)-\epsilon)}$ such that $\mathrm{Tr}\,[\Phi^{\otimes n}(\rho_k^{(n)})E_k^{(n)}] > 1-\epsilon$. For each typical message x, Alice chooses a unique label $k_x \in \{2, \ldots, N_n\}$. However, to all messages x which are atypical she assigns the label $k_x = 1$. Then she associates to each label $k_x \in \{1, \ldots, N_n\}$ a unique state $\tilde{\rho}_{k_x}^{(n)}$ belonging to the above–mentioned set of product states. The states $\tilde{\rho}_{k_x}^{(n)}$, $k_x \in 1, \ldots, N_n$ are the codewords, which are transmitted to Bob through n uses of the channel Φ. The rate of the corresponding code is given by $R_n = (\log(|T_n| + 1))/n$, since there are $(|T_n| + 1)$ codewords. If the codeword $\tilde{\rho}_{k_x}^{(n)}$ is transmitted, then Bob receives the state $\tilde{\sigma}_{k_x}^{(n)} := \Phi^{\otimes n}(\tilde{\rho}_{k_x}^{(n)})$.

To decode the label k_x of the message sent by Alice, Bob does a measurement on $\tilde{\sigma}_{k_x}^{(n)}$ described by positive operators (POVM elements) $E_1^{(n)}, \ldots, E_{N_n}^{(n)}$ (where $\sum_{j=1}^{N_n} E_j^{(n)} \leq \mathbf{I}_n$) and $E_0^{(n)} := \mathbf{I}_n - \sum_{j=1}^{N_n} E_j^{(n)}$. The POVM element $E_{k_x}^{(n)}$ corresponds to the label k_x. The probability of inferring the label k_x correctly is therefore given by $\mathrm{Tr}\left(\tilde{\sigma}_{k_x}^{(n)} E_{k_x}^{(n)}\right)$. If k_x is correctly decoded and found to belong to the set $\{2, \ldots N_n\}$, then Bob can unambiguously infer the corresponding classical message x that was sent sent by Alice. This is because the map $x \to k_x$ was one–to–one for labels k_x in the above set. However, even if the label $k_x = 1$ is correctly inferred, the corresponding classical message x cannot be inferred unambiguously. This is because each atypical message x had been given the same label $k_x = 1$. In this case, Bob fails to decode Alice's message. Nevertheless the average probability of error in decoding still vanishes asymptotically because atypical messages are rarely emitted. Moreover, from *Theorem 1* it follows that for any $\epsilon > 0$, the probability of error for each *typical* message can be made less that ϵ.

The following theorem is the direct part of the HSW theorem generalized to an ergodic source.

Theorem 2 *Consider a memoryless quantum channel given by a completely positive trace-preserving map $\Phi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{K})$, where $\mathcal{H}$ and $\mathcal{K}$ are finite-dimensional Hilbert spaces. Let $\chi(\Phi)$ be the Holevo capacity of the channel. If A is the alphabet of a classical ergodic source of information with probability distribution μ and Kolmogorov–Sinai entropy $H < \chi(\Phi)$, then there exists for any given $\epsilon > 0$, an $n_0 \in \mathbf{N}$, such that for all $n \geq n_0$ there exist a code map $\mathcal{C}_n : A^n \to \mathcal{S}(\mathcal{H}^{\otimes n})$, of rate $R < \chi(\Phi)$, with image in the product states, and a decoding $\mathcal{D}_n : \mathcal{S}(\mathcal{K}^{\otimes n}) \to A^n$, corresponding to POVM elements $\{E_k^{(n)}\}_{k=0}^{N_n}$, such that the average probability of error, p_e, in decoding, satisfies $p_e < 2\epsilon$.*

Proof. By McMillan's theorem (see e.g. [3]), for $\epsilon > 0$ and n large enough, there exists a typical set $T_n \equiv T_\epsilon^{(n)}$ in A^n, such that for all $x \in T_n$, $\mu_n(x) > 2^{-n(H+\epsilon)}$, $\mu_n(T_n) > 1-\epsilon$ and $|T_n| < 2^{n(H+\epsilon)}$, where $|T_n|$ denotes the cardinaliy of the typical set. By the above Theorem 1, there exist product states $\tilde{\rho}_k^{(n)}$ and positive operators $E_k^{(n)}$ with $k = 1, \ldots, N_n$ and $N_n > 2^{n(\chi(\Phi)-\epsilon)}$ such that $\mathrm{Tr}\,[\Phi^{\otimes n}(\tilde{\rho}_k^{(n)}) E_k^{(n)}] > 1 - \epsilon$. Choose ϵ to be so small that $H + \epsilon < \chi(\Phi) - \epsilon$. For this choice, $|T_n| < 2^{n(\chi(\Phi)-\epsilon)}$, and hence the rate R of the code satisfies the bound $R < \chi(\Phi) - \epsilon$. In this case we can define a one to one map $\mathcal{C}_n : T_n \to \mathcal{S}(\mathcal{H}^{\otimes n})$ by $\mathcal{C}_n(x) = \tilde{\rho}_{k_x}^{(n)}$ for some $k_x \in \{2, \ldots, N_n\}$, whereas for all $x \notin T_n$ we define $\mathcal{C}_n(x) := \tilde{\rho}_1^{(n)}$. The average probability of error is given by

$$p_e = \sum_{x \in T_n} \mu_n(x)(1 - \mathrm{Tr}\,[\tilde{\sigma}_{k_x}^{(n)} E_{k_x}^{(n)}]) + \mu_n(T_n^c) < 2\epsilon, \tag{44}$$

where T_n^c denotes the atypical set, i.e., the complement of T_n. □

3.1 Proof of the converse

In this section we prove that it is impossible for Alice to transmit classical messages reliably to Bob through the channel Φ defined in (1) at a rate $R > C(\Phi)$. This is the weak converse of Theorem **??** in the sense that the probability of error does not tend to zero asymptotically as the length of the code increases, for any code with rate $R > C(\Phi)$. To prove the weak converse, suppose that Alice encodes messages labelled by $\alpha \in \mathcal{M}_n$ by product states $\rho_\alpha^{(n)} = \rho_{\alpha,1} \otimes \cdots \otimes \rho_{\alpha,n}$ in $\mathcal{B}(\mathcal{H}^{\otimes n})$. Let the corresponding outputs for the i-th branch of the channel be denoted by $\sigma_{\alpha,i}^{(n)}$, i.e.

$$\sigma_{\alpha,i}^{(n)} = \Phi_i^{(n)}(\rho_\alpha^{(n)}) = \sigma_{\alpha,1}^i \otimes \cdots \otimes \sigma_{\alpha,n}^{i+n-1}, \ \sigma_{\alpha,j}^i = \Phi_i(\rho_{\alpha,j}). \tag{45}$$

Further define

$$\bar{\sigma}_i^{(n)} = \frac{1}{|\mathcal{M}_n|} \sum_{\alpha \in \mathcal{M}_n} \sigma_{\alpha,i}^{(n)} \tag{46}$$

and

$$\bar{\sigma}_{i,j} = \frac{1}{|\mathcal{M}_n|} \sum_{\alpha \in \mathcal{M}_n} \sigma_{\alpha,j}^i \ . \tag{47}$$

Let Bob's POVM elements corresponding to the codewords $\rho_\alpha^{(n)}$ be denoted by $E_\alpha^{(n)}$, $\alpha = 1, \ldots, |\mathcal{M}_n|$. We may assume that Alice's messages are produced uniformly at random from the set $\mathcal{M}_n$. Then Bob's average probability of error is given by

$$\bar{p}_e^{(n)} := 1 - \frac{1}{|\mathcal{M}_n|} \sum_{\alpha \in \mathcal{M}_n} \mathrm{Tr}\ \left[\Phi^{(n)}(\rho_\alpha^{(n)}) E_\alpha^{(n)}\right]. \tag{48}$$

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 $Y^{(n)}$ be the random variable corresponding to Bob's inference of Alice's message. It is defined by the conditional probabilities

$$\mathbf{P}\,[Y^{(n)} = \beta \,|\, X^{(n)} = \alpha] = \frac{1}{L} \sum_{i=0}^{L-1} \mathrm{Tr}\ [\Phi_i^{(n)}(\rho_\alpha^{(n)}) E_\beta^{(n)}]. \tag{49}$$

By Fano's inequality,

$$h(\bar{p}_e^{(n)}) + \bar{p}_e^{(n)} \log(|\mathcal{M}_n| - 1) \geq H(X^{(n)} \,|\, Y^{(n)}) = H(X^{(n)}) - H(X^{(n)} : Y^{(n)}). \tag{50}$$

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

$$\begin{aligned}
H(X^{(n)}:Y^{(n)}) &\leq S\left(\frac{1}{|\mathcal{M}_n|L}\sum_{\alpha\in\mathcal{M}_n}\sum_{i=0}^{L-1}\Phi_i^{(n)}(\rho_\alpha^{(n)})\right) \\
&\qquad -\frac{1}{|\mathcal{M}_n|}\sum_{\alpha\in\mathcal{M}_n} S\left(\frac{1}{L}\sum_{i=0}^{L-1}\Phi_i^{(n)}(\rho_\alpha^{(n)})\right) \\
&\leq \sum_{j=1}^{n}\left[S\left(\frac{1}{|\mathcal{M}_n|}\frac{1}{L}\sum_{\alpha\in\mathcal{M}_n}\sum_{i=0}^{L-1}\Phi_{i+j-1}(\rho_{\alpha,j})\right)\right. \\
&\qquad \left. -\frac{1}{|\mathcal{M}_n|}\sum_{\alpha\in\mathcal{M}_n} S\left(\frac{1}{L}\sum_{i=0}^{L-1}\Phi_{i+j-1}(\rho_{\alpha,j})\right)\right] \\
&= \sum_{j=1}^{n}\frac{1}{|\mathcal{M}_n|}\sum_{\alpha\in\mathcal{M}_n} \\
&\qquad \times S\left(\frac{1}{L}\sum_{i=0}^{L-1}\Phi_{i+j-1}(\rho_{\alpha,j})\,||\,\frac{1}{|\mathcal{M}_n|\,L}\sum_{\alpha,i}\Phi_{i+j-1}(\rho_{\alpha,j})\right) \\
&\leq \frac{1}{L}\sum_{i=0}^{L-1}\sum_{j=1}^{n}\frac{1}{|\mathcal{M}_n|}\sum_{\alpha\in\mathcal{M}_n} \\
&\qquad \times S\left(\Phi_{i+j-1}(\rho_{\alpha,j})\,||\,\frac{1}{|\mathcal{M}_n|}\sum_{\alpha}\Phi_{i+j-1}(\rho_{\alpha,j})\right) \\
&= \frac{1}{L}\sum_{i=0}^{L-1}\sum_{j=1}^{n}\chi_i\left(\left\{\frac{1}{|\mathcal{M}_n|},\rho_{\alpha,j}\right\}_{\alpha\in\mathcal{M}_n}\right) \\
&\leq nC(\Phi).
\end{aligned} \tag{51}$$

Fano's inequality (50) now yields

$$h(\bar{p}_e^{(n)}) + \bar{p}_e^{(n)}\log|\mathcal{M}_n| \geq \log|\mathcal{M}_n| - n\,C(\Phi), \tag{52}$$

Now, if $R = \frac{1}{n}\log|\mathcal{M}_n| > C(\Phi)$, we must have

$$\bar{p}_e^{(n)} \geq 1 - \frac{C(\Phi)+1/n}{R} > 0. \tag{53}$$

□

References

[1] I. Bjelaković, T. Krüger, R. Siegmund-Schultze & A. Szkoła, "The Shannon-McMillan theorem for ergodic quantum lattice systems," *Invent. math.* **155**, 203–222, 2004.

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

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

[4] N.Datta and T. Dorlas, "Product state capacities of a class of channels with memory," *in preparation.*

[5] A. Feinstein, " A new basic theorem of information theory," *IRE Trans. PGIT,* **4**, pp. 2–22, 1954.

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

[7] F. Hiai & D. Petz, "The proper formula for the relative entropy and its asymptotics in quantum probability". *Commun. Math. Phys.* **143**, 257–281, 1991.

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

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

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

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

[12] T. Ogawa and H. Nagaoka, "A New Proof of the Channel Coding Theorem via Hypothesis Testing in Quantum Information Theory," *Proc. 2002 IEEE International Symposium on Information Theory,* pp. 73, 2002.

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

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

[15] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pt. I, pp. 379–423, 1948; pt. II, pp. 623–656, 1948.

[16] A. Winter, "Coding theorem and strong converse for quantum channels," *IEEE Trans. Info. Theory*, **45**, 2481–2485, 1999.