J Stat Phys (2009) 134: 1173–1195
DOI 10.1007/s10955-009-9726-0

Classical Capacity of Quantum Channels with General Markovian Correlated Noise

Nilanjana Datta · Tony Dorlas

Received: 18 July 2008 / Accepted: 5 March 2009 / Published online: 25 March 2009

Abstract The classical capacity of a quantum channel with *arbitrary* Markovian correlated noise is evaluated. For the general case of a channel with long-term memory, which corresponds to a Markov chain which does not converge to equilibrium, the capacity is expressed in terms of the communicating classes of the Markov chain. For an irreducible and aperiodic Markov chain, the channel is forgetful, and one retrieves the known expression (Kretschmann and Werner in Phys. Rev. A 72:062323, 2005) for the capacity.

Keywords Quantum channels with long term memory · Markovian correlated noise · Helstrom's Theorem · Feinstein's Lemma

1 Introduction

Shannon, in his celebrated Noisy Channel Coding Theorem [5, 22, 23], obtained an explicit expression for the channel capacity of discrete, memoryless,[1] classical channels. The first rigorous proof of this fundamental theorem was provided by Feinstein [8]. He used a packing argument (see e.g. [10]) to find a lower bound to the maximal number of codewords that can be sent through the channel *reliably*, i.e., with an arbitrarily low probability of error. More precisely, he proved that for any given $\delta > 0$, and a sufficiently large number, n, of uses of a memoryless classical channel, a lower bound to the maximal number, N_n, of codewords

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

T. Dorlas is on leave from the Dublin Institute of Advanced Studies, School of Theoretical Physics, 10 Burlington Road, Dublin 4, Ireland.

N. Datta (✉) · T. Dorlas
Statistical Laboratory, Centre for Mathematical Sciences, University of Cambridge, Wilberforce Road, Cambridge CB3 0WB, UK
e-mail: n.datta@statslab.cam.ac.uk

T. Dorlas
e-mail: dorlas@stp.dias.ie

that can be transmitted through the channel reliably, is given by

$$N_n \geq 2^{n(H(X:Y)-\delta)}.$$

Here $H(X:Y)$ is the mutual information of the random variables X and Y, corresponding to the input and the output of the channel, respectively. Moreover, he also proved that this result can be extended to channels given by ergodic Markov chains. The ratio R of the number of bits of the message and the number of bits of the code is called the *rate* of the code, i.e.

$$R = \frac{1}{n}\log_2 N_n.$$

Thus, Shannon's bound means that any rate $R < C = \max H(X:Y)$, (the maximum being taken over all possible input distributions) is *achievable*.

Holevo, Schumacher and Westmoreland [13, 21] derived an analogue of Shannon's bound for transmission of classical information in the form of quantum states over a memoryless quantum channel defined by a completely positive map. The analogue of Shannon's bound is usually called the Holevo χ-quantity. The assumption that the channel is memoryless means that the channel map is a tensor product, or, in other words, that the noise is uncorrelated between successive uses of a channel. This is obviously unrealistic: in actual quantum channels, memory effects will play a role. In this paper we consider the transmission of classical information through a class of quantum channels with memory.

The first model of a channel with memory was studied by Macchiavello and Palma [17]. 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. A more general model of a quantum channel with memory was introduced by Bowen and Mancini [4] and also studied by Kretschmann and Werner [15]. In particular, in [15], the capacities of a class of quantum channels with memory, the so-called *forgetful channels* were evaluated. Similar results were obtained by Bjelaković and Boche [2]. Further, in [7], the classical capacity of a class of quantum channels with long-term memory was obtained. The memory of the channel considered in [7] can be viewed as a special case of a general Markovian memory, where the Markov chain is aperiodic but not irreducible, and hence does not converge to equilibrium. Recently, there was a generalization of the result of [7] by Bjelaković and Boche, who in [3] obtained the classical capacities of compound and averaged quantum channels.

Another interesting special case of a channel with long-term memory is that in which the memory is described by a periodic Markov chain. A simple example of this is a channel given by alternating applications of two completely positive trace preserving (CPT) maps Φ_1 and Φ_2, with the first map being Φ_1 or Φ_2 with probability 1/2.

In this paper we study channels with *arbitrary* Markovian correlated noise. This includes, in particular, the above special cases. We show that the capacity in the general case can be expressed in terms of the communicating classes of the underlying Markov chain.

We start the main body of our paper with some preliminaries in Sect. 2. In Sect. 3, the quantum channel is defined and its capacity is stated in the main theorem, Theorem 1, of this paper. In Sect. 4, we prove a special case of the direct part of this theorem, corresponding to a Markov chain which converges to equilibrium and is hence forgetful. This section therefore provides an alternative proof of the result of Kretschmann and Werner [15] for the classical capacity of such a channel. This proof is extended to the case of an arbitrary Markov chain in Sect. 5. In the latter, we employ the idea of adding a preamble to the codewords (as was done in [7]) in order to distinguish between the different communicating classes of the Markov chain. The proof of the (weak) converse part of our main result (Theorem 1) is given in Sect. 5.2.

2 Mathematical Preliminaries

Let $\mathcal{H}$ and $\mathcal{K}$ be given finite-dimensional Hilbert spaces and denote by $\mathcal{B}(\mathcal{H})$ the algebra of linear operators on $\mathcal{H}$. We also consider the tensor product algebras $\mathcal{A}_n = \mathcal{B}(\mathcal{H}^{\otimes n})$ and the infinite tensor product C*-algebra obtained as the strong closure

$$\mathcal{A}_\infty = \overline{\bigcup_{n=1}^{\infty} \mathcal{A}_n}, \tag{1}$$

where we embed $\mathcal{A}_n$ into $\mathcal{A}_{n+1}$ in the obvious way. Similarly, we define $\mathcal{B}_n = \mathcal{B}(\mathcal{K}^{\otimes n})$ and $\mathcal{B}_\infty$. A *state* on an algebra $\mathcal{A}$ is a positive linear functional ϕ on $\mathcal{A}$ with $\phi(\mathbf{1}) = 1$, where $\mathbf{1}$ denotes identity operator. If $\mathcal{A}$ is finite-dimensional then there exists a density matrix ρ_ϕ (i.e., a positive operator with $\mathrm{Tr}\rho_\phi = 1$) such that $\phi(A) = \mathrm{Tr}(\rho_\phi A)$, for any $A \in \mathcal{A}$. We denote the states on $\mathcal{A}_\infty$ by $\mathcal{S}(\mathcal{A}_\infty)$, those on $\mathcal{A}_n$ by $\mathcal{S}(\mathcal{A}_n)$, etc.

3 A Quantum Channel with Classical Memory

Let there be given a Markov chain on a finite state space I with transition probabilities $\{q_{ii'}\}_{i,i'\in I}$ and let $\{\gamma_i\}_{i\in I}$ be an invariant distribution for this chain, i.e.

$$\gamma_{i'} = \sum_{i\in I} \gamma_i q_{ii'}. \tag{2}$$

Moreover, let $\Phi_i : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{K})$ be given completely positive trace-preserving (CPT) maps for each $i \in I$. Then we define a quantum channel with Markovian correlated noise, by the CPT maps $\Phi^{(n)} : \mathcal{S}(\mathcal{A}_n) \to \mathcal{S}(\mathcal{B}_n)$ on the states of $\mathcal{A}_n$ by

$$\Phi^{(n)}(\rho^{(n)}) = \sum_{i_1,\dots,i_n\in I} \gamma_{i_1} q_{i_1 i_2} \cdots q_{i_{n-1} i_n} (\Phi_{i_1} \otimes \cdots \otimes \Phi_{i_n})(\rho^{(n)}) \tag{3}$$

for $A \in \mathcal{B}_n$.

If ϕ is a state on $\mathcal{A}_\infty$ and ϕ_n is the restriction to $\mathcal{A}_n$, and we write $\rho^{(n)} = \rho_{\phi_n}$, then the states $\Phi^{(n)}(\rho^{(n)})$ form a consistent (projective) system of states on $\mathcal{B}_\infty$ because of (2), and therefore define a unique state on $\mathcal{B}_\infty$, which we denote by $\Phi_\infty(\phi)$. Thus

$$(\Phi_\infty(\phi))(A) = \mathrm{Tr}\big(\Phi^{(n)}(\rho_{\phi_n})A\big)$$

for $A \in \mathcal{A}_n$.

Let us consider the transmission of classical information through $\Phi^{(n)}$. Suppose Alice has a set of messages, labelled by the elements of the set $\mathcal{M}_n = \{1, 2, \dots, M_n\}$, which she would like to communicate to Bob, using the quantum channel Φ. To do this, she encodes each message into a quantum state of a physical system with Hilbert space $\mathcal{H}^{\otimes n}$, which she then sends to Bob through n uses of the quantum channel. In order to infer the message that Alice communicated to him, Bob makes a measurement (described by POVM elements) on the state that he receives. The encoding and decoding operations, employed to achieve reliable transmission of information through the channel, together define a quantum error correcting code (QECC). More precisely, a code $\mathcal{C}^{(n)}$ of size N_n is given by a sequence $\{\rho_i^{(n)}, E_i^{(n)}\}_{i=1}^{N_n}$ where each $\rho_i^{(n)}$ is a state in $\mathcal{B}(\mathcal{H}^{\otimes n})$ and each $E_i^{(n)}$ is a positive operator acting in $\mathcal{K}^{\otimes n}$, such that $\sum_{i=1}^{N_n} E_i^{(n)} \leq \mathbf{1}^{(n)}$. Here, $\mathbf{1}^{(n)}$ denotes the identity operator in $\mathcal{B}(\mathcal{K}^{\otimes n})$. Defining $E_0^{(n)} = \mathbf{1}^{(n)} - \sum_{i=1}^{N_n} E_i^{(n)}$, yields a Positive Operator-Valued Measure (POVM) $\{E_i^{(n)}\}_{i=0}^{N_n}$ in $\mathcal{K}^{\otimes n}$. An output $i \geq 1$ would lead to the inference that the state (or codeword) $\rho_i^{(n)}$ was transmitted through the channel $\Phi^{(n)}$, whereas the output 0 is interpreted as a failure of any inference. The average probability of error for the code $\mathcal{C}^{(n)}$ is given by

$$P_e(\mathcal{C}^{(n)}) := \frac{1}{N_n} \sum_{i=1}^{N_n} \big(1 - \mathrm{Tr}\big(\Phi^{(n)}(\rho_i^{(n)}) E_i^{(n)}\big)\big), \tag{4}$$

If there exists an $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$, there exists a sequence of codes $\{\mathcal{C}^{(n)}\}_{n=1}^{\infty}$, of sizes $N_n \geq 2^{nR}$, for which $P_e(\mathcal{C}^{(n)}) \to 0$ as $n \to \infty$, then R is said to be an *achievable* rate.

The *classical capacity* of Φ is defined as

$$C(\Phi) := \sup R, \tag{5}$$

where R is an achievable rate.

Let $\mathcal{C}$ be the set of communicating classes of the Markov chain [19] for which

$$\gamma_C = \sum_{i \in C} \gamma_i > 0. \tag{6}$$

Any other classes can be disregarded. In particular, we can assume that all classes are closed. For $C \in \mathcal{C}$ we define

$$\Phi_C^{(n)}(\rho^{(n)}) := \frac{1}{\gamma_C} \sum_{i_1,\dots,i_n \in C} \gamma_{i_1} q_{i_1 i_2} \cdots q_{i_{n-1} i_n} (\Phi_{i_1} \otimes \cdots \otimes \Phi_{i_n})(\rho^{(n)}), \tag{7}$$

which represents the restriction of the classical memory of the channel to the class C. Notice that the Markov chain restricted to $C \in \mathcal{C}$ is necessarily irreducible, and is either aperiodic or periodic with a single period. In fact,

$$\mathcal{C} = \mathcal{C}_{aper} \cup \mathcal{C}_{per},$$

where $\mathcal{C}_{aper}$ denotes the set of communicating classes in $\mathcal{C}$ which are aperiodic, while $\mathcal{C}_{per}$ denotes the set of communicating classes in $\mathcal{C}$ which are periodic.

If $C \in \mathcal{C}_{aper}$, we define, for any ensemble $\{p_j^{(n)}, \rho_j^{(n)}\}$ of states on $\mathcal{H}^{\otimes n}$, the *mean Holevo quantity* for the class C as

$$\bar{\chi}_C^{(n)}(\{p_j^{(n)}, \rho_j^{(n)}\}) := \frac{1}{n}\left[S\left(\sum_j p_j^{(n)} \Phi_C^{(n)}(\rho_j^{(n)})\right) - \sum_j p_j^{(n)} S\big(\Phi_C^{(n)}(\rho_j^{(n)})\big)\right]. \tag{8}$$

If $C \in \mathcal{C}_{per}$ is periodic, with period L, then there exist subclasses $C^{(0)}, \dots, C^{(L-1)}$ such that the Markov chain cycles through these subclasses, i.e. if $i \in C^{(k)}$ then $i+1 \in C^{(k+1)}$ where the index k is taken modulo L. The subchains $i + mL$, $m = 0, 1, \dots,$ are aperiodic and irreducible Markov chains on $C^{(k)}$ for a fixed k.

We write

$$\Phi_{C,k}^{(n)} = \frac{L}{\gamma_C} \sum_{i \in C^{(k)}} \sum_{i_2,\dots,i_n} \gamma_i q_{i,i_2} \cdots q_{i_{n-1},i_n} \Phi_i \otimes \Phi_{i_2} \otimes \cdots \otimes \Phi_{i_n}, \tag{9}$$

so that

$$\Phi_C^{(n)} = \frac{1}{L} \sum_{k=0}^{L-1} \Phi_{C,k}^{(n)}. \tag{10}$$

We define

$$\bar{\chi}_C^{(n)}(\{p_j^{(n)}, \rho_j^{(n)}\}) = \frac{1}{nL} \sum_{k=0}^{L-1} \chi_{C,k}^{(n)}(\{p_j^{(n)}, \rho_j^{(n)}\}), \tag{11}$$

where for $k \in \{0, 1, \dots, L-1\}$,

$$\chi_{C,k}^{(n)}(\{p_j^{(n)}, \rho_j^{(n)}\}) = S\left(\sum_j p_j^{(n)} \Phi_{C,k}^{(n)}(\rho_j^{(n)})\right) - \sum_j p_j^{(n)} S(\Phi_{C,k}^{(n)}(\rho_j^{(n)})).$$

Our main result is the following theorem. (We use the standard notation $\wedge$ for minimum and $\vee$ for maximum.)

Theorem 1 *The classical capacity of a quantum channel with arbitrary Markovian correlated noise, defined by* (3), *is given by*

$$C(\Phi) = \lim_{n\to\infty} \sup_{\{p_j^{(n)},\rho_j^{(n)}\}} \left[\bigwedge_{C\in\mathcal{C}} \bar{\chi}_C^{(n)}(\{p_j^{(n)}, \rho_j^{(n)}\}) \right]. \tag{12}$$

The existence of the limit in (12) is proved in Lemma 15 of Appendix A.

Before proving Theorem 1, we first consider two special cases in which the Markov chain has a single communicating class, one being aperiodic and irreducible, the other periodic.

4 The Irreducible Case

In this section we assume that the underlying Markov chain is irreducible but not necessarily aperiodic (see e.g. [19]). It follows, in particular, that the invariant distribution, $\{\gamma_i\}_{i\in I}$, is unique. If the chain has period L, we denote the subclasses by $C^{(0)}, \dots, C^{(L-1)}$. We define channels $\Phi_k^{(n)}$ by

$$\Phi_k^{(n)} = L \sum_{i \in C^{(k)}} \gamma_i \sum_{i_2,\dots,i_n \in I} q_{i,i_2} \cdots q_{i_{n-1},i_n} \Phi_i \otimes \Phi_{i_2} \otimes \cdots \otimes \Phi_{i_n}, \tag{13}$$

so that

$$\Phi^{(n)} = \frac{1}{L} \sum_{k=0}^{L-1} \Phi_k^{(n)}. \tag{14}$$

Now, the Markov chain restricted to the subclass $C^{(k)}$ with transition probabilities $q_{ij}^{(L)}$ is irreducible and aperiodic, and therefore ergodic. (Here where $q_{ij}^{(n)}$ denotes the n-step transition probability from the state i to the state j $(i, j \in I)$.) Hence, the Markov chain is ergodic for shifts over L, and satisfies the property of *convergence to equilibrium*, i.e.,

$$q_{ij}^{(nL)} \to L\gamma_j \quad \text{as } n \to \infty.$$

(Note that $\sum_{i \in C^{(k)}} \gamma_i = \frac{1}{L}$.)

In the case of an aperiodic chain, this implies that the correlation in the noise, acting on successive inputs to the channel, dies out after a sufficiently large number of uses of the channel. Hence, in this case the channel belongs to the class of channels introduced and studied by Kretschmann and Werner [15], and referred to as *forgetful channels*.

The (classical) capacity of the channel is defined in the following lemma, which is a special case of Lemma 15 of Appendix A.

Lemma 1 *Given the* (*periodic*) *irreducible channel above*, *define*

$$\bar{\chi}^{(n)}\big(\{p_j^{(n)}, \rho_j^{(n)}\}\big) = \frac{1}{nL}\sum_{k=0}^{L-1} \chi_k^{(n)}\big(\{p_j^{(n)}, \rho_j^{(n)}\}\big),$$

where

$$\chi_k^{(n)}\big(\{p_j^{(n)}, \rho_j^{(n)}\}\big) = S\bigg(\sum_j p_j^{(n)} \Phi_k^{(n)}\big(\rho_j^{(n)}\big)\bigg) - \sum_j p_j^{(n)} S\big(\Phi_k^{(n)}\big(\rho_j^{(n)}\big)\big).$$

Then the limit

$$C(\Phi) = \lim_{n\to\infty} \sup_{\{p_j^{(n)}, \rho_j^{(n)}\}} \bar{\chi}^{(n)}\big(\{p_j^{(n)}, \rho_j^{(n)}\}\big) \tag{15}$$

exists.

We now formulate a quantum Feinstein Lemma, which is the analogue of Theorem 1 in [6] for irreducible Markov chains. It is given by the following theorem.

Theorem 2 (Quantum Feinstein Lemma) *Suppose that $\Phi^{(n)}$ is a quantum channel defined by formula* (3), *where $(q_{i,j})_{i,j\in I}$ is the transition matrix of a* (*periodic*) *irreducible Markov chain, and the maps $\Phi_i : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{K})$ are CPT maps.*

For all $\epsilon > 0$, there exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$, there exist at least $N = N_n \geq 2^{n[C(\Phi)-\epsilon]}$ states $\tilde{\rho}_1^{(n)}, \dots, \tilde{\rho}_N^{(n)}$ on $\mathcal{H}^{\otimes n}$, and positive operators $E_1^{(n)}, \dots, E_N^{(n)}$ on $\mathcal{K}^{\otimes n}$ such that $\sum_{k=1}^N E_k^{(n)} \leq \mathbf{1}$ and

$$\mathrm{Tr}\big[\Phi^{(n)}(\tilde{\rho}_k^{(n)}) E_k^{(n)}\big] > 1 - \epsilon \tag{16}$$

for all $k = 1, \dots, N$.

The proof of this lemma is given in Sect. 4.2. In the case where the period $L \geq 2$, we use the idea of adding a preamble to the codewords (as was done in [7]) to distinguish between the different subclasses of the Markov chain. The construction of the preamble is discussed in the following subsection.

4.1 Construction of a Preamble

To distinguish between the different subclasses, $C^{(k)}$, of the quantum channel $\Phi^{(n)}$, we add a preamble to the input state encoding each message in the set $\mathcal{M}_n$. This is given by an m-fold tensor product of suitable states (as described below). Let us first sketch the idea behind adding such a preamble. Helstrøm [11] showed that two states σ_1 and σ_2, occurring with *a priori* probabilities γ_1 and γ_2 respectively, can be distinguished with an asymptotically vanishing probability of error, if a suitable collective measurement is performed on the m-fold tensor products $\sigma_1^{\otimes m}$ and $\sigma_2^{\otimes m}$, for a large enough $m \in \mathbb{N}$. The optimal measurement is projection-valued. The relevant projection operators, which we denote by Π^+ and Π^-, are the orthogonal projections onto the positive and negative eigenspaces of the difference operator $A_m = \sigma_1^{\otimes m} - \sigma_2^{\otimes m}$. Here we utilize this idea to distinguish between the different subclasses $\Phi_k^{(n)}$. If the preamble is given by a state $\omega^{\otimes m}$, then, by using Helstrøm's result, we can construct a POVM which distinguishes between the average output states $\sigma_k^{(n)} := \Phi_k^{(n)}(\omega^{\otimes m})$ corresponding to the different subclasses $C^{(k)}$. The outcome of this POVM measurement will serve in turn to determine which subclass of the channel is being used for the initial transmission.

We first show that there exists a preamble that can distinguish between the different subclasses, analogous to the branches in [7]. First notice that the corresponding CPT maps $\Phi_k^{(n)}$ *need not all be distinct*! However, by definition of the period, there is no internal periodicity of these maps; otherwise the chain can be contracted to a single such period. This means, that for any two subclasses $k_1 < k_2$ there exists $l \le L - 1$ such that $\Phi_{k_1}^{(l)} \neq \Phi_{k_2}^{(l)}$.

Example Consider the simple period-2 Markov chain with 3 states given by the transition matrix

$$(q_{ij})_{i,j=1}^3 = \begin{pmatrix} 0 & 1 & 0 \\ 2\gamma_1 & 0 & 2\gamma_3 \\ 0 & 1 & 0 \end{pmatrix}.$$

Its equilibrium state is $(\gamma_1, \gamma_2, \gamma_3)$ with $\gamma_2 = \frac{1}{2}$. Given CPT maps Φ_1, Φ_2, Φ_3 we can construct the corresponding channel map $\Phi^{(n)}$. Now suppose that $2\gamma_1\Phi_1 + 2\gamma_3\Phi_3 = \Phi_2$. Then the channel is in fact a memoryless channel: $\Phi^{(n)} = \Phi_2^{\otimes n}$.

We can now choose $\omega = \omega_{k_1,k_2}^{(l)}$ such that

$$f := F(\Phi_{k_1}^{(l)}(\omega), \Phi_{k_2}^{(l)}(\omega)) < 1, \tag{17}$$

where we have defined the *fidelity* of two states as in [18],

$$F(\sigma, \sigma') = \operatorname{Tr} \sqrt{\sigma^{1/2}\sigma'\sigma^{1/2}}. \tag{18}$$

Lemma 2 *Let $\omega_{k_1,k_2}^{(L)} = \omega_{k_1,k_2}^{(l)} \otimes \rho_0^{\otimes(L-l)}$ be a state as defined above, enhanced to a state on $\mathcal{H}^{\otimes L}$ by a tensor product with states $\rho_0 = \frac{1}{d}\mathbf{1}$, where $d = \dim \mathcal{H}$. Then*

$$F\big(\Phi_{k_1}^{(mL)}((\omega_{k_1,k_2}^{(L)})^{\otimes m}), \Phi_{k_2}^{(mL)}((\omega_{k_1,k_2}^{(L)})^{\otimes m})\big) \to 0 \quad \text{as } m \to \infty. \tag{19}$$

Proof Choose $\alpha > 0$ so small that $1 + \alpha < f^{-1}$. Since the subchains $i + mL$ $m = 0, 1, \ldots$ are irreducible and aperiodic Markov chains with transition matrix $q_{ij}^{(L)}$ on $C^{(k)}$ if $i \in C^{(k)}$, there exists R so large that

$$(1-\alpha)L\gamma_j < \sum_{i_2,\dots,i_{R-1}\in C^{(k)}} q_{ii_2}^{(L)}\cdots q_{i_{R-1}j}^{(L)} < (1+\alpha)L\gamma_j \tag{20}$$

for all $i, j \in C^{(k)}$.

Now let $\{\mathcal{E}_r\}_r$ be a POVM on $\mathcal{B}(\mathcal{K}^{\otimes L})$ such that

$$F(\sigma_1,\sigma_2) = \sum_r \sqrt{\operatorname{Tr}(\sigma_1\mathcal{E}_r)\operatorname{Tr}(\sigma_2\mathcal{E}_r)}, \tag{21}$$

(see e.g. (9.74) in [18]) where we denote

$$\sigma_1 = \Phi_{k_1}^{(L)}(\omega_{k_1,k_2}^{(L)}) \quad \text{and} \quad \sigma_2 = \Phi_{k_2}^{(L)}(\omega_{k_1,k_2}^{(L)}). \tag{22}$$

Then we have

$$\begin{aligned}
&F\big(\Phi_{k_1}^{(mR+m)L}((\omega_{k_1,k_2}^{(L)})^{\otimes(mR+m)}),\, \Phi_{k_2}^{(mR+m)L}((\omega_{k_1,k_2}^{(L)})^{\otimes(mR+m)})\big)\\
&\le \sum_{r_1,\dots,r_m}\left[\operatorname{Tr}\left(\Phi_{k_1}^{(mR+m)L}((\omega_{k_1,k_2}^{(L)})^{\otimes(mR+m)})\bigotimes_{i=1}^m(\mathcal{E}_{r_i}\otimes\mathbf{1}^{(RL)})\right)\right.\\
&\quad\left.\times\operatorname{Tr}\left(\Phi_{k_2}^{(mR+m)}((\omega_{k_1,k_2}^{(L)})^{\otimes(mR+m)})\bigotimes_{i=1}^m(\mathcal{E}_{r_i}\otimes\mathbf{1}^{(RL)})\right)\right]^{1/2}\\
&\le \sum_{r_1,\dots,r_m}(1+\alpha)^{m-1}\prod_{i=1}^m\sqrt{\operatorname{Tr}(\sigma_1\mathcal{E}_{r_i})\operatorname{Tr}(\sigma_2\mathcal{E}_{r_i})}\\
&= (1+\alpha)^{m-1}F(\sigma_1\sigma_2)^m \to 0.
\end{aligned} \tag{23}$$

□

We now introduce, for all $k_1 < k_2$, difference operators $A_{k_1,k_2}^{(m)}$, and corresponding projections $\Pi_{k_1,k_2}^{\pm}$ onto their positive and negative eigenspaces, which serve to distinguish the different possibilities, as in [7]. The difference operators are defined by

$$A_{k_1,k_2}^{(m)} = \Phi_{k_1}^{(mL)}((\omega_{k_1,k_2}^{(L)})^{\otimes m}) - \Phi_{k_2}^{(mL)}((\omega_{k_1,k_2}^{(L)})^{\otimes m}). \tag{24}$$

The following lemma was proved in [7]:

Lemma 3 *Suppose that for a given* $\delta > 0$,

$$|\operatorname{Tr}[|A_{k_1,k_2}^{(m)}|] - 2| \le \delta. \tag{25}$$

Then

$$|\operatorname{Tr}[\Pi_{k_1,k_2}^{+}(\Phi_{k_1}^{(m)}(\omega_L^{\otimes m}))] - 1| \le \frac{\delta}{2} \tag{26}$$

and

$$|\operatorname{Tr}[\Pi_{k_1,k_2}^{-}(\Phi_{k_2}^{(m)}(\omega_L^{\otimes m}))] - 1| \le \frac{\delta}{2}. \tag{27}$$

Defining

$$\tilde{\Pi}_k = \bigotimes_{k_1<k_2} \Gamma^{(k)}_{k_1,k_2}, \quad \text{where } \Gamma^{(k)}_{k_1,k_2} = \begin{cases} \mathbf{1}^{(mL)} & \text{if } k_1 \neq k \text{ and } k_2 \neq k, \\ \Pi^-_{k_1,k} & \text{if } k_2 = k, \\ \Pi^+_{k,k_2} & \text{if } k_1 = k, \end{cases} \tag{28}$$

it follows from the fact that $\Pi^+_{k_1,k_2}\Pi^-_{ck_1,k_2} = 0$, that the projections $\tilde{\Pi}_k$ are also disjoint:

$$\tilde{\Pi}_{k_1}\tilde{\Pi}_{k_2} = 0 \quad \text{for } k_1 \neq k_2. \tag{29}$$

The preamble is given by the product over all pairs $k_1 < k_2$:

$$\omega_L = \bigotimes_{k_1<k_2} \omega^{(L)}_{k_1,k_2}. \tag{30}$$

The following lemma then demonstrates that the preamble can distinguish between initial subclasses k:

Lemma 4 *For all* $k = 0, 1, \ldots, L-1$,

$$\lim_{m\to\infty} \mathrm{Tr}\big[\tilde{\Pi}_k \Phi^{(mL)}_k\big(\omega^{(m)}_L\big)\big] = 1. \tag{31}$$

Proof By Lemma 2,

$$F\big(\Phi^{(mL)}_{k_1}(\omega^{\otimes m}_L), \Phi^{(mL)}_{k_2}(\omega^{\otimes m}_L)\big) \to 0 \tag{32}$$

as $m \to \infty$. Using the inequalities [18]

$$\begin{aligned} \mathrm{Tr}(A_1) + \mathrm{Tr}(A_2) - 2F(A_1, A_2) &\le \|A_1 - A_2\|_1 \\ &\le \mathrm{Tr}(A_1) + \mathrm{Tr}(A_2) \end{aligned} \tag{33}$$

for any two positive operators A_1 and A_2, we find that

$$\big|\mathrm{Tr}\big(|A^{(m)}_{k_1,k_2}|\big) - 2\big| \le \delta_m, \tag{34}$$

where $\delta_m \to 0$ as $m \to \infty$, since

$$\mathrm{Tr}\big(|A^{(m)}_{k_1,k_2}|\big) = \|\Phi^{(mL)}_{k_1}(\omega^{\otimes m}_L) - \Phi^{(mL)}_{k_2}(\omega^{\otimes m}_L)\|_1. \tag{35}$$

Replacing m by $m' = m + R$, where $R \in \mathbb{N}$ is large enough so that (20) holds, to separate the different classes, we have for any k,

$$\begin{aligned} 1 &\ge \mathrm{Tr}\left[\tilde{\Pi}_k \Phi^{(m'L)}_k\left(\bigotimes_{k_1<k_2} \omega^{\otimes m'}_L\right)\right] \\ &\ge (1-\alpha)^L \prod_{k_1<k} \mathrm{Tr}\big[\Pi^-_{k_1,k}\big(\Phi^{(mL)}_k(\omega^{\otimes m}_L)\big)\big] \\ &\quad \times \prod_{k_2>k} \mathrm{Tr}\big[\Pi^+_{k,k_2}\big(\Phi^{(mL)}_k(\omega^{\otimes m}_L)\big)\big] \\ &\ge (1-\alpha)^L \left(1 - \frac{\delta_m}{2}\right)^{L-1} \to 1, \end{aligned} \tag{36}$$

since $\delta_m \to 0$ as $m \to \infty$. The last inequality follows from Lemma 3. □

4.2 Proof of Theorem 2 (Quantum Feinstein Lemma)

Given $\delta > 0$, we fix m_0 so large that

$$\mathrm{Tr}\big[\tilde{\Pi}_k \Phi_k^{(m_0 L)}\big(\omega_L^{\otimes m_0}\big)\big] > 1 - \delta \tag{37}$$

for all $k = 0, 1, \ldots, L-1$. The product state $\omega_L^{\otimes m_0}$, defined through (30), is used as a preamble to the code in order to distinguish between different initial subchannels: if $\rho_r^{(n)} \in \mathcal{A}_n$ is a state encoding the r-th classical message, then the r-th codeword is given by $\omega_L^{\otimes m_0} \otimes \rho_r^{(n)}$.

To prove Feinstein's lemma, we follow the same steps as in the proof of Theorem 5.1 of [7] (see also [6]). First we fix l_0 large enough, and an ensemble $\{p_j^{(l_0)}, \rho_j^{(l_0)}\}$ such that

$$\big|C(\Phi) - \bar{\chi}^{(l_0)}(\{p_j^{(l_0)}, \rho_j^{(l_0)}\})\big| < \frac{\epsilon}{6}, \tag{38}$$

where $\bar{\chi}^{(l_0)}(\{p_j^{(l_0)}, \rho_j^{(l_0)}\})$ is given by (1). We may assume that l_0 is incommensurate with L; for definiteness let $l_0 = a_0 L + 1$. In the following we also assume that n is a multiple of l_0: $n = m l_0$.

Next we prove the existence of typical spaces. For this we first need to define limiting states, and prove ergodicity.

The limiting state $\bar{\sigma}_{k,\infty}$ is defined by

$$\bar{\sigma}_{k,\infty}(A) = \mathrm{Tr}(\bar{\sigma}_k^{(m)} A) \tag{39}$$

for $A \in \mathcal{A}_n$, where $n = m l_0$ and where $\bar{\sigma}_k^{(m)}$ is a density matrix on $\mathcal{K}_{l_0}^{\otimes m}$ with $\mathcal{K}_{l_0} = \mathcal{K}^{\otimes l_0}$, given by

$$\bar{\sigma}_k^{(m)} = \Phi_k^{(n)}(\bar{\rho}_{l_0}^{\otimes m}). \tag{40}$$

These states are ergodic:

Lemma 5 *The state $\bar{\sigma}_{k,\infty}$ is strongly clustering and hence completely ergodic for L-shifts, i.e., for any $A, B \in \mathcal{B}(\mathcal{K}^{\otimes m l_0})$,*

$$\lim_{s\to\infty} \bar{\sigma}_{k,\infty}(A \tau^{sLl_0}(B)) = \mathrm{Tr}(\bar{\sigma}_k^{(m)} A)\,\mathrm{Tr}(\bar{\sigma}_k^{(m)} B). \tag{41}$$

Proof Let $A, B \in \mathcal{B}(\mathcal{K}^{\otimes m l_0})$. We may assume that m is a multiple of L. Then

$$\begin{aligned}
&\lim_{s\to\infty} \bar{\sigma}_\infty(A\tau^{sLl_0}(B)) \\
&\quad = \lim_{s\to\infty} \sum_{j_1,\ldots,j_{m+sL}} p_{j_1}^{(l_0)} \cdots p_{j_{m+sL}}^{(l_0)} \sum_{i\in C^{(k)}} \sum_{i_2,\ldots,i_{(m+sL)l_0}\in I} L\gamma_{i_1} q_{i_1,i_2} \cdots q_{i_{(m+sL)l_0-1}, i_{(m+sL)l_0}} \\
&\qquad \times \mathrm{Tr}\big[\big(\Phi_{i_1} \otimes \cdots \otimes \Phi_{i_{(m+sL)l_0}}\big)(\rho_{j_1}^{(l_0)} \otimes \cdots \otimes \rho_{j_{m+sL}}^{(l_0)})(A \otimes \mathbf{1}_{sLl_0})(\mathbf{1}_{sLl_0} \otimes B)\big] \\
&\quad = \lim_{s\to\infty} \sum_{j_1,\ldots,j_m} \sum_{j_1',\ldots,j_m'} \prod_{\alpha=1}^m (p_{j_\alpha}^{(l_0)} p_{j_\alpha'}^{(l_0)}) \sum_{i\in C^{(k)}} \sum_{i_2,\ldots,i_{ml_0+1}\in I} L\gamma_{i_1} q_{i_1,i_2} \cdots q_{i_{ml_0}, i_{ml_0+1}} \\
&\qquad \times \sum_{i_1',\ldots,i'_{(s-m/L)l_0}\in C^{(k)}} q^{(L)}_{i_{ml_0+1}, i_1'} \cdots q^{(L)}_{i'_{(s-m/L)l_0-1}, i'_{(s-m/L)l_0}}
\end{aligned}$$

$$\times \sum_{i_{(sLl_0+1)},\dots,i_{(sL+m)l_0}} q_{i'_{(s-m/L)l_0},i_{sLl_0+1}} \cdots q_{i_{(m+sL)l_0-1},i_{(m+sL)l_0}}$$
$$\times \operatorname{Tr}\big[\Phi_{i_1} \otimes \cdots \otimes \Phi_{i_{ml_0}} (\rho_{j_1}^{(l_0)} \otimes \cdots \otimes \rho_{j_m}^{(l_0)}) A\big]$$
$$\times \operatorname{Tr}\big[\Phi_{i_{sl_0+1}} \otimes \cdots \otimes \Phi_{i_{(m+s)l_0}} (\rho_{j'_1}^{(l_0)} \otimes \cdots \otimes \rho_{j'_n}^{(l_0)}) B\big]. \tag{42}$$

Now, because the subchains are irreducible and aperiodic,

$$\sum_{i'_{n+1},\dots,i'_s} q^{(L)}_{i,i'_{n+1}} \cdots q^{(L)}_{i'_{s-1},i'_s} g(i'_s) \to \sum_{j\in C^{(k)}} L\gamma_j g(j) \tag{43}$$

as $s \to \infty$, if $i \in C^{(k)}$. Therefore

$$\lim_{s\to\infty} \bar{\phi}_\infty(A\tau^{sl_0}(B))$$
$$= \sum_{j_1,\dots,j_m} \prod_{\alpha=1}^m p_{j_\alpha}^{(l_0)} \sum_{i_1,\dots,i_{ml_0}} L\gamma_{i_1} q_{i_1,i_2} \cdots q_{i_{ml_0-1},i_{ml_0}}$$
$$\times \operatorname{Tr}\big[\Phi_{i_1} \otimes \cdots \otimes \Phi_{i_{ml_0}} (\rho_{j_1}^{(l_0)} \otimes \cdots \otimes \rho_{j_m}^{(l_0)}) A\big]$$
$$\times \sum_{j'_1,\dots,j'_m} \prod_{\alpha=1}^m p_{j'_\alpha}^{(l_0)} \sum_{i''_1,\dots,i''_{ml_0}} L\gamma_{i''_1} q_{i''_1,i''_2} \cdots q_{i''_{ml_0-1},i''_{ml_0}}$$
$$\times \operatorname{Tr}\big[\Phi_{i''_1} \otimes \cdots \otimes \Phi_{i''_{ml_0}} (\rho_{j'_1}^{(l_0)} \otimes \cdots \otimes \rho_{j'_m}^{(l_0)}) B\big]$$
$$= \sum_{j_1,\dots,j_m} \prod_{\alpha=1}^m p_{j_\alpha}^{(l_0)} \operatorname{Tr}\big[\Phi^{(ml_0)} (\rho_{j_1}^{(l_0)} \otimes \cdots \otimes \rho_{j_n}^{(l_0)}) A\big]$$
$$\times \sum_{j'_1,\dots,j'_m} \prod_{\alpha=1}^m p_{j'_\alpha}^{(l_0)} \operatorname{Tr}\big[\Phi^{(ml_0)} (\rho_{j'_1}^{(l_0)} \otimes \cdots \otimes \rho_{j'_n}^{(l_0)}) B\big]$$
$$= \operatorname{Tr}(\bar{\sigma}_k^{(m)} A) \operatorname{Tr}(\bar{\sigma}_k^{(m)} B). \tag{44}$$

□

The existence of the entropy is well known, but here we need that it is independent of k:

Lemma 6 *The mean specific entropy*

$$S_M = \lim_{m\to\infty} \frac{1}{ml_0} S\big(\Phi_k^{(ml_0)}(\bar{\rho}_{l_0}^{\otimes m})\big) \quad \textit{and} \quad \bar{\rho}_{l_0} = \sum_j p_j^{(l_0)} \rho_j^{(l_0)} \tag{45}$$

exists and is independent of k.

Proof See Appendix A. □

We now prove the existence of typical projections $\bar{P}_k^{(n)}$ $(k = 0, \dots, L-1)$ for an irreducible channel with period L:

Lemma 7 *Given $\epsilon, \delta > 0$, there exists $m_0 \in \mathbb{N}$ such that for $m \geq m_0$ there are subspaces $\mathcal{N}_{k,\epsilon}^{(n)} \subset \mathcal{K}_{l_0}^{\otimes m}$ $(n = ml_0)$, with projections $\bar{P}_k^{(n)}$ such that*

$$\bar{P}_k^{(n)} \Phi_k^{(n)} (\bar{\rho}_{l_0}^{\otimes m}) \bar{P}_k^{(n)} \leq 2^{-m[S_M - \frac{\epsilon}{4}]}, \tag{46}$$

and

$$\mathrm{Tr}\big(\Phi_k^{(ml_0)} (\bar{\rho}_{l_0}^{\otimes m}) \bar{P}_k^{(n)}\big) > 1 - \delta^2. \tag{47}$$

Proof We follow Hiai and Petz [12]. Let m_1 be so large that for $m \geq m_1$,

$$S_M \leq \frac{1}{ml_0} S\big(\Phi_k^{(ml_0)} (\bar{\rho}_{l_0}^{\otimes m})\big) < S_M + \frac{\epsilon}{8}. \tag{48}$$

Now assume that m_1 is a multiple of L, and let $\Omega = \{\lambda_r\}$ denote the spectrum of $\bar{\sigma}_k^{(m_1)}$, and write π_r for the projection onto the eigenvector with eigenvalue λ_r. For any $p > 0$ and $\mathcal{X} \subset \Omega^p$, put

$$q_{\mathcal{X}} = \sum_{(\lambda_{r_1}, \ldots, \lambda_{r_p}) \in \mathcal{X}} \pi_{r_1} \otimes \cdots \otimes \pi_{r_p}, \tag{49}$$

and define the probability measures ν_p on Ω^p and ν_∞ on $\Omega^{\mathbb{N}}$ by

$$\nu_p(\mathcal{X}) = \mathrm{Tr}(\Phi^{(pm_1 l_0)} (\bar{\rho}_{l_0}^{\otimes (pm_1)}) q_{\mathcal{X}}) \quad \text{and} \quad \nu_\infty(\mathcal{X}) = \bar{\sigma}_\infty(q_{\mathcal{X}}). \tag{50}$$

By Lemma 5, ν_∞ is ergodic and by McMillan's theorem [16] there exists a typical set

$$T_{k,\epsilon}^{(p)} = \big\{(\lambda_{r_1}, \ldots, \lambda_{r_p}) \in \Omega^p \,| \\ 2^{-p(h_{KS}(\nu_\infty) + \epsilon/8)} \leq \nu_p(\{(\lambda_{r_1}, \ldots, \lambda_{r_p})\}) \leq 2^{-p(h_{KS}(\nu_\infty) - \epsilon/8)}\big\}, \tag{51}$$

satisfying

$$\nu_p(T_{k,\epsilon}^{(p)}) > 1 - \delta^2 \tag{52}$$

for p large enough, where $h_{KS}(\nu_\infty)$ is the Kolmogorov-Sinai entropy. Now,

$$h_{KS}(\nu_\infty) = \inf_p \frac{1}{p} H(\nu_p) \leq H(\nu_1) = S(\bar{\sigma}_k^{(m_1)}) < m_1 \left(S_M + \frac{\epsilon}{8}\right), \tag{53}$$

where $H(\nu)$ denotes the Shannon entropy corresponding to the probability measure ν. On the other hand

$$h_{KS}(\nu_\infty) \geq m_1 S_M \tag{54}$$

because, by positivity of the relative entropy,

$$S(\bar{\sigma}_k^{(pm_1)}) \\ = -\mathrm{Tr}\big[\bar{\sigma}_k^{(pm_1)} \log \bar{\sigma}_k^{(pm_1)}\big] \\ \leq -\mathrm{Tr}\Big[\bar{\sigma}_k^{(pm_1)} \log\Big(\bigoplus_{r_1,\ldots,r_p} \big[\mathrm{Tr}\big(\bar{\sigma}^{(pm_1)}\big](\pi_{r_1} \otimes \cdots \otimes \pi_{r_p})\big)\pi_{r_1} \otimes \cdots \otimes \pi_{r_p}\Big)\Big]$$

$$= - \sum_{r_1,\dots,r_p} \mathrm{Tr}[\bar{\sigma}_k^{(pm_1)} \pi_{r_1} \otimes \cdots \otimes \pi_{r_p}] \log \mathrm{Tr}[\bar{\sigma}_k^{(pm_1)} \pi_{r_1} \otimes \cdots \otimes \pi_{r_p}]$$
$$= H(\nu_p). \tag{55}$$

For $n = ml_0$, $m = pm_1$, and $\underline{r} = (r_1, \dots, r_p)$ we now define

$$\overline{T}_{k,\epsilon}^{(p)} = \{\underline{r} : (\lambda_{r_1}, \dots, \lambda_{r_p}) \in T_{k,\epsilon}^{(p)}\}, \tag{56}$$

$$\pi_{k,\underline{r}}^{(m)} = q_{\{(\lambda_{r_1},\dots,\lambda_{r_p})\}}, \tag{57}$$

and

$$\mathcal{N}_{k,\epsilon}^{(n)} = \bigoplus_{\underline{r}\in\overline{T}_{k,\epsilon}^{(p)}} \pi_{k,\underline{r}}^{(m)} (\mathcal{K}_{l_0}^{\otimes m}). \tag{58}$$

Then

$$\mathrm{Tr}\big(\Phi_k^{(ml_0)}(\bar{\rho}_{l_0}^{\otimes m}) \bar{P}_k^{(n)}\big) = \mathrm{Tr}\big(\bar{\sigma}_k^{(n)} \big(\oplus_{\underline{r}\in\overline{T}_{k,\epsilon}^{(n)}} \pi_{k,\underline{r}}^{(m)}\big)\big)$$
$$= \mathrm{Tr}\big[\bar{\sigma}_k^{(pm_1)} q_{\overline{T}_{k,\epsilon}^{(p)}}\big]$$
$$= \nu_p\big(T_{k,\epsilon}^{(p)}\big) > 1 - \delta^2. \tag{59}$$

Moreover,

$$\bar{P}_k^{(n)} \bar{\sigma}_k^{(n)} \bar{P}_k^{(n)} = \bigoplus_{\underline{r}\in\overline{T}_{k,\epsilon}^{(p)}} \pi_{k,\underline{r}}^{(m)} \bar{\sigma}_k^{(n)} \pi_{k,\underline{r}}^{(m)}$$
$$\le \bigoplus_{\underline{r}\in\overline{T}_{k,\epsilon}^{(p)}} 2^{-p(h_{KS}(\nu_\infty)-\epsilon/8)} \pi_{k,\underline{r}}^{(m)}$$
$$\le \bigoplus_{\underline{r}\in\overline{T}_{k,\epsilon}^{(p)}} 2^{-p(m_1 S_M-\epsilon/8)} \pi_{k,\underline{r}}^{(m)}$$
$$\le 2^{-m(S_M-\epsilon/8)} \mathbf{1}. \tag{60}$$

This establishes the result in case m is a large enough multiple of m_1. If m is not a multiple of m_1, we simply pad $\pi_{k,\underline{r}}^{(m)}$ with identity operators on $\mathcal{B}(\mathcal{K}_{l_0}^{\otimes(m-[m/m_1]m_1)})$. □

We need a similar result for the individual states $\rho_{\underline{j}}^{(l_0)}$. This is stated in Lemma 9. To formulate this lemma, we define density matrices Σ_{ml_0} in algebras

$$\mathcal{M}_{ml_0} = \bigoplus_{j_1,\dots,j_m=1}^{J} \mathcal{B}(\mathcal{K}_{l_0}^{\otimes m})$$

by

$$\Sigma_{ml_0} = \bigoplus_{j_1,\dots,j_m} p_{\underline{j}}^{(m)} \Phi^{(ml_0)}\big(\rho_{j_1}^{(l_0)} \otimes \cdots \otimes \rho_{j_m}^{(l_0)}\big), \tag{61}$$

where $p_{\underline{j}}^{(m)} = \prod_{\alpha=1}^{m} p_{j_\alpha}^{(l_0)}$ and $\rho_j^{(l_0)}$, $j \in \{1, 2, \dots, J\}$, belongs to the maximising ensemble (c.f. (38)). In the following we denote $\rho_{\underline{j}}^{(ml_0)} = \rho_{j_1}^{(l_0)} \otimes \cdots \otimes \rho_{j_m}^{(l_0)}$, with $\underline{j} = (j_1, j_2, \dots, j_m)$, for any $m \in \mathbb{N}$.

Lemma 8 *There exists a unique translation-invariant state ψ_∞ on $\mathcal{M}_\infty = \overline{\bigcup_{m=1}^{\infty} \mathcal{M}_{ml_0}}$ such that*

$$\psi_\infty(A) = \mathrm{Tr}\left(\Sigma_{ml_0} A\right) \tag{62}$$

for $A \in \mathcal{M}_{ml_0}$. Moreover, this state is strongly clustering and therefore completely ergodic.

Proof The proof of this lemma is similar to that of Lemma 5. □

Lemma 9 *Given $k \in \{0, \dots, L-1\}$, and a sequence $\underline{j} = (j_1, \dots, j_m) \in \{1, 2, \dots, J\}^m$, let $P_{k,\underline{j}}^{(n)}$ be the projection onto the subspace of $\mathcal{K}^{\otimes n}$ spanned by those eigenvectors of $\Phi_k^{(n)}(\rho_{\underline{j}}^{(n)})$ for which the corresponding eigenvalues $\lambda_{\underline{j},\underline{r}}$ satisfy*

$$\left| \frac{1}{m} \log \lambda_{\underline{j},\underline{r}} + \bar{S}_M \right| < \frac{\epsilon}{4}, \tag{63}$$

where

$$\bar{S}_M = \lim_{m \to \infty} \frac{1}{m} \sum_{\underline{j}} p_{\underline{j}}^{(n)} S\big(\Phi_k^{(n)}(\rho_{j_1}^{(l_0)} \otimes \cdots \otimes \rho_{j_m}^{(l_0)})\big). \tag{64}$$

For any $\delta > 0$ there exists m_2 such that for $m \geq m_2$,

$$\mathbb{E}\left(\mathrm{Tr}\left[\Phi_k^{(ml_0)}\left(\bigotimes_{r=1}^{m} \rho_{j_r}^{(l_0)}\right) P_{k,\underline{j}}^{(n)}\right]\right) > 1 - \delta^2. \tag{65}$$

Note that, just as S_M, $\bar{S}_M$ is also independent of k. It can be equivalently expressed as

$$\bar{S}_M = \lim_{m \to \infty} \frac{1}{mL} \sum_{k=0}^{L-1} \sum_{\underline{j}} p_{\underline{j}}^{(n)} S\big(\Phi_k^{(n)}(\rho_{j_1}^{(l_0)} \otimes \cdots \otimes \rho_{j_m}^{(l_0)})\big).$$

We now proceed with the proof of Feinstein's lemma. This is entirely analogous to the proof in [7], so we can be brief.

Let $N = \tilde{N}(n)$ be the maximal number of states $\tilde{\rho}_1^{(n)}, \dots, \tilde{\rho}_N^{(n)}$ on $\mathcal{H}^{\otimes n}$ for which there exist positive operators $E_1^{(n)}, \dots, E_N^{(n)}$ on $\mathcal{K}^{\otimes n}$ such that

(i) $E_r^{(n)} = \sum_{k=0}^{L-1} \tilde{\Pi}_k \otimes E_{k,r}^{(n)}$ and $\sum_{r=1}^{N} E_{k,r}^{(n)} \leq \bar{P}_k^{(n)}$ for $k = 0, \dots, L-1$, and

(ii) $\frac{1}{L} \sum_{k=0}^{L-1} \mathrm{Tr}[(\tilde{\Pi}_k \otimes E_{k,r}^{(n)}) \Phi_k^{(m_0 L + n)}(\omega_L^{\otimes m_0} \otimes \tilde{\rho}_r^{(n)})] > 1 - \epsilon$, and

(iii) $\frac{1}{L} \sum_{k=0}^{L-1} \mathrm{Tr}[(\tilde{\Pi}_k \otimes E_{k,r}^{(n)}) \Phi_C^{(m_0 L + n)}(\omega_L^{\otimes m_0} \otimes \bar{\rho}^{(n)})] \leq 2^{-n[C(\Phi) - \frac{1}{2}\epsilon]}$.

Remark Note that we can append $\mathbf{1}^{(n - ml_0)}$ to all POVM elements, to reduce the proof to the case $n = ml_0$. In the following we therefore assume $n = ml_0$ for simplicity.

For each k and each $\underline{j} = (j_1, \ldots, j_m)$, we define

$$V_{k,\underline{j}}^{(n)} = \left(\bar{P}_k^{(n)} - \sum_{r=1}^{N} E_{k,r}^{(n)}\right)^{1/2} \bar{P}_k^{(n)} P_{k,\underline{j}}^{(n)} \bar{P}_k^{(n)} \left(\bar{P}_k^{(n)} - \sum_{r=1}^{N} E_{k,r}^{(n)}\right)^{1/2}. \tag{66}$$

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

Put

$$V_{\underline{j}}^{(n)} := \sum_{k=0}^{L-1} \tilde{\Pi}_k \otimes V_{k,\underline{j}}^{(n)}. \tag{67}$$

This is a candidate for an additional measurement operator, $E_{N+1}^{(n)}$, for Bob with corresponding input state $\tilde{\rho}_{N+1}^{(n)} = \omega_L^{(m_0)} \otimes \rho_{\underline{j}}^{(n)}$. Clearly, the condition (i) is satisfied, and we also have

Lemma 10

$$\frac{1}{L} \sum_{k=0}^{L-1} \mathrm{Tr}\big[(\tilde{\Pi}_k \otimes V_{k,\underline{j}}^{(n)}) \Phi_k^{(m_0 L+n)} \big(\omega_L^{(m_0)} \otimes \bar{\rho}_{l_0}^{\otimes m}\big)\big] \le 2^{-n[C(\Phi) - \frac{2}{3}\epsilon]}. \tag{68}$$

Proof In fact we prove the estimate for each k individually:

$$\mathrm{Tr}\big[(\tilde{P}i_k \otimes V_{k,\underline{j}}^{(n)}) \Phi_k^{(m_0 L+n)} \big(\omega_L^{(m_0)} \otimes \bar{\rho}_{l_0}^{\otimes m}\big)\big] \le 2^{-n[C(\Phi) - \frac{2}{3}\epsilon]}. \tag{69}$$

Let $Q_{k,n} = \sum_{r=1}^{N(n)} E_{k,r}^{(n)}$. Note that $Q_{k,n}$ commutes with $\bar{P}_k^{(n)}$ by condition (i). We have

$$\bar{P}_k^{(n)} \bar{\sigma}_k^{(n)} \bar{P}_k^{(n)} \le 2^{-m[S_M - \frac{1}{4}\epsilon]} \mathbf{1} \tag{70}$$

and hence,

$$\bar{P}_k^{(n)} \bar{\sigma}_k^{(n)} \bar{P}_k^{(n)} \le 2^{-n[\frac{1}{m l_0} S(\Phi_k^{(m l_0)}(\bar{\rho}_{l_0}^{\otimes m})) - \frac{1}{4}\epsilon]} \tag{71}$$

for m large enough.

Using this, we get

$$\begin{aligned}
&\mathrm{Tr}(\bar{\sigma}_k^{(n)} V_{k,\underline{j}}^{(n)}) \\
&\quad = \mathrm{Tr}\big[\bar{\sigma}_k^{(n)} (\bar{P}_k^{(n)} - Q_{k,n})^{1/2} \bar{P}_k^{(n)} P_{k,\underline{j}}^{(n)} \bar{P}_k^{(n)} (\bar{P}_k^{(n)} - Q_{k,n})^{1/2}\big] \\
&\quad = \mathrm{Tr}\big[\bar{P}_k^{(n)} \bar{\sigma}_k^{(n)} \bar{P}_k^{(n)} (\bar{P}_k^{(n)} - Q_{k,n})^{1/2} P_{k,\underline{j}}^{(n)} (\bar{P}_k^{(n)} - Q_{k,n})^{1/2}\big] \\
&\quad \le 2^{-n[\frac{1}{m l_0} S(\Phi_k^{(m l_0)}(\bar{\rho}_{l_0}^{\otimes m})) - \frac{1}{4}\epsilon]} \mathrm{Tr}\big[(P_{k,\underline{j}}^{(m l_0)}\big].
\end{aligned} \tag{72}$$

However, by Lemma 9,

$$\mathrm{Tr}(P_{k,\underline{j}}^{(m l_0)}) \le 2^{m[\bar{S}_M + \frac{1}{4}\epsilon]} \le 2^{n[\frac{1}{m l_0} \sum_{\underline{j}} p_{\underline{j}}^{(l_0)} S(\Phi_k^{(n)}(\rho_{\underline{j}}^{(m l_0)}) + \frac{1}{4}\epsilon]}. \tag{73}$$

Hence

$$\mathrm{Tr}(\bar{\sigma}_k^{(n)} V_{k,\underline{j}}^{(n)}) \le 2^{-n[\frac{1}{ml_0}(S(\Phi_k^{(ml_0)}(\bar{\rho}_{l_0}^{\otimes m}))-\sum_{\underline{j}} p_{\underline{j}}^{(ml_0)} S(\Phi_k^{(n)}(\rho_{\underline{j}}^{(ml_0)})))+\frac{1}{2}\epsilon]} \le 2^{-n[C(\Phi)-\frac{2}{3}\epsilon]}. \tag{74}$$

The inequality (69) now follows from the definition of ω_L and (20). □

By maximality of N_n it now follows that the condition (ii) above cannot hold and, upon taking expectations,

Corollary 1

$$\frac{1}{L}\sum_{k=0}^{L-1} \mathbb{E}\big(\mathrm{Tr}\big[(\tilde{\Pi}_k \otimes V_{k,\underline{j}}^{(n)})\Phi_k^{(m_0M+n)}\big(\omega^{(m_0L)} \otimes \tilde{\rho}_r^{(n)}\big)\big]\big) \le 1-\epsilon. \tag{75}$$

We also need the following

Lemma 11 *Assume $\eta > 3\delta$. Then, for n large enough,*

$$\frac{1}{L}\sum_{k=0}^{L-1} \mathbb{E}\big(\mathrm{Tr}\big[(\tilde{\Pi}_k \otimes \bar{P}_k^{(n)} P_{k,\underline{j}}^{(n)} \bar{P}_k^{(n)})\Phi_k^{(m_0L+n)}\big(\omega_L^{(m_0)} \otimes \rho_{\underline{j}}^{(n)}\big)\big]\big) > 1-\eta. \tag{76}$$

Proof We write

$$\begin{aligned}
&\mathbb{E}\big(\mathrm{Tr}\big[(\tilde{\Pi}_k \otimes \bar{P}_k^{(n)} P_{k,\underline{j}}^{(n)} \bar{P}_k^{(n)})\Phi_k^{(m_0L+n)}\big(\omega_L^{(m_0)} \otimes \rho_{\underline{j}}^{(n)}\big)\big]\big) \\
&\quad \ge (1-\alpha)\mathbb{E}\big(\mathrm{Tr}\big[\sigma_{k,\underline{j}}^{(n)} P_{k,\underline{j}}^{(n)}\big]\big) - \mathbb{E}\big(\mathrm{Tr}\big[\sigma_{k,\underline{j}}^{(n)}(\mathbf{1}-\bar{P}_k^{(n)})P_{k,\underline{j}}^{(n)}\big]\big) \\
&\qquad - \mathbb{E}\big(\mathrm{Tr}\big[\sigma_{k,\underline{j}}^{(n)} \bar{P}_k^{(n)} P_{k,\underline{j}}^{(n)}(\mathbf{1}-\bar{P}_k^{(n)})\big]\big).
\end{aligned} \tag{77}$$

By Lemma 9, the first term is $> (1-\alpha)(1-\delta^2) > 1-\delta$ provided α is small enough. The last two terms can be bounded using Cauchy-Schwarz and Lemma 7:

$$\mathbb{E}\big(\mathrm{Tr}\big[\sigma_{k,\underline{j}}^{(n)}(\mathbf{1}-\bar{P}_k^{(n)})P_{k,\underline{j}}^{(n)}\big]\big) \le \big(\mathrm{Tr}\big[\bar{\sigma}_k^{(n)}(\mathbf{1}-\bar{P}_k^{(n)})\big]\big)^{1/2} < \delta \tag{78}$$

and

$$\mathbb{E}\big(\mathrm{Tr}\big[\sigma_{k,\underline{j}}^{(n)} \bar{P}_k^{(n)} P_{k,\underline{j}}^{(n)}(\mathbf{1}-\bar{P}_k^{(n)})\big]\big) \le \delta. \tag{79}$$

□

Lemma 12 *Assume $\eta < \frac{1}{3}\epsilon$ and write*

$$Q_{k,n} = \sum_{r=1}^{N} E_{k,r}^{(n)}. \tag{80}$$

Then for n large enough,

$$\frac{1}{L}\sum_{k=0}^{L-1} \mathrm{Tr}\big[(\tilde{\Pi}_k \otimes Q_{k,n})\Phi_k^{(m_0L+n)}\big(\omega_L^{(m_0)} \otimes \rho_{\underline{j}}^{(n)}\big)\big] \ge (\eta)^2. \tag{81}$$

Proof Define

$$Q'_{k,n} = \bar{P}_k^{(n)} - (\bar{P}_k^{(n)} - Q_{k,n})^{1/2}. \tag{82}$$

By the above corollary,

$$\begin{aligned}
&1-\epsilon \\
&\quad \geq \frac{1}{L}\sum_{k=0}^{L-1}\mathbb{E}\big(\mathrm{Tr}\big[(\tilde{\Pi}_k \otimes (\bar{P}_k^{(n)} - Q'_{k,n})P_{k,\underline{j}}^{(n)}(\bar{P}_k^{(n)} - Q'_{k,n}))\Phi_k^{(m_0M+n)}\big(\omega^{(m_0L)} \otimes \tilde{\rho}_r^{(n)}\big)\big]\big) \\
&\quad = \frac{1}{L}\sum_{k=0}^{L-1}\mathbb{E}\big(\mathrm{Tr}\big[(\tilde{\Pi}_k \otimes \bar{P}_k^{(n)}P_{k,\underline{j}}^{(n)}\bar{P}_k^{(n)})\Phi_k^{(m_0M+n)}\big(\omega^{(m_0L)} \otimes \tilde{\rho}_r^{(n)}\big)\big]\big) \\
&\qquad - \frac{1}{L}\sum_{k=0}^{L-1}\mathbb{E}\big(\mathrm{Tr}\big[\big(\tilde{\Pi}_k \otimes \big(\bar{P}_k^{(n)}P_{k,\underline{j}}^{(n)}Q'_{k,n} + Q'_{k,n}P_{k,\underline{j}}^{(n)}\bar{P}_k^{(n)}\big)\big)\Phi_k^{(m_0M+n)}\big(\omega^{(m_0L)} \otimes \tilde{\rho}_r^{(n)}\big)\big]\big) \\
&\qquad + \frac{1}{L}\sum_{k=0}^{L-1}\mathbb{E}\big(\mathrm{Tr}\big[(\tilde{\Pi}_k \otimes Q'_{k,n}P_{k,\underline{j}}^{(n)}Q'_{k,n})\Phi_k^{(m_0M+n)}\big(\omega^{(m_0L)} \otimes \tilde{\rho}_r^{(n)}\big)\big]\big).
\end{aligned} \tag{83}$$

Since the last term is positive, we have, by Lemma 11,

$$\begin{aligned}
&\frac{1}{L}\sum_{k=0}^{L-1}\mathbb{E}\big(\mathrm{Tr}\big[\big(\tilde{\Pi}_k \otimes \big(\bar{P}_k^{(n)}P_{k,\underline{j}}^{(n)}Q'_{k,n} + Q'_{k,n}P_{k,\underline{j}}^{(n)}\bar{P}_k^{(n)}\big)\big)\Phi_k^{(m_0M+n)}\big(\omega^{(m_0L)} \otimes \tilde{\rho}_r^{(n)}\big)\big]\big) \\
&\quad \geq \epsilon - \eta > 2\eta.
\end{aligned} \tag{84}$$

On the other hand, using Cauchy-Schwarz for each term, the left-hand side is bounded by

$$\frac{2}{L}\sum_{k=0}^{L-1}\mathrm{Tr}\big[(\tilde{\Pi}_k \otimes Q'^2_{k,n})\Phi_k^{(m_0L+n)}\big(\omega_L^{(m_0)} \otimes \rho_{\underline{j}}^{(n)}\big)\big]. \tag{85}$$

To complete the proof, we simply remark that

$$Q_{k,n} \geq (Q'_{k,n})^2. \tag{86}$$

□

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

$$\rho_r^{(m_0ML+n)} = \omega_L^{(m_0)} \otimes \tilde{\rho}_r^{(n)}. \tag{87}$$

For n sufficiently large we then have

$$N = N_{n+m_0L} = \tilde{N}(n) \geq (\eta)^2 2^{n[C(\Phi)-\frac{2}{3}\epsilon]} \geq 2^{(m_0L+n)[C(\Phi)-\epsilon]}. \tag{88}$$

To complete the proof, we need to show that the set $\{E_k^{(n)}\}_{k=1}^N$ satisfies (16). However, this follows immediately from condition (ii). □

5 The General Case

5.1 Proof of the Direct Part of Theorem 1

In the general case we need to distinguish first between classes. This is done by introducing another preamble as for subclasses of periodic classes. The analysis is exactly the same as in [7] and we simply state the necessary lemma:

Lemma 13 *If C and C' are two different periodic classes with periods $L(C)$ and $L(C')$ respectively, then there exists a state $\omega^{(L)}_{C,C'}$ on $\mathcal{H}^{\otimes L}$, where $L = L(C)L(C')$ such that*

$$F\big(\Phi^{(mL)}_{C}((\omega^{(L)}_{C,C'})^{\otimes m}), \Phi^{(mL)}_{C'}((\omega^{(L)}_{C,C'})^{\otimes m})\big) \to 0 \tag{89}$$

as $m \to \infty$.

The difference operators are given by

$$A^{(m)}_{C,C'} = \gamma_C \Phi^{(mL)}_{C}((\omega^{(L)}_{C,C'})^{\otimes m}) - \gamma_{C'} \Phi^{(m)}_{C'}((\omega^{(L)}_{C,C'})^{\otimes m}) \tag{90}$$

and the corresponding projections onto positive and negative eigenvalues are denoted $\Pi^{\pm}_{C,C'}$. We define

$$\tilde{\Pi}_C = \bigotimes_{(C',C'')} \Gamma^{(C)}_{C',C''}, \quad \text{where } \Gamma^{(C)}_{C',C''} = \begin{cases} \mathbf{1}^{(mL)} & \text{if } C' \neq C \text{ and } C'' \neq C, \\ \Pi^{-}_{C',C} & \text{if } C'' = C, \\ \Pi^{+}_{C,C''} & \text{if } C' = C. \end{cases} \tag{91}$$

It follows from the fact that $\Pi^{+}_{C',C''}\Pi^{-}_{C',C''} = 0$, that the projections $\tilde{\Pi}_C$ are also disjoint:

$$\tilde{\Pi}_{C_1}\tilde{\Pi}_{C_2} = 0 \quad \text{for } C_1 \neq C_2. \tag{92}$$

The preamble is given by the product over all pairs C, C':

$$\omega_{\text{cl}} = \bigotimes_{(C,C')} \omega^{(L)}_{C,C'}. \tag{93}$$

The following lemma then demonstrates that the preamble can distinguish between classes C:

Lemma 14 *For all classes C,*

$$\lim_{m\to\infty} \text{Tr}\big[\tilde{\Pi}_C \Phi^{(mL)}_{C}\big(\omega^{(m)}_{\text{cl}}\big)\big] = 1. \tag{94}$$

The proof of the direct part of Theorem 1 now follows the same lines as that of Theorem 2. Note that the maximizing ensemble $\{p^{(l_0)}_j, \rho^{(l_0)}_j\}$ of the minimum

$$\bigwedge_{C\in\mathcal{C}} \bar{\chi}^{(n)}_C(\{p^{(n)}_j, \rho^{(n)}_j\})$$

is in general not the same as that for the individual Holevo quantities $\bar{\chi}^{(n)}_C$. However, the Lemmas 7 and 9 still hold and yield typical spaces with typical projections $\bar{P}^{(n)}_{C,k}$ and $P^{(n)}_{C,k,\underline{j}}$

for each class C. The preamble now consists of $\omega_{\mathrm{cl}}^{\otimes m_{\mathrm{cl}}} \otimes \bigotimes_{C \in \mathcal{C}_{per}} \omega_{C,L}^{\otimes m_C}$. The POVM operators have the form

$$E_r^{(n)} = \sum_{C \in \mathcal{C}} \tilde{\Pi}_C \otimes \sum_{k=0}^{L(C)-1} \tilde{\Pi}_{C,k} \otimes E_{C,k,r}^{(n)}.$$

The remainder of the proof is a carbon copy of that for the irreducible case. □

5.2 Proof of the Converse Part of Theorem 1

In this section we prove that it is impossible for Alice to transmit classical messages reliably to Bob through the channel Φ defined by (3) at a rate $R > C(\Phi)$. This is the (weak) converse part of Theorem 1, 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 states $\rho_\alpha^{(n)}$ in $\mathcal{B}(\mathcal{H}^{\otimes n})$. Let the corresponding outputs for the class C of the channel be denoted by $\sigma_{\alpha,C}^{(n)}$, i.e.

$$\sigma_{\alpha,C}^{(n)} = \Phi_C^{(n)}(\rho_\alpha^{(n)}). \tag{95}$$

Further define

$$\bar{\sigma}_C^{(n)} = \frac{1}{|\mathcal{M}_n|} \sum_{\alpha \in \mathcal{M}_n} \sigma_{\alpha,C}^{(n)}. \tag{96}$$

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

$$\bar{p}_e^{(n)} := 1 - \frac{1}{|\mathcal{M}_n|} \sum_{\alpha \in \mathcal{M}_n} \mathrm{Tr}\left[\Phi^{(n)}(\rho_\alpha^{(n)}) E_\alpha^{(n)}\right]. \tag{97}$$

We also define the average error corresponding to the class C of the channel as

$$\bar{p}_{e,C}^{(n)} := 1 - \frac{1}{|\mathcal{M}_n|} \sum_{\alpha \in \mathcal{M}_n} \mathrm{Tr}\left[\Phi_C^{\otimes n}(\rho_\alpha^{(n)}) E_\alpha^{(n)}\right], \tag{98}$$

so that

$$\bar{p}_e^{(n)} = \sum_{C \in \mathcal{C}} \gamma_C \, \bar{p}_{e,C}^{(n)}. \tag{99}$$

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 $Y_C^{(n)}$ be the random variable corresponding to Bob's inference of Alice's message, when the codeword is transmitted through the class C. It is defined by the conditional probabilities

$$\mathbb{P}[Y_C^{(n)} = \beta \mid X^{(n)} = \alpha] = \mathrm{Tr}[\Phi_C^{(n)}(\rho_\alpha^{(n)}) E_\beta^{(n)}]. \tag{100}$$

By Fano's inequality,

$$h(\bar{p}_{e,C}^{(n)}) + \bar{p}_{C,e}^{(n)} \log(|\mathcal{M}_n| - 1) \geq H(X^{(n)} \mid Y_C^{(n)}) = H(X^{(n)}) - H(X^{(n)} : Y_C^{(n)}). \tag{101}$$

Here $h(\cdot)$ denotes the binary entropy and $H(\cdot)$ denotes the Shannon entropy. By the Holevo bound, for $C \in \mathcal{C}_{aper}$ we have

$$\begin{aligned}
&H(X^{(n)} : Y_C^{(n)}) \\
&\leq S\biggl(\frac{1}{|\mathcal{M}_n|} \sum_{\alpha \in \mathcal{M}_n} \Phi_C^{(n)}(\rho_\alpha^{(n)})\biggr) - \frac{1}{|\mathcal{M}_n|} \sum_{\alpha \in \mathcal{M}_n} S\bigl(\Phi_C^{(n)}(\rho_\alpha^{(n)})\bigr) \\
&= n \bar{\chi}_C \biggl(\biggl\{\frac{1}{|\mathcal{M}_n|}, \rho_\alpha^{(n)}\biggr\}_{\alpha \in \mathcal{M}_n}\biggr),
\end{aligned} \tag{102}$$

where $\bar{\chi}_C(\{\frac{1}{|\mathcal{M}_n|}, \rho_\alpha^{(n)}\}_{\alpha \in \mathcal{M}_n})$ is given by (8).

For $C \in \mathcal{C}_{per}$, with period L,

$$\begin{aligned}
&H(X^{(n)} : Y_C^{(n)}) \\
&\leq S\biggl(\frac{1}{|\mathcal{M}_n|} \sum_{\alpha \in \mathcal{M}_n} \frac{1}{L} \sum_{i \in C} \Phi_{C,i}^{(n)}(\rho_\alpha^{(n)})\biggr) - \frac{1}{|\mathcal{M}_n|} \sum_{\alpha \in \mathcal{M}_n} S\biggl(\frac{1}{L} \sum_{i \in C} \Phi_{C,i}^{(n)}(\rho_\alpha^{(n)})\biggr) \\
&= \frac{1}{|\mathcal{M}_n|} \sum_{\alpha \in \mathcal{M}_n} S\biggl(\frac{1}{L} \sum_{i \in C} \Phi_{C,i}^{(n)}(\rho_\alpha^{(n)}) \biggm\| \frac{1}{|\mathcal{M}_n|} \sum_{\beta \in \mathcal{M}_n} \frac{1}{L} \sum_{i \in C} \Phi_{C,i}^{(n)}(\rho_\beta^{(n)})\biggr) \\
&\leq \frac{1}{|\mathcal{M}_n| L} \sum_{\alpha \in \mathcal{M}_n} \sum_{i \in C} S\biggl(\Phi_{C,i}^{(n)}(\rho_\alpha^{(n)}) \biggm\| \frac{1}{|\mathcal{M}_n|} \sum_{\beta \in \mathcal{M}_n} \Phi_{C,i}^{(n)}(\rho_\beta^{(n)})\biggr) \\
&= \frac{1}{L} \sum_{i \in C} \chi_{C,i}^{(n)}\biggl(\biggl\{\frac{1}{|\mathcal{M}_n|}, \rho_\alpha^{(n)}\biggr\}\biggr) \\
&= n \bar{\chi}_C^{(n)}\biggl(\biggl\{\frac{1}{|\mathcal{M}_n|}, \rho_\alpha^{(n)}\biggr\}\biggr).
\end{aligned} \tag{103}$$

In the above, we use the convexity of the relative entropy $S(\sigma \| \omega) := \mathrm{Tr}\, \sigma(\log \sigma - \log \omega)$, for density matrices σ and ω. See [18, 20].

Therefore, for any class C we have the upper bound

$$H(X^{(n)} : Y_C^{(n)}) \leq n \bar{\chi}_C^{(n)}\left(\left\{\frac{1}{|\mathcal{M}_n|}, \rho_\alpha^{(n)}\right\}\right). \tag{104}$$

Inserting this into Fano's inequality, (101), now yields

$$h(\bar{p}_{C,e}^{(n)}) + \bar{p}_{C,e}^{(n)} \log |\mathcal{M}_n| \geq \log |\mathcal{M}_n| - n \bar{\chi}_C \left(\left\{\frac{1}{|\mathcal{M}_n|}, \rho_\alpha^{(n)}\right\}_\alpha\right). \tag{105}$$

However, since

$$C(\Phi) \geq \bigwedge_{C \in \mathcal{C}} \bar{\chi}_C \left(\left\{\frac{1}{|\mathcal{M}_n|}, \rho_\alpha^{(n)}\right\}_\alpha\right) \tag{106}$$

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

$$\bar{p}_{e,C}^{(n)} \geq 1 - \frac{C(\Phi) + 1/n}{R} > 0. \tag{107}$$

We conclude from (99) and (107) that

$$\bar{p}_e^{(n)} \geq \left(1 - \frac{C(\Phi) + 1/n}{R}\right) \bigwedge_{C \in \mathcal{C}} \gamma_C. \tag{108}$$

□

Remark Note that the strong converse property [9, 24] does not hold for general Markovian channels. For example, for a convex combination of memoryless channels:[2]

$$\Phi^{(n)}(\rho^{(n)}) = \sum_{i=1}^{M} \gamma_i \Phi_i^{\otimes n}(\rho^{(n)}), \tag{109}$$

where $\Phi_i : \mathcal{B}(\mathcal{H}) \to \mathcal{B}()$, Bob's error probability does *not* tend to 1 asymptotically in n for a rate R, such that $C(\Phi) < R < \bar{C}(\Phi)$, where

$$\bar{C}(\Phi) := \bigvee_{i=1}^{M} \chi_i^*,$$

and χ_i^* denotes the Holevo capacity [13, 21] of the memoryless channel Φ_i.

Acknowledgements The authors wish to thank Dr. Igor Bjelaković for pointing out a serious shortcoming in an earlier version of this paper, prompting crucial improvements in the final version. TCD would like to acknowledge the hospitality of the Statistical Laboratory of Cambridge University during his sabbatical. The work was supported by the European Commission through the Integrated Projects SECOQC and FET/QIPC "SCALA".

Appendix A

Lemma 15 *If $\Phi^{(n)}$ is a quantum channel with memory of the form* (3). *Then the limit in* (12) *exists. In particular, the limit in* (15) *exists.*

Proof Denote

$$\bar{\chi}_n = \sup_{\{p_j^{(n)}, \rho_j^{(n)}\}} \bigwedge_{C \in \mathcal{C}} \bar{\chi}_C^{(n)}(\{p_j^{(n)}, \Phi^{(n)}(\rho_j^{(n)})\}). \tag{110}$$

We shall prove that for any $\delta > 0$ there exist n_0 and m_0 such that for all $n' \geq n_0$ and $n \geq m_0 n'$, $\bar{\chi}_n \geq \bar{\chi}_{n'} - \delta$. This proves the lemma because obviously, $0 \leq \bar{\chi}_n \leq \log \dim \mathcal{K}$, and it follows that

$$\liminf_{n \to \infty} \bar{\chi}_n \geq \bar{\chi}_{n'} - \delta$$

and hence $\liminf_{n \to \infty} \bar{\chi}_n \geq \limsup_{n' \to \infty} \bar{\chi}_{n'} - \delta$ where $\delta > 0$ is arbitrary.

[2] A classical version of such a channel was introduced by Jacobs [14] and studied further by Ahlswede [1], who obtained an expression for its capacity.

To prove the statement, let n' be large, and suppose that $\{p_j^{(n')}, \rho_j^{(n')}\}$ is a maximising ensemble for (110), with n replaced by n'. Given $n \geq n'$, put $m = [n/n']$ and $l = n - mn'$. Define the states $\rho_{\underline{j}}^{(n)} = \bigotimes_{r=1}^{m} \rho_{j_r}^{(n')} \otimes \rho_{j_{m+1}}^{(l)}$, where $\rho_j^{(l)}$ is the reduced state on $\mathcal{H}^{\otimes l}$. Then $\bar{\rho}^{(n)} = \otimes_{r=1}^{m} \bar{\rho}^{(n')} \otimes \bar{\rho}^{(l)}$, with $\bar{\rho}^{(n')} := \sum_j p_j^{(n')} \rho_j^{(n')}$. We now write for any class $C \in \mathcal{C}$,

$$\Phi_C^{(n)}(\bar{\rho}^{(n)}) = \sum_{i_1,\ldots,i_{m+1} \in C} \sum_{i'_1,\ldots,i'_{m+1} \in C} \frac{q_{i'_1 i_2}}{\gamma_{i_2}} \cdots \frac{q_{i'_m i_{m+1}}}{\gamma_{i_{m+1}}}$$
$$\times\, \sigma_C^{(n')}(i_1, i'_1) \otimes \cdots \otimes \sigma_C^{(n')}(i_m, i'_m) \otimes \sigma_C^{(l)}(i_{m+1}, i'_{m+1}), \tag{111}$$

where

$$\sigma_C^{(n')}(i, i') = \sum_{i_2,\ldots,i_{n'-1} \in C} \gamma_i q_{i i_2} q_{i_2 i_3} \cdots q_{i_{n'-1} i'}$$
$$\times\, (\Phi_i \otimes \Phi_{i_2} \otimes \cdots \otimes \Phi_{i'})(\bar{\rho}^{(n')}) \tag{112}$$

and similarly for $\sigma_C^{(l)}(i, i')$. Let $\gamma = \bigwedge_{i \in I} \gamma_i$. Using positivity of the density operators and the fact that $q_{ij} \leq 1 \leq \gamma_i/\gamma$, we obtain the simple operator inequality

$$\Phi_C^{(n)}(\bar{\rho}^{(n)}) \leq \frac{1}{\gamma^m} \Phi_C^{(n')}(\bar{\rho}^{(n')}) \otimes \cdots \otimes \Phi_C^{(n')}(\bar{\rho}^{(n')}) \otimes \Phi_C^{(l)}(\bar{\rho}^{(l)}). \tag{113}$$

Inserting this into the definition of $S(\Phi^{(n)}(\bar{\rho}^{(n)}))$ and using the operator monotonicity of the logarithm and the fact that (γ_i) is the equilibrium distribution, i.e. $\sum_{i \in I} \gamma_i q_{ij} = \gamma_j$, we obtain

$$S(\Phi_C^{(n)}(\rho^{(n)})) \geq m S(\Phi_C^{(n')}(\bar{\rho}^{(n')})) + S(\Phi_C^{(l)}(\bar{\rho}^{(l)})) + m \log \gamma. \tag{114}$$

On the other hand, by subadditivity,

$$S(\Phi_C^{(n)}(\rho_{\underline{j}}^{(n)})) \leq \sum_{r=1}^{m} S(\Phi_C^{(n')}(\rho_{j_r}^{(n')})) + S(\Phi_C^{(l)}(\rho_{j_{m+1}}^{(l)})) \tag{115}$$

so that

$$\bar{\chi}_C^{(n)}\big(\{p_{\underline{j}}^{(n)}, \Phi^{(n)}(\rho_{\underline{j}}^{(n)})\}\big) \geq \frac{mn'}{n} \bar{\chi}_{n'} + \frac{m}{n} \log \gamma, \tag{116}$$

for all $C \in \mathcal{C}$. □

Lemma 16 *The mean entropy*

$$S_{M,k} = \lim_{m \to \infty} \frac{1}{m l_0} S\big(\Phi_k^{(m l_0)}(\bar{\rho}_{l_0}^{\otimes m})\big) \tag{117}$$

is independent of k.

Proof We write in a similar way as above

$$\bar{\sigma}^{(m)}(i, i') = \sum_{i_2,\ldots,i_{m l_0 - 1} \in I} \gamma_i q_{i,i_2} \cdots q_{i_{m l_0 - 1}, i'} (\Phi_i \otimes \Phi_{i_2} \otimes \cdots \otimes \Phi_{i_{m l_0 - 1}} \otimes \Phi_{i'})(\bar{\rho}_{l_0}^{\otimes m}), \tag{118}$$

and

$$\Phi_k^{(ml_0)}(\bar{\rho}_{l_0}^{\otimes m}) = L \sum_{i,i' \in C^{(k)}} \sum_{i'' \in C^{(k+1)}} \sum_{i_n \in I} \frac{q_{i',i''}}{\gamma_{i''}} \otimes \bar{\sigma}^{(1)}(i,i')\bar{\sigma}^{(m-1)}(i'',i_n)$$

$$\geq \frac{1}{L\gamma_{\min}^{(k)}} \Phi_k^{(l_0)}(\bar{\rho}_{l_0}) \otimes \Phi_{k+1}^{((m-1)l_0)}(\bar{\rho}_{l_0}^{\otimes(m-1)}), \quad (119)$$

where $\gamma_{\min}^{(k)} = \bigwedge_{i \in C^{(k)}} \gamma_i$. Here we use the assumption that $l_0 = bL + 1$. It follows that

$$S\big(\Phi_k^{(ml_0)}(\bar{\rho}_{l_0}^{\otimes m})\big) \leq S(\Phi_k^{(l_0)}(\bar{\rho}_{l_0}) + S\big(\Phi_{k+1}^{((m-1)l_0)}(\bar{\rho}_{l_0}^{\otimes(m-1)})\big) + \ln(\gamma_{\min}^{(k)} L). \quad (120)$$

This proves that $S_{M,k} \leq S_{M,k+1}$, and by cyclicity they must all be equal. □

References

1. Ahlswede, R.: The weak capacity of averaged channels. Z. Wahrscheinlichk. Verwandte Geb. **11**, 61–73 (1968)
2. Bjelaković, I., Boche, H.: Ergodic classical-quantum channels: structure and coding theorems. quant-ph/0609229
3. Bjelaković, I., Boche, H.: Classical capacities of compound and averaged quantum channels. arXiv:0710.3027
4. Bowen, G., Mancini, S.: Quantum channels with a finite memory. Phys. Rev. A **69**, 01236 (2004)
5. Cover, T.M., Thomas, J.A.: Elements of Information Theory. Wiley, New York (1991)
6. Datta, N., Dorlas, T.C.: A quantum version of Feinstein's lemma and its application to channel coding. In: Proc. of Int. Symp. Inf. Theor. ISIT 2006, Seattle, pp. 441–445 (2006)
7. Datta, N., Dorlas, T.C.: The coding theorem for a class of channels with long-term memory. J. Phys. A, Math. Theory **40**, 8147–8164 (2007)
8. Feinstein, A.: A new basic theorem of information theory. IRE Trans. PGIT **4**, 2–22 (1954)
9. Hayashi, M., Nagaoka, H.: General formulas for capacity of classical-quantum channels. IEEE Trans. Inf. Theory **49**, 1753–1768 (2003)
10. Khinchin, A.I.: Mathematical Foundations of Information Theory. Part II: On the Fundamental Theorems of Information Theory. Dover, New York (1957), Chap. IV
11. Helstrøm, C.W.: Quantum Detection and Estimation Theory. Mathematics in Science and Engineering, vol. 123. Academic Press, London (1976)
12. Hiai, F., Petz, D.: The proper formula for the relative entropy and its asymptotics in quantum probability. Commun. Math. Phys. **143**, 257–281 (1991)
13. Holevo, A.S.: The capacity of a quantum channel with general signal states. IEEE Trans. Inf. Theory **44**, 269–273 (1998)
14. Jacobs, K.: Almost periodic channels. In: Colloqium on Comb. Methods in Prob. Theory. Aarhus (1962)
15. Kretschmann, D., Werner, R.F.: Quantum channels with memory. Phys. Rev. A **72**, 062323 (2005). quant-ph/0502106
16. McMillan, B.: The basic theorems of information theory. Ann. Math. Stat. **24**, 196–219 (1953)
17. Macchiavello, C., Palma, G.M.: Entanglement-enhanced information transmission over a quantum channel with correlated noise. Phys. Rev. A **65**, 050301 (2002)
18. Nielsen, M.A., Chuang, I.L.: Quantum Computation and Quantum Information. Cambridge University Press, Cambridge (2000)
19. Norris, J.R.: Markov Chains. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge (1997)
20. Ohya, M., Petz, D.: Quantum Entropy and Its Use. Springer, Berlin (1993)
21. Schumacher, B., Westmoreland, M.D.: Sending classical information via noisy quantum channels. Phys. Rev. A **56**, 131–138 (1997)
22. Shannon, C.E.: A mathematical theory of communication. Part I. Bell Syst. Tech. J. **27**, 379–423 (1948)
23. Shannon, C.E.: A mathematical theory of communication. Part II. Bell Syst. Tech. J. **27**, 623–656 (1948)
24. Winter, A.: Coding theorem and strong converse for quantum channels. IEEE Trans. Inf. Theory **45**, 2481–2485 (1999)