Quantum capacity of dephasing channels with memory

Quantum capacity of dephasing channels with memory

A D'Arrigo[1,4]**, G Benenti**[2,3] **and G Falci**[1]

[1] MATIS CNR-INFM, Catania and Dipartimento di Metodologie Fisiche e Chimiche per l'Ingegneria, Università degli Studi di Catania, Viale Andrea Doria 6, 95125 Catania, Italy
[2] CNISM, CNR-INFM and Center for Nonlinear and Complex Systems, Università degli Studi dell'Insubria, Via Valleggio 11, 22100 Como, Italy
[3] Istituto Nazionale di Fisica Nucleare, Sezione di Milano, via Celoria 16, 20133 Milano, Italy
E-mail: antonio.darrigo@dmfci.unict.it, giuliano.benenti@uninsubria.it and gfalci@dmfci.unict.it

Abstract. We show that the amount of coherent quantum information that can be reliably transmitted down a dephasing channel with memory is maximized by separable input states. In particular, we model the channel as a Markov chain or a multimode environment of oscillators. While in the first model, the maximization is achieved for the maximally mixed input state, in the latter it is convenient to exploit the presence of a decoherence-protected subspace generated by memory effects. We explicitly compute the quantum channel capacity for the first model while numerical simulations suggest a lower bound for the latter. In both cases memory effects enhance the coherent information. We present results valid for arbitrary input size.

[4] Author to whom any correspondence should be addressed.

Contents

1. Introduction

Quantum communication channels [1, 2] use quantum systems to transfer classical or quantum information. In the first case, we can encode classical bits by means of quantum states. In the latter case, we may want to transfer an unknown quantum state between different units of a quantum system, for instance of a quantum computer, or to distribute entanglement between communicating parties. In both cases, the fundamental question is what is the maximum rate of classical or quantum information that can be faithfully transmitted. Classical and quantum *capacities*, defined as the maximum number of bits/qubits that can be reliably transmitted per channel use, provide the answer to this question.

Quantum channels with memory are the natural theoretical framework for the study of any noisy quantum communication system where correlation times are longer than the time between consecutive uses. This scenario applies to optical fibers which may show a birefringence fluctuating with characteristic time longer than the separation between successive light pulses [3] or to solid state implementations of quantum hardware, where memory effects due to low-frequency impurity noise [4] produce substantial dephasing [5].

Some theoretical results on quantum channels with memory have been already discussed for transmission of both classical and quantum information through a quantum channel. With regard to classical information transmission down a memory channel, it was pointed out that it can be enhanced by using entangled input states [6]–[8], and coding theorems have been recently proved for classes of memory quantum channels [9, 10]. Concerning quantum capacity, a lower bound has been found for some classes of channels with memory [11] and subsequently specific model environments (structured in two parts, one responsible for memory effects and the other acting as a memoryless environment) have been studied [12]–[14]. In particular, coding theorems for quantum capacity have been proved in [14] for the so-called *forgetful channels*, for which memory effects decay exponentially with time.

The problem is formalized by considering the N-uses Hilbert space $\mathcal{H}_N = \mathcal{H}^{\otimes N}$ and defining the system S, described by the reduced density matrix (RDM) ρ for N uses. The input state is $\rho = \sum_{i=1}^{K} p_i \rho_i$, namely states chosen from the ensemble $\{\rho_1, \ldots, \rho_K\}$, with *a priori* probabilities $\{p_1, \ldots, p_K\}$, are sent down the channel. Due to the coupling to further uncontrollable degrees of freedom, the transmission of S may be noisy. The output is therefore described by a linear, completely positive, trace preserving (CPT) map $\mathcal{E}_N(\rho)$, corresponding to N-uses (the single use is defined in $\mathcal{H}$ and described by $\mathcal{E}$). The map $\mathcal{E}_N(\rho)$ can always be

represented starting from an enlarged vector space including a suitably chosen environment E, initially in a pure state: $w_0 \equiv |0\rangle_E \langle 0|$

$$\mathcal{E}_N(\rho) = \text{Tr}_E[U \, (\rho \otimes w_0) \, U^\dagger], \tag{1}$$

where U is a suitable unitary evolution of S + E referring to N uses. The conditional (depending on ρ) evolution of the environment can also be considered. It is described by the environment RDM and allows to define the conjugate CPT map, $w = \text{Tr}_S[U \, (\rho \otimes w_0) \, U^\dagger] =: \tilde{\mathcal{E}}_N(\rho)$.

The *quantum capacity Q* refers to the coherent transmission of quantum information (measured in number of qubits), and it is related to the dimension of the largest subspace of $\mathcal{H}_N$ reliably transmitted down the channel, in the limit of large N. The value of Q can be computed, for memoryless channels, as [15]–[19]

$$Q = \lim_{N \to \infty} \frac{Q_N}{N}, \qquad Q_N = \max_\rho I_c(\mathcal{E}_N, \rho), \tag{2}$$

$$I_c(\mathcal{E}_N, \rho) = S[\mathcal{E}_N(\rho)] - S_e^N(\rho). \tag{3}$$

Here, $S(\rho) = -\text{Tr}[\rho \log_2 \rho]$ is the von Neumann entropy, $S_e^N(\rho) \equiv S[\tilde{\mathcal{E}}_N(\rho)]$ is the *entropy exchange* [20]. The quantity $I_c(\mathcal{E}_N, \rho)$ is called *coherent information* [21] and must be maximized over *all input states ρ*.

The limit $N \to \infty$ in (2) makes the evaluation of Q difficult. On the other hand this regularization is necessary, since in general I_c is not subadditive. Indeed for entangled input states ρ [16], we may have $I_c(\mathcal{E}_N, \rho) > \sum_{k=1}^N I_c(\mathcal{E}, \rho^{(k)})$, where $\rho^{(k)} = \text{Tr}_{S-(k)}(\rho)$ refers to the individual transmission of the kth unit of information, therefore in general it cannot be excluded that $Q_N/N > Q_1$. The regularization is not necessary if the final state w of E can be reconstructed from the final state ρ' of the system. In this case, referred to as *degradable channels* [22]–[25], there exists a CPT map $\mathcal{T}$ such that $\tilde{\mathcal{E}} = \mathcal{T} \circ \mathcal{E}$. It turns out [22] that for degradable channels the coherent information $I_c(\mathcal{E}_N, \rho)$ reduces to a suitable conditional entropy [1], which is subadditive and concave in the input state ρ, and therefore the quantum capacity is given by the 'single-letter' formula $Q = Q_1$.

In this work, we focus on dephasing channels with memory. Dephasing channels are characterized by the property that when N qubits are sent through the channel, the states of a preferential orthonormal basis $\{|j\rangle \equiv |j_1, \ldots, j_N\rangle, \; j_1, \ldots, j_N = 0, 1\}$ are transmitted without errors, implying a conservation law to hold [26]. Therefore, dephasing channels are noiseless from the viewpoint of the transmission of classical information, since the states of the preferential basis can be used for encoding classical information. Of course *superpositions* of basis states may decohere, thus corrupting the transmission of quantum information. Dephasing channels are relevant for systems in which relaxation is much slower than dephasing [4, 27]. When memory effects are taken into account, we have $\mathcal{E}_N \neq \mathcal{E}^{\otimes N}$, i.e. the channel does not act on each carrier *independently*.

We show that the coherent information is maximized by input states separable and diagonal in the reference basis $\{|j\rangle\}$. In particular, we calculate the coherent information for two models of dephasing channels. For a Markov chain, we show that the coherent information is maximized by maximally mixed input states and compute Q. For an environment modeled by a bosonic bath, we propose a coding strategy based on the existence of a decoherence-protected subspace generated by memory effects and use numerical results to suggest a lower bound for Q. It turns out that in both cases memory effects increase the coherent information.

2. The dephasing channel and quantum capacity

The unitary representation of the generalized dephasing channel [22] reads

$$U|j\rangle|0_{\mathsf{E}}\rangle = |j\rangle|\phi_j\rangle, \tag{4}$$

where $|\phi_j\rangle$ are environment states, in general nonmutually orthogonal, describing the conditional evolution. The map $\mathcal{E}_N$ can be written in the Kraus representation [1, 2] as

$$\rho' = \mathcal{E}_N(\rho) = \sum_\alpha A_\alpha\, \rho\, A_\alpha^\dagger, \tag{5}$$

where the system operators $(A_\alpha)_{jl} = \langle\alpha_{\mathsf{E}}|\phi_j\rangle\,\delta_{jl}$ are diagonal in the reference basis (here $\{|\alpha_{\mathsf{E}}\rangle\}$ is an orthonormal basis for the environment). It is easily shown that this channel is degradable [22]. Indeed, for a generic input $\rho = \sum_{j,l}\rho_{jl}|j\rangle\langle l|$, equation (4) yields

$$w = \tilde{\mathcal{E}}_N(\rho) = \sum_j \rho_{jj}|\phi_j\rangle\langle\phi_j|. \tag{6}$$

Since w only depends on the populations ρ_{jj} which are conserved, we can write as well $\tilde{\mathcal{E}}_N = \tilde{\mathcal{E}}_N \circ \mathcal{E}_N$, thus proving degradability.

We now show that for a generalized dephasing channel, the coherent information $I_c(\mathcal{E}_N, \rho)$ is maximized by input states diagonal in the reference basis. To this end, we introduce

$$\rho_k = \frac{1}{2}\left(\rho_{k-1} + \Sigma_z^{(k)}\rho_{k-1}\Sigma_z^{(k)}\right), \quad (k = 1, \ldots, N), \tag{7}$$

where $\rho_0 = \rho$ and the local operator $\Sigma_z^{(k)} = \mathbb{1}^{(1)} \otimes \cdots \otimes \mathbb{1}^{(k-1)} \otimes \sigma_z^{(k)} \otimes \mathbb{1}^{(k+1)} \otimes \cdots \otimes \mathbb{1}^{(N)}$ acts nontrivially only on the kth qubit, by the Pauli operator $\sigma_z^{(k)}$ which has eigenvectors $|j_k\rangle$. We can easily see that ρ_N is the diagonal part of ρ, by using the standard representation of the N-qubit density matrix:

$$\rho = \sum_{\{i_k\}} c_{i_1\ldots i_N}\sigma_{i_1}^{(1)} \otimes \cdots \otimes \sigma_{i_N}^{(N)}, \qquad i_k = 0, x, y, z, \tag{8}$$

where $\sigma_0 = \mathbb{1}$. We now study the action of the operators $\Sigma_z^{(k)}$. First of all $\mathcal{E}_N(\Sigma_z^{(k)}\rho\Sigma_z^{(k)}) = \Sigma_z^{(k)}\mathcal{E}_N(\rho)\Sigma_z^{(k)}$ for any k and ρ, since $\Sigma_z^{(k)}$ commutes with the Kraus operators in (5). Also $S[\Sigma_z^{(k)}\mathcal{E}_N(\rho)\Sigma_z^{(k)}] = S[\mathcal{E}_N(\rho)]$, since the von Neumann entropy is invariant under unitary local transformations. Moreover $\tilde{\mathcal{E}}_N(\Sigma_z^{(k)}\rho\Sigma_z^{(k)}) = \tilde{\mathcal{E}}_N(\rho)$, since the populations of $\Sigma_z^{(k)}\rho\Sigma_z^{(k)}$ are the same as for ρ. We can therefore conclude that $I_c(\mathcal{E}_N, \Sigma_z^{(k)}\rho\Sigma_z^{(k)}) = I_c(\mathcal{E}_N, \rho)$. This latter relation, together with the concavity of the coherent information for degradable channels (a direct consequence of the concavity of the conditional von Neumann entropy) implies that

$$I_c(\mathcal{E}_N, \rho_N) \geqslant I_c(\mathcal{E}_N, \rho_{N-1}) \geqslant \cdots \geqslant I_c(\mathcal{E}_N, \rho_0). \tag{9}$$

Hence, diagonal input states maximize the coherent information. These states are separable, since they can be written in the form

$$\rho_N = \sum_{j_1,\ldots,j_N} q_{j_1\ldots j_N}\rho_{j_1}^{(1)} \otimes \cdots \otimes \rho_{j_N}^{(N)}, \tag{10}$$

with $\rho_{j_k}^{(k)} \equiv |j_k\rangle\langle j_k|$, $(k = 1, \ldots, N)$, $0 \leqslant q_{j_1\ldots j_N} \leqslant 1$ and $\sum_{j_1,\ldots,j_N} q_{j_1\ldots j_N} = 1$.

3. The memory dephasing channel

3.1. Forgetful channels

Interesting results on the quantum capacity of dephasing channels with memory can be obtained for forgetful channels, for which the memory dies out exponentially with time. Forgetfulness is defined in [14], according to a model in which the environment is structured in two parts: a memoryless one and one responsible for memory effects (see also [12]). A key feature of forgetfulness is that it permits, with a negligible error, the mapping of the memory channel itself into a memoryless one. This may be clarified by referring to the double-blocking strategy [14]: we consider blocks of $N + L$ uses of the channel and do the actual coding and decoding for the first N uses, ignoring the remaining L idle uses. The resulting CPT map $\bar{\mathcal{E}}_{N+L}$ acts on density matrices ρ on $H^{\otimes N}$. If we consider M uses of such blocks, the corresponding CPT map $\bar{\mathcal{E}}_{M(N+L)}$ can be approximated by the memoryless setting $(\bar{\mathcal{E}}_{(N+L)})^{\otimes M}$. This is possible because correlations among different blocks decay during the idle uses. This property can be expressed as follows [14]:

$$\|\bar{\mathcal{E}}_{M(N+L)}(\rho) - (\bar{\mathcal{E}}_{(N+L)})^{\otimes M}(\rho)\|_1 \leqslant h\,(M-1)c^{-L}, \tag{11}$$

for any input state ρ in $H^{\otimes MN}$, where $c > 1$, $\|\cdot\|_1$ is the trace distance [1], and h is some constant depending on the memory model (note that c and h are independent of the input state). This equation states that, even though the error committed by replacing the memory channel itself with the corresponding memoryless channel grows with the number M of blocks, it goes to zero exponentially fast with the number L of idle uses in a single block. Equation (11) permits the proof of coding theorems for forgetful quantum memory channels, by mapping them into the corresponding memoryless channels, for which quantum coding theorems hold [14]. In particular, the quantum capacity Q is $\lim_{N \to \infty} Q_N/N$. Equation (11) by itself is a sufficient condition to prove coding theorems. Therefore, in the following we will use the wording forgetful channel for any system satisfying inequality (11), independently of the model from which memory arises. Now we focus on two specific, physically significant models.

3.2. Markovian model

The first model is a quantum channel that maps an arbitrary N-qubit input state ρ onto

$$\rho' = \mathcal{E}_N(\rho) = \sum_{i_1,\ldots,i_N} A_{i_1\ldots i_N} \rho A_{i_1\ldots i_N}^\dagger, \qquad i_k = 0, z, \tag{12}$$

where the Kraus operators $A_{i_1\ldots i_N}$ are defined in terms of the Pauli operators $\sigma_0 = \mathbb{1}$ and σ_z:

$$A_{i_1\ldots i_N} = \sqrt{p_{i_1\ldots i_N}} B_{i_1\ldots i_N}, \qquad B_{i_1\ldots i_N} \equiv \sigma_{i_1}^{(1)} \otimes \cdots \otimes \sigma_{i_N}^{(N)}, \tag{13}$$

with $\sum_{\{i_k\}} p_{i_1\ldots i_N} = 1$ and $\sigma_{i_k}^{(k)}$ acting on the kth qubit[5].

The quantity $p_{i_1\ldots i_N}$ can be interpreted as the probability that the ordered sequence $\sigma_{i_1}^{(1)}, \ldots, \sigma_{i_N}^{(N)}$ of Pauli operators is applied to the N-qubits crossing the channel. We define the single-qubit marginal probability $p_{i_q} = \sum_{\{i_k, k \neq q\}} p_{i_1\ldots i_N}$ and similarly the two-qubit marginal

---

[5] The Kraus operators (13) define a generalized dephasing channel in the sense of equation (4), with

$$U = \sum_{i_1,\ldots,i_N} \sqrt{p_{i_1\ldots i_N}} \sigma_{i_1}^{(1)} \otimes \cdots \otimes \sigma_{i_N}^{(N)} \otimes |i_1 \ldots i_N\rangle_E \langle 0 \ldots 0|.$$

probability $p_{i_{q'}i_q}$ and assume that $\{p_{i_q}\} = \{1 - p_z, p_z\}$ for all $q = 1, \ldots, N$. Under these conditions the maximum of coherent information in model (12) is obtained for the totally unpolarized input state $\rho_{\text{unp}} \equiv (1/2^N) \mathbb{1}^{\otimes N}$. To prove this statement, we construct the same iterative transformation as in (7) but with $\Sigma_x^{(k)} = \mathbb{1}^{(1)} \otimes \cdots \otimes \mathbb{1}^{(k-1)} \otimes \sigma_x^{(k)} \otimes \mathbb{1}^{(k+1)} \otimes \cdots \otimes \mathbb{1}^{(N)}$ instead of $\Sigma_z^{(k)}$, and notice that $\rho_N = \rho_{\text{unp}}$ is obtained starting from an input state ρ_0 diagonal in the reference basis. Moreover, it can be proven that in this case

$$S[\mathcal{E}_N(\Sigma_x^{(k)} \rho_0 \Sigma_x^{(k)})] = S(\Sigma_x^{(k)} \rho_0 \Sigma_x^{(k)}) = S(\rho_0).$$

Since ρ_0 is diagonal and $\mathcal{E}_N$ only changes off-diagonal matrix elements, then $\mathcal{E}_N(\rho_0) = \rho_0$ and $S[\mathcal{E}_N(\Sigma_x^{(k)} \rho_0 \Sigma_x^{(k)})] = S[\mathcal{E}_N(\rho_0)]$. We can also prove that

$$S[\tilde{\mathcal{E}}_N(\Sigma_x^{(k)} \rho_0 \Sigma_x^{(k)})] = S[\tilde{\Sigma}_z^{(k)} \tilde{\mathcal{E}}_N(\rho_0) \tilde{\Sigma}_z^{(k)}] = S[\tilde{\mathcal{E}}_N(\rho_0)].$$

Here $\tilde{\Sigma}_z$ is defined as Σ_z but acts on the environment. Therefore, $I_c(\mathcal{E}_N, \Sigma_x^{(k)} \rho_0 \Sigma_x^{(k)}) = I_c(\mathcal{E}_N, \rho_0)$. Taking again advantage of the concavity of coherent information for degradable channels, we finally obtain

$$I_c(\mathcal{E}_N, \rho_{\text{unp}}) \geqslant I_c(\mathcal{E}_N, \rho_0). \tag{14}$$

We can explicitly compute the quantum capacity when the joint probabilities in equation (13) are described by a Markov chain [6, 11]:

$$p_{i_1, \ldots, i_N} = p_{i_1} p_{i_2 | i_1} \cdots p_{i_N | i_N - 1}, \tag{15}$$

where

$$p_{i_k | i_{k-1}} = (1 - \mu) \, p_{i_k} + \mu \, \delta_{i_k, i_{k-1}}. \tag{16}$$

Here $\mu \in [0, 1]$ measures the partial memory of the channel: it is the probability that the same operator (either $\mathbb{1}$ or σ_z) is applied for two consecutive uses of the channel, whereas $1 - \mu$ is the probability that the two operators are uncorrelated. The limiting cases $\mu = 0$ and $\mu = 1$ correspond to memoryless channels and channels with perfect memory, respectively. In this noise model μ might depend on the time interval between two consecutive channel uses. If the two qubits are sent at a time interval $\tau \ll \tau_c$, where τ_c denotes the characteristic memory timescale for the environment, then the same operator is applied to both qubits ($\mu = 1$), while the opposite limit corresponds to the memoryless case ($\mu = 0$).

The Markov chain model is forgetful, since condition (11) is fulfilled. We first consider a sequence of two blocks of $N + L$ channel uses, for which

$$\rho' = \bar{\mathcal{E}}_{2(N+L)}(\rho) = \sum_I p_I B_I \rho B_I^\dagger, \tag{17}$$

where the index I stands for $i_1, \ldots, i_N, i_{N+L+1}, \ldots, i_{2N+L}$ and the operators B_I are defined in equation (13). The output state ρ' can be approximated by

$$\tilde{\rho}' = (\bar{\mathcal{E}}_{N+L})^{\otimes 2}(\rho) = \sum_I \tilde{p}_I B_I \rho B_I^\dagger, \tag{18}$$

where the factorized probability distribution $\tilde{p}_I \equiv p_{i_1, \ldots, i_N} p_{i_{N+L+1}, \ldots, i_{2N+L}}$. Taking advantage of the strong convexity of trace distance [1], we obtain

$$\|\rho' - \tilde{\rho}'\|_1 \leqslant D(p_I, \tilde{p}_I), \tag{19}$$

where the Kolmogorov distance between the probability distributions $\{p_I\}$ and $\{\tilde{p}_I\}$ is defined as

$$D(p_I, \tilde{p}_I) = \tfrac{1}{2} \sum_I |p_I - \tilde{p}_I|. \tag{20}$$

Using the properties of stationary Markov chains and equation (16) we obtain

$$D(p_I, \tilde{p}_I) \leqslant 2\mu^{L+1}. \tag{21}$$

This implies

$$\|\bar{\mathcal{E}}_{2(N+L)}(\rho) - (\bar{\mathcal{E}}_{(N+L)})^{\otimes 2}(\rho)\|_1 \leqslant 2\mu^{L+1}, \tag{22}$$

from which equation (11) readily follows[6]. The forgetfulness of the Markov chain model allows us to compute the quantum capacity from the regularized coherent information (2) [14].

In order to compute the quantum capacity, we consider the input state ρ_{unp} and evaluate the coherent information $I_c(\mathcal{E}_N, \rho_{\text{unp}})$. In this case $S[\mathcal{E}_N(\rho_{\text{unp}})] = S(\rho_{\text{unp}}) = N$. We now take advantage of the formula $(S_e)_N = S(w)$, where the density operator w has components $w_{i_1...i_N, i'_1...i'_N} = \text{Tr}(A_{i_1...i_N} \rho A^{\dagger}_{i'_1...i'_N})$ [20]. Here w is diagonal and

$$S(W) = -\sum_{\{i_k\}} p_{\{i_k\}} \log_2 p_{\{i_k\}} \equiv H(X_1, \ldots, X_N), \tag{23}$$

where $H(X_1, \ldots, X_N)$ is by definition the Shannon entropy of the collection of random variables $X_1, \ldots, X_N$ (characterized by the joint probabilities $p_{i_1...i_N}$). For a stationary Markov chain, we have [28]

$$\lim_{N \to \infty} \frac{1}{N} H(X_1, \ldots, X_N) = H(X_2|X_1) = p_0 H(q_0) + p_z H(q_z),$$

where $q_{0,z} \equiv (1-\mu)p_{0,z} + \mu$ are the conditional probabilities that the channel acts on two subsequent qubits via the same Pauli operator, and $H(q_0)$, $H(q_z)$ are binary Shannon entropies, defined by $H(q) = -q \log_2 q - (1-q) \log_2 (1-q)$. Therefore, the quantum capacity is given by

$$Q = 1 - p_0 H(q_0) - p_z H(q_z). \tag{24}$$

It is interesting to point out that Q increases for increasing degree of memory of the channel. In particular, for $\mu = 0$, we recover the capacity $Q = Q_1 = 1 - H(p_0)$ of the memoryless dephasing channel, while for perfect memory $(\mu = 1)$ $Q = 1$, that is, the channel is asymptotically noiseless [12]. We also note that the right-hand side of (24) is known [11] to be a lower bound for the quantum capacity of the Markov chain dephasing channel. Our results prove that this bound is tight.

In order to illustrate the convergence of Q_N/N to its limiting value Q, we first compute the entropy exchange for the N-qubit input state ρ_{unp}. It is easy to check that

$$(S_e)_N = p_0 H(q_0) + p_z H(q_z) + (S_e)_{N-1}. \tag{25}$$

Using this recurrence relation we obtain

$$(S_e)_N = (N-1)[p_0 H(q_0) + p_z H(q_z)] + (S_e)_1, \tag{26}$$

---

[6] It is interesting to remind the reader that the Markov chain model can also be formulated in terms of a structured environment [12, 14].
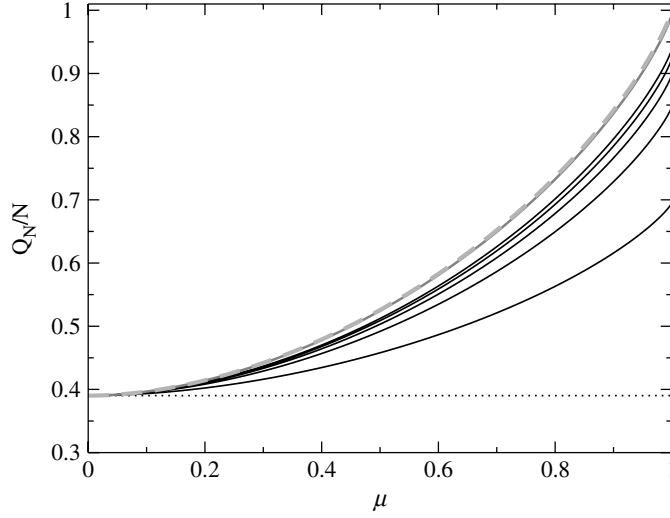
Figure 1. Plot of Q_N/N as a function of μ, for the Markov chain model (15), with $p_0 = 0.85$. From bottom to top: $N = 2, 4, 6, 8, 10$ (black curves), $N = 100, \infty$ (grey curves). The dotted black line gives the memoryless quantum capacity.

where $(S_e)_1 = H(p_0)$. Therefore

$$Q_N = N - (N-1)[p_0 H(q_0) + p_z H(q_z)] - H(p_0). \tag{27}$$

A plot of Q_N/N for various N as a function of the memory factor μ is shown in figure 1.

It is clear that the convergence of Q_N/N is faster when the memory factor is smaller. Indeed, it is easy to prove that

$$\epsilon_N \equiv Q - \frac{Q_N}{N} \tag{28}$$

is a growing function of μ, with $\epsilon_N(\mu = 0) = 0$ and $\epsilon_N(\mu = 1) = H(p_0)/N$. Moreover, for $\mu \ll 1$ we obtain

$$\epsilon_N(\mu) \approx \frac{1}{2\ln 2}\frac{\mu^2}{N}. \tag{29}$$

3.3. Spin-boson model

The second model of dephasing channel is defined by the system (qubits)-environment Hamiltonian

$$H(t) = H_E - \tfrac{1}{2}X_E F(t) + H_C. \tag{30}$$

Here $H_E = \sum_\alpha \omega_\alpha b_\alpha^\dagger b_\alpha$ is a bosonic bath and $X_E = \sum_\alpha (b_\alpha^\dagger + b_\alpha)$ is the environment operator coupled to the qubits. The kth qubit has a switchable coupling to the environment via its Pauli operator $\sigma_z^{(k)}$:

$$F(t) = \lambda \sum_{k=1}^N \sigma_z^{(k)} f_k(t), \tag{31}$$

where $f_k(t) = 1$ when the qubit is inside the channel, and $f_k(t) = 0$ otherwise. Finally,

$$H_C = \sum_\alpha \frac{\lambda^2}{4\omega_\alpha} \sum_{k=1}^{N} \sigma_z^{(k)} \tag{32}$$

is a counterterm [29]. We call τ_p the time each carrier takes to cross the channel and τ the time interval that separates two consecutive qubits entering the channel. The Hamiltonian (30) is expressed in the interaction picture with respect to the qubits. If initially the system and the environment are not entangled, the state of the system at time t is given by the map (1) where

$$U(t) = T e^{-(i/\hbar) \int_0^t ds H(s)}. \tag{33}$$

In particular, we are interested in the final state $\rho' = \rho(t = \tau_N)$, where $\tau_N = \tau_p + (N-1)\tau$ is the transit time across the channel for the N-qubit train. To treat this problem we choose the factorized basis states $\{|j \, \alpha_E\rangle\}$, where—as above—$\{|j\rangle = |j_1, \ldots, j_N\rangle\}$ are the eigenvectors of $\prod_k \sigma_z^{(k)}$. The dynamics preserves the qubit configuration $|j\rangle$ and therefore the evolution operator (33) is diagonal in the system indices:

$$\langle j \, \alpha_E | U(t) | l \alpha_E' \rangle = \delta_{jl} \langle \alpha_E | U(t|j) | \alpha_E' \rangle, \tag{34}$$

where $U(t|j) = \langle j | U(t) | j \rangle$ expresses the conditional evolution operator of the environment alone. Therefore

$$(\rho')_{jl} = (\rho)_{jl} \sum_\alpha \langle \alpha_E | U(t|j) \, w_0 \, U^\dagger(t|l) | \alpha_E \rangle. \tag{35}$$

In this basis representation, the environment only changes the off-diagonal elements of ρ, while populations are preserved. If the environment is initially in the pure state $w_0 \equiv |0\rangle_E \langle 0|$, then the equations (4) and (5) are recovered. At any rate, it is sufficient to consider a purification of w_0 in an enlarged Hilbert space to write our model as a generalized dephasing channel (4).

For a multimode environment of oscillators initially at thermal equilibrium, $w_0 = \exp(-\beta H_E)$, we obtain

$$\sum_\alpha \langle \alpha_E | U(t|j) \, w_0 \, U^\dagger(t|l) | \alpha_E \rangle = \exp \left[-\lambda^2 \int_0^\infty \frac{d\omega}{\pi} S(\omega) \frac{1 - \cos(\omega \tau_p)}{\omega^2} \Big| \sum_{k=1}^{N} (j_k - l_k) e^{i\omega(k-1)\tau} \Big|^2 \right], \tag{36}$$

where $S(\omega)$ is the power spectrum of the coupling operator X_E).

A central question is if and under which conditions a spin-boson environment gives a forgetful channel. Even though we cannot give a rigorous proof, we conjecture on physical grounds that an exponential time decay of the bath symmetrized autocorrelation function $C(t) = 1/2 \langle X_E(t) X_E(0) + X_E(0) X_E(t) \rangle$ is a sufficient condition for forgetfulness. To support this conjecture, we prove inequality (11) in the particular case in which two single channel uses ($N = 1$) are separated by idle times $L\tau$. We consider two qubits ($M = 2$ in equation (11)), prepared in a generic input state ρ. Then we compute the output state ρ' from equation (35), i.e. taking into account memory effects, and the output $\tilde{\rho}'$ in the memoryless limit. We obtain, for a generic monotonic decaying autocorrelation function,

$$\| \rho' - \tilde{\rho}' \|_1 \leqslant 4\lambda^2 g^2 \tau_p^2 C(L\tau), \tag{37}$$
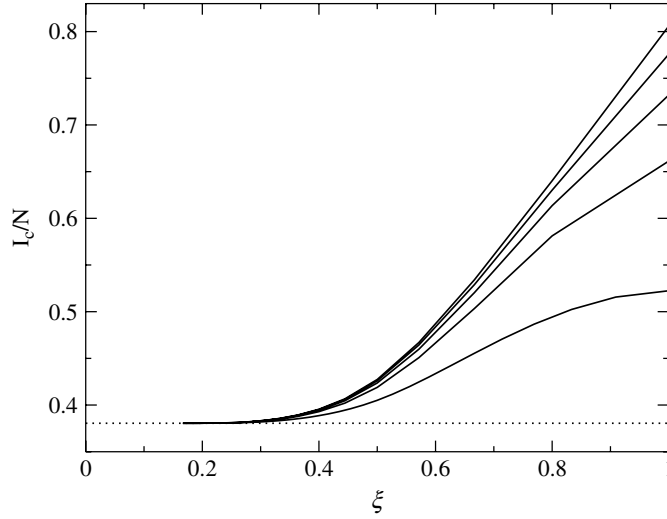
Figure 2. Plot of I_c/N as a function of ξ, for the bosonic bath model (30): Lorentzian power spectrum, $\lambda = 1$, $\tau_c = 1$, $\tau_p = \tau_c$, maximally mixed input state. From bottom to top: $N = 2, 4, 6, 8$ and 10. The dotted line gives the memoryless quantum capacity.

where the dephasing factor g is such that $(\rho')_{01} = g(\rho)_{01}$ and is readily derived from (36) by letting $N = 1$. In particular, we consider a Lorentzian power spectrum $S(\omega) = 2\tau_c/[1 + (\omega\tau_c)^2]$. In this case, the autocorrelation function is $C(\tau) = e^{-\tau/\tau_c}$ and equation (37) is replaced by

$$\|\rho' - \tilde{\rho}'\|_1 \leq 4\lambda^2 g^2 \tau_c^2 (1 - e^{-\tau_p/\tau_c})^2 e^{-L\tau/\tau_c}. \tag{38}$$

Inequality (38) is (11) in the particular case $N = 1$ and $M = 2$ (we can set $h = 4\lambda^2 g^2 \tau_c^2$ by noting that $(1 - e^{-\tau_p/\tau_c})^2 < 1$). We conjecture that (11) also holds for any N and M, since the correlations between blocks of N-qubits decay exponentially with the delay time $L\tau$.

A remarkable feature of model (30) is that in the limit of perfect memory ($\tau_c \to \infty$) there exists for any number N of qubits a decoherence-free subspace $\mathcal{H}_N^{(f)}$, corresponding to a qubit train with an equal number of $|0\rangle$ and $|1\rangle$ states. Since the dimension d of this subspace is such that $\log_2 d \approx N - 1/2 \log_2 N$ at large N, then the channel is asymptotically noiseless, i.e. $Q = 1$. A coding strategy naturally appears when blocks of $\bar{N} \gg 1$ qubits can be sent within the memory timescale τ_c: if the quantum information is encoded in the decoherence-protected subspace $\mathcal{H}_{\bar{N}}^{(f)}$ in such a way that the input state ρ is maximally mixed within this subspace, then a lower bound for the coherent information can be estimated as $I_c(\mathcal{E}_{\bar{N}}, \rho)/\bar{N} \approx \log_2[\dim(\mathcal{H}_{\bar{N}}^{(f)})]/N \approx 1 - \log_2 \bar{N}/(2\bar{N})$. The memoryless dephasing channel instead is recovered in the limit $\tau_c \to 0$ and in this case, the coherent information is maximized by the totally unpolarized input states ρ_{unp} and the channel capacity $Q = Q_1 = 1 - H(p_0)$, where $p_0 = (1 + g)/2$.

Even though we could not compute the channel capacity for generic values of τ, τ_p, and τ_c we show in figure 2 numerical results of the coherent information I_c for a Lorentzian power spectrum $S(\omega)$ and for the input state ρ_{unp} as a function of the degree of memory of the channel, measured by the parameter $\xi \equiv \tau_c/(\tau + \tau_c)$. We fix τ_c, τ_p and vary τ, so that the memoryless and perfect memory limits correspond to $\xi \to 0(\tau \to \infty)$ and $\xi \to 1(\tau \to 0)$. The curves in figure 2 show that memory effects enhance the coherent information I_c/N and that I_c/N grows monotonously with N. Furthermore, these numerical data strongly suggest that I_c/N converges,

for $N \to \infty$, to a limiting value larger than the memoryless capacity Q_1. This value would provide, assuming the above conjectured forgetfulness for the model, a lower bound for the quantum capacity. Therefore, using the previously mentioned double blocking strategy, it is possible to increase the transmission rate if the quantum information is encoded in arbitrarily long blocks, separated by time intervals larger than τ_c.

4. Conclusion

In summary, we have shown that the coherent information in a dephasing channel with memory is maximized by separable input states, computed the quantum capacity Q for a Markov chain noise model and suggested a numerical lower bound for Q in the case of a bosonic bath where memory effects decay exponentially with time. These results also rely on the concept of forgetfulness, which we prove for the first model and strongly support on physical grounds for the second one. It would be relevant to further clarify the connection between the decay of environment autocorrelation functions and forgetfulness. It is important to point out that differently from previous works on quantum memory channels [6], we have carried out the limit in which the number of channel uses $N \to \infty$. It would be interesting to investigate to what extent the results presented in this work could be applied to other physically relevant degradable noise models such as the amplitude damping channel [30]. Another physically relevant question is whether our results could be generalized to environments with algebraically decaying memory effects, which may model typical low-frequency noise in the solid state.

Note added. After completion of our work we became aware of a related paper [31], in which, in particular, the quantum capacity of a Markov chain dephasing channel is provided. Their derivation, not reported in that paper, is based on a method different from ours [32].

Acknowledgments

References

[1] Nielsen M A and Chuang I L 2000 *Quantum Computation and Quantum Information* (Cambridge: Cambridge University Press)
[2] Benenti G, Casati G and Strini G 2007 *Principles of Quantum Computation and Information* (Singapore: World Scientific)
[3] Banaszek K, Dragan A, Wasilewski W and Radzewicz C 2004 *Phys. Rev. Lett.* **92** 257901
[4] Paladino E, Faoro L, Falci G and Fazio R 2002 *Phys. Rev. Lett.* **89** 228304
   Paladino E, Faoro L and Falci G 2003 *Adv. Solid State Phys.* **43** 747
   Paladino E, Mastellone A, D'Arrigo A and Falci G 2005 *Realizing Controllable Quantum States* ed H Takanayanagi and J Nitta (London: World Scientific) p 282

[5] Makhlin Yu and Shnirman A 2004 *Phys. Rev. Lett.* **92** 178301

   Falci G, D'Arrigo A, Mastellone A and Paladino E 2005 *Phys. Rev. Lett.* **94** 167002

   Ithier G *et al* 2005 *Phys. Rev.* B **72** 134519

[6] Macchiavello C and Palma G M 2002 *Phys. Rev.* A **65** 050301(R)

[7] Daems D 2006 *Preprint* quant-ph/0610165

[8] Karimipour V and Memarzadeh L 2006 *Preprint* quant-ph/0611130

[9] Bjelakovic I and Boche H 2006 *Preprint* quant-ph/0609229

[10] Datta N and Dorlas T C 2006 *Preprint* quant-ph/0610049

[11] Hamada H 2002 *J. Math. Phys.* **43** 4382

[12] Bowen G and Mancini S 2004 *Phys. Rev.* A **69** 012306

[13] Giovannetti V 2005 *J. Phys. A: Math. Gen.* **38** 10989

[14] Kretschmann D and Werner R F 2005 *Phys. Rev.* A **72** 062323

[15] Lloyd S 1997 *Phys. Rev.* A **55** 1613

[16] Barnum H, Nielsen M A and Schumacher B 1998 *Phys. Rev.* A **57** 4153

[17] Shor P W 2002 The quantum channel capacity and coherent information *Lecture Notes* MSRI Workshop on
       Quantum Computation

[18] Devetak I 2005 *IEEE Trans. Inform. Theory* **51** 44

[19] Hayden P, Horodecki M, Yard J and Winter A 2007 *Preprint* quant-ph/0702005

[20] Schumacher B 1996 *Phys. Rev.* A **54** 2614

[21] Schumacher B and Nielsen M A 1996 *Phys. Rev.* A **54** 2629

[22] Devetak I and Shor P W 2005 *Commun. Math. Phys.* **256** 287

[23] Caruso F and Giovannetti V 2006 *Phys. Rev.* A **74** 062307

[24] Caruso F, Giovannetti V and Holevo A S 2006 *New J. Phys.* **8** 310

[25] Wolf M M and Pérez-García D 2007 *Phys. Rev.* A **75** 012303

[26] Palma G M, Suominen K-A and Ekert A K 1996 *Proc. R. Soc. Lond.* A **452** 567

   Weiss U 1999 *Quantum Dissipative Systems* 2nd edn (Singapore: World Scientific)

   Falci G, Paladino E and Fazio R 2003 *Decoherence in Josephson Qubits* ed V Tognetti and B Altshuler
       (Amsterdam: IOS Press)

   Paladino E, Sassetti M and Falci G 2004 *Chem. Phys.* **296** 325

[27] Leung D, Vandersypen L, Zhou X, Sherwood M, Yannoni C, Kubinec M and Chuang I 1999 *Phys. Rev.* A
       60 1924

[28] Cover T M and Thomas J A 1991 *Elements of Information Theory* (New York: Wiley)

[29] Weiss U 1999 *Quantum Dissipative Systems* 2nd edn (Singapore: World Scientific)

[30] Giovannetti V and Fazio R 2005 *Phys. Rev.* A **71** 032314

[31] Virmani S and Plenio M 2007 *Preprint* quant-ph/0702059

[32] Virmani S and Plenio M 2007 private communication