DIAS-STP-08-03

CALCULATING A MAXIMIZER FOR QUANTUM MUTUAL INFORMATION

T.C. DORLAS* and C. MORGAN†
Dublin Institute for Advanced Studies,
School of Theoretical Physics,
10 Burlington Road, Dublin 4, Ireland.

January 31, 2008

Abstract

We obtain a maximizer for the quantum mutual information for classical information sent over the quantum amplitude damping channel. This is achieved by limiting the ensemble of input states to antipodal states, in the calculation of the product state capacity for the channel. We also consider the product state capacity of a convex combination of two memoryless channels and demonstrate in particular that it is in general not given by the minimum of the capacities of the respective memoryless channels.

Keywords: product state capacity; maximizing ensemble; memoryless channels.

*dorlas@stp.dias.ie
†cmorgan@stp.dias.ie

1 Introduction

In this paper we obtain the classical capacity of the amplitude damping channel. It is determined by a transcendental equation in a single real variable, which is easily solved numerically. We also consider a convex combination of two memoryless channels and show in particular that the product state capacity of a convex combination of a depolarizing and an amplitude damping channel, which was shown in Ref. [1] to be given by the supremum of the minimum of the corresponding Holevo quantities, is not equal to the minimum of their product state capacities.

1.1 Memoryless channels and the HSW theorem

The transmission of classical information over a quantum channel is achieved by encoding the information as quantum states. A memoryless channel is given by a completely positive trace-preserving map $\mathcal{E} : \mathcal{S}(\mathcal{H}) \to \mathcal{S}(\mathcal{K})$, where $\mathcal{S}(\mathcal{H})$ and $\mathcal{S}(\mathcal{K})$ denote the states on the input and output Hilbert spaces $\mathcal{H}$ and $\mathcal{K}$ respectively. In the case of product-state inputs, the HSW theorem, proved independently by Holevo[2] and by Schumacher and Westmoreland[3], states that the *product state capacity* for classical information sent through a memoryless quantum channel is given by

$$\chi^*(\mathcal{E}) = \max_{\{p_j,\rho_j\}} \chi(\mathcal{E})(\{p_j, \rho_j\}), \tag{1}$$

where the Holevo-χ-quantity is defined by

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

and where S is the von Neumann entropy, $S(\rho) = -\text{trace}\ (\rho\, log\, \rho)$. The maximum is taken over all ensembles of input states ρ_j with probabilities p_j. The capacity for channels with entangled input states has been studied [4], and it has been shown that for certain channels the use of entangled states can enhance the inference of the output state and increase the capacity (e.g. Ref. [5]). We concentrate here on the product state capacity for noisy quantum channels.

Note that, by concavity of the entropy, the maximum in (1) is always attained for an ensemble of *pure states* ρ_j. Moreover, it follows from Carathéodory's theorem (see Ref. [6], Ref. [7], Ref. [8]), that the ensemble can always be assumed to contain no more than d^2 pure states, where $d = \dim(\mathcal{H})$.

In Section 3 we show that, in the case of the amplitude damping channel, the maximum is in fact obtained for an ensemble of two pure states[1]. Moreover, these states are in general not orthogonal as in the example considered by Fuchs[10].

1.2 Convex combination of memoryless channels

In Ref. [1] the product state capacity of a convex combination of memoryless channels was determined. Given a finite collection of memoryless channels $\mathcal{E}_1, \ldots, \mathcal{E}_M$ with common input Hilbert space $\mathcal{H}$ and output Hilbert space $\mathcal{K}$, a convex combination of these channels is defined by the map

$$\mathcal{E}^{(n)}\left(\rho^{(n)}\right) = \sum_{i=1}^{M} \gamma_i \, \mathcal{E}_i^{\otimes n}(\rho^{(n)}), \tag{3}$$

where γ_i, $(i = 1, \ldots, M)$ is a probability distribution over the channels $\mathcal{E}_1, \ldots, \mathcal{E}_M$. Thus, a given input state $\rho^{(n)} \in \mathcal{S}(\mathcal{H}^{\otimes n})$ is sent down one of the memoryless channels with probability γ_i. This introduces long-term memory, and as a result the capacity of the channel $\mathcal{E}^{(n)}$ is no longer given by the maximum of the Holevo quantity. Instead, it was proved in Ref.[1] that it is given by

$$C(\mathcal{E}^{(n)}) = \sup_{\{p_j,\rho_j\}} \left[\bigwedge_{i=1}^{M} \chi_i(\{p_j, \rho_j\}) \right], \tag{4}$$

where $\chi_i = \chi(\mathcal{E}_i)$ is the Holevo quantity for the i-th channel $\mathcal{E}_i$.

2 The amplitude damping channel and the Holevo-χ-quantity.

Acting on the general qubit state $\rho = \begin{pmatrix} a & b \\ \bar{b} & 1-a \end{pmatrix}$, the amplitude damping channel $\mathcal{E}_{amp}$ is given by $\mathcal{E}_{amp}(\rho) = \begin{pmatrix} a + (1-a)\gamma & b\sqrt{1-\gamma} \\ \bar{b}\sqrt{1-\gamma} & (1-a)(1-\gamma) \end{pmatrix}$. The eigenvalues of $\mathcal{E}_{amp}(\rho)$ are easily found to be

$$\lambda_{amp\pm} = \frac{1}{2}\left(1 \pm \sqrt{(1 + 2a(\gamma - 1) - 2\gamma)^2 - 4|b|^2(\gamma - 1)}\right). \tag{5}$$

To maximize the Holevo quantity, given by Eq. (2), for this channel we show that the first term is increased, while keeping the second term fixed, if

[1]The maximizer for this case has also been obtained in [9], but their proof is different.

each pure state ρ_j is replaced by itself and its mirror image in the real b-axis, i.e. if we replace $\rho_j = \begin{pmatrix} a_j & b_j \\ \bar{b}_j & (1-a_j) \end{pmatrix}$ associated with probability p_j, with the states $\rho_j = \begin{pmatrix} a_j & b_j \\ \bar{b}_j & (1-a_j) \end{pmatrix}$ and $\rho_j' = \begin{pmatrix} a_j & -b_j \\ -\bar{b}_j & (1-a_j) \end{pmatrix}$, both with probabilities $p_j/2$.

In general, the states ρ_j must lie inside the Poincaré sphere $\left(a-\frac{1}{2}\right)^2+|b|^2 \leq \frac{1}{4}$ and so the pure states will lie on the boundary $|b|^2 = a(1-a)$.

We first show that the second term in (2) remains unchanged when the states are replaced in the way described above. Indeed, since the eigenvalues (5) depend only on $|b|$, we have $S\left(\mathcal{E}(\rho_j)\right) = S\left(\mathcal{E}(\rho_j')\right)$ and therefore the first term is unchanged. Secondly, by concavity and the fact that $S\left(\sum_j p_j \mathcal{E}(\rho_j')\right) = S\left(\sum_j p_j \mathcal{E}(\rho_j)\right)$,

$$S\left(\sum_j \frac{p_j}{2}\mathcal{E}(\rho_j+\rho_j')\right) \geq S\left(\mathcal{E}\left(\sum_j p_j\rho_j\right)\right).$$

We can conclude that the first term in Eq. (2) is increased with the second term fixed if each state ρ_j is replaced by itself together with its mirror image.

2.1 Convexity of the output entropy

We concentrate here on proving that, in the case of the amplitude damping channel, the second term in the equation for the Holevo-χ-quantity is convex as a function of the parameters a_j when ρ_j is taken to be a pure state, i.e. $b_j = \sqrt{a_j(1-a_j)}$. (Note that $S(a)$ only depends on $|b|$.) Thus $S\left(\mathcal{E}(\rho_j)\right)$ is a function of one variable only, i.e. $S(a_j) = S(\mathcal{E}_{amp}(\rho_{a_j}))$, with $\rho_a = \begin{pmatrix} a & \sqrt{a(1-a)} \\ \sqrt{a(1-a)} & 1-a \end{pmatrix}$ and hence

$$\sigma(a) = \mathcal{E}_{amp}(\rho_a) = \begin{pmatrix} a+(1-a)\gamma & \sqrt{a(1-a)}\sqrt{1-\gamma} \\ \sqrt{a(1-a)}\sqrt{1-\gamma} & (1-a)(1-\gamma) \end{pmatrix}. \quad (6)$$

The eigenvalues of (6) are given by $\lambda_{amp\pm} = \frac{1}{2}(1\pm x)$,
where $x = \sqrt{1-4\gamma(1-\gamma)(1-a)^2}$, and thus $S(a) = H(\frac{1-x}{2})$, where $H(p) = -p\log p - (1-p)\log(1-p)$ is the binary entropy. It is now easy to see that $S''(a) \geq 0$ and hence that $S(a)$ is convex. Writing $\bar{\rho_a} = \sum_j p_j\, \rho_{a_j}$ with $\bar{a} = \sum_j p_j\, a_j$ we have

$$\chi(\{p_j, \rho_j\}) = S(\mathcal{E}_{amp}(\bar{\rho_a})) - \sum_j p_j\, S(a_j) \leq S(\mathcal{E}_{amp}(\bar{\rho_a})) - S(\bar{a}). \quad (7)$$

The capacity is therefore given by

$$\chi(\mathcal{E}_{amp}) = \max_{a\in[0,1]} \left[S\left(\frac{1}{2}(\sigma(a)+\sigma'(a))\right) - S(\sigma(a)) \right]. \tag{8}$$

The maximizing value of a is given by the transcendental equation $\chi'_{AD}(a) = 0$ and can only be computed numerically. It turns out that $a_{\max} \geq \frac{1}{2}$ for all γ. This is in fact easily proved: The determining equation is

$$\chi'_{AD}(a)\ln 2 = -(1-\gamma)\ln\frac{a+\gamma(1-a)}{(1-\gamma)(1-a)} + \frac{4\gamma(1-\gamma)(1-a)}{2x}\ln\frac{1+x}{1-x} = 0. \tag{9}$$

Since $\chi_{AD}(a)$ is concave, the statement follows if we show that $\chi'_{AD}(\frac{1}{2}) > 0$. But, if $a = \frac{1}{2}$, $x = \sqrt{1-\gamma+\gamma^2}$ and $\chi'(\frac{1}{2}) = -(1-\gamma)\ln\frac{1+\gamma}{1-\gamma} + \frac{\gamma(1-\gamma)}{x}\ln\frac{1+x}{1-x} > 0$ because $x > \gamma$ and the function $\frac{1}{2x}\ln\frac{1+x}{1-x} = \frac{\tanh^{-1}(x)}{x}$ is increasing. The resulting capacity is plotted in Figure 1.

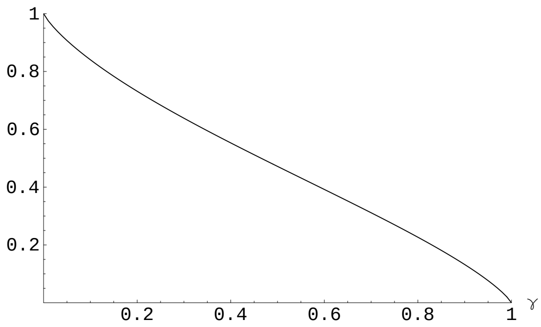

Figure 1: The classical capacity of the amplitude damping channel plotted as a function of γ.

3 Convex combinations of two memoryless channels

Let us now consider a convex combination of two memoryless channels. It was shown in Ref [1] that the product-state capacity is given by (4). Note that we always have: $C(\mathcal{E}^{(n)}) \leq \wedge_{i=1}^{M}\chi_i^*$. We now consider three cases: a convex combination of two depolarizing channels, two amplitude damping channels, and one depolarizing and one amplitude damping channel.

3.1 Two depolarizing channels

In the case of a convex combination of two depolarizing qubit channels $\mathcal{E}_{Dep}(\rho) = (1-\alpha_i)\rho + \alpha_i \left(\frac{I}{2}\right)$ with parameters α_1 and α_2, we have

$$C(\mathcal{E}^{(n)}_{\alpha_1,\alpha_2}) = \chi^*(\alpha_1) \wedge \chi^*(\alpha_2) = \chi^*(\alpha_1 \vee \alpha_2). \tag{10}$$

Indeed, since the maximizing ensemble for both channels is the same, namely two projections onto orthogonal states, this also maximizes the minimum $\chi_1 \wedge \chi_2$. (The product state capacity of a depolarizing qubit channel is well-known of course, and given by $\chi^*_{Dep} = 1 - H\left(\frac{\alpha}{2}\right)$. (In fact, it was proved by King [11], that this is also the capacity of the channel.)

3.2 Two amplitude damping channels

A convex combination of amplitude damping channels is similar. In that case, the maximizing ensemble does depend on the parameter γ, but as can be seen from Figure 2, for any a, $\chi_{AD}(a)$ decreases with γ, so $\chi(\gamma_1) \wedge \chi(\gamma_2) = \chi(\gamma_1 \vee \gamma_2)$ and we have again,

$$C(\mathcal{E}^{(n)}_{\gamma_1,\gamma_2}) = \chi^*(\gamma_1) \wedge \chi^*(\gamma_2) = \chi^*(\gamma_1 \vee \gamma_2). \tag{11}$$

In fact, for $\gamma \le \frac{1}{2}$ this can be seen as follows. The derivative w.r.t. γ is given by:

$$\frac{\partial \chi}{\partial \gamma} = -(1-a) \ln \frac{a + \gamma(1-a)}{(1-\gamma)(1-a)} + \frac{(2\gamma-1)(1-a)^2}{x} \ln \frac{1+x}{1-x}. \tag{12}$$

Clearly, if $\frac{a}{1-a} > 1 - 2\gamma$ both terms are negative. Otherwise, we remark that $x \ge (1-2\gamma)(1-a)$ so that it suffices if $x > y = 1 - 2\gamma - 2a(1-\gamma) > 0$. This is easily checked.

In case $\gamma > \frac{1}{2}$, we need to show that

$$f(a,\gamma) = \ln \frac{a + \gamma(1-a)}{(1-\gamma)(1-a)} - \frac{(2\gamma-1)(1-a)}{x} \ln \frac{1+x}{1-x} \ge 0.$$

Now, if $a = 0$, then $f(0,\gamma) = 0$, and the derivative is given by

$$\frac{\partial f(a,\gamma)}{\partial a} = \frac{1-\gamma}{a+\gamma(1-a)} + \frac{1}{1-a} + \frac{2\gamma-1}{x^3} \ln \frac{1+x}{1-x} - \frac{2(1-a)^2(2\gamma-1)}{x^2}, \tag{13}$$

which can be shown to be positive.

3.3 A depolarizing channel and an amplitude-damping channel

We now investigate the product-state capacity of a convex combination of an amplitude damping and a depolarizing channel. Let χ_1 and χ_2 denote the Holevo quantity of the amplitude damping and depolarizing channels respectively.

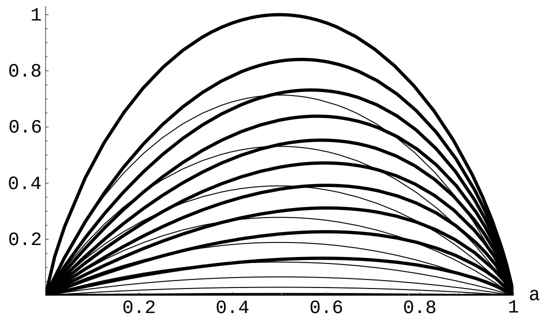

Figure 2: The Holevo χ quantity for the amplitude damping channel and the depolarizing channel plotted as a function of a for different parameter values. The amplitude damping channel is represented in bold.

They are plotted in Figure 2 for $0 \leq \gamma, \alpha \leq 1$. The plot indicates that, for certain values of γ and α the maximizer for the amplitude damping channel lies to the right of the intersection of $\chi_1(a)$ and $\chi_2(a)$ for the depolarizing channel, whereas that for the depolarizing channel lies to the left. Indeed, keeping α fixed, we can increase γ until the maximum of $\chi_{AD}(\gamma)$ lies above the graph of χ_{Dep}. The two graphs then intersect at a value of a intermediate between $\frac{1}{2}$ and the maximizer for χ_{AD}. This proves that the maximum of the minimum of the channels is in general not equal to the minimum of the individual channel capacities.

References

[1] N. Datta and T.C. Dorlas, The coding theorem for a class of quantum channels with long-term memory, *Journal of Physics A, Math. Theor.* **40** (2007) 8147–8164.

[2] A.S. Holevo, The capacity of the quantum channel with general signal states, *IEEE Transactions on Information Theory* **44** (1998) 269–273.

[3] B. Schumacher and M. Westmoreland, Sending classical information via noisy quantum channels, *Phys. Rev. A* **56** (1997) 131–138.

[4] B. Schumacher, Sending entanglement through noisy quantum channel, *Phys. Rev. A* **54** (1996) 2614–2628.

[5] C. Fuchs, C. Bennett and J. Smolin, Entanglement-enhanced classical communication on a noisy quantum channel, arXiv-ph/9611006

[6] E.B. Davies, Information and quantum measurement, *IEEE Transactions on Information Theory* **24** (1978) 596–599.

[7] H.G. Eggleston, *Convexity* (Cambridge University Press, 1958).

[8] B. Grunbaum, *Convex Polytopes* (Interscience Publishers, 1967).

[9] V. Gioannetti and R. Fazio, Information-capacity description of spin-chain correlations *Phys. Rev. A* **71** (2005) 032314.

[10] C. Fuchs, Nonorthogonal quantum states maximize classical information capacity, *Phys. Rev. Lett.* **79** (1997) 1162–1165.

[11] C. King, The capacity of the quantum depolarizing channel, *IEEE Transactions on Information Theory* **49** (2003) 221–229.