Fortschr. Phys. **51**, No. 4–5, 414 – 420 (2003) / **DOI** 10.1002/prop.200310056

Classical information transfer over noisy quantum channels

Chiara Macchiavello[1,*] and **G. Massimo Palma**[2]

[1] Dipartimento di Fisica "A.Volta" and INFM-Unità di Pavia, Via Bassi 6, 27100 Pavia, Italy
[2] Dipartimento di Scienze Fisiche ed Astronomiche and NEST-INFM, Via Archirafi 36, 90123 Palermo, Italy

Received 15 May 2002, accepted 22 May 2002
Published online 30 April 2003

We study the role of entanglement in the transmission of classical information over noisy quantum channels. We show that entanglement leads to an enhancement of the mutual information when the transmission channel is not memoryless.

1 Introduction

Entanglement is now regarded as one of the key ingredients - if not *the* ingredient - of all the known quantum information processing algorithms and protocols. In most cases the information which is processed more efficiently with a supply of entanglement is of quantum nature, like e.g. in the case of teleportation [1]. Here however we will address the question whether entanglement is a useful resource in the transmission of classical information, and in particular we will analyse the role of entanglement in the context of the classical capacity of some quantum channels. The classical capacity of a quantum channel [2] is, roughly speaking, the amount of classical information that can be reliably transmitted by using quantum states in the presence of a noisy environment. Its study has attracted much interest recently. One of the main objects of such an interest is indeed the role played by entanglement: if we encode the information into quantum states there may be the possibility that by entangling multiple uses of the channel a larger amount of classical information per use can be reliably transmitted. This property is known as superadditivity (a more precise definition will be provided later in the text). Attention sofar has been paid to memoryless channels, i.e. to channels in which independent noise acts on each use. The absence of superadditivity was first proved analytically for the case of two entangled uses of the depolarising channel [3] and then extended to a broader class of memoryless channels [4]. It has been later shown [5] that entanglement is a precious resource to increase the information transmission when the noise introduced by the channel exhibits some correlations among subsequent uses.

Following the approach of [5], we consider in this paper the case of the depolarising quantum channel. We will first review the case of the memoryless depolarising channel and then show how entanglement can be exploited when correlations are introduced. Finally, we want to point out that a different related problem, which we will not consider here, is the entanglement-assisted classical capacity. In [6] it was shown that prior entanglement between sender and receiver can increase the classical capacity of some noisy memoryless quantum channels.

* Corresponding author E-mail: chiara@unipv.it

2 The mutual information

In the following each use of the channel will be a qubit, i.e a quantum state belonging to a two-dimensional Hilbert space. In the simplest scenario the transmitter can send one qubit at a time along the channel. In this case the message to be transmitted is encoded in codewords that are tensor products of the states of the individual qubits. Quantum mechanics however allows also the possibility to entangle multiple uses of the channel, and thus to encode the classical information to be transmitted into entangled states. For this more general strategy it has been shown that the amount of reliable information which can be transmitted per use of the channel is given by [2]

$$C_n = \frac{1}{n} \mathrm{sup}_{\mathcal{E}} I_n(\mathcal{E}) \ , \tag{1}$$

where $\mathcal{E} = \{P_i, \pi_i\}$ with $P_i \geq 0, \sum P_i = 1$ is the input ensemble of states π_i, transmitted with *a priori* probabilities P_i, of n – generally entangled – qubits and $I_n(\mathcal{E})$ is the mutual information

$$I_n(\mathcal{E}) = S(\rho) - \sum_i P_i S(\rho_i) \ . \tag{2}$$

In the above equations the index n stands for the number of uses of the channel and

$$S(\chi) = -\mathrm{Tr}(\chi \log \chi) \tag{3}$$

is the von Neumann entropy of the density operator χ, $\rho_i = \Phi(\pi_i)$ are the density operators describing the output states and $\rho = \sum_i P_i \rho_i$. Logarithms are taken to base 2. The advantage of the expression (2) is that it includes an optimization over all possible POVMs at the output, including collective ones. Therefore no explicit maximization procedure for the decoding at the output of the channel is needed.

As mentioned above, the interest for the possibility of using entangled states as channel inputs is motivated by the fact that it cannot generally be excluded that $I_n(\mathcal{E})$ is superadditive in the presence of entanglement, i.e. we might have $I_{n+m} > I_n + I_m$ and therefore $C_n > C_1$. In this scenario the classical capacity C of the channel is defined as

$$C = \lim_{n \to \infty} C_n \ . \tag{4}$$

3 The memoryless depolarising channel

We will briefly review here the case of the memoryless depolarising quantum channel, following the approach of [3], and then analyse in the next section the case where correlations among subsequent uses of the channel are present. The action of transmission channels is generally described in terms of Kraus operators [7] A_i, satisfying $\sum_i A_i^\dagger A_i = \mathbb{1}$, such that if we send through the channel a qubit in a state described by the density operator π the corresponding output state $\Phi(\pi)$ is given by the map

$$\pi \longrightarrow \Phi(\pi) = \sum_i A_i \pi A_i^\dagger \ . \tag{5}$$

In this section we will consider memoryless quantum channels, i.e. channels whose action on arbitrary signals π_s, consisting of n qubits (including entangled ones) can be written as

$$\Phi(\pi_s) = \sum_{i_1 \cdots i_n} (A_{i_n} \otimes \cdots \otimes A_{i_1}) \pi_s \left(A_{i_1}^\dagger \otimes \cdots \otimes A_{i_n}^\dagger \right) \ . \tag{6}$$

An interesting class of Kraus operators acting on individual qubits can be expressed in terms of the Pauli operators $\sigma_{x,y,z}$

$$A_i = \sqrt{p_i}\sigma_i \ , \tag{7}$$

with $\sum_i p_i = 1$, $i = 0, x, y, z$ and $\sigma_0 = \mathbb{1}$. A noise model for these Kraus operators is for instance the application of a random rotation of an angle π around axis $\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}}$ with probability p_x, p_y, p_z on the qubit state or the identity operator with probability p_0. We will consider in this paper the case of two uses of the channel, and therefore the actions will be

$$A^u_{k_1,k_2} = \sqrt{p_{k_1}}\sqrt{p_{k_2}}\sigma_{k_1}\sigma_{k_2} \ . \tag{8}$$

The depolarising channel is characterised by $p_x = p_y = p_z = p/3$, and therefore it can be described as a funtion of one parameter only.

We will now show how to optimise the mutual information I_2 for two uses of the channel over the input ensemble. Notice that the mutual information (2) consists of two terms. We can start to maximise the second one, namely the opposite of the Von Neumann entropy for each output state ρ_i. We can then consider an input state π_1 and parametrise it as follows

$$|\pi_1\rangle = \cos\vartheta|00\rangle + \sin\vartheta|11\rangle \tag{9}$$

where $|0\rangle\,,|1\rangle$ are the eigenstates of the σ_z operators for each qubit. It is convenient to use the Bloch representation [9] to describe the action of the channel on the states

$$\pi = \frac{1}{4}\left\{\mathbb{1}\otimes\mathbb{1} + \mathbb{1}\otimes\sum_k \beta^{(2)}_k\sigma_k + \sum_k \beta^{(1)}_k\sigma_k\otimes\mathbb{1} + \sum_{kl}\chi_{kl}\sigma_k\otimes\sigma_l\right\} \tag{10}$$

where the Bloch vectors and tensor are defined respectively as $\beta_i = \mathrm{Tr}(\pi\sigma_i), \chi_{ij} = \mathrm{Tr}(\pi\sigma_i\sigma_j)$.

For convenience, in the following we will express the action of the depolarising channel in terms of the so called shrinking factor [10] $\eta = 1 - 4p/3$. It is then straightforward to verify that the action of the memoryless depolarising channel corresponds to the following transformations of the Pauli operators

$$\begin{aligned}
&\sum_{k_1,k_2} A_{k_1,k_2}\mathbb{1}\otimes\sigma_j A^\dagger_{k_1,k_2} = \eta\mathbb{1}\otimes\sigma_j \\
&\sum_{k_1,k_2} A_{k_1,k_2}\sigma_j\otimes\mathbb{1} A^\dagger_{k_1,k_2} = \eta\sigma_j\otimes\mathbb{1} \\
&\sum_{k_1,k_2} A_{k_1,k_2}\sigma_k\otimes\sigma_j A^\dagger_{k_1,k_2} = \eta^2\sigma_k\otimes\sigma_j \ .
\end{aligned} \tag{11}$$

As we can see, the consequence of the above transformations is that the components of the Bloch vectors $\beta^{(i)}_k$ are shrunk isotropically by the shrinking factor η. The eigenvalues of the state (9) after the action of the channel are given by

$$\lambda_{1,2} = \frac{1}{4}(1 - \eta^2) \ , \tag{12}$$

$$\lambda_{3,4} = \frac{1}{4}\left[1 + \eta^2 \pm 2\eta\sqrt{\cos^2 2\vartheta + \eta^2\sin^2 2\vartheta}\right] \ . \tag{13}$$

The Von Neumann entropy $S(\rho_1)$ can now be minimized as a function of ϑ by maximising the term under square root in the expression for $\lambda_{3,4}$, namely for $\vartheta = 0, \pi/2$. The first term $S(\rho)$ in the mutual information is maximised for four orhtogonal equiprobable input states. Therefore, the mutual information is maximized for equiprobable orthogonal input states π_i corresponding to the same minimum output Von Neumann entropy. This set of orthogonal input states can be chosen for example as

$$\begin{aligned}|\pi_1\rangle &= \cos\vartheta|00\rangle + \sin\vartheta|11\rangle\\ |\pi_2\rangle &= \sin\vartheta|00\rangle - \cos\vartheta|11\rangle\\ |\pi_3\rangle &= \cos\vartheta|01\rangle + \sin\vartheta|10\rangle\\ |\pi_4\rangle &= \sin\vartheta|01\rangle - \cos\vartheta|10\rangle\ .\end{aligned} \tag{14}$$

Actually, the above states have the same eigenvalues and therefore the same Von Neumann entropy. We can conclude then that the mutual information is maximised with orthogonal product input states ($\vartheta = 0$), for all values of the shrinking factor η. Notice that an input ensemble of maximally entangled states corresponds to the worst performance of the channel. We want to stress that the choice of an input state of the form (14) is not a restriction, due to the isotropy of the memoryless depolarising channel. Actually, it corresponds to the Schmidt decomposition of a state of two qubits: due to isotropy, the basis can be chosen arbitrarily. Therefore, any choice of basis different from the eigenvectors of the σ_z operator would lead to the same result as a function of ϑ.

We can therefore conclude that no superadditivity is present for two uses of the memoryless depolarising channel, i.e. $C_2 = C_1$. It was also recently shown more generally that for this channel $C = C_1$ [4].

4 The depolarising channel with memory

We will now focus our attention to quantum channels where correlated noise acts on consecutive uses, i.e. to channels with memory effects. We will generalise the action of Pauli channels, introduced in the previous section for the memoryless case, by considering Kraus operators of the following form

$$A_{k_1\ldots k_n} = \sqrt{p_{k_1\ldots k_n}}\sigma_{k_1}\ldots\sigma_{k_n}\ , \tag{15}$$

with $\sum_{k_1\ldots k_n} p_{k_1\ldots k_n} = 1$. The quantity $p_{k_1\ldots k_n}$ can be interpreted as the probability that a given random sequence of rotations of an angle π along axis $k_1\ldots k_n$ is applied to the sequence of n qubits sent through the channel. In the particular case of a memoryless channel $p_{k_1\ldots k_n} = p_{k_1}p_{k_2}\ldots p_{k_n}$. An interesting generalization is described by Markov chains defined as

$$p_{k_1\ldots k_n} = p_{k_1}p_{k_2|k_1}\cdots p_{k_n|k_{n-1}} \tag{16}$$

where $p_{k_n|k_{n-1}}$ can be interpreted as the conditional probability that a rotation of the angle π around axis k_n is applied to the n-th qubit given that a π rotation around axis k_{n-1} was applied on the $n-1$-th qubit. Here we will consider the case of two consecutive uses of a channel with partial memory, i.e. we will assume $p_{k_n|k_{n-1}} = (1-\mu)p_{k_n} + \mu\delta_{k_n,k_{n-1}}$. This means that with probability μ the same rotation is applied to both qubits while with probability $1-\mu$ the two rotations are uncorrelated. The parameter μ describes the degree of memory of the channel.

From the physical point of view, this noise model can describe situations where time correlations are present in the system. For instance, μ could depend on the time lapse between the two channel uses. If the two qubits are sent at a very short time interval one after the other, the properties of the channel, which determine the direction of the random rotations, will be unchanged and it is therefore reasonable to assume that the action on both qubits will take the form

$$A_k^c = \sqrt{p_k}\sigma_k\sigma_k\ . \tag{17}$$

If on the other hand, the time interval between the channel uses is such that the channel properties have changed then the channel is memoryless and its actions will be given in Eq. (8). An intermediate case, as mentioned above, is described by actions of the form

$$A^i_{k_1,k_2} = \sqrt{p_{k_1}\left[(1-\mu)p_{k_2} + \mu\delta_{k_2|k_1}\right]}\sigma_{k_2}\sigma_{k_2} \ . \tag{18}$$

We will now show how to optimise the mutual information for this kind of channels. We first want to note that the Bell states

$$\begin{aligned} |\Phi_\pm\rangle &= \frac{1}{\sqrt{2}}\{|00\rangle \pm |11\rangle\} \\ |\Psi_\pm\rangle &= \frac{1}{\sqrt{2}}\{|01\rangle \pm |10\rangle\} \end{aligned} \tag{19}$$

are eigenstates of the operators A^c_k and therefore will pass undisturbed through the channel. If used as equiprobable signal states they maximise I_2, as we will have $I_2 = 2$. Furthermore it is immediate to verify that the value $I_2 = 2$ cannot be achieved by any ensemble of tensor product input states. This situation is reminiscent of the so called noiseless codes, where collective states are used to encode and protect quantum information against collective noise [8].

In the following we will concentrate our attention to the case of the depolarizing channel, i.e. $p_0 = 1-p$ and $p_x = p_y = p_z = p/3$ as in the previous section. We will consider an ensemble of four orthogonal input states parametrised as in Eqs. (14). Although it is not a priori certain that this is the optimal choice for all values of μ we know that it maximizes I_2 with $\vartheta = 0$ for $\mu = 0$ (memoryless case shown in the previous section), and with $\vartheta = \frac{\pi}{4}$ for $\mu = 1$ (fully correlated noise). We will therefore optimize the ansatz (14) by looking for the value $\vartheta(\mu)$ which maximizes I_2 as a function of μ.

We will now show that there is a threshold value μ_t for which

$$I_2\left[\vartheta = \frac{\pi}{4}, \mu_t\right] = I_2\left[\vartheta = 0, \mu_t\right] \ . \tag{20}$$

Below the threshold value

$$I_2\left[\vartheta = 0, \mu < \mu_t\right] > I_2\left[\vartheta = \frac{\pi}{4}, \mu < \mu_t\right] \tag{21}$$

while above

$$I_2\left[\vartheta = 0, \mu > \mu_t\right] < I_2\left[\vartheta = \frac{\pi}{4}, \mu > \mu_t\right] . \tag{22}$$

It is straightforward to verify that the Pauli operators are transformed as follows for $\mu = 1$

$$\begin{aligned} \sum_{k_1,k_2} A_{k_1,k_2}\mathbb{1}\otimes\sigma_j A^\dagger_{k_1,k_2} &= \eta\mathbb{1}\otimes\sigma_j \\ \sum_{k_1,k_2} A_{k_1,k_2}\sigma_j\otimes\mathbb{1}A^\dagger_{k_1,k_2} &= \eta\sigma_j\otimes\mathbb{1} \\ \sum_{k_1,k_2} A_{k_1,k_2}\sigma_k\otimes\sigma_j A^\dagger_{k_1,k_2} &= \delta_{kj}\sigma_k\otimes\sigma_j + (1-\delta_{kj})\eta\sigma_k\otimes\sigma_j \ . \end{aligned} \tag{23}$$

It is interesting to note that also for $\mu = 1$, as in the memoryless case, the components of the Bloch vectors $\beta^{(i)}_k$ of the input states are shrunk isotropically by the shrinking factor η. The difference between the two cases is the action on the Bloch tensor χ. The input state $|\pi_1\rangle$ is transformed by the action of the depolarizing cannel with partial memory defined in equation (18) into the output state

$$\rho_1 = \tfrac{1}{4}\{\mathbb{1}\otimes\mathbb{1} + \eta\cos 2\vartheta(\mathbb{1}\otimes\sigma_z + \sigma_z\otimes\mathbb{1}) + \left[\mu + (1-\mu)\eta^2\right]\left[\sigma_z\otimes\sigma_z + \sin 2\vartheta(\sigma_x\otimes\sigma_x - \sigma_y\otimes\sigma_y)\right]\}. \quad (24)$$

The corresponding eigenvalues are:

$$\lambda_{1,2} = \tfrac{1}{4}(1-\mu)(1-\eta^2) \quad (25)$$

$$\lambda_{3,4} = \tfrac{1}{4}\left\{1 + \mu + \eta^2(1-\mu) \pm 2\sqrt{\eta^2\cos^2 2\vartheta + \left[\eta^2(1-\mu)+\mu\right]^2\sin^2 2\vartheta}\right\}. \quad (26)$$

As in the memoryless case, the first two eigenvalues are degenerate and do not depend on ϑ. Moreover, the same eigenvalues are obtained for the output states ρ_2, ρ_3, ρ_4. As discussed in the previous section, the Von Neumann entropy $S(\rho_i)$ is minimized as a function of ϑ when the term under square root in the expression for $\lambda_{3,4}$ is maximum and the mutual information is then maximized for equiprobable states π_i with minimum Von Neumann entropy. Therefore for $\eta^2 > \left[\eta^2(1-\mu)+\mu\right]^2$ the mutual information is maximal for uncorrelated states ($\vartheta = 0$), while for $\eta^2 < \left[\eta^2(1-\mu)+\mu\right]^2$ it is maximal for the Bell states. The threshold value μ_t is a function of the shrinking factor and for $0 < \eta < 1$ takes the form

$$\mu_t = \frac{\eta}{1+\eta}\,. \quad (27)$$

Therefore, for channels with $\mu < \mu_t$ the most convenient choice within the ansatz (14) corresponds to uncorrelated states, while for $\mu > \mu_t$ to maximally entangled states. At the threshold value any set of states of the form (14) leads to the same value for the mutual information. It is interesting to notice that, within the ansatz (14), for any value of μ the mutual information is optimized by either maximally entangled or completely unentangled states. We have used sofar the z axis as the axis of quantisation for the system; however, due to the symmetry of the channel, we would have obtained the same results using x or y as the axis of quantisation.

Notice that sofar we have restricted our attention to input states of the form (14). We will now show that the product states that are less deteriorated when transmitted through the channel are the eigenstates of $\sigma_{z1}\sigma_{z2}$ or $\sigma_{y1}\sigma_{y2}$ or $\sigma_{x1}\sigma_{x2}$. This suggests that no different choice of product signal states can achieve a higher I_2 than our ansatz (14). From Eqs. (12) and (23) it follows that the output density operator corresponding to an arbitrary input product state takes the form

$$\Phi(\pi) = \tfrac{1}{4}\left[\mathbb{1}\otimes\mathbb{1} + \eta\left(\mathbb{1}\otimes\sum_i \beta_{2i}\sigma_{2i} + \sum_i \beta_{1i}\sigma_{1i}\otimes\mathbb{1}\right) + \left(\mu + (1-\mu)\eta^2\right)\sum_i \beta_{1i}\beta_{2i}\sigma_{1i}\otimes\sigma_{2i} + \left(\mu\eta + (1-\mu)\eta^2\right)\sum_{i\neq j}\beta_{1i}\beta_{2j}\sigma_{1i}\otimes\sigma_{2j}\right], \quad (28)$$

A measure of the degree of purity of the state at the output of the channel is given by Tr$[\rho^2]$. It is straightforward to show that for the above state we have

$$\mathrm{Tr}\left[\Phi(\pi)^2\right] = \tfrac{1}{4}\left[1 + 2\eta^2 + \left(\mu + (1-\mu)\eta^2\right)^2\sum_i \beta_{1i}^2\beta_{2i}^2 + \left(\mu\eta + (1-\mu)\eta^2\right)^2\sum_{i\neq j}\beta_{1i}^2\beta_{2j}^2\right]. \quad (29)$$

The above expression is maximised when both Bloch vectors point in the same x, y or z direction. It is straightforward to verify that these states maximise also the fidelity, defined as $F = \mathrm{Tr}[\pi\Phi(\pi)]$,

$$F = \tfrac{1}{4}\left[1 + 2\eta + \left(\mu + (1-\mu)\eta^2\right)\sum_i \beta_{1i}^2\beta_{2i}^2 + \left(\mu\eta + (1-\mu)\eta^2\right)\sum_{i\neq j}\beta_{1i}^2\beta_{2j}^2\right]. \quad (30)$$

Moreover, we have numerical evidence that for any value of μ and η the input product states that maximise the mutual information are still of this form. Therefore, no better choice of product states leads to a higher mutual information than the one achieved by the ansatz (14).

5 Conclusions

In conclusion, we have investigated the problem of the classical capacity of noisy quantum channels and addressed in particular the case where correlations are present. This problem is of great interest not only from the theoretical viewpoint but also from the experimental one as time correlated noise is not rare in real physical quantum transmission channels. We have shown that, while in the case of the memoryless depolarising channel the best performance in terms of mutual information is achieved by using product input states, in the presence of collective noise the transmission of classical information can be enhanced by employing maximally entangled states as carriers of information instead of product states. In particular, for channels with correlated noise our results show that a higher mutual information can indeed be achieved above a certain memory threshold, which depends on the characteristics of the channel, by entangling two consecutive uses of the channel. This result broadens the class of situations in which the use of entanglement enhances the efficiency in communications and information processing. Work is in progress to understand whether the onset of the threshold value in the degree of memory can be identified for a more general class of quantum transmission channels.

Acknowledgements This work was supported in part by the EC under the contracts IST-1999-11053-EQUIP and IST-2000-29681-ATESIT.

References

[1] C. H. Bennett et al., Phys. Rev. Lett. **70**, 1895 (1993).
[2] B. Schumacher and M. D. Westmoreland Phys. Rev. A **56**, 131 (1997).
A. S. Holevo, IEEE Trans. Inf. Theory **44**, 269 (1998); (also quant-ph/9611023).
C. Macchiavello, G. M. Palma, and A. Zeilinger (eds.), Quantum computation and quantum information theory (World Scientific, Singapore, 2001).
[3] D. Bruß, L. Faoro, C. Macchiavello, and G. M. Palma, J. Mod. Opt. **47**, 325, 2000; (also quant-ph/9903033).
[4] C. King, J. Math. Phys. **43**, 4641 (2002); IEEE Trans. Inf. Th. **49**, 221 (2003).
[5] C. Macchiavello and G. M. Palma Phys. Rev. A **65**, R050301 (2002).
[6] C. H. Bennett, P. W. Shor, J. A. Smolin, and A. V. Thapliyal, Phys. Rev. Lett. **83**, 3081 (1999).
[7] K. Kraus, States, Effects, and Operations: Fundamental Notions of Quantum Theory (Springer, Berlin, 1983).
M. Nielsen and I. Chuang, Quantum computation and quantum information (Cambridge University Press, 2000).
J. Preskill, Lecture notes for Physics 229: Quantum Information and Computation (Caltech, 1998).
A. Peres Quantum Theory: concepts and methods (Kluwer Academic Publishers, 1995).
[8] G. M. Palma, K.-A. Suominen, and A. K. Ekert, Proc. Roy. Soc. London A **452**, 567 (1996); for a recent review see J. Kempe, D. Bacon, D. A. Lidar, and K. B. Whaley, Phys. Rev. A **63**, 2307 (2001).
[9] J. Schlienz and G. Mahler, Phys. Rev. A **52**, 4396 (1995).
[10] D. Bruss, A. Ekert, and C. Macchiavello, Phys. Rev. Lett. **81**, 2598 (1998).