Entanglement-enhanced classical capacity of two-qubit quantum channels with memory: the exact solution

D. Daems

Quantum Information and Communication, Ecole Polytechnique, CP 165, Université Libre de Bruxelles, 1050 Brussels, Belgium

The maximal amount of information which is reliably transmitted over two uses of general Pauli channels with memory is proven to be achieved by maximally entangled states beyond some memory threshold. In particular, this proves a conjecture on the depolarizing channel by Macchiavello and Palma [Phys. Rev. A 65, 050301(R) (2002)]. Below the memory threshold, for arbitrary Pauli channels, the two-use classical capacity is only achieved by a particular type of product states.

The transmission of information over long distances in devices like optical fibers or the storage of information in some type of memory are tasks of quantum information processing that can be described by quantum channels. A major problem in quantum information theory is the evaluation of the classical capacity of quantum communication channels, which represents the amount of classical information which can be reliably transmitted by quantum states in the presence of a noisy environment. Early works in this direction were mainly devoted to memoryless channels for which consecutive signal transmissions through the channel are not correlated

Recently, much attention was given to quantum channels with memory in the hope that by entangling multiple uses of the channel, a larger amount of classical information per use could be reliably transmitted. For bosonic continuous variable memory channels, entangled states are shown to enhance the channel capacity except in the absence of input energy constraints. Moreover, when the memory is modelled as a correlated noise, for each value of the noise correlation parameter, there exists an optimal degree of entanglement that maximizes the channel capacity. For qubit channels with memory it was shown that maximally entangled states enhance the two-use channel capacity with respect to product states if the correlation is stronger than some critical value. This was conjectured for the depolarizing channel with memory in the hope that by entangling multiple uses of the channel, a larger amount of classical information per use could be reliably transmitted. For bosonic continuous variable memory channels, entangled states are shown to enhance the channel capacity except in the absence of input energy constraints. Moreover, when the memory is modelled as a correlated noise, for each value of the noise correlation parameter, there exists an optimal degree of entanglement that maximizes the channel capacity.

For qubit channels with memory it was shown that maximally entangled states enhance the two-use channel capacity with respect to product states if the correlation is stronger than some critical value. This was conjectured for the depolarizing channel with memory in the hope that by entangling multiple uses of the channel, a larger amount of classical information per use could be reliably transmitted.

The action of \( E \) on a state \( \rho \) is described by a completely positive map

\[
\rho \rightarrow E(\rho) = \sum_k A_k \rho A_k^\dagger, \quad \sum_k A_k^\dagger A_k = \text{id.} \tag{1}
\]

The amount of classical information which is reliably transmitted by quantum states through the channel is given by the Holevo-Schumacher-Westmoreland bound

\[
\chi(\mathcal{E}) = \max_{\{p_i, \rho_i\}} \left( S \left( \sum_i p_i E(\rho_i) \right) - \sum_i p_i S(E(\rho_i)) \right), \tag{2}
\]

where \( S(\rho) = -\text{Tr}(\rho \log_2 \rho) \) is the von Neumann entropy and the maximum is taken over all ensembles of input states \( \rho_i \) with \( a \text{ priori} \) probabilities \( p_i \). The \( n \)-use classical capacity of the channel is this amount of reliably transmitted information per use

\[
C_n(\mathcal{E}) = \frac{1}{n} \chi(\mathcal{E}), \tag{3}
\]

whereas the classical capacity is defined as \( C = \sup_n C_n \).

Here we focus on the case of two uses of a single qubit channel with memory considered in Refs. \( \text{[5],[7]} \). An intriguing open question we shall address is whether for some Pauli channels with memory the capacity could be achieved by progressively entangling two uses of the channel, as occurs for some Gaussian channels where no threshold of correlations is present. We prove here that the states which optimize the transmission of classical information over two uses of any Pauli channel with memory modelled as a correlated noise are particular product states below some memory threshold, and maximally entangled states above that threshold.

The action of \( n \) uses of a transmission channel on an initial state \( \rho \) is described by a completely positive map \( \mathcal{E} \) which can be represented as an operator-sum

\[
\rho \rightarrow \mathcal{E}(\rho) = \sum_k A_k \rho A_k^\dagger, \quad \sum_k A_k^\dagger A_k = \text{id.} \tag{4}
\]

The amount of classical information which is reliably transmitted by quantum states through the channel is given by the Holevo-Schumacher-Westmoreland bound

\[
\chi(\mathcal{E}) = \max_{\{p_i, \rho_i\}} \left( S \left( \sum_i p_i \mathcal{E}(\rho_i) \right) - \sum_i p_i S(\mathcal{E}(\rho_i)) \right), \tag{5}
\]

where \( S(\rho) = -\text{Tr}(\rho \log_2 \rho) \) is the von Neumann entropy and the maximum is taken over all ensembles of input states \( \rho_i \) with \( a \text{ priori} \) probabilities \( p_i \). The multiple uses of the channel, a larger amount of classical information per use could be reliably transmitted.

An intriguing open question we shall address is whether for some Pauli channels with memory the capacity could be achieved by progressively entangling two uses of the channel, as occurs for some Gaussian channels where no threshold of correlations is present. We prove here that the states which optimize the transmission of classical information over two uses of any Pauli channel with memory modelled as a correlated noise are particular product states below some memory threshold, and maximally entangled states above that threshold.

The action of \( n \) uses of a transmission channel on an initial state \( \rho \) is described by a completely positive map \( \mathcal{E} \) which can be represented as an operator-sum

\[
\rho \rightarrow \mathcal{E}(\rho) = \sum_k A_k \rho A_k^\dagger, \quad \sum_k A_k^\dagger A_k = \text{id.} \tag{6}
\]

The amount of classical information which is reliably transmitted by quantum states through the channel is given by the Holevo-Schumacher-Westmoreland bound

\[
\chi(\mathcal{E}) = \max_{\{p_i, \rho_i\}} \left( S \left( \sum_i p_i \mathcal{E}(\rho_i) \right) - \sum_i p_i S(\mathcal{E}(\rho_i)) \right), \tag{7}
\]

where \( S(\rho) = -\text{Tr}(\rho \log_2 \rho) \) is the von Neumann entropy and the maximum is taken over all ensembles of input states \( \rho_i \) with \( a \text{ priori} \) probabilities \( p_i \). The multiple uses of the channel, a larger amount of classical information per use could be reliably transmitted.

The amount of classical information which is reliably transmitted by quantum states through the channel is given by the Holevo-Schumacher-Westmoreland bound

\[
\chi(\mathcal{E}) = \max_{\{p_i, \rho_i\}} \left( S \left( \sum_i p_i \mathcal{E}(\rho_i) \right) - \sum_i p_i S(\mathcal{E}(\rho_i)) \right), \tag{8}
\]

where \( S(\rho) = -\text{Tr}(\rho \log_2 \rho) \) is the von Neumann entropy and the maximum is taken over all ensembles of input states \( \rho_i \) with \( a \text{ priori} \) probabilities \( p_i \). The multiple uses of the channel, a larger amount of classical information per use could be reliably transmitted.

The amount of classical information which is reliably transmitted by quantum states through the channel is given by the Holevo-Schumacher-Westmoreland bound

\[
\chi(\mathcal{E}) = \max_{\{p_i, \rho_i\}} \left( S \left( \sum_i p_i \mathcal{E}(\rho_i) \right) - \sum_i p_i S(\mathcal{E}(\rho_i)) \right), \tag{9}
\]

where \( S(\rho) = -\text{Tr}(\rho \log_2 \rho) \) is the von Neumann entropy and the maximum is taken over all ensembles of input states \( \rho_i \) with \( a \text{ priori} \) probabilities \( p_i \). The multiple uses of the channel, a larger amount of classical information per use could be reliably transmitted.

The amount of classical information which is reliably transmitted by quantum states through the channel is given by the Holevo-Schumacher-Westmoreland bound

\[
\chi(\mathcal{E}) = \max_{\{p_i, \rho_i\}} \left( S \left( \sum_i p_i \mathcal{E}(\rho_i) \right) - \sum_i p_i S(\mathcal{E}(\rho_i)) \right), \tag{10}
\]

where \( S(\rho) = -\text{Tr}(\rho \log_2 \rho) \) is the von Neumann entropy and the maximum is taken over all ensembles of input states \( \rho_i \) with \( a \text{ priori} \) probabilities \( p_i \). The multiple uses of the channel, a larger amount of classical information per use could be reliably transmitted.

The amount of classical information which is reliably transmitted by quantum states through the channel is given by the Holevo-Schumacher-Westmoreland bound

\[
\chi(\mathcal{E}) = \max_{\{p_i, \rho_i\}} \left( S \left( \sum_i p_i \mathcal{E}(\rho_i) \right) - \sum_i p_i S(\mathcal{E}(\rho_i)) \right), \tag{11}
\]

where \( S(\rho) = -\text{Tr}(\rho \log_2 \rho) \) is the von Neumann entropy and the maximum is taken over all ensembles of input states \( \rho_i \) with \( a \text{ priori} \) probabilities \( p_i \). The multiple uses of the channel, a larger amount of classical information per use could be reliably transmitted.
$\rho_*$ can be identified \[3]. In a nutshell, the argument amounts to constructing from $\rho_*$ an ensemble of input states $\sigma_i \otimes \sigma_j \rho_* \sigma_i \otimes \sigma_j$ which each have the same output entropy. On the other hand, for such an ensemble taken with equal a priori probabilities, one can show that the output state is maximally mixed. To optimize the transmission of information in Pauli channels with memory, this search can be restricted to pure input states $\rho_* = |\Psi\rangle \langle \Psi| = \rho_{\Psi}$. To date, the optimality of some input states has been conjectured \[5\] for the depolarizing channel ($q_0 = 1 - p, q_1 = q_2 = q_3 = p/3$) and proven \[6\] only in one particular instance of Pauli channel with memory ($q_0 = q_3 = p, q_1 = q_2 = \frac{p}{3} - p$). To study the nature of the optimal states for arbitrary Pauli channels, we consider the two-qubit pure state obtained from the general superposition

$$|\Psi\rangle = c_{00}|00\rangle + c_{11}e^{i\varphi_{11}}|11\rangle + c_{10}e^{i\varphi_{10}}|10\rangle + c_{01}e^{i\varphi_{01}}|01\rangle.$$ (7)

The normalization implies the relation $c_{00}^2 + c_{11}^2 + c_{10}^2 + c_{01}^2 = 1$. This constraint is taken into account here by expressing the pertaining parameters in terms of three angles $\theta, \phi$ and $\psi$ as follows

$$c_{00} = \cos\frac{\phi + \psi}{2} \cos\frac{\theta}{2},$$
$$c_{11} = \sin\frac{\phi - \psi}{2} \sin\frac{\theta}{2},$$
$$c_{10} = \cos\frac{\phi - \psi}{2} \sin\frac{\theta}{2}.$$ (8)

The density matrix $\rho_{\Psi}$ can be expressed in terms of the tensor products of Pauli matrices as

$$\rho_{\Psi} = \frac{1}{4} \sum_{n,k=0}^{3} w_{nk} \sigma_n \otimes \sigma_k,$$ (9)

with the real coefficients $w_{nk}$ given by

$$w_{nk} = Tr(\rho_{\Psi} \sigma_n \otimes \sigma_k).$$ (10)

Note that $w_{00} = Tr(\rho_{\Psi}) = 1$ and, by the Schwartz inequality, $|w_{nk}| \leq 1$. The purity $\Pi(\rho_{\Psi}) = Tr(\rho_{\Psi}^2)$ is unity for a pure state which is accounted for by the relation

$$\sum_{k=1}^{3} w_{kk}^2 + \sum_{n \neq k=0}^{3} w_{nk}^2 = 3.$$ (11)

In addition, the following inequality holds for any permutation of the indexes 1, 2, 3

$$w_{jj}^2 + w_{kk}^2 - w_{nn}^2 \leq 1.$$ (12)

For instance, from the explicit expression of \[10\] for each $w_{kk}$ in terms of the angles and phases entering \[11\]-\[13\], one arrives at

$$\alpha_\pm \equiv w_{11}^2 + w_{22}^2 \pm w_{33}^2,$$
$$= \frac{1}{2} \left\{ \cos^2(\varphi_{10} - \varphi_{01})[\sin \phi + \sin \psi]^2 + \cos^2 \varphi_{11}[\sin \phi - \sin \psi]^2 \right\} \pm \left( \cos \theta \cos \phi \cos \psi - \sin \phi \sin \psi \right)^2.$$ (13)

This yields the tight bounds $\alpha_- \leq 1$, which proves one version of \[12\], and $\alpha_+ \leq 3$. Another quantity that will be relevant below is

$$\beta \equiv \sum_{n=1}^{3} (w_{n0}^2 + w_{0n}^2),$$
$$= 2 \left( 1 - \sin^2 \theta [\sin^2 \psi \cos^2 \varphi + \sin^2 \phi \sin^2 \varphi] \right),$$ (14)

with the notation $\varphi \equiv (\varphi_{10} + \varphi_{01} - \varphi_{11})/2$. Clearly, the bound $\beta \leq 2$ is tight.

The interest of the decomposition \[3\] is that, in addition to the channel-independent relations \[11\]-\[14\], the action of the channel takes on the simple form

$$E(\sigma_n \otimes \sigma_k) = \varepsilon_{nk} \sigma_n \otimes \sigma_k,$$ (15)

with the channel parameter $\varepsilon_{nk} \equiv \sum_{i,j=0}^{3} P_{ij} s_{in}s_{jk} \in [-1,1]$. This stems from the identity $\sigma_j \sigma_k \sigma_j = s_{jk} \sigma_k$ where $s_{nk} = +1$ if either $n = k$ or $n = 0$ or $k = 0$, and $s_{nk} = -1$ otherwise. With the joint probability \[13\], the channel parameters read

$$\varepsilon_{kk'} = (1 - \mu) \varepsilon_{k} \varepsilon_{k'}, \mu \varepsilon_{k''},$$ (16)

where $k''$ is the index of the matrix $\sigma_{k''}$ to which $\sigma_k \sigma_{k'}$ is proportional (i.e., $k'' = 0$ if $k = k'$, $k'' = k'$ if $k = 0$, $k'' = k$ if $k' = 0$ and $k'' = \{1,2,3\}/\{k,k'\}$ otherwise). Notice that $\varepsilon_{nk} = \varepsilon_{kn}$. We define the channel parameter

$$\varepsilon_n = \sum_{k=0}^{3} q_k s_{kn},$$ (17)

which implies that $\varepsilon_0 = \varepsilon_{00} = 1$ and $\varepsilon_{k0} = \varepsilon_k$. The ordering of the channel parameters \[10\]-\[17\] will turn out to be crucial. For that purpose, we introduce the
minimized. It reads

$$R$$

have to be maximized. As whereas in

$$A$$

non-zero indexes \(l\) (large), \(m\) (medium) and \(s\) (small) by

$$|\varepsilon_l| \geq |\varepsilon_m| \geq |\varepsilon_s|. \quad (18)$$

The following properties then hold for any value of \(\mu\)

$$\varepsilon_{il}^2 \geq \varepsilon_{kk'}^2 \quad \forall k, k' \neq 0 \quad (19)$$

$$\varepsilon_{il}^2 \geq \varepsilon_{kk'}^2 \quad \forall k \neq k', \quad (20)$$

The roots of the pertaining characteristic equation \(\lambda^4 - \lambda^3 + a_2 \lambda^2 + a_1 \lambda + a_0 = 0\) are given by

$$\lambda_{\eta, \nu} = \frac{1}{4}(1 + \eta R + \nu Q_\eta) \quad \eta, \nu = \pm 1, \quad (22)$$

where

$$R = \sqrt{1 - 4a_2 + \omega(\alpha_1)} \quad \text{and} \quad Q_\eta = \sqrt{2 - 4a_2 + \eta(4a_2 - 8a_1 - 1)}R - \omega(\alpha_1).$$

The function \(\omega\) need not be specified here. Owing to the symmetric structure of the roots in \(\eta, \nu = \pm 1\), it can be shown that extrema in the eigenvalues can be achieved if and only if the coefficients \(a_1\) are each extremal. To minimize the output entropy, the quantities \(R, \eta\) and \(Q_\eta\) have to be maximized. As \(a_2\) is positive and enters both \(R\) and \(Q_\eta\) with a negative sign, this coefficient has to be minimized. It reads \(a_2 = \frac{1}{8}(3 - A - B - C)\) with

$$A = \varepsilon_{ll}^2 w_{ll}^2 + \varepsilon_{mm}^2 w_{mm}^2 + \varepsilon_{ss}^2 w_{ss}^2 \quad (23)$$

$$B = \varepsilon_{ll}^2 (w_{ll}^2 + w_{ll}^2) + \varepsilon_{mm}^2 (w_{mm}^2 + w_{mm}^2) + \varepsilon_{ss}^2 (w_{ss}^2 + w_{ss}^2)$$

$$C = \varepsilon_{ll}^2 (w_{ll}^2 + w_{ll}^2) + \varepsilon_{mm}^2 (w_{mm}^2 + w_{mm}^2) + \varepsilon_{ss}^2 (w_{ss}^2 + w_{ss}^2).$$

The optimal states are those that maximize \(A + B + C\). Their identification rests on two elements: the constraints imposed on the weights \(w_{nk}^2\) associated with the decomposition \(9\) and, on the other hand, the ordering of the channel parameters \(\varepsilon_{kk}\) and \(|\varepsilon_l|\) which is not covered by \(10 - 20\) and depends on \(\mu\) through \(10\). Firstly, the sum of all the weights \(w_{nk}^2\) involved in \(A + B + C\) is equal to 3 by the pure state identity \(11\). The weights featured in \(A\) sum to \(\alpha_+\) by its definition \(13\), those entering \(B\) sum to \(\beta\) by \(13\) and those of \(C\) sum thus to \(3 - \alpha_+ - \beta\). From the parametrization \(7 - 8\) of the pure state \(\rho_\Phi\), we derived the bounds \(\alpha_+ \leq 3\) and \(\beta \leq 2\). These bounds are tight but cannot be achieved by the same pure state (since \(\alpha_+ + \beta \leq 3\)). They imply that the 3 weights involved in \(A\) can be saturated for some states whereas in \(B\) at most 2 of the 6 weights can be equal to unity. Note that the positivity of \(a_2\) also stems from the purity identity \(11\) together with the property \(\varepsilon_{nk}^2 \leq 1\) since \(\varepsilon_{kk'} \leq \varepsilon_{ll}^2 \leq |\varepsilon_l|\) for all \(k, k' \neq 0\) and \(\varepsilon_{kk'} \leq |\varepsilon_l| \leq 1\) for all \(k \neq k'\) (as \(\varepsilon_{kk'} = 1\) otherwise).

In order to identify the states \(\rho_\Phi\), whose output entropy \(S(\mathcal{E}(\rho_\Phi))\) is minimal, the eigenvalues of \(\mathcal{E}(\rho_\Phi)\) are to be considered. In terms of the decomposition \(9\) and of the mapping \(10\), the channel \(4\) reads explicitly \(\mathcal{E}(\rho_\Phi) = \sum_{f, f'} f \rho f' f'\) with

$$\langle f s | \mathcal{E}(\rho_\Phi) | f s \rangle = \frac{1}{4} - \frac{1}{4} \varepsilon_{ll}^2 w_{ll}^2 + \frac{1}{4} \varepsilon_{mm}^2 w_{mm}^2 + \frac{1}{4} \varepsilon_{ss}^2 w_{ss}^2 \quad (21)$$

Secondly, the degree of correlation modifies the positions of the channel parameters \(\varepsilon_{kk}\) with respect to \(|\varepsilon_l|\). When \(\mu\) goes from 0 to 1, each \(\varepsilon_{kk} = (1 - \mu) \varepsilon_{kk}^2 + \mu\) increases from \(\varepsilon_{kk}^2 = 1\). Since \(\varepsilon_{kk}^2 \leq |\varepsilon_l| \leq 1\) and, by definition, \(|\varepsilon_{kk}| = |\varepsilon_l|\), there are values of \(\mu\) where \(\varepsilon_{kk}\) crosses \(|\varepsilon_l|\). The ordering \(10\) also entails that \(\varepsilon_{ll} \geq \varepsilon_{mm} \geq \varepsilon_{ss}\) for any \(\mu\).

On combining these two aspects we are led to distinguish several intervals of the memory parameter. For \(0 \leq \mu \leq \mu_{ml}\) one has \(\varepsilon_{ll}^2 \geq \varepsilon_{mm}^2\). Applying the inequality \(10\) to the definitions \(24\) yields \(A + C \leq \varepsilon_{ll}^2(3 - \beta).\) Similarly, \(20\) leads to \(B \leq \varepsilon_{ll}^2\beta.\) Hence, we obtain

$$A + B + C \leq 3\varepsilon_{ll}^2 + \beta(\varepsilon_{ll}^2 - \varepsilon_{ll}^2) \leq 3\varepsilon_{ll}^2 + \beta(\varepsilon_{ll}^2 - \varepsilon_{mm}^2) \leq \varepsilon_{ll}^2 + 2\varepsilon_{ll}^2 \quad (24)$$

On the second line use was made of the relation \(\varepsilon_{mm}^2 \leq \varepsilon_{ll}^2\) to introduce the factor \(\varepsilon_{ll}^2 - \varepsilon_{mm}^2\) which is positive in this region of \(\mu\). Hence the corresponding term is majorized by taking the upper bound \(\beta = 2\). The bound \(24\) is tight and achieved if and only if \(w_{ll}^2 = w_{mm}^2 = w_{ss}^2 = 1\) which characterizes the optimal states. The threshold \(\mu_{ml} = (|\varepsilon_l| - \varepsilon_{ll})/(1 - \varepsilon_{ll})\) is the value of \(\mu\) such that \(\varepsilon_{ll}^2 = \varepsilon_{ll}^2\).

This result can be understood as follows. In this interval of the memory parameter, \(\varepsilon_{ll}^2, \varepsilon_{ll}^2\) and \(\varepsilon_{ll}^2\) are larger than any other \(\varepsilon_{kk}^2\). These channel parameters are associated with precisely three weights: \(w_{ll}^2, w_{mm}^2\) which are featured in \(B\) and \(w_{ll}^2\) which is featured in \(A\). The optimum is thus \(A + B + C = 2\varepsilon_{ll}^2 + \varepsilon_{ll}^2\) and it is reached only for \(w_{ll}^2 = w_{mm}^2 = w_{ss}^2 = 1\). Recalling \(14\), the optimal states are product states of the form

$$\rho_{\Phi*} = \frac{1}{4}(\sigma_0 + \zeta \sigma_1 \otimes \sigma_0 + \xi \sigma_1 \otimes \psi_{l, \xi}, \zeta, \xi \geq \pm 1, \quad (25)$$

where \(\psi_{l, \xi}\) is a single qubit eigenstate of \(\sigma_1, i.e., \sigma_1 \otimes \psi_{l, \xi} = \xi \otimes \psi_{l, \xi}\). For low correlations the optimal states
are therefore not any product states but those which correspond to the eigenstates associated with the channel parameter $\varepsilon_l$ of largest absolute value. The eigenvalues of $\mathcal{E}(\rho_\psi)$, required to calculate $C_\varepsilon(\mathcal{E})$ from (5), are

$$\lambda_{\varepsilon_l} = \frac{1}{4}(1 + \eta \varepsilon_l + \nu[1 + \eta] \varepsilon_l) \quad \eta, \nu = \pm 1. \quad (26)$$

The second inequality rests on the facts that $w_{11}^2 \leq 1$ and $e_{mm}^2 \geq e_l^2$. Hence, one might have been tempted to consider a state which saturates $w_{22}^2$ instead of both $w_{00}^2$ and $w_{11}^2$ whose prefactor in $B$ is $e_l^2$. However, a state characterized, for instance, by $w_{00}^2 = w_{11}^2 = w_{22}^2 + w_{00}^2 = 1$ does not exist as it would violate (12). The derivation of (27) also proves that there is no other optimal state than (26). Indeed, if the second and third term on the second line of (27) do not vanish identically, then the optimum is not reached. The optimality requires both $w_{00}^2 = 0$ and $w_{11}^2 = w_{22}^2 = 1$, and therefore $w_{22}^2 = 1$ since $e_{mm}^2 \geq e_l^2$. The threshold $\mu_*$ is the value of $\mu$ for which $e_{ss}^2 + e_{mm}^2 = 2e_l^2$. With the notation $\delta_k = 1 - e_k^2$, it reads

$$\mu_* = \frac{-\delta_m e_m^2 - \delta_s e_s^2 + \sqrt{2\delta_l^2 (e_m^2 + e_s^2) - (\delta_m - \delta_s)^2}}{\delta_m + \delta_s}. \quad (28)$$

For $\mu_* \leq \mu \leq 1$, the ordering of the largest channel parameters is changed to $e_l^2 \leq \frac{1}{2}(e_{mm}^2 + e_{ss}^2)$. This yields

$$A + B + C \leq 2e_l^2 + e_l^2 + w_{22}^2 (e_{mm}^2 + e_{ss}^2 - 2e_l^2)$$

$$\leq e_l^2 + e_{mm}^2 + e_{ss}^2. \quad (29)$$

The first line comes from the second one of (27) where the third term which is still negative or zero in the interval of $\mu$ considered here has been upper bounded by taking $w_{00}^2 + w_{22}^2 - w_{ss}^2 = 1$. On the other hand, the term $w_{ss}^2(e_{mm}^2 + e_{ss}^2 - 2e_l^2)$ is now positive and upper bounded by setting $w_{ss}^2 = 1$ which gives the final result. The bound (29) is thus tight and achieved if and only if $w_{22}^2 = w_{mm}^2 = w_{ll}^2 = 1$. By (9), the optimal input states are the maximally entangled density matrices

$$\rho_{\psi_*} = \frac{1}{4}(\sigma_0 \otimes \sigma_0 + \eta \sigma_1 \otimes \sigma_1 + \nu \sigma_2 \otimes \sigma_2 + \xi \sigma_3 \otimes \sigma_3), \quad (30)$$

which entails that $|\Psi_*\rangle$ correspond to the Bell states $\frac{1}{\sqrt{2}}(|00\rangle \pm |11\rangle)$ for $\pm \eta = \pm \nu = \xi = 1$ and $\frac{1}{\sqrt{2}}(|01\rangle \pm |10\rangle)$ for $\pm \eta = \pm \nu = \xi = -1$. The eigenvalues of $\mathcal{E}(\rho_{\psi_*})$ are

$$\lambda_{\eta, \nu} = \frac{1}{4}(1 + \eta \varepsilon_{33} + \nu \varepsilon_{11} + \eta \varepsilon_{22}) \quad \eta, \nu = \pm 1. \quad (31)$$

Illustration: For $q_0 = 0.2$, $q_1 = 0.1$, $q_2 = 0.3$, $q_3 = 0.4$, the channel parameters (16) are $\varepsilon_0 = -0.4$, $\varepsilon_2 = 0$, $\varepsilon_3 = 0.2$, so that $l = 1$, $m = 3$ and $s = 2$. Hence, up to $\mu_* \approx 0.39$ the optimal states (26) are the product states associated with eigenstates of $\sigma_1$. Notice that $\sigma_1$ is not the most probable transformation. This shows the relevance of the channel parameters: $\varepsilon_l \equiv q_0 + q_1 - q_2 - q_3$ dominates because the rotations $\sigma_2$, $\sigma_3$ add up and are not compensated by $\sigma_0$, $\sigma_1$.

In conclusion, for two uses of arbitrary Pauli channels with memory modelled as a correlated noise, the amount of classical information which can be reliably transmitted per use is proven to be $C_\varepsilon(\mu) = 1 - \frac{1}{2} \sum_{\eta, \nu = \pm 1} \lambda_{\eta, \nu} \log_2 \lambda_{\eta, \nu} + \lambda_{\eta, \nu} \lambda_{\eta, \nu} \mu$ given by (26) for $0 \leq \mu \leq \mu_*$ and by (31) for $\mu_* \leq \mu \leq 1$. Below $\mu_*$, the capacity is achieved by the tensor product of the single qubit density matrices pertaining to the eigenstates of the Pauli matrix $\sigma_1$ whose associated channel parameter $\varepsilon_l$ has the largest absolute value. Above the memory threshold, the two-use classical capacity is reached by maximally entangled states. Entanglement is thus a useful resource to enhance the transmission of classical information for this general class of quantum channels with memory. The author is grateful to N. J. Cerf and E. Karpov for stimulating discussions.

[1] A. S. Holevo, IEEE Trans. Inf. Theory 44, 269 (1998); B. Schumacher and M. D. Westmoreland, Phys. Rev. A 56, 131 (1997).
[2] V. Giovannetti, S. Guha, S. Lloyd, L. Maccone, J. H. Shapiro and H. P. Yuen, Phys. Rev. Lett. 92, 027902 (2004).
[3] A. Sen(De), U. Sen, B. Gromek, D. Bruß and M. Lewenstein, Phys. Rev. Lett. 95, 260503 (2005).
[4] G. G. Amosov, Probl. Inf. Transm. 42, 67 (2006).
[5] Ch. Macchiavello and G. M. Palma, Phys. Rev. A 65, 050301(R) (2002).
[6] Ch. Macchiavello, G. M. Palma and S. Virmani, Phys. Rev. A 69, 010303(R) (2004).
[7] G. Bowen and S. Mancini, Phys. Rev. A 69, 012306 (2004).
[8] G. Bowen, I. Devetak and S. Mancini, Phys. Rev. A 71, 034310 (2005).
[9] D. Kretschmann and R. F. Werner, Phys. Rev. A 72, 062323 (2005)
[10] V. Giovannetti and S. Mancini, Phys. Rev. A 71, 062304 (2005).
[11] N. J. Cerf, J. Clavareau, Ch. Macchiavello and J. Roland, Phys. Rev. A 72, 042330 (2005).
[12] G. Ruggeri, G. Soliani, V. Giovannetti and S. Mancini, Europhys. Lett. 79, 719 (2005).
[13] E. Karpov, D. Daems and N. J. Cerf, Phys. Rev. A 74, 032320 (2006).
[14] M. Abramowitz and A. Stegun, *Handbook of mathematical functions*, (Dover, New York, 1972).