Transition of D-Level Quantum Systems Through Quantum Channels with Correlated Noise

A. Fahmi\textsuperscript{1,}\textsuperscript{*} and M. Golshani\textsuperscript{1,2,}\textsuperscript{†}

\textsuperscript{1} Institute for Studies in Theoretical Physics and Mathematics (IPM) P. O. Box 19395-5531, Tehran, Iran
\textsuperscript{2} Department of Physics, Sharif University of Technology, P. O. Box 11365-9161, Tehran, Iran

Entanglement and entanglement-assisted are useful resources to enhance the mutual information of the Pauli channels, when the noise on consecutive uses of the channel has some partial correlations. In this Paper, we study quantum-communication channels in \(d\)-dimensional systems and derive the mutual information of the quantum channels for maximally entangled states and product states coding with correlated noise. Then, we compare fidelity between these states. Our results show that there exists a certain fidelity memory threshold which depends on the dimension of the Hilbert space \(d\) and the properties of noisy channels. We calculate the classical capacity of a particular correlated noisy channel and show that in order to achieve Holevo limit, we must use \(d\) particles with \(d\) degrees of freedom. Our results show that entanglement is a useful means to enhance the mutual information. We choose a special non-maximally entangled state and show that in the quasi-classical depolarization and quantum depolarization channels, maximum classical capacity in the higher memory channels is given by the maximally entangled state. Hence, our results show that for high error channels in every degree of memory, maximally entangled states have better mutual information.

PACS numbers: 03.67.Hk, 05.40.Ca

I. INTRODUCTION

One of the remarkable by-products of the development of quantum mechanics in recent years is quantum information and quantum computation theories. Classical and quantum information theories have some basic differences. Among these differences are superposition principle, uncertainty principle, and non-local effects. The non-locality associated with the entanglement in quantum mechanics is one of the most subtle and intriguing phenomena in nature \cite{1}. Its potential usefulness has been demonstrated in a variety of applications, such as quantum teleportation, quantum cryptography, and quantum dense coding \cite{1}. On the other hand, quantum entanglement is a fragile feature, which can be destroyed by interaction with the environment. This effect, which is due to decoherence \cite{2}, is the main obstacle to the practical implementations of quantum computing and quantum communication. Several strategies have been devised against decoherence. Quantum error correction codes, fault-tolerant quantum computation \cite{1}, and decoherence free subspaces \cite{3} are among them. One of the main problems in the quantum communication is the decoherence effects in the quantum channels.

Recently, the study of quantum channels has attracted a lot of attention. Early works in this direction were devoted, mainly, to memoryless channels for which consecutive signal transmissions through the channel are not correlated. The capacities of some of these channels were determined \cite{4,5} and it was proved that in most cases their capacities are additive for single uses of the channel. For Gaussian channels under Gaussian inputs, the multiplicativity of the output purities was proved in \cite{6} and the additivity of the energy-constrained capacity, even in the presence of classical noise and thermal noise, was proved in \cite{7}, under the assumption that successive uses of the channel are represented by the tensor product of the operators representing a single use of the channel, i.e., the channel is memoryless. In a recent letter, Bartlett et al. \cite{8} showed that it is possible to communicate with perfect fidelity, and without a shared reference frame, at a rate that asymptotically approaches one encoded qubit per transmitted qubit. They proposed a method to encode a qubit, using photons in a decoherence-free subspace of the collective noise model. Boileau et al. \cite{9} and they gave a realistic robust scheme for quantum communication, with polarized entangled photon pairs \cite{10}. In the last few years much attention has been given to bosonic quantum channels \cite{11}.

Recently, Macchiavello et al. \cite{12,13}, considered a different class of channels, in which correlated noise acts on consecutive uses of channels. They showed that higher mutual information can be achieved above a certain memory threshold, by entangling two consecutive uses of the channel. These types of channels and their extension to the bosonic case, has attracted a lot of attention in recent years \cite{14}. K. Banaszek et al. \cite{15} implemented the suggestion of Macchiavello et al. experimentally. They showed how entanglement could be used to enhance classical communication in a noisy channel. In their setting, the introduction of entanglement between two photons is required in order to maximize the amount of information that can be encoded in their joint polarization degree of freedom, and they obtained experimental classical capacity with entangled states and showed that it is more than 2.5 times the theoretical upper limit, when no quantum correlations are allowed. Hence, recently some people have shown that provided the sender and receiver share prior entanglement, a higher amount of classical information is transmitted over Pauli channels in the presence of memory, as compared to the product and the entangled state coding \cite{16}.

On the other hand, there exists a steadily growing interest in entanglement in higher dimensions, since it allows for the realization of new types of quantum communication protocols...
Those provide more security in quantum communication and quantum key distribution for $d$-level systems, against individual attacks, in the sense that a slightly higher error rate is acceptable. Recently, some setups have been realized by using orbital angular momentum states of photons. Also, transmitting states that belong to finite dimensional Hilbert space through quantum channels associated with a larger Hilbert space and channel capacity in the higher dimensional systems have taken a lot of attention. These considerations encourage us to study various aspects of quantum information theories at higher dimensions, e.g., quantum coding, quantum superdense coding, and quantum key distribution at higher dimensions, through a correlated noisy channel.

In this paper, we consider the effect of Pauli channels with correlated noise on the $d$-dimensional systems. We compare $d$-dimensional maximally entangled and product states with each other and find a certain memory threshold which depends on the dimension of Hilbert space ($d$) and the properties of the noisy channel. For such states, our results show that in order to reach a higher fidelity between input and output states, we must use $d$ particles in the $d$-dimensional systems, and fidelity memory threshold ($\mu^d$) goes to zero for higher dimensional systems. In the following, we calculate the explicit form of mutual information of a particular correlated noisy channel and show that in order to achieve Holevo limit we must make use of $d$ particles with $d$ degree of freedom. Then, entanglement is a useful means to enhance the mutual information. We choose a special non-maximally entangled state and show that in the quantum depolarizing, quasi-classical depolarizing and very high error channels, maximum classical capacity in the higher memory channels is given by a maximally entangled state. Our results show that for high error channels, maximally entangled states have better mutual information in every degree of memory.

This paper is organized as follows: In Sec. II we briefly review some properties of quantum channels with correlated noise in the two dimensional systems. In Sec. III, we extend quantum communication with correlated noise to $d$-dimensional systems and calculate fidelity between the input and output states in the Pauli channels. In Sec. IV, we calculate the mutual information for the maximally entangled and the product states, for the Pauli channel of systems with $d$-degree of freedom. In Sec. V, we discuss optimal properties of quasi-classical depolarizing, quantum depolarizing and very high error channels. In Sec. VI, we discuss some applications of quantum channels with correlated noise in the $d$-dimensional systems. In the remaining part of this manuscript (Appendixes A,B and C) we derive some useful relations.

II. ENTANGLEMENT-ENHANCED INFORMATION TRANSMISSION OVER A QUANTUM CHANNEL WITH CORRELATED NOISE

Encoding classical information into quantum states of physical systems gives a physical implementation of the constructs of information theory. The majority of research into quantum communication channels has focused on the memoryless case, although there have been a number of important results obtained for quantum channels with correlated noise operators or more general quantum channels. The action of transmission channels is described by Kraus operators $A_i$, which satisfy the $\sum_i A_i^\dagger A_i = I$, the equality holds when the map is trace-preserving. An interesting class of Kraus operators acting on the individual qubits can be expressed in terms of the Pauli operators $\sigma_{x,y,z}$

$$A_i = \sqrt{p_0} \sigma_i,$$

with $\sum_i p_i = 1$, $i = 0, x, y, z$, and $\sigma_0 = I$. A noise model for these actions is, for instance, the application of a random rotation by angle $\pi$ around the axis $x, y, z$ with the probabilities $p_x, p_y, p_z$, respectively, and the identity operation with probability $p_0$. If we send an arbitrary signal $\rho$, consisting of $n$ qubits (including the entangled ones) through the channel, the corresponding output state is given by:

$$\rho \rightarrow \mathcal{E} (\rho) = \sum_{i_1 \cdots i_n} (A_{i_1} \otimes \cdots \otimes A_{i_n}) \rho (A_{i_1}^\dagger \otimes \cdots \otimes A_{i_n}^\dagger).$$

In the case of Pauli channels a more general situation is described by action operators of the following form:

$$A_{k_1 \cdots k_n} = \sqrt{p_{k_1 \cdots k_n}} \sigma_{k_1} \cdots \sigma_{k_n},$$

with $\sum_{k_1 \cdots k_n} p_{k_1 \cdots k_n} = 1$. The quantity $p_{k_1 \cdots k_n}$ can be interpreted as the probability that a given random sequence of rotations by angle $\pi$ along the axis $k_1 \cdots k_n$ is applied to the sequence of $n$ qubits, sent through the channel. For a memoryless channel $p_{k_1 \cdots k_n} = p_{k_1} p_{k_2} \cdots p_{k_n}$. An interesting generalization is described by a Markov chain defined as:

$$p_{k_1 \cdots k_{n-1}} = p_{k_1} p_{k_2} p_{k_3} \cdots p_{k_n},$$

where $p_{k_n | k_{n-1}}$ can be interpreted as the conditional probability that a $\pi$ rotation around the axis $k_n$ is applied to the $n$-th qubit, given that a $\pi$ rotation around axis $k_{n-1}$ was applied on the $n-1$-th qubit. Here we consider the case of two consecutive uses of a channel with partial memory, i.e., we shall assume $p_{k_n | k_{n-1}} = (1 - \mu) p_{k_n} + \mu \delta_{k_n,k_{n-1}}$. This means that with probability $\mu$ the same rotation is applied to both qubits, while with probability $1 - \mu$ the two rotations are uncorrelated. In the Macchiavello and co-workers' noise model, the degree of memory $\mu$ could depend on the time laps between the two channel uses. If two qubits are sent within a very short time interval, 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:

$$A_k^\varepsilon = \sqrt{p_k} \sigma_k \sigma_k.$$

If on the other hand, the time interval between the channel uses is such that the channel properties are changed, then the actions will be:

$$A_{k_1,k_2}^\mu = \sqrt{p_{k_1}} \sqrt{p_{k_2}} \sigma_{k_1} \sigma_{k_2}.$$
An intermediate case, as mentioned above, is described by actions of the form:

\[ A_{kn-1,kn}^i = \sqrt{p_{kn-1}(1-\mu)p_{kn} + \mu \delta_{kn,kn-1}} \sigma_{kn-1} \sigma_{kn}. \]  

(6)

They have shown that the transmission of classical information can be enhanced by employing maximally entangled states as carriers of information, rather than the product states. Hence, they obtained a threshold in the degree of memory above which a higher amount of classical information is transmitted with entangled signals [12].

III. FIDELITY BETWEEN D-DIMENSIONAL INPUT AND OUTPUT STATES IN THE PAULI CHANNELS WITH CORRELATED NOISE

In this section, we would like to compute the fidelity between input and output of d-dimensional systems with correlated noise for maximally entangled states and product states. At first, it seems that for generalization of the previous works [12] to d-dimensional systems, we must use two particles with d-degrees of freedom [25], but as we shall see in the next section, it is more appropriate to use d particles each having d degrees of freedom.

Let us consider first the effect of the Pauli channels with correlated noise on d-particles in d-dimensional states. Maximally entangled states \(|\psi_{ls,i,s'}\rangle\) are a set of \((d^2)\) orthonormal Hilbert spaces of d qudits:

\[ |\psi_{ls,i,s'}\rangle = \frac{1}{\sqrt{d}} \sum_{j=0}^{d-1} e^{2\pi i \frac{d}{d} j} |j + l_i \rangle |j + l_{i'}\rangle \]

\[ \langle \psi_{ls,i,s'} | \psi_{ls,i,s'} \rangle = \delta_{s,s'} \prod_{i=1}^{d-1} \delta_{l_i,l_i'} \]  

(7)

where \([l_i,s]\) are used for summarizing the representation of the states with \(l_i,s = 0,...,d-1\), \(i = 1,...,d-1\), and kets must be taken \textit{modulo} d here. Quantum channel effects on the qudits are described by \(U_{m,n}\), generalizing the Pauli matrices for qudits to d-dimensional systems, which are a group of qubit error operators defined as:

\[ U_{m,n} = \sum_{k=0}^{d-1} e^{-\frac{2\pi i k m}{d}} |k+m\rangle \langle k|. \]  

(8)

Here, \(m\) labels the shift errors (extending the bit flip \(\sigma_z\)) and \(n\) labels the phase errors (extending the phase flip \(\sigma_x\)). We could consider a simple set of error operators which are used in multidimensional quantum error-correction codes [26], in this approach a mixture of pure states can give the same results [27]. The effect of the coupling between the qudits and environment can be absorbed in a dielectric coefficient (for example, photons of atmosphere or optical fiber). The dielectric constant has a spatial and temporal dependence, leading to an overall time dependent unitary transformation of the polarization state of a single qudit \(A(l)m,n\), as the net effect of the channel. The effect of this noise on the state of qudits can be considered by:

\[ \rho_{[l,i,s]} \rightarrow \mathcal{E}(\rho_{[l,i,s]}) = \sum_{[m,n]} A_{[m,n]}^\mu \rho_{[m,n]} A_{[m,n]}^\mu \dagger. \]

Here \(\rho_{[l,i,s]} = |\psi_{[l,i,s]}\rangle \langle \psi_{[l,i,s]}|\) and \(A_{[m,n]}^\mu\) are extending Kraus operators, which satisfy the relation \(\sum_{[m,n]} A_{[m,n]}^\mu A_{[m,n]}^\mu \dagger = 1\), where \([m,n]\) with \(i = 0,...,d-1\) and \(m_i,n_i = 0,...,d-1\), summarize the representation of all particles.

We consider the shortest variational time of the channel (fiber) under thermal and mechanical fluctuations as \(\tau_{fluc}\). If the time lapse \((\tau_{lap})\) between the two channels used is small compared to \(\tau_{fluc}\), the effects of the channel on various qudits can be considered with correlated noise. For example, in the experiment of K. Banaszek \textit{et al.} [15] \(\tau_{lap} \approx 6\) ns, which is much smaller than the mechanical fluctuations of the fiber. For simplicity, we consider only two types of sending particles: (I) \(d\) particles in Alice’s hand are sent at the same time \((\tau_{lap} \ll \tau_{fluc}\) for each pair); (II) \(d\) particles are sent with a time delay \((\tau_{fluc} \ll \tau_{lap}\) for each pair). Although we can consider the general case of the transmission of quantum states, its calculation is very complicated and does not clarify any physical properties. Similar to the previous cases, Kraus operators \(A_{[m,n]}^\mu\) in the presence of partial memory are described by:

\[ A_{[m,n]}^\mu = \left(1 - \mu \right) \prod_{i=0}^{d-1} P_{m_i,n_i} + \mu \prod_{i=0}^{d-1} \delta_{m_i,m_{i+1}} \delta_{n_i,n_{i+1}} \right)^\dagger \right) U_{m_0,n_0}...U_{m_{d-1},n_{d-1}} \]

(9)

where \(P_{m_i,n_i}\) can be interpreted as the probability that error \(U_{m,n}\) is applied on the \(i\)-th qudit, and \(\mu\) could be considered as the degree of memory of the the channel, which can depend on the time lapse between the two channels used. For the early realization of channels with correlated noise, we would like to compare the similarity between the input and the output states.
Fidelity can be considered as a measure for comparison. It will be shown that above a certain memory threshold $\mu^f_t$, we have higher fidelity for maximally entangled states.

The fidelity between the input and the output states can be expressed as: $F_{[l,s]} = \langle \psi_{[l,s]} | \mathcal{E}(\rho_{[l,s]}) | \psi_{[l,s]} \rangle$, which for maximally entangled states (7) becomes:

$$F_{max-en}^{[l,s]} = (1 - \mu) \frac{1}{d^2} \sum_{j_1,j_2,m=0}^{d-1} \left[ \sum_{n=0}^{d-1} P_{m,n} e^{\frac{2\pi i}{d} (j_1 - j_2) n} \right]^d + \mu.$$ 

We consider a $(d \times d)$ probability parameter $P_{m,n}$, represented by the matrix $P$, which is a generalization of the form given in [19]:

$$P = \begin{bmatrix} p & q & \cdots & q \\ r & t & \cdots & t \\ \vdots & \vdots & \ddots & \vdots \\ r & t & \cdots & t \end{bmatrix}$$ (10)

where $p$ represents the probability of the transition without any errors, and $r$ and $q$ represent the probabilities of the shift and phase errors, respectively, and $t$ represents the probability of mutual shift and phase errors, where $p, q, r, t$ are real variables satisfying the normalization condition $p + (d - 1)q + (d - 1)r + (d - 1)2t = 1$. We consider $d$-dimensional unitary operators, which were considered by Gottlesman et al. [26, 28]. The explicit form of $F_{max-en}^{[l,s]}$ is:

$$F_{max-en}^{[l,s]} = \frac{(1 - \mu)}{d} (p - q) \times \{(p - q)^{d-1}(d - 1) + (p - q + qd)^{d-1}\}$$

In the above relation, for highly noisy channels $(p = q)$ and for very low noisy channels $(p \gg q)$ with $d \to \infty$, the fidelity memory threshold goes to zero $\mu^f_t \to 0$; in other words, similar to the previous works of [19], we can show that at higher dimensions, the use of maximally entangled states are very suitable for quantum communication.

IV. HOLEVO LIMIT FOR THE $D$-DIMENSIONAL PAULI CHANNELS WITH CORRELATED NOISE

In this section, we would like to compute the amount of information that can be transmitted through a quantum channel

The mutual information $I(\mathcal{E}(\rho_{[l,s]}))$ of a general quantum channel $\mathcal{E}$ is given by:

$$I(\mathcal{E}(\rho_{[l,s]})) = \sum_{n=0}^{d-1} P_{m,n} e^{\frac{2\pi i}{d} (j_1 - j_2) n}$$
Hence, it is the maximum value which is called the Holevo-Schumacher-Westmoreland channel capacity, and is defined as \[\chi(\mathcal{E}, \pi_{[i,s]}) = \max_{\rho_{[i,s]}} I(\mathcal{E}(\rho_{[i,s]}), \pi_{[i,s]}) \] (15)
where \(S(\omega) = -Tr(\omega \log_2 \omega)\) is the von Neumann entropy of the density operator \(\omega\) and the maximization is performed over all input probability distribution \(\pi_{[i,s]}\) and density matrix \(\rho_{[i,s]}\). Note that this bound incorporates maximization over all positive operator value measures (POVM) measurements at the receiver, including the collective ones over multiple uses of the channel. In what follows, we shall derive \(I(\mathcal{E}(\rho_{[i,s]}))\) for maximally entangled and product states. We know that entropy function does not change by any unitary transformations (\[4\]):

\[
\rho_{[i,s]}\rightarrow \Lambda_{[i,s]} \rho_{[i,s]} \Lambda_{[i,s]}^\dagger
\]
On the other hand, with respect to the relation \(Tr_{[i,s,i\neq j]}(\rho_{[i,s]}) = \frac{1}{d} I_j\), we show that the Holevo limit can be attained by setting \(\pi_{[i,s]} = \frac{1}{d}\) (with \(l, s = 0, \ldots, d-1\)). The quantum state \(\mathcal{E}(\rho_{[i,s]})\) in the first term can be written as \[30\]:

\[
\frac{1}{d^2} \sum_{[i,s]} \mathcal{E}(\rho_{[i,s]}) = \mathcal{E}\left(\frac{1}{d}I_0 \otimes \cdots \frac{1}{d}I_{d-1}\right)
\]
To get the final result, we take suitable bases for density matrix representation.

If we consider maximally entangled states as input states Eq. \(7\) for \(s, l_i = 0\) (with \(i = 1, \ldots, d - 1\)), which is represented by \(|\psi_{[0,0]}^{max-en}\rangle\), for channels with correlated noise, the output states are given by:

\[E_{max-en}(\rho_{[0,0]}) = \frac{(1-\mu)}{d} \sum_{j_1, j_2=0}^{d-1} \left[ \sum_{m_0,n_0=0}^{d-1} e^{\frac{2\pi i}{d} (j_1 - j_2)n_0} P_{m_0,n_0} |j_1 + m_0\rangle \langle j_2 + m_0| \right] \otimes^d + \mu \rho_{[0,0]}
\]
\[= \frac{(1-\mu)}{d} \sum_{j_1, j_2=0}^{d-1} \left[ (p - q) |j_1\rangle \langle j_2| + (r - t) \sum_{m=1}^{d-1} |j_1 + m\rangle \langle j_2 + m| + (q - t) \delta_{j_1,j_2} |1\rangle \langle j_2| + t \delta_{j_1,j_2} I \right] \otimes^d
\]
(16)

In the above relation, \([\sigma]^{\otimes k}\) represents the tensor product of \(k, \sigma\) matrices and we have used probability parameters suggested by Eq. \(10\). Furthermore, we would like to calculate the entropy function for the above output density matrix. For simplicity of our calculations, we apply some unitary transformations (C-NOT operator) on the above density matrix (it is known that entropy function does not change by any unitary transformation). As we have calculated in Appendix A, after C-NOT unitary transformations, the mutual information is given by the following equation:

\[I^{max-en}(\mathcal{E}(\rho_{[i,s]})) = d \log_2 d + \sum_{[ki=0]}^{d-1} \Lambda^0_{[ki]} \log_2 \Lambda^0_{[ki]}
\]
\[+ \sum_{[ki=0]}^{d-1} \Lambda^1_{[ki]} \log_2 \Lambda^1_{[ki]}
\]
(17)
In the above relation, \([\kappa_i]\) represents a set of variables with \(i = 1, \ldots, d - 1\), and \(\Lambda_{[ki]}\) given by:

\[\Lambda^0_{[ki]} = dB_{[ki]} + A_{[ki]}, \quad \Lambda^1_{[ki]} = A_{[ki]}, \quad l = 1, \ldots, d - 1
\]
channels without memory), the product states in the computational
iminitly entangled states are the most suitable states for quan-
talgebra, the mutual information of the above density matrix is
with $x = p + (d - 1)q$ and $y = qd$, and after some simple
algebra, the mutual information of the above density matrix is

\begin{equation}
I_{\rho_0}^{\text{pr}}(\mathcal{E}_m) = d \log_2 d + [(1 - \mu)(x + qd)^d + \mu(x + qd)] \log_2 [(1 - \mu)(x + qd)^d + \mu(x + qd)] \\
+ (d - 1) [(1 - \mu)(y + dt)^d + \mu(y + dt)] \log_2 [(1 - \mu)(y + dt)^d + \mu(y + dt)] \\
+ \sum_{k=0}^{d-1} [(d - 1)^{d-k} \binom{d}{k} - (d - 1)\delta_{k,0}] [(1 - \mu)(x + qd)^k(y + dt)^{d-k}] \\
\times \log_2 [(1 - \mu)(x + qd)^k(y + dt)^{d-k}].
\end{equation}

In the next section, we shall restrict ourselves to a quasi-
classical depolarizing channel, a quantum depolarizing chan-
el, and a very high error channel and discuss their properties.

V. DISCUSSION ON OPTIMIZATION

In this section, we consider quasi-classical depolarizing
channels and quantum depolarizing channels and suggest
a non-maximally entangled state that interpolates between the
product state and the maximally entangled state and show that
the mutual information is monotonously modified when this
state goes from a product state to a maximally entangled state.
With an overview on the output density matrix of the channel,
we see that the mutual information in the $\mu = 0$ case (chan-
nels without memory), the product states in the computational
basis [29] are the most suitable states; on the other hand, in the
$\mu = 1$ case (channels with completely correlated noise), max-
imally entangled states are the most suitable states for quan-
tum communications. These two mutual informations cross
over each other at the point $\mu_t$. It must be shown that mu-
tual information for each non-maximally entangled state does
not cross mutual information of the product states for $\mu < \mu_t$
and the maximally entangled state for $\mu > \mu_t$. In the follow-
ing, we discuss various types of non-maximal input states and
show that they have aforementioned properties and in the spe-
cial case of high error channels $p = q = r = t$, we prove the
above properties.

A. Quasi-Classical Depolarizing Channel

A quasi-classical depolarizing channel is given by the fol-
lowing probability parameters, which are same as the prob-
bility parameters in Eq. (10), with $p = q$ and $r = t$:

\begin{equation}
p_{m,n} = p_m = \begin{cases} 
  p = q & m = 0, \\
  r = t = 1 - dp & d(d - 1), \\
  \end{cases}
\end{equation}

Similar to two qubit [12] and two qudit cases [25], we define $p_*$ to denote a chosen input state giving minimal output en-
tropy, when transmitted through the channel $\mathcal{E}$. We consider

show that any state from pure states $H$ state. Therefore, we can restrict our search to the whole space of non-maximally entangled states with different values of the optimization parameter $\theta$, in the $d = 4$ dimensional systems. Solid lines represent the mutual information for maximally entangled and product states and the dashed lines are used for non-maximally entangled states with $\cos^2 \alpha = \frac{1}{4}$, $n = 1.25, 2$.

$U_{m,n_i}$ as an irreducible representation of a compact group with $U_{m,n_i} U_{m,n_j} = e^{2\pi i (m_n_i - m_n_j) / d} U_{m,n_i} U_{m,n_j}$, and quantum channels that are covariant with respect to this compact group. We consider averaging the operator $F$

$$F(\rho) = \frac{1}{d^d} \sum_{[n_i=0]}^{d-1} (U_{0,n_0}...U_{0,n_{d-1}}) \rho(U_{0,n_0}^\dagger...U_{0,n_{d-1}}^\dagger).$$

(22)

It is not complicated to show that $F(\rho)$ does not affect the output of the quantum channel, in the sense that:

$$E_2 \circ F = E_2, \quad S(E(\rho_s)) = S(\rho_s)$$

Thus, if $\rho_s$ is an optimal state, then $F(\rho_s)$ is also an optimal state. Therefore, we can restrict our search to the whole space $H^{\otimes d}$ to $F(H^{\otimes d})$. Finally, using (9), it is straightforward to show that any state from $F(H^{\otimes d})$ is a convex combination of pure states $|\omega_{t,s}\rangle = |\omega_{t,s}\rangle$ where:

$$|\omega_{t,s}\rangle = \sum_{j=0}^{d-1} a_j e^{i\theta_j} |j\rangle |j + t_0\rangle...|j + l_{d-1}\rangle,$$

$$a_j \in \mathbb{R}, \quad \sum_{j=0}^{d-1} a_j^2 = 1.$$  

(23)

Restricting our search to the states of the form (23), we reduce the number of real optimization parameters from $(2d)^d$ to $2d$, which can still be a large number. In order to reduce this number to 1, we consider the following ansatz (25):

$$|\psi(\alpha)\rangle = \cos \alpha |0\rangle^{\otimes d} + \frac{\sin \alpha}{\sqrt{d-1}} \sum_{j=1}^{d-1} |j\rangle^{\otimes d}.$$  

(24)

interpolating between the product state ($\cos \alpha = 0$) and the maximally entangled state ($\cos^2 \alpha = 1/d$). Using the one-parameter family of input states $\rho_\alpha = |\psi(\alpha)\rangle \langle \psi(\alpha)|$, in Fig. (1), we show the mutual information $I(E_2(\rho_\alpha))$ for different values of $\alpha$ (Appendix B). The mutual information is monotonously modified when $\alpha$ goes from a product state to a maximally entangled state, whereas the crossover point $\mu_t$ stays intact. However, we cannot guarantee that no other configuration of the parameters $A_\alpha$ and $\phi_\alpha$ minimizes the entropy $S(E_2(\rho_\alpha))$, and this provides the maximum of the mutual information.

As another example, we consider other states that continuously interpolate between the product basis and the maximally entangled basis (at least in the special case [31]). We consider, e.g., the single state $|\phi_{[0,\theta]}\rangle = \sum_{j=0}^{d-1} A_j |j\rangle^{\otimes d}$, as an input state. With some revision in the coefficients in the [31], for the $d = 4$ case (complex coefficients), these coefficients would be [32]:

$$A_0 = \frac{1}{2} (1 + e^{i\theta} \cos \theta), \quad A_1 = -i A_2 = -A_3 = \frac{1}{2} e^{i\theta} \sin \theta.$$  

(25)

We apply the above coefficients in the output density matrix (30) and derive the mutual information (its explicit form is derived in Appendix B). In Fig. (2), we plot the mutual information for various amounts of $\theta_t$ versus its memory coefficient $\mu_t$. Fig. (2) shows that the use of product states for $\mu < \mu_t$ and maximally entangled states for $\mu > \mu_t$ are more appropriate for communication, although, the crossover point $\mu_t$ doesn't stay fixed for various $\theta$. 

FIG. 1: Mutual information $I(E_2(\rho_\alpha))$ as a function of the memory parameter $\mu$ for the quasi-depolarizing channel with $p = 0.15$, for different values of the optimization parameter $\alpha$, in the $d = 4$ dimensional systems. Solid lines represent the mutual information for maximally entangled and product states and the dashed lines are used for non-maximally entangled states with $\cos^2 \alpha = \frac{1}{4}$, $n = 1.25, 2$.

FIG. 2: Mutual information $I(E_2(\rho_\alpha))$ as a function of the memory parameter $\mu$ for the quasi-depolarizing channel with $p = 0.2$ for the state of Eq. (25) and different values of the optimization parameter $\theta$, in the $d = 4$ dimensional systems. Solid lines represent mutual information for maximally entangled and product states and the dashed lines are used for non-maximally entangled states with $\cos^2 \theta = \frac{1}{4}$, $n = 2, 4$. 
parameter are used for non-maximally entangled states with \( \cos^2 \alpha = \frac{7}{9} \), \( n = 1.25, 2 \).

**B. Quantum Depolarizing Channel**

In this subsection, we would like to discuss the depolarizing channels and to show how to derive an explicit expression for the depolarizing channel with correlated noise. This channel is given by the following probability parameters which are the same as in Eq. (10), with \( q = r = t \):

\[
P_{m,n} = \begin{cases} 
  p, & m = n = 0, \\
  q = r = t = \frac{1 - p}{d^2 - 1}, & \text{otherwise}.
\end{cases}
\]  

Although a discussion for quantum depolarizing channels similar to the case of quasi-classical depolarizing channels would be complicated, but, we consider (24) as an input state for quantum depolarizing channels and discuss the output density matrix and mutual information for various amounts of \( \alpha \). The output density matrix and the explicit form of the mutual information are derived in the Appendix C. In Fig. (3), we present the mutual information \( I(\mathcal{E}_2(p_\alpha)) \) for different values of \( \alpha \). Similar to quasi-classical depolarizing channels, the mutual information is monotonously modified when \( \alpha \) goes from a product state to a maximally entangled state. In Fig. (4), we plot the mutual information of the quantum depolarizing channel for the state coefficients that were suggested in the Eq. (25) and an explicit form of them is derived in Appendix C.

**C. Optimal Quantum Communications in the Very High Error Channels**

Here, we would like to compute the maximum amount of information that can be transmitted through a noisy channel, defined by Eq. (29). We shall compare the capacity of the quantum channel for maximally entangled states (7) and the product states (12). Mutual information \( I(\mathcal{E}(\rho_{\alpha,1})) \) of the quantum channel for the special cases \( q = r = t \) and \( p = q, r = t \) are calculated. Here, we consider very high error channels for which \( p = q = r = t \). Concerning Pauli channel effects (9), to optimize the information transmission of the channel, we must have input states that minimize the output entropy (30, 33). For maximally entangled states we have:

\[
I_{\text{max-en}}(\mathcal{E}(\rho_{\alpha,1})) = d \log_2 d
+ (1 - d^{-d})(1 - \mu) \times \log_2 \{(1 - \mu)d^{-d}\}
+ \{(1 - \mu)d^{-d} + \mu\} \log_2 \{(1 - \mu)d^{-d} + \mu\}
\]

This shows that if noises are completely correlated (\( \mu = 1 \)), then the Holevo limit \( \chi_{\text{max-en}}(\mathcal{E}) = \max I_{\text{max-en}}(\mathcal{E}) = d \log_2 d \) can be achieved. For the product states \( I^{\text{pro}}(\mathcal{E}) \), we have, in a similar manner:

\[
I^{\text{pro}}(\mathcal{E}) = d \log_2 d
+ \{(1 - \mu)d^{-d} + \mu\} \log_2 \{(1 - \mu)d^{-d} + \mu d^{-1}\}
+ (1 - d^{1-d})(1 - \mu) \log_2 \{(1 - \mu)d^{-d}\}
\]

In Figures (5) we compare \( I_{\text{max-en}}(\mathcal{E}) \) and \( I^{\text{pro}}(\mathcal{E}) \) schematically, as a function of \( \mu \). These figures show that at high error rates and for channels without memory (\( \mu = 0 \)), the mutual
information is equal to its minimum values \( I^{\max-en}(\mathcal{E}) = I^{\text{purity}}(\mathcal{E}) = 0 \); on the other hand, in the qudit-environment interaction (no matter how strong), there exist quantum maximally entangled states, which are invariant under this interaction, and the mutual information can attain its maximum value. Hence, for every degree of memory, these states have better classical information capacity than the product states. Here, we show that maximally entangled states \((\mathcal{E})\) optimize mutual information transition. In the channels with high errors, any output density matrix can be transformed to the following form:

\[
\mathcal{E}(\rho) = (1 - \mu) \frac{1}{d^d} \mathbf{1}^\otimes d + \mu \mathbf{1}
\]

(with \( Tr(\mathbf{1}) = 1, Tr(\mathcal{E}(\rho) = 1) \). Optimal mutual information is obtained by minimizing the output entropy, and for this, we must have a pure state at the output channel.

An indication of the optimality of mutual information and minimality of entropy is give by \( Tr(\rho_{\text{pur}})^2 = 1 \) in the output states. Thus, for \( \mathcal{E}(\rho) \) we have:

\[
Tr(\mathcal{E}(\rho))^2 = (1 - \mu)^2 \frac{1}{d^d} + \mu^2 Tr(\mathbf{1})^2 + 2 \mu(1 - \mu) \frac{1}{d^d}.
\]

The left hand side of the above relation is going to be maximum for any amount of \( \mu \) if \( \sigma \) is pure and this happens if the input state is a maximally entangled state. The optimization of the Holevo quantity can be achieved by going to an appropriate bases that diagonalizes \( \sigma \). If we assume that \( \sigma \) has \( k \) nonzero diagonal elements, then, the entropy is given by:

\[
S(\mathcal{E}(\rho)) = k(1 - \mu)d^{-d} + \mu \frac{1}{k} \log_2 \{(1 - \mu)d^{-d} + \mu \frac{1}{k}\} + (d^d - k)(1 - \mu)d^{-d} \log_2 \{(1 - \mu)d^{-d}\}.
\]

The minimum value of the above relation can be obtained for \( k = 1 \). In other words, \( \sigma \) must be a pure state, and this happens if the input state is a maximally entangled state.

VI. CONCLUSION

Quantum channels with correlated noise open a new landscape to quantum communication processes. One of the main applications is in the standard quantum cryptography BB84 [34], if Alice and Bob use appropriate states that are suggested in [35], then parties can distillate secure keys (by error correction and privacy amplification) in the higher amount of quantum bit error rate (QBER). In other words, \( I_{AE}(QBER) > I_{BE}(QBER) \). At this stage QBER is a function of the amount of memory \( QBER = QBER(\mu) \). Hence, this approach can be extended to quantum key distribution protocols where the key is carried by quantum states in a space of arbitrary dimension \( d \), using two (or \( d + 1 \)) mutually unbiased bases, where for the high memory channels we have very low error rates. This procedure ensures that any attempt by any eavesdropper Eve to gain information about the sender’s state induces errors in the transmission, which can be detected by legitimate parties [19]. In other words, for arbitrary dimension we can derive the mutual information of \( I_{AB}(\mu), I_{AE}(\mu), \) and \( I_{BE}(\mu) \) and show that in the higher error channels (no matter how strong), there exist quantum maximally entangled states, which are invariant under this interaction. On the other hand, if we are interested in other quantum key distribution protocols (such as the EPR protocol [36]), we must see a qubit in a decoherence-free (DF) subspace of the collective noise for key distribution [9]. Hence for use of the total dimension of Hilbert space, we must revise the EPR protocol for this new approach [27]. Another application of the above extension can be quantum coding, quantum superdense coding at the higher dimensions, and quantum teleportation in the Pauli channels with correlated noise. Although errors in the memory channels can be considered as a subset of collective noise, which are considered in the DF approach, some experimental results [15] show that in some special cases the use of these states is appropriate, because in the memorial channels we make use of all of the maximally entangled states.

To summarize, we have studied quantum communication channels with correlated noise in \( d \)-dimensional systems and have generalized memory channels for \( d \)-level systems and have shown that there exists a memory threshold \( \mu^f \), which goes to zero for very high noisy channels. We derived the mutual information of the quantum channels for maximally entangled states and the product states for channels with correlated noise. Then, we calculated the classical capacity of a particular correlated noisy channel and show that for attaining the Holevo limit we must use \( d \) particles with \( d \) degrees of freedom. Furthermore, we chose a special non-maximally entangled state and showed that in the quantum depolarizing and quasi-classical depolarizing channels, maximum classical capacity in the higher memory channels is given by a maximally entangled state.
ACKNOWLEDGMENTS We would like to thank S. Falahli for his help. This work was supported under project: ARAM.

VII. APPENDIX A: MUTUAL INFORMATION OF MAXIMALLY ENTANGLED STATES IN THE CHANNEL’S WITH CORRELATED NOISE

We evaluate the action of the channel given by \( \mathcal{E} \) on the maximally entangled state \( \left| \psi_{0,0} \right> \). Before going further, we would like to define the generalized controlled-NOT gates in \( d \)-dimensional systems. We define the \( d \)-dimensional unitary operator \( U_{1,0} = \sigma_d = \sum_{k=0}^{d-1} (k + 1) \mod d) \left| k \right> \) with \( \sigma_d^2 = I \) \[8\]. This operator should be calculated to a mod-\( d \) adder \[9\]. We will have:

\[
\sigma_d |j> = |j + 1>, \quad \mod d \\
\sigma_d |j> = |j - 1>, \quad \mod d. \tag{28}
\]

For the case of qubits \( (d = 2) \), it corresponds to Pauli matrix \( \sigma_x = \sigma_2 \). For every unitary operator \( \sigma_d (\sigma_d^\dagger) \), the controlled gate \( C_{t,l} \) \( (C_{t,l}^\dagger) \), which acts on the site \( l \) conditioned on the site \( t \) is naturally defined as:

\[
C_{t,l} |i> \otimes |j> = |i> \otimes (\sigma_d^\dagger)^l |j> \\
C_{t,l}^\dagger |i> \otimes |j> = |i> \otimes (\sigma_d)^l |j>. \tag{29}
\]

The difference with the qubit case should be taken into account. In the qubit case the controller operator just acts when the value of the site \( t \) bit is equal to 1, whereas in the \( d \)-dimensional case, this operator acts \( i \) times whenever the value of the \( t \)-th qubit is equal to \( i \).

For simplicity, we consider \( |\psi_{0,0}> \) as an input state and apply some unitary transformation \( \text{(C-NOT operators)} \) on the output density matrix \( (16) \). C-NOT operates between first qubit (as the controller) and the remaining qubits (as the target).

The density matrix has two types of sentences, I) the state of first qubit is \( |j_1> |j_2>, \) (include special case of \( j_1 = j_2 \)) II) the state of first qubit is identity matrix \( I_0 \).

In the first case, if the density matrix is represented by \( |j_1> |j_2> \otimes \sigma_1 \otimes \sigma_i \otimes ... \sigma_d-1 \), after C-NOT operations, the density matrix changes to:

\[
\bigotimes_{i=1}^{d-1} C_{0,i}^\dagger \left( |j_1> |j_2> \otimes \sigma_1 \otimes \sigma_i \otimes ... \sigma_d-1 \right) \bigotimes_{i=1}^{d-1} C_{0,i} = |j_1> |j_2> \otimes \theta_1 \otimes ... \otimes \theta_d-1. \tag{30}
\]

In the above relation, \( \bigotimes_{i=1}^{d-1} C_{0,i} \) is the summary form of \( C_{0,1} \otimes C_{0,2} \otimes ... \otimes C_{0,d-1} \) and \( \sigma_i \) are \( |j_1> |j_2> \) or the identity matrix \( I_1 \) and \( \theta_i \) are \( |0> \otimes |0> \) or identity matrix \( I_1 \). Under C-NOT operations, we have:

\[
C_{0,i}^\dagger |j_1> |j_2> \otimes |j_1|j_2| > C_{0,i} \rightarrow |j_1> |j_2> \otimes |0> |0>, \tag{31}
\]

\[
C_{0,i}^\dagger |j_1> |j_2> \otimes (\delta_{j_1,j_2} I_1) C_{0,i} \rightarrow |j_1> |j_2> \otimes (\delta_{j_1,j_2} I_1) \text{ with } i = 1, ..., d - 1. \tag{32}
\]

In a similar manner, for the second case of first qudit, which is in the form \( \sum_{j=0}^{d-1} I_0 \otimes \sigma_1 \otimes ... \otimes \sigma_i \otimes ... \sigma_d-1 \) (where \( \sigma_i \) were previously defined), after C-NOT operations, we have:

\[
\bigotimes_{i=1}^{d-1} C_{0,i}^\dagger (I_0 \otimes \sigma_1 \otimes ... \otimes \sigma_i \otimes ... \sigma_d-1) \bigotimes_{i=1}^{d-1} C_{0,i} = \bigotimes_{i=1}^{d-1} I_0 \otimes \eta_i \otimes ... \otimes \eta_{d-1} \tag{33}
\]

The density matrix after C-NOT operations is given by:

\[
\mathcal{E}_{\text{max-en}}(\rho_{0,0}) = \bigotimes_{i=1}^{d-1} C_{0,i}^\dagger \mathcal{E}_{\text{max-en}}(\rho_{0,0}) C_{0,i} = \frac{(1 - \mu)}{d} \bigotimes_{j_1,j_2=0}^{d-1} |j_1> |j_2> [(x - y)|0> |0> + y I] \otimes (d-1)
\]

\[
+ \frac{(1 - \mu)}{d} \sum_{j_1,j_2=0}^{d-1} |j_1> |j_2> [m_0 |m_0> + y I] \otimes (d-1)
\]

\[
+ \frac{\mu}{d} \sum_{j_1,j_2=0}^{d-1} |j_1> |j_2> |0> |0> \otimes (d-1) - \frac{(1 - \mu)}{d} (x - y) I \otimes [(x - y)|0> |0> + y I] \otimes (d-1)
\]
\[ -\frac{(1-\mu)}{d} y I \otimes \sum_{j=0}^{d-1} [(x-y)|j\rangle \langle j| + y I]^{\otimes(d-1)} \]
\[ + \frac{(1-\mu)}{d} (x-y + d(q-t)) I [(x-y + d(q-t))|0\rangle \langle 0| + (y+dt) I]^{\otimes(d-1)} \]
\[ + \frac{(1-\mu)}{d} (y+dt) I \otimes \sum_{j=0}^{d-1} [(x-y + d(q-t))|j\rangle \langle j| + (y+dt) I]^{\otimes(d-1)}. \]

with \(x\) and \(y\) defined as \(x = p - q\) and \(y = r - t\). After some simple calculations, the mutual information would be:

\[ I_{\text{max}}(\mathcal{E}_{\text{max-en}}(\rho_{[i,n]})) = d \log_2 d + \sum_{[k_i]=0}^{d-1} A_{[k_i]}^0 \log_2[A_{[k_i]}^0] \]
\[ + (d-1) \sum_{[k_i]=0}^{d-1} A_{[k_i]} 1 \log_2[A_{[k_i]} 1]. \]

\[ A_{[k_i]} = -\frac{(1-\mu)}{d}(x-y) \prod_{i=1}^{d-1} [(x-y)\delta_{0,k_i} + y] - \frac{(1-\mu)}{d} y \sum_{j=0}^{d-1} \prod_{i=1}^{d-1} [(x-y)\delta_{j,k_i} + y] \]
\[ + \frac{(1-\mu)}{d} (x-y + d(q-t)) \prod_{i=1}^{d-1} [(x-y + d(q-t))\delta_{0,k_i} + (y+dt)]^{\otimes(d-1)} \]
\[ + \frac{(1-\mu)}{d} (y+dt) \sum_{j=0}^{d-1} \prod_{i=1}^{d-1} [(x-y + d(q-t))\delta_{j,k_i} + (y+dt)] \]
\[ B_{[k_i]} = \frac{(1-\mu)}{d} x \prod_{i=1}^{d-1} [(x-y)\delta_{0,k_i} + y] + \frac{(1-\mu)}{d} y \sum_{m=1}^{d-1} \prod_{i=1}^{d-1} [(x-y)\delta_{m_0,k_i} + y] + \frac{\mu}{d} \prod_{i=1}^{d-1} \delta_{0,k_i}. \]

VIII. APPENDIX B: QUASI-CLASSICAL DEPOLARIZING CHANNEL

In this appendix, we would like to consider state \(|\Psi_{[0,0]}\rangle = \sum_{j=0}^{d-1} A_j |j\rangle^{\otimes d}\), with \(A_j \in \mathbb{C}\) as the input state and calculate

\[ \mathcal{E}(\rho_{[0,0]}) = \left\{ \begin{array}{l}
\mu a \sum_{j_1,j_2=0}^{d-1} A_{j_1} A_{j_2}^* \left[ a|j_1\rangle \langle j_2| + b \sum_{m=1}^{d-1} |j_1 + m\rangle \langle j_2 + m| \right] \\
+ (1-\mu)(a-d-b^d) \sum_{j=0}^{d-1} |A_j|^2 |j\rangle \langle j| + (1-\mu)b^d I \right\} |0\rangle \langle 0|^{\otimes(d-1)} \\
+ (1-\mu)(a-b) \sum_{j=0}^{d-1} |A_j|^2 |j\rangle \otimes [(a-b)|0\rangle \langle 0| + (1-\mu)b I |0\rangle \langle 0|^{\otimes(d-1)} \right\} 
\]
\[(1 - \mu) b \sum_{l=0}^{d-1} |l\rangle \langle l| \otimes \sum_{j=0}^{d-1} |A_{j+l}|^2 [(a - b)|j\rangle \langle j| + (1 - \mu) bI^\otimes(d-1)\]
\[= \left\{ (1 - \mu)(a^d - b^d) \sum_{j=0}^{d-1} |A_j|^2 |j\rangle \langle j| + (1 - \mu) b^d I \right\} |0\rangle \langle 0|^\otimes(d-1). \quad (36)\]

In the above relation \(a = pd\) and \(b = qd\) and the density matrix has two types of sentences which are orthogonal to each other. The matrix elements of the first part (only the first qudit which is the tensor product to \(|0\rangle \langle 0|^\otimes(d-1)\) is given by:

\[E_{k,k} = (1 - \mu) \left[ (a^d - b^d) |A_k|^2 + b^d \right] + \mu \left[ (a - b) |A_k|^2 + b \right], \quad \text{for} \quad k \geq 1, \quad k \neq k'. \quad (37)\]

The explicit form of the matrix elements for the state suggested in the Eq. (24) would be:

\[t = E_{0,0} = (1 - \mu) \left[ (a^d - b^d) \cos^2 \alpha + b^d \right] + \mu \left[ (a - b) \cos^2 \alpha + b \right], \quad \text{for} \quad k \geq 1, \quad k \neq k'. \quad (38)\]

\[s = E_{0,k} = E_{k,0} = \mu \frac{e^{\cos \alpha \sin \alpha}}{\sqrt{d - 1}} \]

\[A_{|k,i\rangle} = (1 - \mu) (a - b) \left[ \cos^2 \alpha \delta_{0,k_0} + \frac{\sin^2 \alpha}{d - 1} (1 - \delta_{0,k_0}) \right] \prod_{j=1}^{d-1} [(a - b) \delta_{0,k_j} + b] \]
\[+ (1 - \mu) b \sum_{j=0}^{d-1} \left[ \cos^2 \alpha (\delta_{0,k_0+j} + \delta_{d,k_0+j}) + \frac{\sin^2 \alpha}{d - 1} [1 - (\delta_{0,k_0+j} + \delta_{d,k_0+j})] \right] \times \prod_{j=0}^{d-1} [(a - b) \delta_{j,k_j} + b] \]
\[= \left\{ (1 - \mu)(a^d - b^d) \left[ \cos^2 \alpha \delta_{0,k_0} + \frac{\sin^2 \alpha}{d - 1} (1 - \delta_{0,k_0}) \right] + (1 - \mu) b^d \right\} \prod_{j=1}^{d-1} \delta_{0,k_j}. \quad (39)\]

We apply the above density matrix to the calculation of the mutual information, as a function of \(\alpha, \mu, d, a,\) and \(b.\) It is given by:

\[I(\xi, \alpha, \mu, d, a, b) = d \log_2 d + (d - 2) \lambda^0 \log_2 \lambda^0 \]
\[+ \lambda^1 \log_2 \lambda^1 + \lambda^2 \log_2 \lambda^2 + \sum_{|k_i, 0\rangle} A_{|k_i\rangle} \log_2 A_{|k_i\rangle} \]

In the above relation \(\lambda^0\) [with the degeneracy of \((d - 2)\)] and \(\lambda^{1,2}\) are the eigenvalues of the first part of the density matrix and are given by:

\[\lambda^0 = c - r, \]
\[\lambda^{1,2} = \frac{1}{2} \left\{ t + c + (d - 2)r \pm \left[ t^2 + 4(d - 1)s \right. \right. \]
\[- 2tc + c^2 - 2(d - 2)(c - t)r + (d - 2)^2 r^2 \left. \right]^{1/2} \} \]

In the following, we consider \(|\Psi_{[0,0]}\rangle = \sum_{j=0}^{d-1} A_j |j\rangle^\otimes d\) with coefficients that are suggested in Eq. (25) for the case \(d = 4,\) as the input state, and use the output density matrix (36). Then the mutual information as a function of \(\theta, \mu, d, a,\) and \(b,\) is given by:

\[I(\xi, \theta, \mu, d, a, b) = d \log_2 d + 2 \lambda^0 \log_2 \lambda^0 \]
\[+ \lambda^1 \log_2 \lambda^1 + \lambda^2 \log_2 \lambda^2 + \sum_{|k_i, 0\rangle} D_{|k_i\rangle} \log_2 D_{|k_i\rangle} \]

In the above relation \(\lambda^0\) (with the degeneracy of 2) and \(\lambda^{1,2}\) are eigenvalues of the first part of the density matrix and \(D_{|k_i\rangle}\) is given by:
\[ D_{[k_i]} = \frac{(1-\mu)}{4}(a-b) \left[ (1 + 3 \cos^2 \theta)\delta_{0,k_0} + \sin^2 \theta(1-\delta_{0,k_0}) \right] \prod_{i=1}^{d-1} ((a-b)\delta_{0,k_i} + b) \]

\[ + \frac{(1-\mu)}{4} b \sum_{j=0}^{d-1} \left[ (1 + 3 \cos^2 \theta)(\delta_{0,k_0+j} + \delta_{d,k_0+j}) + \sin^2 \theta[1 - (\delta_{0,k_0+j} + \delta_{d,k_0+j})] \right] \times \prod_{i=1}^{d-1} ((a-b)\delta_{j,k_i} + b) \]

\[ - \left\{ \frac{(1-\mu)}{4} (a^d - b^d) \left[ (1 + 3 \cos^2 \theta)\delta_{0,k_0} + \sin^2 \theta(1-\delta_{0,k_0}) \right] + (1-\mu)b^d \right\} \prod_{i=1}^{d-1} \delta_{0,k_i}, \quad (41) \]

\[
\lambda^0 = -w + z + f, \\
\lambda^{1,2} = w + 2z + f \pm \sqrt{3v^2 + 4w^2 - 2wz + z^2}^{1/2} \quad (42) \\
\]

\[
w = \frac{\mu}{4} \sin^2 \theta, \quad v = \frac{(a-b)}{2} \cos \theta \sin \theta \\
z = \left\{ \frac{\mu(a-b)}{4} + (1-\mu) \frac{(a^4-b^4)}{4} \right\} \cos^2 \theta \\
f = \mu \left\{ \frac{(a-b)}{4} + b \right\} + (1-\mu) \left[ \frac{(a^4-b^4)}{4} + b^4 \right] \quad (43) \\
\]

**IX. APPENDIX C: DEPOLARIZING CHANNEL**

Similar to Appendix B, the output density matrix and the corresponding mutual information for the depolarizing channel with correlated noise, is given by:

\[
\mathcal{E}(\rho_{[0,0]}) = \begin{cases} 
\left[ (1-\mu)(x-y)^d + \mu x \right] \sum_{j_1,j_2=0}^{d-1} A_{j_1} A_{j_2}^* \langle j_1 | j_2 \rangle + \mu y \sum_{j_1,j_2=0}^{d-1} \sum_{m=1}^{d-1} A_{j_1-m} A_{j_2-m}^* \langle j_1 | j_2 \rangle \\
+ (1-\mu)(x^d - y^d - (x-y)^d) \sum_{j=0}^{d-1} |A_j|^2 \langle j | j \rangle + (1-\mu)y^d \mathbb{I} \right] |0\rangle \langle 0 | \otimes (d-1) \\
+ (1-\mu)(x-y) \sum_{j=0}^{d-1} |A_j|^2 |j \rangle \otimes [(x-y)|0\rangle \langle 0 | + y \mathbb{I}) \otimes (d-1) \\
+ (1-\mu)y \sum_{l=0}^{d-1} |l\rangle \langle l | \otimes \sum_{j=0}^{d-1} |A_{j+l}|^2 [(x-y)|j\rangle \langle j | + y \mathbb{I}) \otimes (d-1) \\
- \left\{ (1-\mu)(x^d - y^d) \sum_{j=0}^{d-1} |A_j|^2 \langle j | j \rangle + (1-\mu)y^d \mathbb{I} \right\} |0\rangle \langle 0 | \otimes (d-1) \\
\end{cases} \\
\]

In the above relation \( x = p + (d-1)q \) and \( y = qd \) and the density matrix has two types of sentences which are orthogonal to each other. The matrix elements of the first part (only the first qudit, which is the tensor product to \(|0\rangle \langle 0 | \otimes (d-1)\) is given by:

\[
E_{k,k} = (1-\mu) \left[ (x^d - y^d)|A_k|^2 + y^d \right] + \mu \left[ (x-y)|A_k|^2 + y \right], \\
E_{k,k'} = \left[ (1-\mu)(x-y)^d \right] + \mu x \left[ A_k A_{k'}^* + \mu y \sum_{m=1}^{d-1} A_{k-m} A_{k'-m}^* \right]. \quad (44) \\
\]

The explicit form of the matrix elements for the state that are suggested in Eq. (24) would be:

\[
t = E_{0,0} = \left[ (1-\mu)(x-y) + \mu x \cos^2 \alpha + \mu y \sin^2 \alpha \right] + \left[ (1-\mu)(x^d - y^d - (x-y)^d) \right] \cos^2 \alpha + (1-\mu)y^d, \\
\]

The output density matrix and the corresponding mutual information for the depolarizing channel with correlated noise, is given by:

\[
\mathcal{E}(\rho_{[0,0]}) = \begin{cases} 
\end{cases} \\
\]
\[ s = E_{0,k} = E_{k,0} = \left[(1 - \mu)(x - y)^d + \mu x \right] \frac{\cos \alpha \sin \alpha}{\sqrt{d - 1}} + \mu y \left[ \frac{\sin^2 \alpha}{d - 1} (d - 2) + \frac{\cos \alpha \sin \alpha}{\sqrt{d - 1}} \right], \]
\[ k \geq 1, \]
\[ c = E_{k,k} = \left[(1 - \mu)(x - y)^d + \mu x \right] \frac{\sin^2 \alpha}{d - 1} + \mu y \left[ \frac{\sin^2 \alpha}{d - 1} (d - 2) + \cos^2 \alpha \right] + \left[(1 - \mu)(x^d - y^d) - (x - y)^d \right] \frac{\sin^2 \alpha}{d - 1} \]
\[ k \neq k' \geq 1. \]

We apply the above density matrix to calculate the mutual information as a function of \( \alpha, \mu, d, x, \) and \( y. \) It is given by:
\[
I(\mathcal{E}, \alpha, \mu, d, a, b) = d \log_2 d + (d - 2) \lambda^0 \log_2 \lambda^0 + \lambda^1 \log_2 \lambda^1 + \lambda^2 \log_2 \lambda^2 + \sum_{[k_i = 0]}^{d-1} A_{[k_i]} \log_2 A_{[k_i]}.
\]

In the above relation \( \lambda^0 \) (with the degeneracy of \( (d - 2) \)) and \( \lambda^{1,2} \) are eigenvalues of the first part of the density matrix which are given by:
\[
\lambda^0 = c - r,
\]
\[
\lambda^{1,2} = \frac{1}{2} \left\{ t + c + (d - 2)r \pm \left[ t^2 + 4(d - 1)s + 2tc + c^2 - 2(d - 2)(c - t)r + (d - 2)r^2 \right]^{1/2} \right\}.
\]

Similar to the previous section, we consider \( |\Psi_{[0,0]}\rangle = \sum_{j=0}^{d-1} A_j |j\rangle \otimes d \) with coefficients that are suggested in Eq. (25) for the case \( d = 4 \), as the input state and calculate the mutual information for the quantum depolarizing channel as a function of \( \theta, \mu, d, x, \) and \( y \) for states that are suggested in (25). The mutual information would be given by Eq. (40), by replacing \( x \rightarrow x \) and \( b \rightarrow y \) and a similar expression for \( D_{[k_i]} \) and \( \lambda^{0,1,2} \), with the following variables for \( w, v, z, \) and \( f \):
\[
w = \frac{1}{4} \left[ (1 - \mu)(x^4 - y^4) + \mu \right] \sin^2 \theta,
\]
\[
v = \frac{1}{2} \left[ (1 - \mu)(x^4 - y^4) + \mu (x - y) \right] \cos \theta \sin \theta
\]
\[
z = \left\{ \frac{\mu(x - y)}{4} + (1 - \mu) \left( \frac{x^4 - y^4}{4} \right) \right\} \cos^2 \theta
\]
\[
f = \frac{x - y}{4} + y + (1 - \mu) \left[ \frac{x^4 - y^4}{4} + y^4 \right].
\]

[1] M. A. Nielsen and I. L. Chuang, in Quantum Computation and Quantum Information (Cambridge University Press, Cambridge, U. K., 2000).
[2] P. W. Shor, Phys. Rev. A 52, 2493 (1995); E. Knill and R. Laflamme, Phys. Rev. A 55, 900 (1997); D. Gottesman, Phys. Rev. A 54, 1862 (1996); H. Bennett et al., Phys. Rev. A 54, 3824 (1996). [3] E. Knill et al., Science 279, 342 (1998); P. Zanardi and M. Rasetti, Phys. Rev. Lett. 79, 3306 (1997); D. A. Lidar, I. L. Chuang, and K. B. Whaley, Phys. Rev. Lett. 81, 2594 (1998); D. A. Lidar et al., Phys. Rev. A, 61, 052307 (2000); J. Kempe et al., Phys. Rev. A, 63, 042307 (2001); D. Bacon et al., Phys. Rev. Lett. 85, 1758 (2002).
[4] 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).
A. Serafini, J. Eisert and M. M. Wolf, Phys. Rev. A 71, 012320 (2005).

T. Hiroshima, Phys. Rev. A 73, 012320 (2006).

S. D. Bartlett, T. Rudolph, and R. W. Spekkens, Phys. Rev. Lett. 91, 027901 (2003).

J. C. Boileau et al., Phys. Rev. Lett. 92, 017901 (2004).

J. C. Boileau et al., Phys. Rev. Lett. 93, 220501 (2004).

V. Giovannetti et al., Phys. Rev. Lett. 92, 27902 (2004); V. Giovannetti et al., Phys. Rev. Lett. 91, 47901 (2003); V. Giovannetti et al., Phys. Rev. A 70, 22328 (2004); V. Giovannetti et al., Phys. Rev. A 70, 32315 (2004); V. Giovannetti et al., Phys. Rev. A 70, 12307 (2004); V. Giovannetti et al., Phys. Rev. A 68, 62323 (2003); V. Giovannetti et al., Phys. Rev. A 69, 62307 (2004).

C. Macchiavello and G. M. Palma, Phys. Rev. A 65, 050301(R) (2002); C. Macchiavello, G. M. Palma, and S. Virmani, Phys. Rev. A 69, 010303(R) (2004).

G. Bowen and S. Mancini, Phys. Rev. A, 69, 012306 (2004); D. Bruß, L. Faoro, C. Macchiavello and G.M. Palma, J.Mod.Opt. 47, 325, 2000; (also quant-ph/0003033); C. King, quant-ph/0103156; K. Matsumoto, T. Shimono and A. Winter, quant-ph/0206148; K. M. R. Audenaert and S. L. Braunstein, quant-ph/0303045.

V. Giovannetti and S. Mancini, Phys. Rev. A, 71, 062304 (2004); D. Kretschmann and R. Werner, Phys. Rev. A, 72, 062323 (2005); N. J. Cerf, J. Clavareau, C. Macchiavello and J. Roland, Phys. Rev. A, 72, 042330 (2005); quant-ph/0412089; V. Giovannetti, J. Phys. A, 38, 10989 (2005).

J. Ball, A. Dragan, and K. Banaszek, Phys. Rev. A, 69, 042324 (2004); K. Banaszek et al., Phys. Rev. Lett., 89, 257901, (2004).

N. Arshed and A. H. Toor, Phys. Rev. A 73, 014304 (2006); A. Fahmi, quant-ph/0605024.

C. Brukner, M. Žukowski, and A. Zeilinger, Phys. Rev. Lett. 89, 197901 (2002); R. W. Spekkens, and T. Rudolph, Phys. Rev. Lett. 89, 227901 (2001).

H. Bechmann-Pasquinucci and A. Peres, Phys. Rev. Lett. 85, 3313, (2000).

N. J. Cerf et al., Phys. Rev. Lett. 88, 127902 (2002).

A. Mair, et al., Nature (London) 412, 313 (2001); A. Váziri, G. Weihs, and A. Zeilinger, Phys. Rev. Lett. 89, 240401 (2002); A. Váziri, et al. Phys. Rev. Lett. 91, 227902 (2003); G. Molina-Terriza, et al. Phys. Rev. Lett. 92, 167903 (2004); J. Leach, et al. Phys. Rev. Lett. 88, 257901 (2002); J. Leach, et al. Phys. Rev. Lett. 92, 013601 (2004); A. T. O’Neill, et al. Phys. Rev. Lett. 88, 033601 (2002); S. Barreiro and J. W. R. Tabosa, Phys. Rev. Lett. 90, 133001 (2003).

G. G. Amosov, S. Mancini, and V. I. Manko, J. Phys. A, 39, 3375-3380 (2006).

A. Vaziri, et al. (quant-ph/0008022); S. D. Bartlett, et al. (quant-ph/0109066); N. D. Mermin, (quant-ph/0105117); V. Karimipour, et al., Phys. Rev. A 65, 052331 (2002).