

Quantum channels with a finite memory

Garry Bowen^{1,*} and Stefano Mancini^{2,†}¹*Centre for Quantum Computation, Clarendon Laboratory, University of Oxford, Oxford OX1 3PU, United Kingdom*²*Dipartimento di Fisica, Università di Camerino, I-62032 Camerino, Italy*

(Received 2 May 2003; published 13 January 2004)

In this paper we study quantum communication channels with correlated noise effects, i.e., quantum channels with memory. We derive a model for correlated noise channels that includes a channel memory state. We examine the case where the memory is finite, and derive bounds on the classical and quantum capacities. For the entanglement-assisted and unassisted classical capacities it is shown that these bounds are attainable for certain classes of channel. Also, we show that the structure of any finite-memory state is unimportant in the asymptotic limit, and specifically, for a *perfect* finite-memory channel where no information is lost to the environment, achieving the upper bound implies that the channel is asymptotically noiseless.

DOI: 10.1103/PhysRevA.69.012306

PACS number(s): 03.67.Hk, 03.65.Ud, 89.70.+c

I. INTRODUCTION

Quantum communication through noisy channels has, to date, mainly concentrated on quantum channels that are memoryless. A memoryless channel is defined as a noisy channel where the noise acts independently on each symbol transmitted through the channel. In classical information theory the discrete memoryless channel is well understood in terms of the capacity of the channel, and the capacity of such channels is invariant under inclusion of feedback or shared random correlations [1]. The existing capacity theorems for quantum channels have also concentrated on the memoryless case [2–5]. The calculation of capacities for classical channels with correlated noise, or *memory* channels, has had much more limited success. One type of classical memory channel for which the capacity is known is the channel with Markov correlated noise. In this paper, we examine the quantum extension of the channel with Markov correlated noise. In particular, we examine a model of a correlated noise channel which utilizes unitary operations between the transmitted states, an environment, and a shared memory state, and determine some of the characteristics of such a class of quantum channels.

The fact that entangled alphabet states may increase the classical capacity of a particular correlated noise channel has been shown [6]. For the corresponding memoryless channel the capacity is known to be additive, and hence entangled input states cannot increase the classical capacity of the memoryless channel over product state encoding.

The phrase *finite memory* is used to describe one aspect of the model, a memory state of finite dimension. The finiteness of the memory is defined only in terms of the model used to describe the correlated noise of the channel, and is not necessarily a physical consideration. The correlations between errors may be considered either temporally over each use of a single channel, or spatially between uses of many parallel channels. The dimension of the memory is determined by the number of Kraus operators in the single channel expansion

and the correlation length of the channel, which may be defined as the maximum number of channel uses for which the noise is not conditionally independent. Any channel with a finite correlation length may be generated by a channel with a finite memory, according to this model. Although a physical interpretation of the model is not necessary to achieve the goal of determining capacity theorems, it may give an understanding of the physical motivation. Over short times the environment with which the transmitted state acts may be assumed to be arbitrarily large, with interactions between components of the vast environment essentially making the recovery of the information impossible. The memory, however, may be interpreted as a subspace of the environment which does not “decay” over the time scale of separate uses of the channel, and is therefore dependent on the previous state of the channel.

A physical example of a memory channel is the recent proposal by Bose [7], which uses unmodulated spin chains to transmit quantum information. In this case the initial state introduced to the chain by Alice acts as both the input state and a part of the memory state for further uses of the channel, as it is assumed Alice replaces each transmitted state with a new spin state after each use of the channel, while the remaining elements of the spin chain constitute both the physical channel, the memory of the channel, and the output state, which may be removed from the chain for future decoding by Bob.

II. A MODEL FOR QUANTUM MEMORY CHANNELS

The Kraus representation theorem [8] is an elegant and powerful method of representing quantum dynamics in two different ways, as a sum over operators acting on the state, or alternatively as a unitary evolution of a state and environment. The unitary interaction model provides an intuitive understanding of open quantum systems, as well as providing a method of calculation. In deriving our model of a quantum memory channel, we try to preserve the useful aspects that such a unitary representation provides.

A. Unitary representation of memoryless channels

A quantum channel is defined as a completely positive, trace preserving map from the set of density operators to

*Electronic address: gab30@cam.ac.uk

†Electronic address: stefano.mancini@unicam.it

itself. Any such map may be represented as a unitary operation between the system state and an environment with a known initial state. For a single channel use the output state is given by

$$\rho'_Q = \text{Tr}_E[U_{QE}(\rho_Q \otimes |0_E\rangle\langle 0_E|)U_{QE}^\dagger], \quad (1)$$

with ρ_Q the input state, and ρ'_Q the output state. For a sequence of transmissions through the channel,

$$\rho'_Q = \text{Tr}_E[U_{n,E_n} \cdots U_{1,E_1}(\rho_Q \otimes |0_{E_1} \cdots 0_{E_n}\rangle\langle 0_{E_1} \cdots 0_{E_n}|)U_{1,E_1}^\dagger \cdots U_{n,E_n}^\dagger] \quad (2)$$

$$= (\Lambda_n \otimes \cdots \otimes \Lambda_1)\rho_Q, \quad (3)$$

where the state ρ_Q now represents a (possibly entangled) input state across the n channel uses, and the environment state is a product state $|0_{E_1} \cdots 0_{E_n}\rangle = |0_{E_1}\rangle \otimes \cdots \otimes |0_{E_n}\rangle$.

B. A unitary model for memory channels

One model of a quantum memory channel is where each state going through the channel acts with a unitary interaction *on the same channel memory state*, as well as an independent environment. The backaction of the channel state on the message state therefore gives a memory to the channel. The general model thus includes a channel memory M , and the independent environments for each qubit E_i . Hence,

$$\begin{aligned} \rho'_Q &= \text{Tr}_{ME}[U_{n,ME_n} \cdots U_{1,ME_1}(\rho_Q \otimes |M\rangle\langle M| \otimes |0_{E_1} \cdots 0_{E_n}\rangle\langle 0_{E_1} \cdots 0_{E_n}|) \\ &\quad \times U_{1,ME_1}^\dagger \cdots U_{n,ME_n}^\dagger] \\ &= \text{Tr}_M[\Lambda_{n,M} \cdots \Lambda_{1,M}(\rho_Q \otimes |M\rangle\langle M|)], \end{aligned} \quad (4)$$

where ρ_Q and ρ'_Q are the input and the output state of the channel while the trace over the environment is over all environment states. If the unitaries factor into independent unitaries acting on the memory and the combined state and environment, that is, $U_{n,ME_n} = U_{n,E_n}U_M$, then the memory traces out and we have a memoryless channel. If the unitaries reduce to $U_{n,M}$, we can call it a *perfect* memory channel, as no information is lost to the environment.

The question remains as to what channels can be modeled by Eq. (4)? From the Kraus representation theorem [8], we know that for any block of length n , then any channel acting on the n states may be modeled with a memory plus environment of dimension at most d^{2n} , for d the dimension of the channel. However, the unitary operation may not be factorable into a product of operators acting in the form of Eq. (4).

C. Examples of finite-memory channels

A naive “memory” channel can be considered by the two qubit channel given by the Kraus operators, $A_0 = \frac{1}{2}(\mathbb{I} \otimes \mathbb{I})$, and, $A_1 = \frac{1}{2}(\sigma_Z \otimes \sigma_Z)$, and can be modeled by using a memory state that is initially in the $|0\rangle\langle 0|$ state, and is the target of a controlled-NOT (CNOT) operation by only two qu-

bits before being reset to the initial state. However, this channel is essentially just a memoryless channel in the higher dimension space, transmitting qudits of dimension four, and cannot therefore be considered useful as a model of a memory channel. This channel also does not fit into the model of Eq. (4) as the memory is erased separately after every two qubits. All such channels which may be factored into memoryless channels for some finite number of uses may therefore be described using the existing properties known for memoryless channels.

A simple example of a perfect memory channel is an extension of the qubit dephasing channel. For this channel CNOT gates operate between the qubits going through the channel and a target memory state, initially given as $|0_M\rangle\langle 0_M|$, which replaces the environment. The output states have the same reduced density matrices as if they had passed through a memoryless dephasing channel, but the states are also correlated across channel uses, that is, a product input state does not necessarily give a product output state. We call this channel the correlated dephasing channel.

Using the unitary SWAP gate to model a channel simply acts as a shift by a single state. Since for this “shift channel” the SWAP gate the unitaries act to increment the index for the position of the input states, then on a block of n inputs only the last input state is not recoverable. Hence the transmission rate for intact states for blocks of size n is simply $1 - 1/n$, which approaches a noiseless channel in the limit $n \rightarrow \infty$.

D. Channels with Markovian correlated noise

An important class of channels that may be represented by the memory channel model are channels with Markovian correlated noise. A Markovian correlated noise channel of length n , is of the form,

$$\begin{aligned} \Lambda^{(n)}\rho &= \sum_{i_0, \dots, i_n} P_{i_n|i_{n-1}} P_{i_{n-1}|i_{n-2}} \cdots P_{i_1|i_0} P_{i_0} \\ &\quad \times (A_{i_n} \otimes \cdots \otimes A_{i_0})\rho(A_{i_n}^\dagger \otimes \cdots \otimes A_{i_0}^\dagger), \end{aligned} \quad (5)$$

where the set A_{i_k} are Kraus operators for single uses of the channel on state k [6]. The motivation for looking at Markovian channels is that the properties of typical sequences generated from Markovian sources are well understood, and the typical sequences of errors generated in Eq. (5) will be directly related to these typical sequences.

The correlated dephasing channel may be described using the memory to correlate the dephasing error for each qubit, that is, the probability of the k th qubit undergoing a phase error is determined exactly by whether an error occurred on the previous qubit. Thus, for the correlated dephasing channel with error operators $A_{0_n} = \mathbb{I}^{(n)}$ and $A_{1_n} = \sigma_Z^{(n)}$, acting on the n th qubit, the conditional probabilities are given by $P_{k_n|i_{n-1}} = \delta_{jk}$, with an initial probability of error given by $P_0 = P_1 = 1/2$. This channel may be generated using the unitary operation

$$U_{i,M}|\phi^{(i)}\rangle|0_M\rangle = |\phi^{(i)}\rangle|0_M\rangle, \quad (6)$$

$$U_{i,M}|\phi^{(i)}\rangle|1_M\rangle = \sigma_Z^{(i)}|\phi^{(i)}\rangle|1_M\rangle, \quad (7)$$

with an initial memory state $|M\rangle = 1/\sqrt{2}(|0\rangle + |1\rangle)$. The equivalence of the controlled phase gate in Eqs. (6) and (7) to the use of a CNOT with a memory initially in the $|0_M\rangle\langle 0_M|$ state is obtained by noting $U_{CPHASE} = (\mathbb{I} \otimes H)U_{CNOT}(\mathbb{I} \otimes H)$ for H a qubit Hadamard rotation on the memory state. The channel is asymptotically noiseless, as all states with an even number of $|1\rangle$'s are invariant, and therefore this subspace may be mapped onto by simply adding a single ancilla qubit. The encoding map from $\mathcal{H}^{\otimes n}$ to $\mathcal{H}^{\otimes(n+1)}$, may then transform states with even numbers of $|1\rangle$'s to the same state tensored with $|0\rangle$, and those states with odd numbers of $|1\rangle$'s to these states tensored with $|1\rangle$. The coded subspace of $\mathcal{H}^{\otimes(n+1)}$ is then noiseless. The rate of transmission through $n+1$ uses of the channel is therefore $n/(n+1)$, which approaches unity in the limit $n \rightarrow \infty$.

For a general channel with Markovian correlated noise, that is, $p_{j|j-1} = p_{j|(j-1)(j-2)\dots i}$ for all $i < j$, the channel may be generated using the model given in Eq. (4), where the unitary operator is given by,

$$U_{i,ME_i}|\phi^{(i)}\rangle|j_M\rangle|0_{E_i}\rangle = \sum_k \sqrt{p_{kj}}A_k^{(i)}|\phi^{(i)}\rangle|k_M\rangle|j_{E_i}\rangle. \quad (8)$$

The initial memory state determined by the values of the initial probability vector for the error operators $[p_0, p_1, \dots, p_m]$, for a family of m operators, by the relationship

$$\vec{\alpha} = \Gamma^{-1}\vec{p} \quad (9)$$

for Γ the transition matrix with entries $p_{j|i}$, and $\vec{\alpha}$ the squares of the amplitudes for the initial memory state $|M\rangle = \sum_j \sqrt{\alpha_j}|j_M\rangle$. This is, of course, provided that the transition matrix is not singular. For a singular matrix we may utilize a different unitary operation V on the initial use of the channel, which will not change the asymptotic behavior of the channel. We may also utilize a mixed initial memory state $\rho_M = \sum_j \alpha_j |j_M\rangle\langle j_M|$, instead of the pure state, without affecting the behavior of the channel.

The derivation of the specific model for the correlated dephasing channel in Eqs. (6) and (7) differs from the prescription given in Eq. (8), in that it does not require the extra environment. The unitary operation on the initial states produces orthogonal outputs, whereas in the general case the states for each $|j_M\rangle$ in Eq. (8) may not necessarily be orthogonal without the environment. If the output state for a given $|j_M\rangle$ in Eq. (8) is orthogonal to all other outputs generated by different initial memory states, then the final environment state for this particular output can ‘‘overlap’’ and does not need to be orthogonal to the other environment states. This occurs in the correlated dephasing channel, and results in the channel requiring no environment at all. However, it shall be shown that the behavior of these two different channel constructions is identical, as the actual size of the memory becomes irrelevant in the asymptotic limit, provided it is finite.

The noisy channel described by Macchiavello and Palma [6] may be described in the context of Eq. (8), with the error operators given by the identity $A_{0_n} = \mathbb{I}^{(n)}$ and the Pauli spin matrices $A_{1_n} = \sigma_X^{(n)}$, $A_{2_n} = \sigma_Y^{(n)}$, and $A_{3_n} = \sigma_Z^{(n)}$, and the transition matrix elements defined as $p_{k|j} = (1-\mu)p_k + \mu\delta_{jk}$, where μ is a correlation parameter. The steady-state probabilities for this transition matrix are given by the uniform distribution.

III. CAPACITIES FOR FINITE-MEMORY CHANNELS

Having generated a model for the channel, we must address whether such a model is instructive in obtaining capacity theorems for the channels the model represents. The existence of a unitary representation of an interaction with an environment does allow the extension of results from memoryless channels which rely on similar arguments, such as the coherent information bound and the quantum Fano inequality [9].

A. Results for classical capacities

An upper bound on the classical information that may be sent through the memory channel is given by the Holevo bound [10]. The maximum mutual information generated between sender and receiver, per channel use, for n channels is then given by,

$$S_{\max}^{(n)} = \max_{\{p_i, \rho^i\}} \frac{1}{n} \left[S \left(\sum_i p_i \text{Tr}_M [\Lambda_M^{(n)}(\rho_Q^i \otimes \rho_M)] \right) - \sum_i p_i S(\text{Tr}_M [\Lambda_M^{(n)}(\rho_Q^i \otimes \rho_M)]) \right], \quad (10)$$

where for each n , $\Lambda_M^{(n)} = \Lambda_{n,M} \cdots \Lambda_{1,M}$, is a channel, and the asymptotic limit is achieved by taking $n \rightarrow \infty$. The ensemble of states $\rho_Q^i = \rho_A^i$ is a set of states generated by the sender, Alice, for unassisted communication, or $\rho_Q^i = \rho_{AB}^i$ is a set of shared entangled states between sender, Alice, and receiver, Bob, for entanglement assisted communication, with the requirement that $\rho_B^i = \rho_B$. To reduce the number of subscripts, the use of the notation $\rho^i \equiv \rho_Q^i$ for the signal states shall be used for the rest of this section.

The argument for achieving this upper bound does not extend easily to the memory channel case. The problem lies in the fact that the coding for the channel cannot be broken up into blocks of n uses, as the memory state may be entangled with the previous block and thus may not be identical for each block.

The bound in Eq. (10) is achievable for a class of regular Markovian correlated noise channels. The channels are assumed to be representable by unitary Kraus operators (and are therefore unital), and have initial error probability distributions equal to the steady-state probabilities. The asymptotic use of the channel may be segmented into approximate channels of length n . That is, by tracing out all other states for each length n segment, we obtain a channel where for a total length $l \gg n$ we have $\Lambda^{(l)} \approx \Lambda^{(n)} \otimes \dots$

$\otimes \Lambda^{(n)}$. From the theory of Markov chains, we know that the approximate channel for a product state input is given for a single use by

$$\Lambda^{(1)}\rho \approx \sum_{i_n} \tilde{p}_{i_n} A_{i_n} \rho A_{i_n}^\dagger = \text{Tr}_M \Lambda_M^{(1)}(\rho \otimes \tilde{\rho}_M), \quad (11)$$

where $\tilde{p}_{i_n} = \tilde{p}_i$ are the steady-state probabilities, $\tilde{\rho}_M$ is the memory density matrix with the \tilde{p}_i on the diagonal, and n is taken to be suitably large. The derivations required for this approximation are shown in the Appendix. Similarly, with n large, two uses the channel are approximated by,

$$\begin{aligned} \Lambda^{(2)}\rho &\approx \sum_{i_n, i_{n-1}} p_{i_n | i_{n-1}} \tilde{p}_{i_{n-1}} A_{i_n, i_{n-1}} \rho A_{i_n, i_{n-1}}^\dagger \\ &= \text{Tr}_M \Lambda_M^{(2)}(\rho \otimes \tilde{\rho}_M), \end{aligned} \quad (12)$$

with $A_{i_n, i_{n-1}} = A_{i_n} \otimes A_{i_{n-1}}$, and ρ a possibly entangled input state across the two transmissions through the channel. This construction may be extended for arbitrary lengths n . In the case that the initial distribution p_{i_0} is equal to the steady-state distribution $p_{i_0} = \tilde{p}_i$, the approximations in Eqs. (11) and (12) become exact. This is true for all lengths n , with $\text{diag } \rho_M = \text{diag } \tilde{\rho}_M$ always, where $\text{diag } \rho$ is the density matrix formed from the diagonal elements of ρ . Therefore the achievable rate is obtained immediately from the *Holevo-Schumacher-Westmoreland theorem* [3,4].

The correlated dephasing channel gives an easy example of the achievability of the capacity, as for this channel any initial distribution is a steady-state probability. A rate equal to the unassisted classical capacity is achieved using the orthogonal states $\{|0\rangle, |1\rangle\}$ with *a priori* probability of $p_i = 1/2$ for this channel, and hence the limit is achieved in this case when $n = 1$. The entanglement assisted capacity for this channel $C_E = 2$ is, however, only achieved in the asymptotic limit as the block size $n \rightarrow \infty$.

In the case that the initial error probabilities differ from the steady state, much of the derivation above is still applicable. From the convergence properties of regular Markovian sequences, we know that $\text{diag } \rho_M \rightarrow \tilde{\rho}_M$ as n becomes large, where $\tilde{\rho}_M$ is the diagonal density matrix with eigenvalues equivalent to the steady-state probabilities. Similarly, for any $\delta > 0$ there exists an n for which the total probability of the atypical sequences of Kraus operators is less than δ . This follows from the behavior of regular Markovian sources in the Shannon theory [1]. The contribution to the state $\Lambda^{(n)}\rho$ when the initial probabilities are not the steady-state probabilities may therefore be small enough such that the bounds on the total probability of error may be made arbitrarily small asymptotically, although at present this remains an open question.

For any channel with a finite memory where the capacity equals the upper bound it may be seen that the exact nature of the memory has little effect on the asymptotic behavior of the channel. The correlated dephasing channel, where two possible constructions exist each with a different sized

memory state, is an example. To analyze the behavior we assume that Bob has access to the memory after the block is sent, and as such he can measure the information in M as well, then reset the memory to a given initial state before the next block. This gives an achievable rate,

$$\begin{aligned} R \equiv \lim_{n \rightarrow \infty} \max_{\{p_i, \rho^i\}} \frac{1}{n} &\left[S \left(\sum_i p_i \Lambda^{(n)}(\rho^i \otimes \rho_M) \right) \right. \\ &\left. - \sum_i p_i S(\Lambda^{(n)}(\rho^i \otimes \rho_M)) \right] \end{aligned} \quad (13)$$

$$\begin{aligned} &\leq \lim_{n \rightarrow \infty} \max_{\{p_i, \rho^i\}} \frac{1}{n} \left[S \left(\sum_i p_i \text{Tr}_M[\Lambda^{(n)}(\rho^i \otimes \rho_M)] \right) \right. \\ &\quad \left. - \sum_i p_i S(\text{Tr}_M[\Lambda^{(n)}(\rho^i \otimes \rho_M)]) + 2 \log_2 d_M \right] \end{aligned} \quad (14)$$

$$= S_{\max}^{(n)} + \frac{2}{n} \log_2 d_M \quad (15)$$

where Eq. (14) follows from Eq. (13) by strong subadditivity and the factor $2 \log_2 d_M$ is an upper bound on entropy of the memory state living in a space of dimension d_M . The bound for the rate R of a channel generated from tracing both the environment and the memory, is then sandwiched by the terms, $n S_{\max}^{(n)} + 2 \log_2 d_M \geq nR \geq n S_{\max}^{(n)}$, which would approach the channel capacity for the channel including access to the memory, as $n \rightarrow \infty$, for any finite-memory channel. The channel capacity is thus only affected by the loss of information to the environment, and the loss of information into the memory state may be seen to vanish in the asymptotic limit. For a perfect memory channel the channel will be asymptotically noiseless, as was shown for the examples of the shift channel and the correlated dephasing channel.

B. Results for quantum capacities

The quantum capacities are determined by the maximum asymptotic rates at which intact quantum states may be transmitted through a noisy quantum channel. The coherent information bound [9,11] on the quantum capacity applies directly to the case of memory channels. The role the memory plays in the coherent information bound may be seen by examining the converse to the bound, the quantum Fano inequality, which is shown in the following section.

There exist a number of quantum capacities dependent on available additional resources. Primarily there is the unassisted quantum capacity Q , the capacities assisted by classical side channels Q_1, Q^{FB}, Q_2 , denoting forward, backward (feedback), and two way classical communication, respectively, and, the entanglement assisted quantum capacity Q_E , achievable when sender and receiver share unlimited amounts of entanglement prior to communication taking place. For memoryless channels the situation is slightly simplified by the equivalence $Q_1 = Q$ [12,13], whether this holds for channels with memory is not yet known.

The entanglement assisted quantum capacity is simply related to the entanglement assisted classical capacity by the use of quantum dense coding and quantum teleportation, giving the equality $C_E = 2Q_E$ [14]. The actual nature of the channel does not affect this relationship.

C. The quantum Fano inequality

The quantum Fano inequality [9] is used to give a converse to any quantum coding theorems. The inequality describes the loss in fidelity of the transmitted states that occurs due to the exchange of entropy to the environment during transmission through the channel.

Taking a state ρ_Q with a purification in terms of a reference system R , such that, $\rho_Q = \text{Tr}_R |\psi_{QR}\rangle\langle\psi_{QR}|$, we define the entanglement fidelity as $F = \langle\psi_{QR}|\rho'_{QR}|\psi_{QR}\rangle$, where ρ'_{QR} is the total output state following the transmission of ρ_Q through the noisy channel. The quantum Fano inequality may be applied to the finite-memory channel by simply noting that the entropy exchange to the environment E may be rewritten as,

$$S(\rho'_E) = S(\rho'_{MQR}) \leq S(\rho'_M) + S(\rho'_{QR}). \quad (16)$$

It is assumed here that the memory state is initially pure, as it does not affect the derivation compared to a mixed memory state. This is because any finite-memory state may be purified with another finite reference system. This is also equivalent to applying the Fano inequality using $S(\rho'_{ME})$ as the environment, and then utilizing the Araki–Lieb inequality to obtain $S(\rho'_{QR}) \geq S(\rho'_E) - S(\rho'_M)$.

This leads to a Fano inequality for channels with a finite memory,

$$S_E \leq \log_2 d_M + h(F) + (1-F)\log_2(d^2 - 1) \quad (17)$$

for S_E is the entropy exchange with the environment, F is the entanglement fidelity, $h(F) = -F\log_2 F - (1-F)\log_2(1-F)$ the binary entropy of the entanglement fidelity, d the dimension of \mathcal{H}_Q , and d_M the dimension of the memory. For a single channel use, this inequality may be weak, but in the case of multiple uses the inequality can become stronger. This is due to the average entropy exchange for a large number of channel uses N being given by

$$\begin{aligned} \frac{1}{N} S_E &\leq \frac{1}{N} [\log_2 d_M + h(F) + (1-F)\log_2(d^{2N} - 1)] \\ &\approx 2(1-F)\log_2 d, \end{aligned} \quad (18)$$

where the first two terms in the sum on the right-hand side may be made arbitrarily small, given large enough N . This may be interpreted as the fact that a high entanglement fidelity over many uses of the channel necessarily implies a low average entropy exchange with the environment. In the asymptotic limit the particular channel construction used, and the exact nature of any finite-memory state, are both “irrelevant” in terms of the bounds on the channel capacity.

IV. CONCLUSION

A model for a class of quantum channels with memory has been proposed. The class of channels that may be described by this model include the set of channels with Markovian correlated quantum noise. For these types of channels it has been shown that the memory state required to generate the channel is finite. These finite-memory channels have similar asymptotic behavior to the quantum memoryless channels, in that they may be essentially described by the loss of information to an initial product state environment after a unitary interaction with the states transmitted through the channel. The size of the memory state is finite, and so the effect on loss from the channel is vanishing in the asymptotic limit. The simplest demonstration is the case of perfect memory channels where no information at all is lost to the environment and so achievement of the upper bound on the capacity for this class of channels will asymptotically give a noiseless quantum channel.

It has also been demonstrated that Holevo–Schumacher–Westmoreland coding can achieve the capacity bound for channels with Markov correlated noise, where the Kraus operators are unitary, providing the initial error probabilities are equal to the steady-state probabilities for the regular Markov chain.

The unitary representation of the channel also allows for derivations of bounds on the quantum capacity using the coherent information, and application of the quantum Fano inequality to finite-memory channels.

ACKNOWLEDGMENTS

While this work was undertaken, G.B. was supported by the Oxford–Australia Trust, the Harmsworth Trust, and Universities UK.

APPENDIX EVOLUTION OF THE CHANNEL AND MEMORY STATE

1. Derivation of the channel from the unitary construction

Here it is shown that the diagonal elements of the memory state determine the error operators for the next transmitted state. For the memory state $\rho_M = \sum_{jl} \lambda_{jl} |j_M\rangle\langle l_M|$, the channel for the next transmitted state is given by

$$\begin{aligned} &\sum_{jl} \lambda_{jl} \text{Tr}_{ME} [U_{QME} |\phi_Q\rangle |j_M\rangle |0_E\rangle \langle 0_E| \langle l_M| \langle \phi_Q| U_{QME}^\dagger] \\ &= \sum_{jl} \lambda_{jl} \text{Tr}_M \left[\sum_{km} \sqrt{p_{k|j} p_{m|l}} \delta_{jl} A_k |\phi_Q\rangle |k_M\rangle \langle m_M| \langle \phi_Q| A_m^\dagger \right] \\ &= \sum_{jl} \lambda_{jl} \delta_{jl} \sum_{kmn} \sqrt{p_{k|j} p_{m|l}} A_k |\phi\rangle \langle n| k_M \langle m_M| n \rangle \langle \phi| A_m^\dagger \\ &= \sum_j \lambda_{jj} \sum_k p_{k|j} A_k |\phi_Q\rangle \langle \phi_Q| A_k^\dagger. \end{aligned}$$

The error operators are determined by the diagonal elements of the memory only, the off-diagonal matrix elements have no effect on the channel.

2. Convergence of the diagonal elements of the memory state

The exact nature of the memory state itself depends on the states transmitted through the channel. Perhaps surprisingly, however, the diagonal elements of the state are *independent* of the transmitted states. To show this we note for a memory initially in the state $\rho_M = \sum_{j,l} \lambda_{jl} |j_M\rangle\langle l_M|$, the new memory state after one iteration of the channel is

$$\begin{aligned} & \sum_{j,l} \lambda_{jl} \text{Tr}_{QE} [U_{QME} |\phi_Q\rangle |j_M\rangle |0_E\rangle\langle 0_E| \langle l_M| \langle \phi_Q| U_{QME}^\dagger] \\ &= \sum_{j,l} \lambda_{jl} \text{Tr}_Q \left[\sum_{kmn} \sqrt{p_{k|j} p_{m|l}} \delta_{jl} A_k |\phi_Q\rangle |k_M\rangle\langle m_M| \langle \phi_Q| A_m^\dagger \right] \\ &= \sum_{j,l} \lambda_{jl} \delta_{jl} \sum_{kmn} \sqrt{p_{k|j} p_{m|l}} \langle n| A_k |\phi\rangle |k_M\rangle\langle m_M| \langle \phi| A_m^\dagger |n\rangle \end{aligned}$$

$$\begin{aligned} &= \sum_j \lambda_{jj} \sum_{km} \sqrt{p_{k|j} p_{m|j}} \langle \phi| A_m^\dagger A_k |\phi\rangle |k_M\rangle\langle m_M| \\ &= \sum_j \lambda_{jj} \left[\sum_k p_{k|j} \langle \phi| A_k^\dagger A_k |\phi\rangle |k_M\rangle\langle k_M| \right. \\ & \quad \left. + \sum_{k \neq m} \sqrt{p_{k|j} p_{m|j}} \langle \phi| A_m^\dagger A_k |\phi\rangle |k_M\rangle\langle m_M| \right] \end{aligned}$$

if $m=k$ then the unitaries give $A_m^\dagger A_k = \mathbb{I}$, therefore the diagonal elements of ρ_M undergo the process of a Markov chain. The off-diagonal elements do not necessarily vanish, but they do not affect the error operators acting on the transmitted states. Only the diagonal elements of the memory state affect the behavior of the channel.

-
- [1] T.M. Cover and J.A. Thomas, *Elements of Information Theory* (Wiley, New York, 1991).
 [2] S. Lloyd, Phys. Rev. A **55**, 1613 (1997).
 [3] B. Schumacher and M.D. Westmoreland, Phys. Rev. A **56**, 131 (1997).
 [4] A.S. Holevo, IEEE Trans. Inf. Theory **44**, 269 (1998).
 [5] C.H. Bennett, P.W. Shor, J.A. Smolin, and A.V. Thapliyal, IEEE Trans. Inf. Theory **48**, 2637 (2002).
 [6] C. Macchiavello and G.M. Palma, Phys. Rev. A **65**, 050301 (2002).
 [7] S. Bose, Phys. Rev. Lett. **91**, 207901 (2003).
 [8] K. Kraus, *States, Effects, and Operations: Fundamental No-*

- tions of Quantum Theory* (Springer-Verlag, Berlin, 1983).
 [9] B. Schumacher, Phys. Rev. A **54**, 2614 (1996).
 [10] A.S. Holevo, Probl. Peredachi Inf. **9**, 3 (1973).
 [11] B. Schumacher and M.A. Nielsen, Phys. Rev. A **54**, 2629 (1996).
 [12] C.H. Bennett, D.P. DiVincenzo, J.A. Smolin, and W.K. Wootters, Phys. Rev. A **54**, 3824 (1996).
 [13] H. Barnum, E. Knill, and M.A. Nielsen, IEEE Trans. Inf. Theory **46**, 1317 (2000).
 [14] C.H. Bennett, P.W. Shor, J.A. Smolin, and A.V. Thapliyal, Phys. Rev. Lett. **83**, 3081 (1999).