von Neumann capacity of noisy quantum channels

C. Adami[1,2,3] and N. J. Cerf[1,3]
[1]*W. K. Kellogg Radiation Laboratory, California Institute of Technology, Pasadena, California 91125*
[2]*Computation and Neural Systems, California Institute of Technology, Pasadena, California 91125*
[3]*Institute for Theoretical Physics, University of California, Santa Barbara, California 93106*

(Received 15 October 1996; revised manuscript received 2 June 1997)

We discuss the capacity of quantum channels for information transmission and storage. Quantum channels have dual uses: they can be used to transmit *known* quantum states which code for classical information, and they can be used in a purely quantum manner, for transmitting or storing quantum entanglement. We propose here a definition of the *von Neumann* capacity of quantum channels, which is a quantum-mechanical *extension* of the Shannon capacity and reverts to it in the classical limit. As such, the von Neumann capacity assumes the role of a classical or quantum capacity depending on the usage of the channel. In analogy to the classical construction, this capacity is defined as the maximum *von Neumann mutual entropy* processed by the channel, a measure which reduces to the capacity for classical information transmission through quantum channels (the ''Kholevo capacity'') when *known* quantum states are sent. The quantum mutual entropy fulfills all basic requirements for a measure of information, and observes quantum data-processing inequalities. We also derive a quantum Fano inequality relating the *quantum loss* of the channel to the fidelity of the quantum code. The quantities introduced are calculated explicitly for the quantum depolarizing channel. The von Neumann capacity is interpreted within the context of superdense coding, and an extended Hamming bound is derived that is consistent with that capacity. [S1050-2947(97)04511-3]

PACS number(s): 03.65.Bz, 89.70.+c

I. INTRODUCTION

The problem of transmission and storage of quantum states has received a considerable amount of attention recently, owing to the flurry of activity in the field of quantum computation [1] sparked by Shor's discovery of a quantum algorithm for factoring [2]. In anticipation of physical realizations of such computers (which still face major conceptual challenges), it is necessary to extend to the quantum regime the main results of Shannon's information theory [3], which provides limits on how well information can be compressed, transmitted, and preserved. In this spirit, the quantum analog of the noiseless coding theorem was obtained recently by Schumacher [4]. However, noisy quantum channels are less well understood, mainly because quantum noise is of a very different nature than classical noise, and the notion of ''quantum information'' is still under discussion. Yet, important results have been obtained concerning the correction of errors induced by the decoherence of quantum bits (qubits) via suitable quantum codes. These error-correcting codes [5–12] work on the principle that quantum information can be encoded in blocks of qubits (codewords) such that the decoherence of any qubit can be corrected by an appropriate code, much like the classical error-correcting codes. Therefore, it is expected that a generalization of Shannon's fundamental theorem to the quantum regime should exist, and efforts toward such a proof have appeared recently [13–15]. The capacity for the transmission of *classical* information through quantum channels was recently obtained by Hausladen *et al.* [16] for the transmission of pure states, and by Kholevo [17] for the general case of mixed states.

When discussing quantum channels, it is important to keep in mind that they can be used in two very different modes. On the one hand, one may be interested in the capacity of a channel to transmit or else store, an *unknown* quantum state in the presence of quantum noise. This mode is unlike any use of a channel we are accustomed to in classical theory as, strictly speaking, classical information is not transmitted in such a use (no measurement is involved). Rather, such a capacity appears to be a measure of how much *entanglement* can be transmitted (or maintained) in the presence of noise induced by the interaction of the quantum state with a ''depolarizing'' environment. On the other hand, a quantum channel can be used for the transmission of *known* quantum states (classical information), and the resulting capacity (i.e., the classical information transmission capacity of the quantum channel) represents the usual bound on the rate of arbitrarily accurate information transmission. In this paper, we propose a definition for the *von Neumann* capacity of a quantum channel, which encompasses the capacity for processing quantum as well as classical information. This definition is based on a quantum-mechanical extension of the usual Shannon mutual entropy to a von Neumann mutual entropy, which measures quantum as well as classical correlations. Still, a natural separation of the von Neumann capacity into classical and purely quantum pieces does not appear to be straightforward. This reflects the difficulty in separating classical correlation from quantum entanglement (the ''quantum separability'' problem, see, e.g., Ref. [18] and references therein). It may be that there is no unambiguous way to separate classical from purely quantum capacity for all channels and all noise models. The von Neumann capacity we propose, as it does not involve such a separation, conforms to a number of ''axioms'' for such a measure among which are positivity, subadditivity, concavity (convexity) in the input (output), and dual data-processing inequalities. We also show that the von Neumann capacity naturally reverts to the capacity for classical information transmission through noisy quantum channels of Kholevo [17] (the Kholevo capacity) if the unknown states are measured just before transmission, or, equivalently, if the quantum states are *prepared*. In such a use, thus, the purely quantum piece of the von Neumann capacity vanishes. We stop short of proving that the von Neumann capacity can be achieved by quantum coding, i.e., we do not prove the quantum equivalent of Shannon's noisy coding theorem for the total capacity. We do, however, provide an example where the von Neumann capacity appears achievable: the case of noisy superdense coding.

In Sec. II we recapitulate the treatment of the *classical* communication channel in a somewhat novel manner, by insisting on the deterministic nature of classical physics with respect to the treatment of information. This treatment paves the way for a formal discussion of quantum channels along the lines of Schumacher [13] in Sec. III, which results in a proposal for the definition of a von Neumann capacity for transmission of entanglement and/or correlation that parallels the classical construction. We also prove a number of properties of such a measure, such as subadditivity, concavity or convexity, forward and/or backward quantum data-processing inequalities, and derive a quantum Fano inequality relating the loss of entanglement in the channel to the fidelity of the code used to protect the quantum state. This proof uses an inequality of the Fano-type obtained recently by Schumacher [13]. In Sec. IV we demonstrate that the von Neumann capacity reduces to the recently obtained Kholevo capacity [17] if the quantum states are *known*, i.e., measured and kept in memory, before sending them on. In Sec. V we then apply these results directly to a specific example, the quantum depolarizing channel [19]. This generic example allows a direct calculation of all quantities involved. Specifically, we calculate the entanglement and/or correlation processed by the channel as a function of the entropy of the input and the probability of error of the channel. We also show that this capacity reverts to the well-known capacity for classical information transmission in a depolarizing channel if *known* quantum states are transmitted through the channel. In Sec. VI, finally, we interpret the von Neumann capacity in the context of superdense coding, and derive a quantum Hamming bound consistent with it.

II. CLASSICAL CHANNELS

The information theory of classical channels is well known since Shannon's seminal work on the matter [3]. In this section, rather than deriving any new results, we expose the information theory of classical channels in the light of the *physics* of information, in preparation of the quantum treatment of channels that follows. Physicists are used to classical laws of physics that are *deterministic*, and therefore do not consider noise to be an intrinsic property of channels. In other words, randomness, or a stochastic component, does not exist *per se*, but is a result of incomplete measurement. Thus, for a physicist there are no noisy channels, only incompletely monitored ones. As an example, consider an information transmission channel where the sender's information is the face of a coin before it is flipped, and the receiver's symbol is the face of the coin after it is flipped. Information theory would classify this as a useless channel, but for a physicist it is just a question of knowing the initial

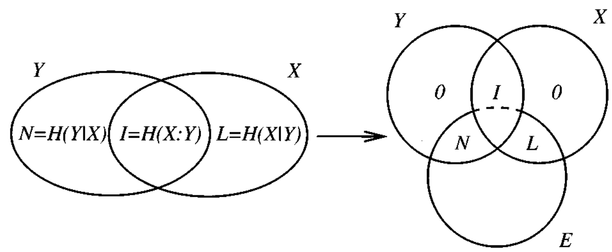

FIG. 1. Entropy Venn diagram for the classical channel XY, and its ''physical'' extension including the environment.

conditions of the channel *and the environment* well enough. From this, he can calculate the trajectory of the coin, and, by examining the face at the received side, infer the information sent by the sender. Classical physics, therefore, demands that all *conditional* probability distributions can be made to be *peaked*, if the environment, enlarged enough to cover all interacting systems, is monitored. In other words, $p_{i|j}=1$ or 0 for all i,j: if the outcome j is known, i can be inferred with certainty. As a consequence, *all* conditional entropies can be made to vanish for a closed system.

According to this principle, let us then construct the classical channel. Along with the ensemble of source symbols X (symbols $x_1,\ldots,x_N$ appearing with probabilities $p_1,\ldots,p_N$), imagine an ensemble of received symbols Y. The usual noisy channel is represented by the diagram on the left in Fig. 1: the conditional entropy $H(X|Y)$ represents the loss L in the channel, i.e., the uncertainty of inferrring X from Y, whereas $H(Y|X)$ stands for noise N in the output, which is unrelated to the error rate of the channel. A channel for which $L=0$ is called a *lossless* channel (no transmission errors occur), whereas $N=0$ characterizes a *deterministic* channel (the input unambiguously determines the output). On the right-hand side in Fig. 1, we extend the channel to include the environment. All conditional entropies are zero, and the noise and loss are simply due to correlations of the source or received ensembles with an environment, i.e., $L=H(X{:}E|Y)$ and $N=H(Y{:}E|X)$. The capacity of the classical channel is obtained by maximizing the mutual entropy between source and received symbols [the information $I=H(X{:}Y)$ processed by the channel] over all input distributions:

$$C=\max_{p(x)} I. \tag{2.1}$$

If the output of the channel Y is subjected to *another* channel (resulting in the output Z, say), it can be shown that the information processed by the combined channel, $H(X{:}Z)$, cannot possibly be larger than the information processed in the *first* leg, $H(X{:}Y)$. In other words, any subsequent processing of the output cannot possibly increase the transmitted information. This is expressed in the data-processing inequality (see, e.g., Ref. [20])

$$H(X{:}Z)\leqslant H(X{:}Y)\leqslant H(X). \tag{2.2}$$

By the same token, a *reverse* data-processing inequality can be proven, which implies that the information processed in the *second* leg of the channel, $H(Y{:}Z)$, must exceed the information processed by the total channel, $H(X{:}Z)$:

$$H(X{:}Z) \leq H(Y{:}Z) \leq H(Z). \tag{2.3}$$

This inequality together with Eq. (2.2) reflects microscopic time-reversal invariance: any channel used in a forward manner can be used in a backward manner.

As far as coding is concerned, the troublesome quantity is the loss L, while the noise N is unimportant. Indeed, for a message of length n, the typical number of input sequences for every output sequence is 2^{nL}, making decoding impossible. The principle of error correction is to embed the messages into *codewords*, that are chosen in such a way that the conditional entropy of the ensemble of codewords *vanishes*, i.e., on the level of message transmission the channel is lossless. Not surprisingly, there is then a relationship between the channel loss L and the probability of error p_c of a *code c* that is composed of s codewords:

$$L \leq H_2[p_c] + p_c \log_2(s-1), \tag{2.4}$$

where $H_2[p]$ is the dyadic Shannon entropy

$$H_2[p] = H_2[1-p] = -p\ \log_2\ p - (1-p)\log_2(1-p). \tag{2.5}$$

Equation (2.4) is the Fano inequality (see, e.g., Ref. [20]), which implies, for example, that the loss vanishes if the probability of error of the code vanishes. Note that the noise of the channel itself in general is not zero in this situation. Let us now turn to quantum channels.

III. QUANTUM CHANNELS

A. Information theory of entanglement

Quantum channels have properties fundamentally different from the classical channel just described, owing to the superposition principle of quantum mechanics and the noncloning theorem that ensues [21]. First and foremost, the input quantum state, after interaction with an environment, is lost, having become the output state. Any attempt at copying the quantum state before decoherence will result in a classical channel, as we will see later. Thus a joint probability for input and output symbols does not exist for quantum channels. However, this is not essential, as the quantity of interest in quantum communication is *not* the state of an isolated quantum system (a *product state*), but the degree of entanglement between one quantum system and another, parametrized by their mutual entropy as shown below. A single nonentangled pure quantum system (such as an isolated spin-$\frac{1}{2}$ state) carries no entropy, and is of no interest for quantum communication as it can be arbitrarily recreated at any time. Entangled composite systems (such as Bell states), on the other hand, are interesting because the entanglement can be used for communication. Let us very briefly recapitulate the quantum information theory of entanglement [22–25].

For a composite quantum system AB, we can write relations between von Neumann entropies that precisely parallel those written by Shannon for classical entropies. Specifically, we can define the conditional entropy of A (conditional on the knowledge of B),

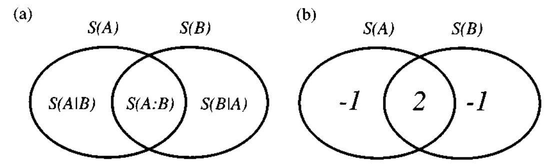

FIG. 2. (a) Entropy Venn diagram for a bipartite entangled quantum system AB, depicting $S(AB)$ (total area), marginal entropies [$S(A)$ or $S(B)$], conditional [$S(A|B)$ or $S(B|A)$] and mutual [$S(A{:}B)$] entropies. (b) Entropy diagram for a fully entangled Bell state.

$$S(A|B) = S(AB) - S(B), \tag{3.1}$$

via a suitable definition of a *conditional amplitude matrix* $\rho_{A|B}$. The latter matrix can have eigenvalues larger than unity, revealing its nonclassical nature and allowing conditional quantum entropies to be *negative* [22]. Similarly, we can define a *mutual amplitude matrix* $\rho_{A:B}$, giving rise to a mutual von Neumann entropy

$$S(A{:}B) = S(A) + S(B) - S(AB), \tag{3.2}$$

which exceeds the usual bound obtained for mutual Shannon entropies by a factor of two:

$$S(A{:}B) \leq 2\ \min[S(A), S(B)]. \tag{3.3}$$

The latter equation demonstrates that quantum systems can be more strongly correlated than classical ones: they can be *supercorrelated*. These relations can be conveniently summarized by entropy Venn diagrams [Fig. 2(a)], as is usual in classical information theory. The extension to the quantum regime implies that negative numbers can appear which are classically forbidden.[1] As an example, we show in Fig. 2(b) the quantum entropies of the Bell states (which are fully entangled states of two qubits). These notions can be extended to multipartite systems, and will be used throughout the paper.

The degree of entanglement of a bipartite pure quantum state is customarily indicated by the marginal entropy of one of its parts, i.e., the von Neumann entropy of the density matrix obtained by tracing the joint density matrix over the degrees of freedom of the other part (the entropy of entanglement, see Ref. [10]). However, since the parts of an entangled system do not possess a state on their own, it takes up to twice the marginal entropy of one of the parts to specify (in bits) the state of entanglement. For example, it takes up to two bits to specify the entanglement between two qubits (there are four Bell-basis states). Thus we propose to measure the entanglement of pure states by the mutual entropy between the two parts, which takes values between 0 (for nonentangled systems) and $2S$ (for entangled systems of marginal entropy S each). In order to avoid confusion with

[1] In classical entropy Venn diagrams, negative numbers can only appear in the mutual entropy of three or more systems.

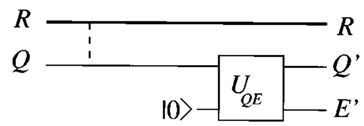

FIG. 3. Quantum network representation of a noisy quantum channel. R purifies the mixed state Q; the corresponding entanglement is indicated by a dashed line.

the previously defined entropy of entanglement, we propose to call this quantity the *mutual entanglement* (or simply the von Neumann mutual entropy), and denote it by the symbol I_Q:

$$I_Q = S(A{:}B). \tag{3.4}$$

For pure entangled states, the mutual entanglement I_Q is just twice the entropy of entanglement, demonstrating that either is a good measure for the *degree* of entanglement, but not necessarily for the absolute amount. Estimating the entanglement of *mixed* states, on the other hand, is more complicated, and no satisfying definition is available (see Ref. [10] for the most established ones). The quantum mutual entropy for mixed states does *not* represent pure quantum entanglement, but rather classical *and* quantum correlation that is difficult to separate consistently. For reasons that become more clear in the following, we believe that the mutual entanglement I_Q between two systems is the most straightforward generalization of the mutual information I of classical information theory, and will serve as the vehicle to define a quantum (including classical) *von Neumann* capacity for quantum channels.

B. Explicit model

In constructing a general quantum channel formally, we follow Schumacher [13]. A quantum mixed state Q suffers entanglement with an environment E so as to lead to a new mixed state Q' with possibly increased or decreased entropy. In order to monitor the entanglement transmission, the initial mixed state Q is *purified* by considering its entanglement with a *reference* system R,

$$|RQ\rangle = \sum_i \sqrt{p_i}\,|r_i, i\rangle, \tag{3.5}$$

where $|r_i\rangle$ are the R eigenstates. Indeed, this can always be achieved via a Schmidt decomposition. Then the mixed state Q is simply obtained as a partial trace of the pure state QR:

$$\rho_Q = \mathrm{Tr}_R[\rho_{QR}] = \sum_i p_i |i\rangle\langle i|. \tag{3.6}$$

Also, the interaction with the environment,

$$QRE \xrightarrow{U_{QE}\otimes 1_R} Q'R'E', \tag{3.7}$$

now can be viewed as a channel to transmit the entanglement between QR to the system $Q'R'$. Here, U_{QE} is the unitary operation entangling QR with the environment E, which is initially in a pure state. This construction is summarized in Fig. 3.

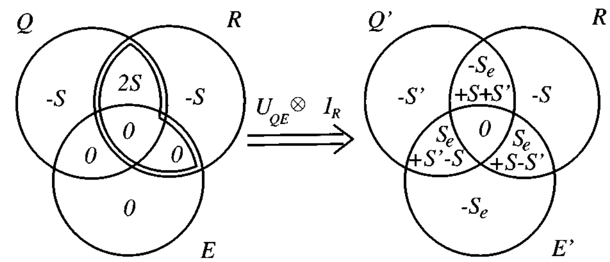

FIG. 4. Unitary transformation entangling the pure environment $|E\rangle$ with the pure system $|QR\rangle$. The reference system R is not touched by this transformation, which implies that no entropy can be exchanged across the double solid lines in the diagram on the left.

The evolution of entropies in such a channel is depicted in Fig. 4, where the entropy of the reference state [which is the same as the entropy of Q *before* entanglement, $S(Q) = S(R)$] is denoted by S,

$$S = -\sum_i p_i \log_2 p_i, \tag{3.8}$$

while the entropy of the quantum state Q' after entanglement $S(Q') = S'$, and the entropy of the environment $S(E) = S_e$. The latter was termed "exchange entropy" by Schumacher [13].

Note that, as for any tripartite pure state, the entropy diagram of the entangled state $Q'R'E'$ is uniquely fixed by three parameters, the marginal entropies of Q', R', and E' respectively, i.e., the numbers S, S', and S_e. Also, in any pure entangled diagram involving three systems, the ternary mutual entropy [the center of the ternary diagram, $S(Q'{:}R'{:}E')$] is always zero [23–25].

To make contact with the classical channel of Sec. II, let us define the *quantum loss* L_Q:[2]

$$L_Q = S(R'{:}E'|Q') = S_e + S - S'. \tag{3.9}$$

This represents the difference between the entropy acquired by the environment, S_e, and the entropy change of Q, that is, $(S' - S)$, and thus stands for the loss of entanglement in the quantum transmission. It plays a central role in error correction as shown below and in Sec. III D. The entropy diagram in terms of S, S_e, and L_Q is depicted in Fig. 5. From this diagram we can immediately read off inequalities relating the loss L_Q and the entropies S and S_e by considering triangle inequalities for quantum entropies [26], namely,

$$0 \le L_Q \le 2S, \tag{3.10}$$

[2]Here we follow the nomenclature that "quantum" always means "quantum including classical," rather than "purely quantum," in the same sense as the von Neumann entropy is not just a purely quantum entropy. This nomenclature is motivated by the difficulty to separate classical from quantum entanglement.

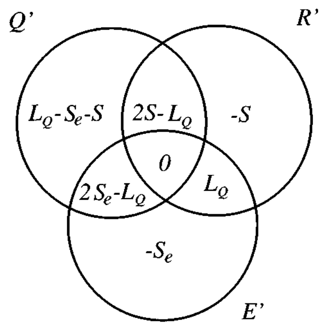

FIG. 5. Entropy diagram summarizing the entropy relations between the entangled systems Q', R', and E'.

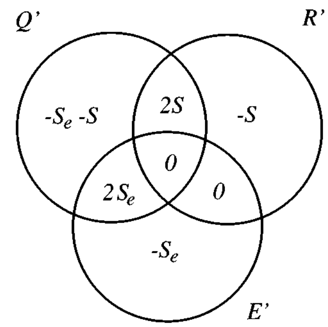

FIG. 6. Entanglement between Q', R', and E' in the lossless quantum channel.

$$0 \leqslant L_Q \leqslant 2S_e, \tag{3.11}$$

which can be combined to

$$0 \leqslant L_Q \leqslant 2\,\min(S,S_e). \tag{3.12}$$

We find therefore that the initial mutual entanglement $2S$ is split, through the action of the environment, into a piece shared with Q' [i.e., $S(Q'{:}R')=2S-L_Q$], and a piece shared with the environment (the remaining loss L_Q) according to the relation

$$S(R'{:}Q')+S(R'{:}E'|Q')=S(R'{:}E'Q')=S(R{:}Q), \tag{3.13}$$

or, equivalently,

$$I_Q+L_Q=2S. \tag{3.14}$$

Finally, we are ready to propose a definition for the von Neumann capacity. Again, in analogy with the classical construction, the von Neumann capacity C_Q would be the mutual entanglement processed by the channel (mutual von Neumann entropy), maximized over the density matrix of the input channel, i.e.,

$$C_Q=\max_{\rho_Q} I_Q, \tag{3.15}$$

where $I_Q=S(R'{:}Q')=S(R{:}Q')$ is the entanglement processed by the channel,

$$I_Q=2S-L_Q. \tag{3.16}$$

From the bound (3.10), we find that the entanglement processed by the channel is non-negative, and bounded from above by the initial entanglement $2S$. An interesting situation arises when the entanglement processed by the channel saturates this upper bound. This is the case of the *lossless* quantum channel, where $L_Q=0$.

It was shown recently by Schumacher and Nielsen [14] that an error-correction procedure meant to restore the initial quantum state (and thus the initial entanglement $2S$) can only be successful when $L_Q=0$. From Fig. 5 we can see that when $L_Q=0$, Q' is entangled *separately* with the reference state and the environment, leading to the diagram represented in Fig. 6. For this reason alone it is possible to recover the initial entanglement between Q and R via interaction with an ancilla A (that can be viewed as a *second* environment in a chained channel). The latter effects a transfer of the entanglement betwee Q' and E' to entanglement between E' and A. This operation can be viewed as an *incomplete* measurement of Q' by A which only measures the environment E', while keeping intact the entanglement of Q' with R. It was shown in [14] that $L_Q=0$ is in fact a necessary *and* sufficient condition for this to be feasible. Such a transfer of entanglement corresponds to the quantum equivalent of error correction, and will be discussed with reference to the quantum Fano inequality in Sec. III D.

C. Axioms for quantum information

In the following, we present a number of reasonable *axioms* for a quantum mutual information, and show that I_Q defined above has the required properties. These are: (i) non-negativity, (ii) concavity in ρ_Q (for a fixed channel), (iii) convexity in ρ_Q' (for fixed ρ_Q), and (iv) subadditivity. These requirements for a quantum mutual entropy (*entanglement and/or correlation processed by the channel*) are very natural, and reflect the kind of requirements that are put on classical channels. The non-negativity of I_Q is simply a consequence of the subadditivity of quantum entropies. (Just like the mutual Shannon entropy, the mutual quantum entropy is a non-negative quantity).

Concavity of quantum information in ρ_Q [axiom (ii)] reflects that the information processed by a channel with a mixture of quantum states $\rho_Q=\Sigma_i w_i\rho_Q^i$ (with $\Sigma_i w_i=1$) as input should be larger than the average information processed by channels that each have state ρ_Q^i as input, i.e.,

$$I_Q(\rho_Q)\geqslant\sum_i w_i I_Q(\rho_Q^i). \tag{3.17}$$

This is the quantum analog of the concavity of the Shannon mutual information $H(X{:}Y)$ in the input probability distribution $p(x)$ for a fixed channel, i.e., fixed $p(y|x)$. The proof uses the fact that, if the quantum operation achieved by the channel is fixed, we have

$$\begin{aligned}\rho'_{QE}&=U_{QE}\left(\sum_i w_i\rho^i\otimes|0\rangle\langle 0|\right)U^{\dagger}_{QE}\\&=\sum_i w_i U_{QE}(\rho^i\otimes|0\rangle\langle 0|)U^{\dagger}_{QE}\\&=\sum_i w_i\rho'^{i}_{QE}.\end{aligned}\tag{3.18}$$

Therefore, using

$$\begin{aligned}I_Q(\rho_Q)&=S(R{:}Q')=S(R)+S(Q')-S(RQ')\\&=S(Q'E')+S(Q')-S(E')\\&=S(Q'|E')+S(Q'),\end{aligned}\tag{3.19}$$

the concavity of the quantum information in the input results from the concavity of $S(Q'|E')$ in ρ'_{QE} and from the concavity of $S(Q')$ in ρ'_Q [27].

Convexity of the processed information in ρ'_Q [axiom (iii)] states that, if the superoperator that takes a fixed ρ_Q into ρ'_Q is such that

$$\rho'_Q=\sum_j w_j\rho'^{j}_{Q},\tag{3.20}$$

then

$$I_Q(\rho_Q\to\rho'_Q)\leqslant\sum_j w_j I_Q(\rho_Q\to\rho'^{j}_{Q}).\tag{3.21}$$

Thus the processed information of a channel that is a superposition of channels (each used with probability w_j) that result in ρ'_Q cannot exceed the average of the information for each channel. There is a similar property for classical channels: the mutual information $H(X{:}Y)$ is a convex function of $p(y|x)$ for a fixed input distribution $p(x)$. The proof follows from noting that, if the input is fixed, we have

$$\rho'_{RQ}=\sum_j w_j\rho'^{j}_{RQ}.\tag{3.22}$$

Then, expressing the quantum information as

$$I_Q(\rho_Q\to\rho'_Q)=S(R{:}Q')=S(R)-S(R|Q'),\tag{3.23}$$

and noting that $S(R)$ is constant, the concavity of $S(R|Q')$ in ρ'_{RQ} implies the convexity of the quantum information in the output.

Finally, the subadditivity of quantum information [axiom (iv)] is a condition which ensures that the information processed by a joint channel with input $\rho_{Q_1Q_2}$ is smaller than or equal to the information processed in parallel by two channels with input $\rho_{Q_1}=\mathrm{Tr}_{Q_2}(\rho_{Q_1Q_2})$ and $\rho_{Q_2}=\mathrm{Tr}_{Q_1}(\rho_{Q_1Q_2})$, respectively. Thus, if R is the reference system purifying the joint input Q_1Q_2, Q_1 is purified by RQ_2 while Q_2 is purified by RQ_1 (see Fig. 7).

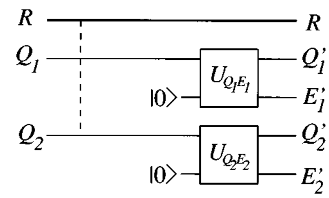

FIG. 7. Parallel channels as quantum network, in the derivation of the subadditivity of mutual von Neumann entropies. The entanglement between Q_1, Q_2, and the reference is indicated by a dashed line.

The subadditivity of von Neumann mutual entropies for such a channel can be written as

$$S(R{:}Q'_1Q'_2)\leqslant S(RQ_2{:}Q'_1)+S(RQ_1{:}Q'_2),\tag{3.24}$$

which can be read as

$$I_{12}\leqslant I_1+I_2,\tag{3.25}$$

with the corresponding identifications, and mirrors the classical inequality

$$H(X_1X_2{:}Y_1Y_2)\leqslant H(X_1{:}Y_1)+H(X_2{:}Y_2)\tag{3.26}$$

for two independent channels taking $X_1\to Y_1$ and $X_2\to Y_2$.

To prove inequality (3.24), we rewrite the quantum information of each channel using Eq. (3.19) and the fact that E_1 and E_2 are initially in a *product* state. Equation (3.24) then becomes

$$\begin{aligned}S(Q'_1Q'_2|E'_1E'_2)+S(Q'_1Q'_2)\leqslant{}&S(Q'_1|E'_1)+S(Q'_1)+S(Q'_2|E'_2)\\&+S(Q'_2).\end{aligned}\tag{3.27}$$

Subadditivity of *conditional* entropies, i.e.,

$$\begin{aligned}S(Q'_1Q'_2|E'_1E'_2)&=S(Q'_1|E'_1E'_2)+S(Q'_2|E'_1E'_2)\\&\quad-\underbrace{S(Q'_1{:}Q'_2|E'_1E'_2)}_{\geqslant 0}\\&\leqslant S(Q'_1|E'_1E'_2)+S(Q'_2|E'_1E'_2)\\&\leqslant S(Q'_1|E'_1)-\underbrace{S(Q'_1{:}E'_2|E'_1)}_{\geqslant 0}\\&\quad+S(Q'_2|E'_2)-\underbrace{S(Q'_2{:}E'_1|E'_2)}_{\geqslant 0}\\&\leqslant S(Q'_1|E'_1)+S(Q'_2|E'_2),\end{aligned}\tag{3.28}$$

together with the subadditivity property of ordinary (marginal) von Neumann entropies, proves Eq. (3.24). The terms that are ignored in the above inequality are positive due to strong subadditivity. This property of subadditivity of the information processed by quantum channels can be straightforwardly extended to n channels.

An alternative definition for the quantum information processed by a channel, called ''coherent information,'' was proposed by Schumacher and Nielsen [14] and Lloyd [15].

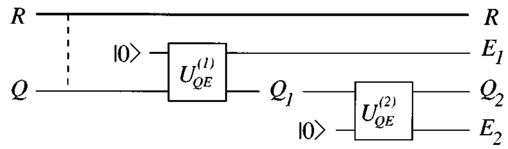

FIG. 8. Chaining of channels in the derivation of the data-processing inequality. The output Q_1 is subjected to a second channel by entangling with an environment E_2 independent from E_1, to give output Q_2.

This quantity $I_e = S(R'|E') = S - L_Q$ is not positive [axiom (i)], and violates axioms (ii) and (iv), which leads to a *violation* of the reverse data-processing inequality, while the forward one is respected [14] (as opposed to the von Neumann mutual entropy which observes both, see below). The coherent information attempts to capture the purely quantum piece of the processed information while separating out any classical components. This separation appears to be at the origin of the shortcomings mentioned above.

D. Inequalities for quantum channels

From the properties of the mutual entanglement I_Q derived above, we can prove data-processing inequalities for I_Q which reflect probability conservation, as well as the Fano inequality which relates the loss of a channel to the fidelity of a code.

1. Data processing

Assume that starting with the entangled state QR, entanglement with environment E_1 produces the mixed state Q_1. This output is used again as an input to another channel, this time entangling Q_1 with E_2 to obtain Q_2 (see Fig. 8).

The quantum analog of the (forward) data-processing inequality (2.2) that holds for mutual informations in classical channels involves the mutual entanglements $S(R{:}Q_1)$ and $S(R{:}Q_2)$, and asserts that the mutual entanglement between reference and output cannot be increased by any further processing:

$$S(R{:}Q_2) \leq S(R{:}Q_1) \leq 2S. \tag{3.29}$$

That such an inequality should hold is almost obvious from the definition of the mutual entanglement, but a short proof is given below. This proof essentially follows Ref. [14], and is based on the property of strong subadditivity applied to the system RE_1E_2:

$$S(R{:}E_2|E_1) = S(R{:}E_1E_2) - S(R{:}E_1) \geq 0. \tag{3.30}$$

For the channel $Q \to Q_1$, we see easily (see Fig. 5) that

$$S(R{:}E_1) = S(R{:}Q_1E_1) - S(R{:}Q_1|E_1) = 2S - S(R{:}Q_1). \tag{3.31}$$

Similarly, considering E_1E_2 as the environment for the overall channel $Q \to Q_2$, we find

$$S(R{:}E_1E_2) = 2S - S(R{:}Q_2). \tag{3.32}$$

Plugging Eqs. (3.31) and (3.32) into the positivity condition (3.30), we obtain the quantum data processing inequality, Eq. (3.29), as claimed.

The *reverse* quantum data-processing inequality implies that the entanglement processed by the second leg of the channel, $S(RE_1{:}Q_2)$, must be larger than the entanglement processed by the entire channel:

$$S(R{:}Q_2) \leq S(RE_1{:}Q_2) \leq S(RE_1E_2{:}Q_2) = 2S(Q_2). \tag{3.33}$$

The proof relies on strong subadditivity applied to $Q_2E_1E_2$:

$$S(Q_2{:}E_1|E_2) = S(Q_2{:}E_1E_2) - S(Q_2{:}E_2) \geq 0\,. \tag{3.34}$$

For treating the channel $Q_1 \to Q_2$ (i.e., the second leg), we have to purify the input state of Q_1, that is consider RE_1 as the reference. Thus we have

$$S(Q_2{:}RE_1) = 2S(Q_2) - S(Q_2{:}E_2). \tag{3.35}$$

For the overall channel $Q \to Q_2$, we have

$$S(Q_2{:}R) = 2S(Q_2) - S(Q_2{:}E_1E_2). \tag{3.36}$$

These two last equations, together with Eq. (3.34), result in the reverse quantum data-processing inequality Eq. (3.33).

From Eq. (3.29) we immediately obtain an inequality relating the loss of entanglement after the first stage L_1 (we drop the index Q that indicated the quantum nature of the loss in this discussion), with the overall loss L_{12}:

$$0 \leq L_1 \leq L_{12}. \tag{3.37}$$

Physically, this implies that the loss L_{12} cannot decrease from simply chaining channels, just as in the classical case. As emphasized earlier, the loss L_1 corresponds to the share of initial entanglement that is irretrievably lost to the environment. Indeed, if the environment cannot be accessed (which is implicit by calling it an environment) the decoherence induced by the channel cannot be reversed. Only if $L_1 = 0$ can this be achieved [14]. In view of this fact, it is natural to seek for a quantum equivalent to the classical Fano inequality (2.4).

2. Fano inequality

To investigate this issue, let us consider the chained channel above, where error correction has taken place via transfer of entanglement with a second environment. Let us also recall the definition of the *entanglement fidelity* of Schumacher [13], which is a measure of how faithfully the dynamics of the channel has preserved the initial entangled quantum state QR:

$$F_e(QR, Q'R) = \langle QR|\rho_{Q'R}|QR\rangle \equiv F_e^{QQ'}\,. \tag{3.38}$$

Since this entanglement fidelity does not depend on the reference system [13], we drop R from F_e from here on, as indicated in Eq. (3.38).

Naturally, the entanglement fidelity can be related to the probability of error of the channel. The quantum analog of the classical Fano inequality should relate the fidelity of the *code* (in our example above the fidelity between QR and

Q_2R, the error-corrected system) to the loss of the error-correcting channel L_{12}. The derivation of such an inequality is immediate using the Fano-type inequality derived by Schumacher [13], which relates the entropy of the environment of a channel $S(E')$ to the fidelity of entanglement,

$$S(E')\leq H_2[F_e^{QQ'}]+(1-F_e^{QQ'})\log_2(d_Qd_R-1), \tag{3.39}$$

where d_Q and d_R are the Hilbert-space dimensions of Q and R, respectively, and $H_2[F]$ is again the dyadic Shannon entropy. Let us apply this inequality to an error-correcting channel (decoherence plus error correction), i.e., the chained channel considered above. In that case, the environment is E_1E_2, and the entanglement fidelity is now between Q and Q_2, i.e., the fidelity of the *code*, and we obtain

$$S(E_1E_2)\leq H_2[F_e^{QQ_2}]+(1-F_e^{QQ_2})\log_2(d-1). \tag{3.40}$$

Here, $d=d_Rd_{Q_2}$ can be viewed as the Hilbert space dimension of the code (this is more apparent in superdense coding discussed in Sec. VI). To derive the required relationship, we simply note that

$$S(E_1E_2)\geq L_{12}/2 \tag{3.41}$$

[this is Eq. (3.11) applied to the composite channel]. This relates the fidelity of the code $F_e^{QQ_2}$ to the loss L_{12}, yielding the Fano inequality for a quantum code

$$L_{12}\leq 2\{H_2[F_e^{QQ_2}]+(1-F_e^{QQ_2})\ln(d-1)\}. \tag{3.42}$$

As we noted throughout the construction of quantum channels, a factor of 2 also appears in the quantum Fano inequality, commensurate with the fact that the loss can be twice the initial entropy. Inequality (3.42) puts an upper limit on the fidelity of a code for any nonvanishing loss L_{12}.

IV. CLASSICAL USE OF QUANTUM CHANNEL

In recent papers [16,17], the capacity for the transmission of *classical* information through quantum channels has been discussed. Essentially, this capacity is equal to the maximal accessible information χ in the system, known as the Kholevo bound [28].

What we show in the following is that the mutual entanglement introduced in Sec. III, i.e., the quantum mutual entropy $S(R:Q')$ between the decohered quantum state Q' and the reference state R, reduces to χ if the quantum state is measured before it is transmitted, or, equivalently, if Q is prepared by a classical *preparer* X. Let the system QR be purified again via a Schmidt decomposition as in Eq. (3.5). If we measure Q in its eigenbasis, we can write

$$|RXQ\rangle=\sum_i \sqrt{p_i}|r_ix_ii\rangle, \tag{4.1}$$

where x_i are the eigenstates of X (if X is in state x_i, Q is in state i, etc.). (Figure 9 summarizes the relationship between the respective entropies.) Naturally then, tracing over R, we obtain

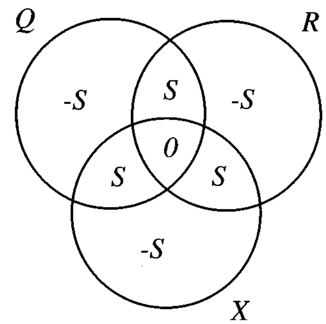

FIG. 9. Entanglement between Q, R, and the ancilla (or preparer) X after measurement of the initial state of Q by X, but prior to entanglement with the environment. The initial state of Q (before decoherence) is kept in memory, as it were, by X via classical correlation with Q.

$$\rho_{XQ}=\sum_i p_i|x_i\rangle\langle x_i|\otimes\rho_i\,, \tag{4.2}$$

with $\rho_i=|i\rangle\langle i|$, and similarly for ρ_{RQ}.

Thus, X and Q are *classically* correlated: each state of the preparer X represents a state of Q, or alternatively, X reflects (keeps in memory) the initial quantum state of Q. If the entropy of the quantum system Q before transmission is S (just like in the previous section), the mutual entropy between R and Q (as well as between X and Q) is also S, unlike the value $2S$ found in the quantum use. Decoherence now affects Q by entangling it with the environment, just like earlier. Thus

$$\rho_{XQ}\to\rho_{XQ'}=\sum_i p_i|x_i\rangle\langle x_i|\otimes\rho_i'\,, \tag{4.3}$$

where

$$\rho_i'=\mathrm{Tr}_E\{U_{QE}(\rho_i\otimes|0\rangle\langle 0|)U_{QE}^{\dagger}\}, \tag{4.4}$$

and we assumed again that the environment E is in a fixed "0" state before interacting with Q. Now our proof proceeds as before, only that the loss in the classical channel obeys different inequalities. The requirement that the entangling operation U_{QE} does not affect X or R now implies

$$S(X':E'Q')=S(X:Q)=S(R:Q)=S \tag{4.5}$$

(see Fig. 10).

Applying the chain rule to the left-hand side of Eq. (4.5) leads to

$$S(X':E'Q')=S(X:Q')+S(X:E'|Q'). \tag{4.6}$$

The quantum mutual entropy between the preparer and the quantum state after decoherence, $S(X:Q')$, can be shown to be equal to the Kholevo bound χ (see Ref. [29]). With $L=S(X:E'|Q')$ (the classical loss of the channel) we thus conclude from Eqs. (4.5) and (4.6) that

$$S=\chi+L. \tag{4.7}$$

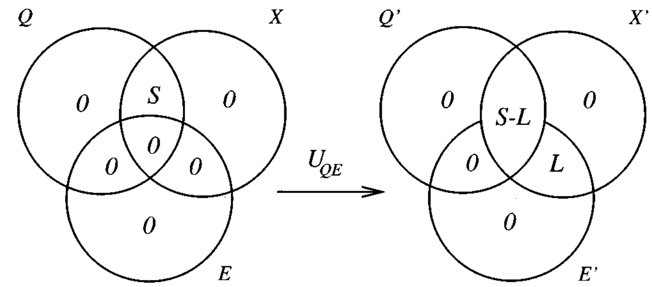

FIG. 10. Unitary transformation entangling the ''preparer'' (or alternatively, the classical ''memory'') X with the pure environment E and the quantum system Q. Neither the reference R nor the preparer X are affected by this operation. As the ternary Venn diagram between Q', E' and X' is not pure in this case, mutual entropy between Q' and X' *can* be shared by E'.

Note that $S(X{:}Q')$ is equal to $S(R{:}Q')$, the mutual entanglement I_Q introduced earlier, as $S(X)=S(R)$ and $S(XQ')=S(RQ')$. Thus

$$I_Q \equiv S(R{:}Q')=\chi, \tag{4.8}$$

if known quantum states are sent through the channel, as advertised. It was shown recently by Kholevo [17] that the maximum of the latter quantity indeed plays the role of channel capacity for classical information transmission,

$$C=\max_{p_i}\left[S(\rho')-\sum_i p_i S(\rho_i')\right]\equiv \max_{p_i}\chi, \tag{4.9}$$

where $\{p_i\}$ is a probability distribution of symbols at the source, and ρ_i' are the (not necessarily orthogonal) quantum states received at the output, with the probability distribution $\{p_i\}$ and $\rho'=\Sigma_i p_i \rho_i'$. Thus the quantity C_Q that we propose as a capacity for entanglement and/or correlation transmission reverts to the capacity for information transmission C if the unknown quantum states are *measured* before transmission. This represents solid evidence in favor of our interpretation. Let us now calculate the quantities introduced here for a specific simple model of quantum noise.

V. QUANTUM DEPOLARIZING CHANNEL

The quantum depolarizing channel is an idealization of a quantum storage and transmission process in which the stored quantum state can undergo bit-flip and phase errors. This is not the most general one-qubit channel,[3] but appears to be sufficient to examine a number of interesting aspects of quantum communication.

A. Quantum use

Imagine a quantum state

$$|\Psi\rangle=\alpha|0\rangle+\beta|1\rangle, \tag{5.1}$$

[3]A more general depolarizing channel could be constructed by allowing each of the possible errors a different probability.

where the basis states of the qubit can be taken to be spin-$\frac{1}{2}$ states polarized in the z direction, for example. (Specifically, we use the convention $\sigma_z|1\rangle=|1\rangle$.) The depolarizing channel is constructed in such a way that, due to an interaction with an environment, the quantum state survives with probability $1-p$, but is depolarized with probability $p/3$ by either a pure bit flip, a pure phase error, or a combination of both:

$$\begin{aligned}
|\Psi\rangle &\xrightarrow{1-p} |\Psi\rangle,\\
|\Psi\rangle &\xrightarrow{p/3} \sigma_x|\Psi\rangle=\alpha|1\rangle+\beta|0\rangle,\\
|\Psi\rangle &\xrightarrow{p/3} \sigma_z|\Psi\rangle=-\alpha|0\rangle+\beta|1\rangle,\\
|\Psi\rangle &\xrightarrow{p/3} \sigma_x\sigma_z|\Psi\rangle=-\alpha|1\rangle+\beta|0\rangle,
\end{aligned} \tag{5.2}$$

where the σ's are Pauli matrices. Such an arbitrary quantum state Ψ can, without loss of generality, considered to be a state Q that is entangled with a reference state R, such that the marginal density matrix of Q can be written as

$$\rho_Q=q|0\rangle\langle 0|+(1-q)|1\rangle\langle 1|, \tag{5.3}$$

with entropy $S(\rho_Q)=-\mathrm{Tr}\rho_Q \log_2 \rho_Q = H_2[q]$ and q a probability $(0\leqslant q\leqslant 1)$. In other words, the coefficients α and β need not be complex numbers. Conversely, we can start with such a mixed state at the input, and consider QR as a *pure* quantum state that this mixed state obtains from. For example,

$$|QR\rangle=\sqrt{1-q}|10\rangle-\sqrt{q}|01\rangle. \tag{5.4}$$

Naturally then, the mixed state Eq. (5.3) is obtained by simply tracing over this reference state. Pure states with real coefficients such as Eq. (5.4) are not general, but suffice for the depolarizing channel as R is always traced over.

Let us now construct a basis for QR that interpolates between completely independent and completely entangled states, and allows us to choose the initial entropy of Q with a single parameter q. We thus introduce the orthonormal *q-basis* states

$$\begin{aligned}
|\Phi^-(q)\rangle &= \sqrt{1-q}|00\rangle-\sqrt{q}|11\rangle,\\
|\Phi^+(q)\rangle &= \sqrt{q}|00\rangle+\sqrt{1-q}|11\rangle,\\
|\Psi^-(q)\rangle &= \sqrt{1-q}|10\rangle-\sqrt{q}|01\rangle,\\
|\Psi^+(q)\rangle &= \sqrt{q}|10\rangle+\sqrt{1-q}|01\rangle.
\end{aligned} \tag{5.5}$$

Note that, for $q=0$ or 1, these states are product states, while for $q=\frac{1}{2}$ they are completely entangled, and $\Psi^{\pm}(\frac{1}{2})$ and $\Phi^{\pm}(\frac{1}{2})$ are just the usual Bell basis states. The possibility of quantum decoherence of these states is introduced by entangling them with an environment in a pure state, taken to be of the same Hilbert space dimension as QR for simplicity, i.e., a four-dimensional space for the case at hand. This is the minimal realization of a depolarizing channel.

Let us assume that QR (for definiteness) is initially in the state $|\Psi^-(q)\rangle$, and the environment in a superposition

$$|E\rangle=\sqrt{1-p}|\Psi^-(q)\rangle+\sqrt{p/3}[|\Phi^-(q)\rangle+|\Phi^+(q)\rangle+|\Psi^+(q)\rangle]. \tag{5.6}$$

The environment and QR are then entangled by means of the unitary operator $U_{QRE}=U_{QE}\otimes 1_R$, with

$$U_{QE}=1\otimes P_{\Psi^-}(q)+\sigma_x\otimes P_{\Phi^-}(q)+(-i\sigma_y)\otimes P_{\Phi^+}(q) + \sigma_z\otimes P_{\Psi^+}(q), \tag{5.7}$$

where $P_\Phi(q)$ and $P_\Psi(q)$ stand for projectors projecting onto q-basis states. Note that the Pauli matrices act only on the first bit of the q-basis states, i.e., the entanglement operation only involves Q and E. Depending on the entanglement between Q and R, however, this operation also affects the entanglement between R and E. Thus we obtain the state

$$\begin{aligned}|Q'R'E'\rangle&=U_{QRE}|QR\rangle|E\rangle\\&=\sqrt{1-p}|\Psi^-_{QR}(q),\Psi^-_E(q)\rangle\\&\quad+\sqrt{p/3}[|\Phi^-_{QR}(q),\Phi^-_E(q)\rangle\\&\quad+|\Phi^+_{QR}(1-q),\Phi^+_E(q)\rangle+|\Psi^+_{QR}(1-q),\Psi^+_E(q)\rangle]\end{aligned} \tag{5.8}$$

on account of the relations

$$\sigma_x|\Psi^-_{QR}(q)\rangle=|\Phi^-_{QR}(q)\rangle, \tag{5.9}$$

$$(-i\sigma_y)|\Psi^-_{QR}(q)\rangle=|\Phi^+_{QR}(1-q)\rangle, \tag{5.10}$$

$$\sigma_z|\Psi^-_{QR}(q)\rangle=|\Psi^+_{QR}(1-q)\rangle, \tag{5.11}$$

and with obvious notation to distinguish the environment (E) and quantum system (QR) basis states. The (partially depolarized) density matrix for the quantum system is obtained by tracing over the environment:

$$\begin{aligned}\rho_{Q'R'}&=\mathrm{Tr}_E(|Q'R'E'\rangle\langle Q'R'E'|)\\&=(1-p)P_{\Psi^-}(q)\\&\quad+p/3[P_{\Phi^-}(q)+P_{\Phi^+}(1-q)+P_{\Psi^+}(1-q)].\end{aligned} \tag{5.12}$$

Its eigenvalues can be obtained to calculate the entropy:

$$S_e(p,q)\equiv S(Q'R')=H\left[\frac{2p}{3}(1-q),\frac{2pq}{3},\frac{1}{2}\left(1-\frac{2p}{3}+\Delta\right),\frac{1}{2}\left(1-\frac{2p}{3}-\Delta\right)\right], \tag{5.13}$$

with $H[p_1,\ldots,p4]$ the Shannon entropy, and

$$\Delta=[(1-2p/3)^2-16/3\,p(1-p)q(1-q)]^{1/2}. \tag{5.14}$$

By tracing over the reference state we obtain the density matrix of the quantum system after the interaction $\rho_{Q'}$, and its respective entropy

$$S'(p,q)\equiv S(Q')=H_2\left[q+\frac{2p}{3}(1-2q)\right]. \tag{5.15}$$

Together with the entropy of the reference state (which is unchanged since R was not touched by the interaction), $S(R')=S(R)=H_2[q]$; this is enough to fill in the ternary entropy diagram reflecting the dynamics of the channel, Fig. 5. We thus find the mutual entanglement processed by the channel,

$$I_Q=S(Q':R)=2H_2[q]-L_Q(p,q), \tag{5.16}$$

where the loss is

$$L_Q(p,q)=H_2[q]-H_2\left[q+\frac{2p}{3}(1-2q)\right]+S_e(p,q). \tag{5.17}$$

The mutual entanglement is plotted in Fig. 11, as a function of the error probability p of the channel and of the parameter q which determines the initial entropy.

The mutual entanglement is maximal when the entropy of the source is maximal (as in the classical theory), i.e., $q=\frac{1}{2}$. Then

$$C_Q=\max_q I_Q=2-S_e(p,\tfrac{1}{2})=2-H_2[p]-p\log_2 3. \tag{5.18}$$

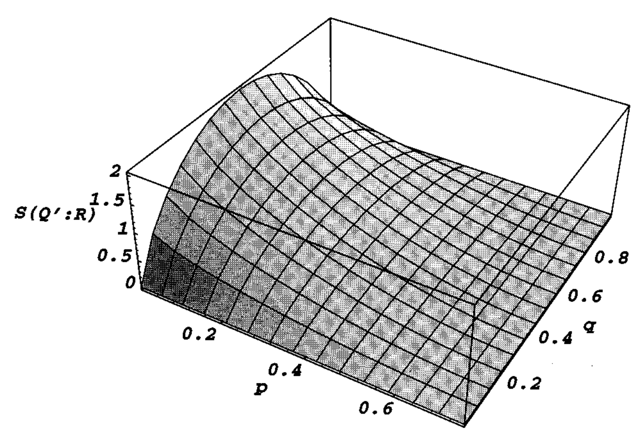

FIG. 11. Mutual entanglement between the depolarized state Q' and the reference system $R'=R$, as a function of error p and parameter q. Note that the channel is 100% depolarizing at $p=\frac{3}{4}$. The concavity in q [according to axiom (ii)] as well as the convexity in p [axiom (iii)] are apparent.

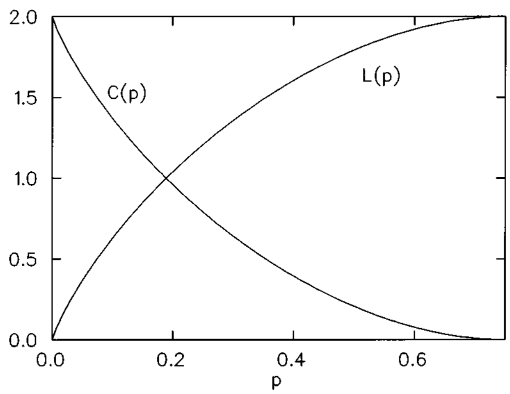

FIG. 12. Maximal entanglement transfer $C(p)$ and minimal loss $L(p)$ as a function of the error probability p.

In that case, the maximal rate of entanglement transfer is two bits (error-free transfer, $p=0$). The capacity only vanishes at $p=\frac{3}{4}$, i.e., the 100% depolarizing channel. This is analogous to the vanishing of the classical capacity of the binary symmetric channel at $p=\frac{1}{2}$. As an example of such a channel, we shall discuss the transmission of the entanglement present in a Bell state (one out of four fully entangled qubit pairs) through a superdense coding channel in Sec. VI A. The maximal mutual entanglement and minimal loss implied by Eq. (5.18) are plotted in Fig. 12 as a function of p. This error rate p can be related to the fidelity of the channel by

$$F_e^{Q'Q}=1-p+\frac{p}{3}(1-2q)^2, \tag{5.19}$$

where $F_e^{Q'Q}$ is Schumacher's fidelity of entanglement introduced earlier. Note that this implies that the Fano inequality Eq. (3.39) is saturated at $q=\frac{1}{2}$ for any p.

B. Classical use

Now, instead of using the channel to transmit entanglement (sending unknown quantum states), one could equally well use it to send classical information (known quantum states) as outlined in Sec. IV. Here we calculate the capacity for the transmission of classical information through the quantum depolarizing channel, and verify that the result is equal to the value obtained by Calderbank and Shor [6] using the Kholevo theorem.

Before entanglement with the environment, let us measure the mixed state Q via an ancilla X, after which Q and X are classically correlated, with mutual entropy $H_2[q]$. Note that this operation leads to an entangled triplet QRX at the outset, as in Fig. 9, with $S=H_2[q]$. We now proceed with the calculation as before. The basis states for the system $|QXR\rangle$ are then simply

$$\begin{aligned}
|\Phi_X^-(q)\rangle &= \sqrt{1-q}|000\rangle-\sqrt{q}|111\rangle,\\
|\Phi_X^+(q)\rangle &= \sqrt{q}|000\rangle+\sqrt{1-q}|111\rangle,\\
|\Psi_X^-(q)\rangle &= \sqrt{1-q}|110\rangle-\sqrt{q}|001\rangle,\\
|\Psi_X^+(q)\rangle &= \sqrt{q}|110\rangle+\sqrt{1-q}|001\rangle,
\end{aligned} \tag{5.20}$$

where we used the index X on the basis states to distinguish them from the two-qubit basis states introduced earlier. The entanglement operation is as before, with a unitary operator acting on Q and E only. Because of the additional trace over the ancilla X, however, we now find for the density matrix $\rho_{Q'R'}$:

$$\begin{aligned}
\rho_{Q'R'} =& (1-2p/3)[(1-q)|10\rangle\langle 10|+q|01\rangle\langle 01|]\\
&+2p/3[(1-q)|00\rangle\langle 00|+q|11\rangle\langle 11|].
\end{aligned} \tag{5.21}$$

Consequently, for the mutual information transmitted through the channel, we find

$$I=S(Q':R)=H_2[q]-L(p,q), \tag{5.22}$$

with the (classical) loss of information

$$\begin{aligned}
L(p,q)=&H\left[\frac{2p}{3}(1-q),\frac{2p}{3}q,\left(1-\frac{2p}{3}\right)(1-q),\left(1-\frac{2p}{3}\right)q\right]\\
&-H_2\left[q+\frac{2p}{3}(1-2q)\right].
\end{aligned} \tag{5.23}$$

Maximizing over the input distribution as before, we obtain

$$C=\max_q S(Q':R)=1-H_2[2p/3], \tag{5.24}$$

the result derived recently for the depolarizing channel simply from using the Kholevo theorem [6]. Note that Eq. (5.24) is just the Shannon capacity of a binary symmetric channel [20], with a bit-flip probability of $2p/3$ (of the three quantum error syndromes, only two are classically detectable as bit flips).

VI. INTERPRETATION

A. Quantum capacity and superdense coding

The interpretation of the capacity suggested here as a quantum-mechanical extension of the classical construction can be illustrated in an intuitive manner with the example of the depolarizing channel introduced above. The idea is that I_Q reflects the capacity for transmission of quantum mutual entropy (entanglement and/or classical information) but that the amount transferred in a particular channel depends on how this channel is used. A particularly elegant channel that uses I_Q to its full extent is the noisy superdense coding channel. There, the entanglement between sender and receiver is used to transmit two bits of classical information by sending just *one* quantum bit [30,22]. In a general superdense coding scheme, the initial state QR is one of a set of entangled states specified by classical bits C. This situation can be related to our previous discussion by noting that all entropies appearing there are to be understood as *conditional* on the classical bits C that are to be sent through the channel as shown in Fig. 13. The von Neumann capacity introduced above is then just

$$I_Q=S(R:Q'|C). \tag{6.1}$$

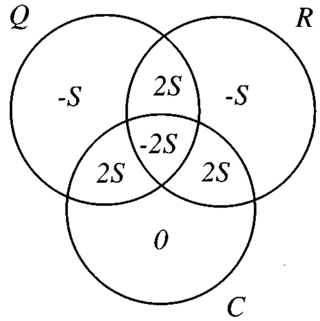

FIG. 13. Quantum Venn diagram for the noisy superdense coding channel before decoherence. Conditionally on the classical bits C,QR is in a pure entangled state described by a Venn diagram of the form $(-S,2S,-S)$. Note that no information about C is contained in R or Q *alone*, i.e., $S(C{:}R)=S(C{:}Q)=0$.

It is not immediately obvious that this von Neumann capacity is equal to the *classical* capacity between preparer (usually termed Alice) and the receiver (Bob). However, it is not difficult to prove [using the fact that $S(R{:}Q)=S(R{:}C)=S(Q{:}C)=0$] that Eq. (6.1) is in fact equal to the maximal amount of classical information about C extractable from RQ' (after Q decohered), which is[4]

$$\chi=S(RQ'{:}C). \tag{6.2}$$

Thus, in this example the amount of entanglement processed in a channel can be viewed as the amount of *classical* information about the preparer of the entangled state QR. This amount of information can reach *twice* the entropy of Q (2 bits in standard superdense coding), which is classically impossible. (The superdense coding and teleportation channels will be discussed in detail elsewhere).

Having established this relation between superdense coding and the general quantum channels treated here, let us imagine that the qubit that is sent through the channel (and which is ''loaded'' with entanglement) is subject to the depolarizing noise of Sec. V. Indeed, if $p=0$ the two classical bits can be decoded perfectly, achieving the value of the capacity. It has been argued recently [22] that this can be understood by realizing that besides the qubit that is sent forwards in time in the channel, the entanglement between sender and receiver can be viewed as an antiqubit sent *backwards* in time (which is equivalent to a qubit sent forwards in time if the appropriate operations are performed on it in the future). Thus, the quantum mechanics of superdense coding allows for the time-delayed (error-free) transmission of information, which shows up as excessive capacity of the respective channel. On the other hand, it is known that (for unencoded qubits) superdense coding becomes impossible if $p\approx 0.189$, which happens to be the precise point at which $I_Q=1$. This is related to the fact that at this point the purification of noisy pairs becomes impossible. However, the capacity of this channel is not zero. While no information can be retrieved from the past in this case, the single qubit that is sent through the channel still carries information, indeed, it shares one bit of mutual entropy with the qubit stored by the receiver. Clearly, this is still a quantum channel: if it were classical, the transmission of one bit could not take place with unit rate and perfect reliability, due to the noise level $p=0.189$. As the receiver possesses both this particle (Q') and the one that was shared earlier (R), he can perform joint measurements (in the space $Q'R$) to retrieve at least one of the two classical bits.

An extreme example is the *dephasing* channel, which is a depolarizing channel with only σ_z-type errors that affect the phase of the qubit. As is well known, classical bits are unaffected by this type of noise, while quantum superpositions are dephased. The channel becomes useless (for the storage of superpositions) at $p=0.5$, yet measuring the qubit yields one *classical* bit in an error-free manner. A calculation of $\max_q S(R{:}Q')$ for this channel indeed yields

$$I_Q(p)=2-H_2[p]. \tag{6.3}$$

In this limiting case thus, it appears possible to separate the classical $(I=1)$ from the purely quantum capacity. However, it might well be possible that this cannot be achieved in general. Below, we show that such an excessive von Neumann capacity (as in superdense coding) is consistent with a commensurate quantum Hamming bound.

B. Quantum Hamming bounds

Classically, the Hamming bound [20] is an upper bound on the number s of codewords (bit strings of length n) for a code to correct t errors:

$$s\sum_{i=0}^{t}\binom{n}{i}\leqslant 2^n. \tag{6.4}$$

This is a necessary (but not sufficient) condition for error-free coding, which reflects the necessary space to accommodate all the codewords and associated descendants for all error syndromes. For s codewords coding for k bits $(s=2^k)$, we can consider the asymptotics of Eq. (6.4) in the limit of infinitely long messages $(n\to\infty)$, and find that the rate of error-free transmission is limited by

$$R\leqslant -\frac{1}{n}\log_2\sum_{i=0}^{pn}\binom{n}{i}\left(\frac{1}{2}\right)^i\left(\frac{1}{2}\right)^{n-i}, \tag{6.5}$$

where $R=k/n$ is the transmission rate and $p=t/n$ is the asymptotic probability of error. Using

$$\begin{aligned}\lim_{n\to\infty}-\frac{1}{n}\log_2&\left\{\sum_{i=0}^{pn}\binom{n}{i}r^i(1-r)^{n-i}\right\}\\&=p\log_2\frac{p}{r}+(1-p)\log_2\frac{1-p}{1-r}\\&\equiv H(p,1-p\|r,1-r),\end{aligned} \tag{6.6}$$

where $H(p,1-p\|r,1-r)$ is the *relative* entropy between the probability distributions p and r, we can write

[4]That the quantum mutual entropy between a preparer and a quantum system is an upper bound to the amount of classical information obtainable by measuring the quantum system (the Kholevo bound) is shown in Ref. [29].

$$R \leqslant H(p,1-p\|1/2,1/2) = 1 - H_2(p). \tag{6.7}$$

The relative entropy thus turns out to be just the classical capacity of the binary symmetric channel, and measures the *distance* of the error-probability of the channel relative to the ''worst case,'' i.e., $p=\frac{1}{2}$. Note that relative entropies are positive semidefinite.

For quantum channels, the standard quantum Hamming bound for nondegenerate (orthogonal) codes is written as [8–10]

$$2^k \sum_{i=0}^{t} 3^i \binom{n}{i} \leqslant 2^n, \tag{6.8}$$

which expresses that the number of orthogonal states identifying the error syndromes on the 2^k different messages must be smaller than 2^n, the dimension of the Hilbert space of the quantum state Q (n qubits). In the limit of large n, this translates into an upper bound for the rate of nondegenerate quantum codes

$$R \leqslant -\frac{1}{n}\log_2\left\{\sum_{i=0}^{pn}\binom{n}{i}\left(\frac{3}{4}\right)^i\left(\frac{1}{4}\right)^{n-i}\right\} - 1. \tag{6.9}$$

which can (as in the classical case) be written in terms of a relative entropy

$$\begin{aligned} R &\leqslant H(p,1-p\|3/4,1/4) - 1 = 1 - S_e(p) \\ &= 1 - H_2[p] - p\log_2 3. \end{aligned} \tag{6.10}$$

Thus the usual quantum Hamming bound limits the rate of nondegenerate quantum codes by the capacity based on coherent information proposed in Refs. [14,15], which is thought of as the purely quantum piece of the capacity. Note that the positivity of the relative relative entropy does *not* in this case guarantee such a capacity to be positive, which may just be a reflection of the inseparability of the von Neumann capacity.

The quantum Hamming bound shown above relies on coding the error syndromes only into the quantum state Q that is processed, or, in the case of superdense coding, sent through the noisy channel. As we noted earlier, however, a quantum system that is entangled does not, as a matter of principle, have a state on its own. Thus the entangled reference system R *necessarily* becomes part of the quantum system, even if it is not subject to decoherence. Thus, the Hilbert space available for coding automatically becomes as large as $2n$, the combined Hilbert space of Q and R. This is most obvious again in superdense coding, where the decoding of the information explicitly involves joint measurements of the decohered Q' *and* the reference R, shared between sender and receiver (in a noise-free manner). The corresponding *extended* quantum Hamming bound therefore can be written by remarking that while the coding space is $2n$, only n qubits are sent through the channel, and thus

$$2^k \sum_{i=0}^{t} 3^i \binom{n}{i} \leqslant 2^{2n}. \tag{6.11}$$

Proceeding as before, the rate of such quantum codes is limited by

$$R \leqslant H(p,1-p\|3/4,1/4) = 2 - S_e(p), \tag{6.12}$$

the von Neumann capacity C_Q for the depolarizing channel proposed in this paper, Eqs. (3.15) and (5.18). The latter is always positive, and represents the distance between the error probability p of the channel and the worst-case error $p=\frac{3}{4}$ (corresponding to a 100% depolarizing channel), in perfect analogy with the classical construction.

VII. CONCLUSIONS

We have shown that the classical concept of information transmission capacity can be extended to the quantum regime by defining a von Neumann capacity as the maximum mutual von Neumann entropy between the decohered quantum system and its reference. This mutual von Neumann entropy, that describes the amount of information—classical and/or quantum—processed by the channel, obeys axioms that any measure of information should conform to. As for any quantum extension, the von Neumann capacity reverts to its classical counterpart when the information is classicized (i.e., it reverts to the Kholevo capacity when measured or prepared states are sent), and ultimately to the Shannon capacity if all quantum aspects of the channel are ignored (i.e., if orthogonal states are sent and measured). Thus the von Neumann capacity of a channel can only vanish when the classical capacity is also zero, but it can be excessive as entanglement allows for superdense coding. In order to take advantage of this, however, both the quantum system that decoheres *and* the reference system it is entangled with need to be accessible. In practical quantum channels this appears to be impossible, and the rate of practical codes must then be considerably smaller than the von Neumann capacity. Yet, because of the inseparability of entangled states, a consistent definition of channel capacity *has* to take into account the full Hilbert space of the state. Whether a capacity can be defined *consistently* that characterizes the purely quantum component of a channel is still an open question.

ACKNOWLEDGMENTS

We would like to thank John Preskill for discussions on the depolarizing channel, as well as Howard Barnum and Michael Nielsen for discussions during the Quantum Computation and Quantum Coherence Program at the ITP in Santa Barbara, where most of this work was done. This research was supported in part by NSF Grant Nos. PHY 94-12818 and PHY 94-20470 at the Kellogg Radiation Laboratory, and Grant No. PHY 94-07194 at the ITP in Santa Barbara, as well as by a grant from DARPA/ARO through the QUIC Program (Grant No. DAAH04-96-1-3086).

[1] For reviews, see D. P. DiVincenzo, Science **270**, 255 (1995); S. Lloyd, Sci. Am. (Int. Ed.) **273** (4), 140 (1995); A. Ekert and R. Josza, Rev. Mod. Phys. **68**, 733 (1996).

[2] P. W. Shor, in *Proceedings of the 35th Annual Symposium on Foundations of Computer Science*, edited by S. Goldwasser (IEEE Computer Society, New York, 1994), pp. 124–134.

[3] C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication* (University of Illinois Press, Chicago, 1949).

[4] B. Schumacher, Phys. Rev. A **51**, 2738 (1995); B. Schumacher and R. Josza, J. Mod. Opt. **41**, 2343 (1994).

[5] P. W. Shor, Phys. Rev. A **52**, 2493 (1995).

[6] A. R. Calderbank and P. W. Shor, Phys. Rev. A **54**, 1098 (1996).

[7] A. M. Steane, Proc. R. Soc. London Ser. A **452**, 2551 (1996); Phys. Rev. Lett. **77**, 793 (1996).

[8] R. Laflamme, C. Miquel, J. P. Paz, and W. H. Zurek, Phys. Rev. Lett. **77**, 198 (1996).

[9] A. Ekert and C. Macchiavello, Phys. Rev. Lett. **77**, 2585 (1996).

[10] C. H. Bennett, D. DiVincenzo, J. A. Smolin, and W. K. Wootters, Phys. Rev. A **54**, 3824 (1996).

[11] E. Knill and R. Laflamme, Phys. Rev. A **55**, 900 (1997).

[12] A. R. Calderbank, E. M. Rains, P. W. Shor, and N. J. A. Sloane, IEEE Trans. Inf. Theory (to be published).

[13] B. Schumacher, Phys. Rev. A **54**, 2614 (1996).

[14] B. Schumacher and M. A. Nielsen, Phys. Rev. A **54**, 2629 (1996);H. Barnum, M. A. Nielsen, and B. Schumacher, e-print quant-ph/9702049.

[15] S. Lloyd, Phys. Rev. A **55**, 1613 (1997).

[16] P. Hausladen, R. Josza, B. Schumacher, M. Westmoreland, and W. K. Wootters, Phys. Rev. A **54**, 1869 (1996).

[17] A. S. Kholevo, e-print quant-ph/9611023.

[18] R. Horodecki and M. Horodecki, Phys. Rev. A **54**, 1838 (1996).

[19] C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin, and W. K. Wootters, Phys. Rev. Lett. **76**, 722 (1996).

[20] R. B. Ash, *Information Theory* (Dover, New York, 1965).

[21] W. K. Wootters and W. H. Zurek, Nature (London) **299**, 802 (1982); D. Dieks, Phys. Lett. A **92**, 271 (1982).

[22] N. J. Cerf and C. Adami, e-print quant-ph/9512022; and in *Proceedings of the Second International Symposium on Fundamental Problems in Quantum Physics*, edited by M. Ferrero and A. van der Merwe (Kluwer, Dordrecht, 1997).

[23] N. J. Cerf and C. Adami, in *Proceedings of the 4th Workshop on Physics and Computation, Boston University, 22-24 Nov. 1996*, edited by T. Toffoli, M. Biafore, and J. Leao (New England Complex Systems Institute, Cambridge, MA, 1996), pp. 65–71.

[24] N. J. Cerf and C. Adami, e-print quant-ph/9605002.

[25] C. Adami and N. J. Cerf (unpublished).

[26] H. Araki and E. H. Lieb, Commun. Math. Phys. **18**, 160 (1970).

[27] A. Wehrl, Rev. Mod. Phys. **50**, 221 (1978).

[28] A. S. Kholevo, Probl. Inform. Transmission **9**, 110 (1973).

[29] N. J. Cerf and C. Adami, e-print quant-ph/9611032.

[30] C. H. Bennett and S. J. Wiesner, Phys. Rev. Lett. **69**, 2881 (1992).