Information, Relative Entropy of Entanglement, and Irreversibility

L. Henderson and V. Vedral
Centre for Quantum Computation, Clarendon Laboratory, University of Oxford, Parks Road, Oxford OX1 3PU, United Kingdom
(Received 9 September 1999)

Previously proposed measures of entanglement, such as entanglement of formation and assistance, are shown to be special cases of the relative entropy of entanglement. The difference between these measures for an ensemble of mixed states is shown to depend on the availability of classical information about particular members of the ensemble. Based on this, relations between relative entropy of entanglement and mutual information are derived.

PACS numbers: 03.67.–a, 03.65.Bz

In quantifying entanglement, a number of measures have been proposed. For bipartite pure states, ρ_{AB}, the Von Neumann entropy of the reduced density matrix of either subsystem, ρ_A or ρ_B, has been found to be a good and unique measure [1,2]. Relative entropy of entanglement has been proposed as a measure which extends to mixed states [3,4]. Loosely speaking, it quantifies how "far" an entangled state is from the set of disentangled states. Entanglement of mixed states has also been characterized by the "entanglement of formation," and by the "entanglement of distillation" [5]. Rather surprisingly, use of entanglement in mixed states is not reversible in the sense that all the entanglement required to construct a particular mixed state cannot be distilled out again, so the entanglement of formation is greater than the entanglement of distillation [4]. In this Letter, we clarify the role of classical information about the identity of particular members of an ensemble of mixed states, and show that the loss of such information is responsible for the difference between the entanglement of formation and the entanglement of distillation. We provide a unifying framework for entanglement measures by showing how previously proposed measures are special cases of the relative entropy of entanglement. This gives a strong physical argument for using quantum relative entropy as a unique way to understand entanglement in general.

Suppose that Alice and Bob share a state described by the density matrix ρ_{AB}. The state ρ_{AB} has an infinite number of different decompositions $\varepsilon = \{|\psi^i_{AB}\rangle\langle\psi^i_{AB}|, p_i\}$, into pure states $|\psi^i_{AB}\rangle$, with probabilities p_i [6]. We denote the mixed state ρ_{AB} written in decomposition ε by

$$\rho^{\varepsilon}_{AB} = \sum_i p_i |\psi^i_{AB}\rangle\langle\psi^i_{AB}|. \tag{1}$$

Measures of entanglement are associated with formation and distillation of pure and mixed entangled states [1,5]. The basis of formation is that Alice and Bob would like to create an ensemble of n copies of the nonmaximally entangled state, ρ_{AB}, using only local operations, classical communication, and a number, m, of maximally entangled pairs. It is customary to assume that the only "cost" in communication is due to the use of entanglement resources, or sending information down a quantum channel, while classical communication costs nothing. Entanglement of formation is the asymptotic conversion ratio, $\frac{m}{n}$, in the limit of infinitely many copies. It is given by the average entanglement of the pure states, minimized over all decompositions, $E_F(\rho_{AB}) = \min_{\varepsilon} \sum_i p_i S(\rho^i_B)$, where ρ^i_B is the reduced density matrix for subsystem A of the pure state $|\psi^i_{AB}\rangle\langle\psi^i_{AB}|$. Distillation is the reverse process, where Alice and Bob share an ensemble of n copies of the nonmaximally entangled state, ρ_{AB}, and would like to extract the largest number of maximally entangled pairs using only local operations and classical communications. The entanglement of distillation, $E_D(\rho_{AB})$, is the number of maximally entangled singlets per copy of ρ_{AB} which can be distilled from an asymptotically large ensemble of copies.

Relative entropy of entanglement of the mixed state is defined as $E_{RE}(\rho_{AB}) = \min_{\sigma_{AB}\in D} S(\rho_{AB}\|\sigma_{AB})$, where $S(\rho\|\sigma) = Tr(\rho \log\rho - \rho \log\sigma)$ is the quantum relative entropy [4]. The minimum is taken over D, the set of completely disentangled or "separable" states. A state is separable if it can be written as a convex combination of product states $\sigma = \sum_i p_i \sigma^i_A \otimes \sigma^i_B$, with $\sum_i p_i = 1$. The relative entropy of entanglement provides an upper bound for the distillable entanglement [4]. The known relationships between the different measures of entanglement for mixed states are $E_D(\rho_{AB}) \leq E_{RE}(\rho_{AB}) \leq E_F(\rho_{AB})$ [4]. Equality holds for pure states, where all the measures reduce to the Von Neumann entropy, $S(\rho_A) = S(\rho_B)$.

Formation of an ensemble of n nonmaximally entangled *pure* states, $\rho_{AB} = |\psi_{AB}\rangle\langle\psi_{AB}|$, is achieved by the following protocol. Alice first prepares the states she would like to share with Bob locally. She then uses Schumacher compression [7] to compress subsystem B into $nS(\rho_B)$ states. Subsystem B is then teleported to Bob using $nS(\rho_B)$ maximally entangled pairs. Bob decompresses the states he receives and so ends up sharing n copies of ρ_{AB} with Alice. The entanglement of formation is therefore $E_F(\rho_{AB}) = S(\rho_B)$. For pure states, this process requires no classical communication in the asymptotic limit [8]. The reverse process of distillation is accomplished using the Schmidt projection method [1], which allows $nS(\rho_B)$ maximally entangled pairs to be distilled in the limit as n becomes very large. No classical communication between the separated parties is required. Therefore pure states are fully interconvertible in the asymptotic limit.

The situation for mixed states is more complex. When any mixed state, denoted by Eq. (1), is created, it may be imagined to be part of an extended system whose state is pure. The pure states $|\psi_{AB}^i\rangle$ in the mixture may be regarded as correlated to orthogonal states $|m_i\rangle$ of a memory M. The extended system is in the pure state $|\psi_{MAB}\rangle = \sum_i \sqrt{p_i}\,|m_i\rangle|\psi_{AB}^i\rangle$. If we have no access to the memory system, we trace over it to obtain the mixed state in Eq. (1). In fact, the lack of access to the memory is of a completely general nature. It may be due to interaction with another inaccessible system, or it may be due to an intrinsic loss of information. Our results are universally valid and do not depend on the nature of the information loss. We will see that the amount of entanglement involved in the different entanglement manipulations of mixed states depends on the accessibility of the information in the memory at different stages. Note that a unitary operation on $|\psi_{MAB}\rangle$ will convert it into another pure state $|\phi_{MAB}\rangle$ with the same entanglement, and tracing over the memory yields a different decomposition of the mixed state. Reduction of the pure state to the mixed state may be regarded as due to a projection-valued measurement on the memory with operators $\{E_i = |m_i\rangle\langle m_i|\}$.

Consider first the protocol of formation by means of which Alice and Bob come to share an ensemble of n mixed states ρ_{AB}. Alice first creates the mixed states locally by preparing a collection of n states in a particular decomposition, $\varepsilon = \{|\psi_{AB}^i\rangle\langle\psi_{AB}^i|, p_i\}$ by making np_i copies of each pure state $|\psi_{AB}^i\rangle$. At the same time we may imagine a memory system entangled to the pure states to be generated, which keeps track of the identity of each member of the ensemble. We consider first the case where the state of subsystems A and B together with the memory is pure. Later, we will consider the situation in which Alice's memory is decohered. There are then three ways for her to share these states with Bob. First of all, she may simply compress subsystem B to $nS(\rho_B)$ states, and teleport these to Bob using $nS(\rho_B)$ maximally entangled pairs. The choice of which subsystem to teleport is made so as to minimize the amount of entanglement required, so that $S(\rho_B) \leq S(\rho_A)$. The teleportation in this case would require no classical communication in the asymptotic limit, just as for pure states [8]. The state of the whole system which is created by this process is an ensemble of pure states $|\psi_{MAB}\rangle$, where subsystems M and A are on Alice's side and subsystem B is on Bob's side. In terms of entanglement resources, however, this process is not the most efficient way for Alice to send the states to Bob. She may do it more efficiently by using the memory system of $|\psi_{MAB}\rangle$ to identify blocks of np_i members in each pure state $|\psi_{AB}^i\rangle$, and applying compression to each block to give $np_iS(\rho_B^i)$ states. Then the total number of maximally entangled pairs required to teleport these states to Bob is $n\sum_i p_iS(\rho_B^i)$, which is clearly less than $nS(\rho_B)$, by concavity of the entropy. The amount of entanglement required clearly depends on the decomposition of the mixed state ρ_{AB}. In order to decompress these states, Bob must also be able to identify which members of the ensemble are in which state. Therefore Alice must also send him the memory system. She now has two options. She may either teleport the memory to Bob, which would use more entanglement resources. Or she may communicate the information in the memory classically, with no further use of entanglement. When Alice uses the minimum entanglement decomposition, $\varepsilon = \{|\psi_{AB}^i\rangle\langle\psi_{AB}^i|, p_i\}$, this process, originally introduced by Bennett *et al.* [5], makes the most efficient use of entanglement, consuming only the entanglement of formation of the mixed state, $E_F(\rho_{AB}) = \sum_i p_iS(\rho_B^i)$. We may think of the classical communication between Alice and Bob in one of two equivalent ways. Either Alice may measure the memory locally to decohere it, and then send the result to Bob classically, or she may send the memory through a completely decohering quantum channel. Since Alice and Bob have no access to the channel, the state of the whole system which is created by this process is the mixed state

$$\rho_{ABM}^{\varepsilon} = \sum_i p_i|\psi_{AB}^i\rangle\langle\psi_{AB}^i| \otimes |m_i\rangle\langle m_i|\,, \qquad (2)$$

where Bob is classically correlated to the AB subsystem. Bob is then able to decompress his states using the memory to identify members of the ensemble.

Once the collection of n pairs is shared between Alice and Bob, it is converted into an ensemble of n mixed states ρ_{AB} by destroying access to the memory which contains the information about the state of any particular member of the ensemble. It is the loss of this information which is responsible for the fact that entanglement of distillation is lower than entanglement of formation, since it is not available to parties carrying out the distillation. (The relation between classical information and distillable entanglement was previously discussed by Eisert *et al.* [9], in a different context.) If Alice and Bob, who do have access to the memory, were to carry out the distillation, they could obtain as much entanglement from the ensemble as was required to form it. In the case where Alice and Bob share an ensemble of the pure state $|\psi_{MAB}\rangle$, they would simply apply the Schmidt projection method [1]. The relative entropy of entanglement gives the upper bound to distillable entanglement, $E_{RE}(|\psi_{(MA):B}\rangle\langle\psi_{(MA):B}|) = S(\rho_B)$, which is the same as the amount of entanglement required to create the ensemble of pure states, as described above. Here MA

and B are spatially separated subsystems on which joint operations may not be performed. In our notation, we use a colon to separate the local subsystems.

On the other hand, if Alice used the least entanglement for producing an ensemble of the mixed state ρ_{AB}, together with classical communication, the state of the whole system is an ensemble of the mixed state ρ^{ε}_{ABM}, and the process is still reversible. Because of the classical correlation to the states $|\psi^i_{AB}\rangle$, Alice and Bob may identify blocks of members in each pure state $|\psi^i_{AB}\rangle$, and apply the Schmidt projection method to them, giving $np_iS(\rho^i_B)$ maximally entangled pairs, and hence a total entanglement of distillation of $\sum_i p_iS(\rho^i_B)$. The relative entropy of entanglement again quantifies the amount of distillable entanglement from the state ρ^{ε}_{ABM} and is given by $E_{RE}(\rho^{\varepsilon}_{A:(BM)}) = \min_{\sigma_{ABM}\in D} S(\rho^{\varepsilon}_{ABM}||\sigma_{ABM})$. The disentangled state which minimizes the relative entropy is $\sigma_{ABM} = \sum_i p_i\sigma^i_{AB} \otimes |m_i\rangle\langle m_i|$, where σ^i_{AB} is obtained from $|\psi^i_{AB}\rangle\langle\psi^i_{AB}|$ by deleting the off-diagonal elements in the Schmidt basis. This is the minimum because the state ρ_{MAB} is a mixture of the orthogonal states $|m_i\rangle|\psi^i_{AB}\rangle$, and for a pure state $|\psi^i_{AB}\rangle$, the disentangled state which minimizes the relative entropy is σ^i_{AB}. The minimum relative entropy of the extended system is then

$$S(\rho^{\varepsilon}_{ABM}||\sigma_{ABM}) = \sum_i p_iS(\rho^i_B).$$

This relative entropy, $E_{RE}(\rho^{\varepsilon}_{A:(BM)})$, has previously been called "entanglement of projection" [10], because the measurement on the memory projects the pure state of the full system into a particular decomposition. The minimum of $E_{RE}(\rho^{\varepsilon}_{A:(BM)})$ over all decompositions is equal to the entanglement of formation of ρ_{AB}. However, Alice and Bob may choose to create the state ρ_{AB} by using a decomposition with higher entanglement than the entanglement of formation. The maximum of $E_{RE}(\rho^{\varepsilon}_{A:(BM)})$ over all possible decompositions is called the "entanglement of assistance" of ρ_{AB} [11]. Because $E_{RE}(\rho^{\varepsilon}_{A:(BM)})$ is a relative entropy, it is invariant under local operations and nonincreasing under general operations, properties which are conditions for a good measure of entanglement [4]. However, unlike $E_{RE}(\rho_{AB})$ and $E_F(\rho_{AB})$, it is not zero for completely disentangled states. In this sense, the relative entropy of entanglement, $E_{RE}(\rho^{\varepsilon}_{A:(BM)})$, defines a class of entanglement measures interpolating between the entanglement of formation and entanglement of assistance. Note that an upper bound for the entanglement of assistance, E_A, can be shown using concavity [11] to be $E_A(\rho_{AB}) \le \min[S(\rho_A), S(\rho_B)]$. This bound can also be shown from the fact that the distillable entanglement from any decomposition, $E_{RE}(\rho^{\varepsilon}_{A:(BM)}) \le E_A(\rho_{AB})$ cannot be greater than the entanglement of the original pure state.

We may also derive relative entropy measures that interpolate between the relative entropy of entanglement and the entanglement of formation by considering nonorthogonal measurements on the memory. First of all, the fact that the entanglement of formation is in general greater than the upper bound for entanglement of distillation, emerges as a property of the relative entropy, namely, that it cannot increase under the local operation of tracing one subsystem [12],

$$E_F(\rho_{AB}) = \min_{\sigma_{ABM}\in D} S(\rho_{ABM}||\sigma_{ABM}) \ge \min_{\sigma_{AB}\in D} S(\rho_{AB}||\sigma_{AB}).$$

In general, the loss of the information in the memory may be regarded as a result of an imperfect classical channel. This is equivalent to Alice making a nonorthogonal measurement on the memory, and sending the result to Bob. In the most general case, $\{E_i = A_iA_i^+\}$ is a positive operator valued measure performed on the memory. The decomposition corresponding to this measurement is composed of mixed states, $\xi = \{q_i, Tr_M(A_i\rho_{MAB}A_i^+)\}$, where $q_i = Tr(A_i\rho_{MAB}A_i^+)$. The relative entropy of entanglement of the state ρ^{ξ}_{MAB}, when ξ is a decomposition of ρ_{AB} resulting from a nonorthogonal measurement on M, defines a class of entanglement measures interpolating between the relative entropy of entanglement and the entanglement of formation of the state ρ_{AB}. In the extreme case where the measurement gives no information about the state ρ_{AB}, $E_{RE}(\rho^{\varepsilon}_{A:(BM)})$ becomes the relative entropy of entanglement of the state ρ_{AB} itself. In between, the measurement gives partial information. So far, we have shown that the measures interpolating between entanglement of assistance and entanglement of formation result from making orthogonal measurements on preparations of the pure state $|\psi_{MAB}\rangle$ in different bases. We note that they may be equally achieved by using the preparation associated with entanglement of assistance, and making increasingly nonorthogonal measurements.

The loss of entanglement may be related to the loss of information in the memory. As we have seen, there are two stages at which distillable entanglement is lost. The first is in the conversion of the pure state $|\psi_{MAB}\rangle$ into a mixed state ρ_{ABM}. This happens because Alice uses a *classical* channel to communicate the memory to Bob. The second is due to the loss of the memory, M, taking the state ρ_{ABM} to ρ_{AB}. The amount of information lost may be quantified by the difference in mutual information between the respective states. Mutual information is a measure of correlations between the memory M and the system AB, giving the amount of information about AB which may be obtained from a measurement on M. The quantum mutual information between M and AB is defined as $I_Q(\rho_{M:(AB)}) = S(\rho_M) + S(\rho_{AB}) - S(\rho_{MAB})$. The mutual information loss in going from the pure state $|\psi_{MAB}\rangle$ to the mixed state in Eq. (2) is $\Delta I_Q = S(\rho_{AB})$. There is a corresponding reduction in the relative entropy of entanglement, from the entanglement of the original pure state, $E_{RE}(|\psi_{(MA):B}\rangle\langle\psi_{(MA):B}|)$, to the entanglement of the mixed state $E_{RE}(\rho^{\varepsilon}_{A:(BM)})$ for all decompositions ε arising as the result of an orthogonal measurement on the memory. It is possible to prove, using the nonincrease of relative entropy under local operations [12], that when the mutual information loss is added to the relative entropy of entanglement of the mixed state $E_{RE}(\rho^{\varepsilon}_{A:(BM)})$, the result is greater than the relative entropy of entanglement of the original pure state, $E_{RE}(|\psi_{(MA):B}\rangle\langle\psi_{(MA):B}|)$ [13]. The strongest case, which occurs when $E_{RE}(\rho^{\varepsilon}_{A:(BM)}) = E_F(\rho_{AB})$, is

$$E_{RE}(|\psi_{(MA):B}\rangle\langle\psi_{(MA):B}|) \leq E_F(\rho_{AB}) + S(\rho_{AB}). \quad (3)$$

A similar result may be proved for the second loss, due to the loss of the memory [13]. Again the mutual information loss is $\Delta I_Q = S(\rho_{AB})$. The relative entropy of entanglement is reduced from $E_{RE}(\rho^{\varepsilon}_{A:(BM)})$, for any decomposition ε resulting from an orthogonal measurement on the memory, to $E_{RE}(\rho_{AB})$, the relative entropy of entanglement of the state ρ_{AB} with no memory. When the mutual information loss is added to $E_{RE}(\rho_{AB})$, the result is greater than $E_{RE}(\rho^{\varepsilon}_{A:(BM)})$. In this case, the result is strongest for $E_{RE}(\rho^{\varepsilon}_{A:(BM)}) = E_A(\rho_{AB})$:

$$E_A(\rho_{AB}) \leq E_{RE}(\rho_{AB}) + S(\rho_{AB}). \quad (4)$$

Notice that if ρ_{AB} is a pure state, then $S(\rho_{AB}) = 0$, and equality holds. Inequalities (3) and (4) provide lower bounds for $E_F(\rho_{AB})$ and $E_{RE}(\rho_{AB})$, respectively. They are of a form typical of irreversible processes in that restoring the information in M is not sufficient to restore the original correlations between M and AB. In particular, they express that the loss of entanglement between Alice and Bob at each stage must be accompanied by an even greater reduction in mutual information between the memory and subsystems AB.

In summary, the relative entropy of entanglement of the state ρ_{AB} depends only on the density matrix ρ_{AB}, and gives an upper bound to the entanglement of distillation. The other measures of entanglement, which are given by relative entropies of an extended system, all depend on how the information in the memory is used, or how the density matrix is decomposed. There are numerous decompositions of any bipartite mixed state into a set of states ρ_i with probability p_i. The average entanglement of states in each decomposition is given by the relative entropy of entanglement of the system extended by a memory whose orthogonal states are classically correlated to the states of the decomposition. This correlation records which state ρ_i any member of an ensemble of mixed states $\rho_{AB}^{\otimes n}$ is in. It is available to parties involved in formation of the mixed state, but is not accessible to parties carrying out distillation. When the classical information is fully available, different decompositions give rise to different amounts of distillable entanglement, the highest being entanglement of assistance and the lowest, entanglement of formation. If access to the classical record is reduced, the amount of distillable entanglement is reduced. In the limit where no information is available, the upper bound to the distillable entanglement is given by the relative entropy of entanglement of the state ρ_{AB} itself, without the extension of the classical memory. Our work shows that relative entropy of entanglement provides a unifying measure for all cases, elucidating the role of classical information and the appearance of irreversibility in manipulations of mixed state entanglement.

[1] C.H. Bennett, H.J. Bernstein, S. Popescu, and B. Schumacher, Phys. Rev. A **53**, 2046 (1996).
[2] S. Popescu and D. Rohrlich, Phys. Rev. A **56**, R3319 (1997).
[3] V. Vedral, M.B. Plenio, M.A. Rippin, and P.L. Knight, Phys. Rev. Lett. **78**, 2275 (1997).
[4] V. Vedral and M.B. Plenio, Phys. Rev. A **57**, 1619 (1998).
[5] C.H. Bennett, D.P. DiVincenzo, J.A. Smolin, and W.K. Wootters, Phys. Rev. A **54**, 3824 (1996).
[6] L.P. Hughston, R. Jozsa, and W.K. Wootters, Phys. Lett. A **183**, 14 (1993).
[7] B. Schumacher, Phys. Rev. A **51**, 2738 (1995).
[8] H. Lo and S. Popescu, Phys. Rev. Lett. **83**, 1459 (1999).
[9] J. Eisert *et al.*, quant-ph/9907021, 1999.
[10] R. Garisto and L. Hardy, Phys. Rev. A **60**, 827 (1999).
[11] D.P. DiVincenzo *et al.*, quant-ph/9803033, 1998.
[12] G. Lindblad, Commun. Math. Phys. **40**, 147 (1975).
[13] L. Henderson and V. Vedral (to be published).