Minimal Conditions for Local Pure-State Entanglement Manipulation

Daniel Jonathan* and Martin B. Plenio†

Blackett Laboratory, Imperial College, London SW7 2BZ, United Kingdom

(Received 15 March 1999)

We find a minimal set of necessary and sufficient conditions for the existence of a local procedure that converts a finite pure state into one of a set of possible final states. This result provides a powerful method for obtaining optimal local entanglement manipulation protocols for pure initial states. As an example, we determine analytically the optimal distillable entanglement for arbitrary finite pure states. We also construct an explicit protocol achieving this bound.

PACS numbers: 03.67.Hk

The existence of nonlocal correlations, or entanglement, between parts of a composite quantum system is at the heart of quantum information theory and its applications [1]. In recent years, much effort has been expended on the problem of how to define and quantify the entanglement of a given state in physically meaningful ways. One very fruitful approach, first pursued by Bennett and coworkers [2–4], is to regard entanglement in terms of the *limitations* that exist to the manipulation of a composite system when each subsystem is operated on locally. A paradigmatic situation is as follows: suppose Alice and Bob each possess part of a quantum system, which is prepared in a state ρ. Qualitatively, the existence of entanglement implies that some transformations of ρ which are in principle possible cannot be realized if Alice and Bob are allowed to perform only *local* operations on their respective subsystems, and to exchange classical communication. In this paper, transformations of this type will be referred to as "local transformations," or "LQCC" for short.

A quantitative way of expressing this fact is in terms of so-called *entanglement monotones* (EMs) [5]. These are functions $\varepsilon(\rho)$ of the quantum state that can, on average, never increase under LQCC [6]. There are many known EMs, for example the entanglements of distillation [2–4] and formation [4,7], and the relative entropy of entanglement [8] (in fact, any reasonable measure of entanglement must by definition be an entanglement monotone, and vice versa). Despite their different physical interpretations, they all share a common feature: a transformation which, on average, increases *any single* EM cannot be realized locally. In other words, they provide *necessary* conditions any local transformation T must satisfy.

A natural question that presents itself is then: what are *sufficient* conditions for T to be local? In other words, we would like to have a set $\{\varepsilon_i\}$ of entanglement monotones such that, if the average $\langle \varepsilon_i(T[\rho]) \rangle \leq \varepsilon_i(\rho)$ for all i, then T is local. Ideally, this set should also be *minimal*, in the sense that these conditions should not be redundant [9]. An important result in this direction was recently presented by Nielsen [10], who found sufficient conditions for the locality of transformations that take one given *pure* state to another with 100% probability. In the present Letter, we extend Nielsen's theorem to the case where the transformation need not be deterministic, that is, when T may lead to several possible final states. We demonstrate that, for this case, a set of EMs recently introduced by Vidal [5] is in fact minimal in the sense described above. They therefore provide us with a powerful universal tool for finding optimal local entanglement manipulation protocols. We apply it to the problem of *entanglement concentration* (or *purification*), which concerns understanding to which extent distant parties can extract a maximally entangled state from a nonmaximally entangled one using only LQCC [2–4,11]. This is one of the central problems of quantum information theory, and is of crucial importance for all applications, such as teleportation [12], which require the existence of maximally entangled states between distant parties. With the help of our generalization of Nielsen's theorem, and of results from the well-known simplex optimization method of linear programming theory [13], we are able to analytically determine the *optimal* purification protocol for the case where Alice and Bob share a given pure state $|\psi\rangle$. By "optimal" we mean the following: assume that Alice and Bob locally manipulate their shared state until they obtain either a maximally entangled state (of some dimension), or a completely disentangled one. We determine the strategy that awards them, on average, with the largest amount of distilled entanglement, which we find to be

$$\langle E \rangle_{\max} = \sum_{j=1}^{N} (\alpha_j - \alpha_{j+1}) j \ln j \,, \tag{1}$$

where $\alpha_1 \geq \cdots \geq \alpha_N$ are the nonzero Schmidt coefficients of $|\psi\rangle$.

It is important to stress that our results pertain to any finite shared state. Until now (see also *note added*), the problem of finding the best purification protocol in the sense above had been completely solved (for pure states and some particular mixed states), only in the *asymptotic limit*, where Alice and Bob share $N \to \infty$ identical copies of the same state [2,4]. This limit has fundamental significance in quantum information and communication theory—for example, for deriving bounds on channel capacities [4]. Nevertheless, it is clear that in practice

Alice and Bob will always share only a *finite*, in general small, amount of entanglement. Thus, as a number of authors [5,9,10,14] have stressed, it is also important to understand entanglement transformations in this regime, with the asymptotic results emerging in the suitable limit.

Suppose then that Alice and Bob share a pure state $|\psi\rangle$ of a bipartite quantum system, with ordered Schmidt decomposition $|\psi\rangle = \sum_{i=1}^{N} \sqrt{\alpha_i}\,|i_A\rangle|i_B\rangle$ [15]. Vidal [5] has shown that each of the following set of functions of the α_i constitutes an entanglement monotone:

$$E_l(|\psi\rangle) = \sum_{i=l}^{N} \alpha_i, \qquad 1 \le l \le N\,. \tag{2}$$

We can use these monotones to describe the following theorem due to Nielsen [10]: let $|\eta\rangle$ be another pure bipartite state. Then there exists a local transformation that takes $|\psi\rangle$ to $|\eta\rangle$ with 100% certainty iff $E_l(|\eta\rangle) \le E_l(|\psi\rangle)$, $2 \le l \le N$. In other words, the $\{E_l\}$ form a sufficient set of monotones for this kind of transformation. In fact, since they also uniquely determine the Schmidt components of $|\psi\rangle$ and $|\eta\rangle$, which completely and minimally characterize such transformations (Sec. 5.1 of [9]), it follows that $\{E_l\}$ is actually a *minimal* set of EMs in this case.

Quantum mechanics is not, however, concerned only with deterministic transformations. As long as Alice and Bob do not lose or discard information about their system, the most general transformation they can apply on $|\psi\rangle$ will produce one of m possible pure states $|\eta_i\rangle$, with probability p_i. We demonstrate now that Vidal's monotones also provide necessary and sufficient conditions for these general transformations to be realized locally.

Theorem 1: Let 2 distant parties share a pure state $|\psi\rangle$; let $\{|\eta_j\rangle\}_{j=1}^{m}$ be a set of m other pure bipartite states. Then a transformation T of $|\psi\rangle$ that outputs state $|\eta_j\rangle$ with probability p_j ($\sum_j p_j = 1$) can be realized using LQCC iff the N entanglement monotones E_l do not increase on average, that is, iff

$$\sum_{j=1}^{m} p_j E_l(|\eta_j\rangle) \le E_l(|\psi\rangle), \qquad 1 \le l \le N\,. \tag{3}$$

Proof: Necessity follows from the definition of an entanglement monotone, and is proven for the E_l functions in [5]. To prove sufficiency, assume Eq. (3) is satisfied. We will construct an explicit local strategy that realizes the transformation T. First of all, it is clear that we need to consider only the special case where all target states $|\eta_j\rangle$ have the same Schmidt basis as $|\psi\rangle$ (which we can refer to as the "standard" basis). The general case then follows from the following simple facts: (i) Any two states with the same Schmidt components are interconvertible by a local unitary operation, so that to realize T one needs only to generate, with probability p_j, a state $|\eta_j'\rangle$ with the same Schmidt coefficients as $|\eta_j\rangle$ in the standard basis. (ii) If two or more target states $|\eta_{j_1}\rangle, \cdots, |\eta_{j_n}\rangle$ have *exactly* the same Schmidt components, one can generate the state $|\eta_{j_1}'\rangle$ with probability $\sum_{k=1}^{n} p_{j_k}$, and then "roll a classical die" with relative probabilities $(p_{j_i}/\sum_{k=1}^{n} p_{j_k})$ to decide which one of the $|\eta_{j_k}\rangle$ to transform to.

Suppose then that the target states can all be written in the ordered Schmidt form $|\eta_j\rangle = \sum_{i=1}^{N} \sqrt{\mu_{ji}}\,|i_A\rangle|i_B\rangle$ (note that the number of nonzero Schmidt components of $|\eta_j\rangle$ cannot be greater than N [14]).

Let us now define the *average target state* $|\bar{\eta}\rangle$ as

$$|\bar{\eta}\rangle \equiv \sum_{i=1}^{N} \sqrt{\gamma_i}\,|i_A\rangle|i_B\rangle, \qquad \gamma_i = \sum_{j=1}^{m} p_j \mu_{ji}\,. \tag{4}$$

It can be seen that $\gamma_i \ge \gamma_{i+1}$, so

$$E_l(|\bar{\eta}\rangle) = \sum_{i=l}^{N} \sum_{j=1}^{m} p_j \mu_{ji} = \sum_{j=1}^{m} p_j E_l(|\eta_j\rangle) \le E_l(|\psi\rangle)\,, \tag{5}$$

where we have used condition (3). We can therefore apply Nielsen's theorem, which implies that there exists a local protocol L for *deterministically* converting from $|\psi\rangle$ to $|\bar{\eta}\rangle$. Let us now define the following set of positive operators on Alice's subspace:

$$A_j = \sum_{i=1}^{N} \sqrt{\frac{p_j \mu_{ji}}{\gamma_i}}\,|i_A\rangle\langle i_A|, \qquad 1 \le j \le m\,. \tag{6}$$

We can see that, for $1 \le j \le m$,

$$A_j \otimes \mathbf{1}_B |\bar{\eta}\rangle \sum_{i=1}^{N} \sqrt{p_j \mu_{ji}}\,|i_A\rangle|i_B\rangle = \sqrt{p_j}\,|\eta_j\rangle\,, \tag{7}$$

$$\sum_{j=1}^{m} A_j^{\dagger} A_j = \sum_{i=1}^{N} \left(\frac{\sum_{j=1}^{m} p_j \mu_{ji}}{\gamma_i} \right) |i_A\rangle\langle i_A| = \mathbf{P}\,, \tag{8}$$

where $\mathbf{P}$ is the projector $\sum_{i=1}^{N} |i_A\rangle\langle i_A|$. Together with the complement $\mathbf{1}_A - \mathbf{P}$, the set $\{A_j\}_{j=1}^{m}$ constitutes therefore a local POVM which, if applied to $|\bar{\eta}\rangle$, outputs state $|\eta_j\rangle$ with probability p_j. The combination of this POVM with the deterministic protocol L realizes the required transformation T.

This result can be directly extended to the case where the target states may be mixed. In this case, Eq. (3) still holds (substituting ρ_j for $|\eta_j\rangle$), as long as we extend the definition of E_l using the "convex roof" rule [5]

$$E_l(\rho_j) = \min_{\rho_j = \sum_i q_{ij} |\eta_{ij}\rangle\langle\eta_{ij}|} \sum_{i=1}^{m} q_{ij} E_l|\eta_{ij}\rangle\,, \tag{9}$$

where the minimum is taken over all realizations of ρ_j.

Theorem 1 provides a powerful tool for optimizing local quantum transformations according to a wide range of criteria. For example, in Ref. [5], the author seeks to determine the local transformation that maximizes the probability of converting one given pure state to another. He obtains an upper bound on this probability from the existence of the monotones E_l, and then constructs an explicit protocol realizing the bound. Theorem 1 justifies this result, showing that a similar strategy will work for *any* optimization problem involving an initial pure

state; in other words, the optimum given the constraints expressed in Eq. (3) will always be achievable.

We can immediately apply this result to the problem of optimally concentrating the entanglement of a finite bipartite pure state. This situation has already been considered by Lo and Popescu [14], who have obtained the local protocol that gives the greatest probability of converting a given pure state $|\psi\rangle$ to a maximally entangled state of any *given* number of levels. However, it may well be that Alice and Bob merely wish to concentrate their entanglement, without regard to what maximally entangled state they end up with. In this case, a reasonable question to ask is the following: out of all such local concentration protocols, which one leads, on average, to the largest amount of shared distilled entanglement?

The problem may be formally posed as follows: let Alice and Bob share a single pure state $|\psi\rangle = \sum_{i=1}^{N} \sqrt{\alpha_i}\,|i_A\rangle\,|i_B\rangle$, whose entanglement they wish to concentrate using LQCC. Following the notation of Ref. [14], let us define $|\phi_j\rangle = (1/\sqrt{j}\,)\sum_{i=1}^{j} |i_A\rangle\,|i_B\rangle$ as a maximally entangled state of j levels (note that $|\phi_1\rangle$ is a product state). Consider the set of local transformations that generate $|\phi_j\rangle$, $1 \le j \le N$, with probability p_j [16]. If we choose to measure the amount of entanglement in $|\phi_j\rangle$ by the von Neumann entropy of $tr_B|\phi_j\rangle\langle\phi_j|$, namely $\ln j$, then the average amount of distilled entanglement obtained from such a procedure is

$$\langle E\rangle = \sum_{j=1}^{N} p_j \ln j\,. \tag{10}$$

Our problem is to maximize this quantity over all probability distributions for the p_i that are consistent with the constraints in Eq. (3). Theorem 1 then guarantees the existence of a local protocol leading to this optimal distribution.

It is easily seen that, for $l \le j$,

$$E_l(|\phi_j\rangle) = \frac{j - l + 1}{j}\,, \tag{11}$$

and that it vanishes otherwise. In this case, the constraints in Eq. (3) read

$$\sum_{j=l}^{N} p_j \left(\frac{j - l + 1}{j}\right) \le \sum_{j=l}^{N} \alpha_j, \qquad 1 \le l \le N\,. \tag{12}$$

This is a linear optimization problem with linear inequality constraints, a kind widely studied in many fields of science and engineering. It can be solved using the techniques of *linear programming* theory, a branch of applied linear algebra that is familiar to most engineers, though not so well known among physicists. We will not attempt to explain the terminology and results from this theory that are required for our solution; instead, we refer the reader to textbooks (e.g., [13]). Our main result is

Theorem 2: The optimal entanglement concentration procedure for a single pure bipartite state $|\psi\rangle$ with Schmidt coefficients $\alpha_1 \ge \cdots \alpha_N > 0$ is one that produces a maximally entangled state $|\phi_j\rangle$ of $j \le N$ levels with probability $p_j^{\text{opt}} = j(\alpha_j - \alpha_{j+1})$. The corresponding optimal average distilled entanglement is $\langle E\rangle_{\max} = \sum_{j=1}^{N}(\alpha_j - \alpha_{j+1}) j \ln j$.

Proof: First, it is easy to check that this probability distribution satisfies (actually, *saturates*) all the inequalities in Eq. (12). In matrix form, we have $\mathbf{B}\vec{p} = \vec{q}$, where $\vec{p}$, $\vec{q}$ are vectors with components $p_j = j(\alpha_j - \alpha_{j+1})$; $q_l = \sum_{j=l}^{N} \alpha_j$, and $\mathbf{B}$ is an upper triangular $N \times N$ matrix with components $b_{lj} = \frac{j+1-l}{j}$ for $j \ge l$, and 0 otherwise. In the parlance of linear programming theory, this is a *basic, feasible* solution to the problem, with all the *slack variables* assuming the value zero. We can then apply the *simplex algorithm* to check whether this is the optimal solution or, if not, to find a better one. A sufficient condition for optimality [[13], Eqs. (2.36) and (2.37)] is that the following inequalities are all satisfied

$$z_k \equiv \sum_{i=1}^{N} c_i \beta_{ik} \ge 0, \qquad 1 \le k \le N\,, \tag{13}$$

where $c_i = \ln i$ is the coefficient of p_i in Eq. (10), and β_{ik} are the elements of the inverse of $\mathbf{B}$. It is easy to show that the only nonzero β_{ik} are

$$\begin{aligned} &\beta_{k-2,k} = k - 2; \qquad \beta_{k-1,k} = -2(k-1)\,; \\ &\beta_{kk} = k\,. \end{aligned} \tag{14}$$

The conditions in Eq. (13) are then trivially satisfied for $k = 1, 2$. For $k \ge 3$, we have

$$\begin{aligned} z_k = {} & (k-2)\ln(k-2) + k\ln(k) \\ & - 2(k-1)\ln(k-1)\,. \end{aligned} \tag{15}$$

The remaining inequalities follow from the convexity of $x \ln x$ for $x > 0$. The distribution $p_j^{\text{opt}} = j(\alpha_j - \alpha_{j+1})$ is therefore optimal.

In the remainder of this article, we examine some aspects of theorem 2. First of all, note that, though theorem 1 provides an explicit local protocol realizing the optimal probability distribution given above, it is a complicated one, involving a series of local measurements and subsequent conditioned local rotations by Alice and Bob. We can, however, also explicitly construct a simpler optimal protocol, involving only a *single* local generalized measurement (such a simple protocol always exists for *any* local transformation on a bipartite pure state [14]). Consider the positive operators

$$O_j = \sum_{i=1}^{j} \sqrt{\frac{\alpha_j - \alpha_{j+1}}{\alpha_i}}\,|i_A\rangle\langle i_A| \otimes \mathbf{1}_B\,. \tag{16}$$

It is easily seen that

$$O_j|\psi\rangle = \sqrt{\alpha_j - \alpha_{j+1}} \sum_{i=1}^{j} |i_A\rangle\,|i_B\rangle = \sqrt{p_j^{\text{opt}}}\,|\phi_j\rangle\,, \tag{17}$$

$$\sum_{j=1}^{N} O_j^{\dagger} O_j = \sum_{i=1}^{N} \frac{\sum_{j=i}^{N}(\alpha_j - \alpha_{j+1})}{\alpha_i}\,|i_A\rangle\langle i_A| = \mathbf{P}\,, \tag{18}$$

where we have interchanged $\sum_{j=1}^{N}\sum_{i=1}^{j} \leftrightarrow \sum_{i=1}^{N}\sum_{j=i}^{N}$, and where $\mathbf{P}$ is as in Eq. (7). The set $\{O_j, \mathbf{1} - \mathbf{P}\}$ corresponds thus to a single local POVM measurement that optimally concentrates the entanglement of state $|\psi\rangle$.

Although it is optimal, the protocol provided by theorem 2 is also in general *irreversible,* i.e., it is impossible to recover the original state with 100% probability. This follows since, in general, $\langle E\rangle_{\max} < S$, where S is the entropy of entanglement of $|\psi\rangle$. Note that, since the monotones E_l are all conserved in this process [the inequalities in Eq. (12) are all saturated], this set is *not* sufficient to indicate the reversibility of a local transformation. It can be shown, however, that our protocol does become reversible in the asymptotic limit where Alice and Bob share $N \to \infty$ copies of identical pure states (in which case $\langle E\rangle_{\max} \to S$). This result, which recovers the one obtained by Bennett *et al.* [2] can be derived from expression (1) for $\langle E\rangle_{\max}$ using the saddle point method. It can also be checked that, for any finite pure state, our protocol is always more efficient than the one suggested in [2]. This is not surprising, as their protocol is state independent, while ours is state dependent.

The solution provided by theorem 2 has an intuitive appeal: the optimal protocol for concentrating entanglement is one that first maximizes p_N, that is, the likelihood of obtaining the most entangled state possible; then, given this, it maximizes p_{N-1}, and so forth. Although this seems very reasonable, it is not at all obvious that it should be the case: for instance, it could have conceivably been more advantageous not to attempt to obtain $|\phi_N\rangle$, if this choice had sufficiently increased the likelihood of generating $|\phi_{N-1}\rangle$ [i.e., enough to increase the final average in Eq. (10)]. In fact, it can be readily seen that a *different* optimal solution may be obtained if Alice and Bob choose to use a different entanglement measure to "weigh" each probability in Eq. (10). As a simple example: if they use the trivial "indicator" measure that assigns a value 0 to a disentangled state, and 1 to *any* entangled state [8], then the optimal solution is the one that maximizes p_2. (This follows from the fact that, for any $j > 2$, $|\phi_j\rangle$ may be locally converted to $|\phi_2\rangle$ with 100% efficiency [14]). This solution will in general *not* maximize p_N [14], so it differs from the one found in theorem 2. Ultimately, the choice of which measure to use (and in the finite-state regime, there are many possibilities [9]) depends on Alice and Bob's particular needs. Whatever the choice, however, the techniques of theorems 1 and 2 always determine the optimal protocol.

In summary, we have presented a general method for determining the locality of transformations on a given pure bipartite state, based on the nonincrease of a minimal set of entanglement monotones. We have then used this method to determine the optimal strategy for locally concentrating the entanglement in such a state. We believe that a similar approach will also prove fruitful for more general problems involving mixed and/or multiparticle states [17].

We would like to thank P. L. Knight, S. Bose, E. Le Ru and L. Hardy for helpful discussions. D. J. also thanks Mr. Vee for helpful distractions. We acknowledge the support of the Brazilian agency Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPQ), The ORS Award Scheme, the United Kingdom Engineering and Physical Sciences Research Council, The Leverhulme Trust, the European Science Foundation, and the European Union.

Note added—After this work was completed, Hardy called our attention to his simultaneous work [18], in which Eq. (1) is also obtained using entirely different methods.

*Email address: d.jonathan@ic.ac.uk
†Email address: m.plenio@ic.ac.uk

[1] See, e.g., Phys. World **11** (1998); M. B. Plenio and V. Vedral, Contemp. Phys. **39**, 431 (1998).
[2] C. H. Bennett *et al.*, Phys. Rev. A **53**, 2046 (1996).
[3] C. H. Bennett *et al.*, Phys. Rev. Lett. **76**, 722 (1996).
[4] C. H. Bennett *et al.*, Phys. Rev. A **54**, 3824 (1996).
[5] G. Vidal, Phys. Rev. Lett. **83**, 1046 (1999).
[6] For a rigorous definition, see [5,8,9].
[7] W. K. Wootters, Phys. Rev. Lett. **80**, 2245 (1998).
[8] V. Vedral *et al.*, Phys. Rev. Lett. **78**, 2275 (1997); V. Vedral and M. B. Plenio, Phys. Rev. A **57**, 1619 (1998).
[9] G. Vidal, quant-ph/9807077.
[10] M. Nielsen, Phys. Rev. Lett. **83**, 436 (1999).
[11] N. Gisin, Phys. Lett. A **210**, 151 (1996).
[12] C. H. Bennett *et al.*, Phys. Rev. Lett. **70**, 1895 (1993).
[13] G. R. Walsh, *An Introduction to Linear Programming* (J. Wiley & Sons, New York, 1985), Chap. 2.
[14] H. K. Lo and S. Popescu, quant-ph/9707038.
[15] An "ordered" Schmidt decomposition is one where the Schmidt coefficients appear in decreasing order. All Schmidt decompositions in this paper are of this kind.
[16] Here, as in the proof of theorem 1, we consider for simplicity that all procedures lead to states with the same Schmidt basis.
[17] For example, it can be shown that the convex roof extension of $\langle E\rangle_{\max}$ to mixed states is itself an entanglement monotone under all LQCC.
[18] L. Hardy, quant-ph/9903001.