The asymptotic entanglement cost of preparing a quantum state

Patrick M Hayden[1], Michał Horodecki[2] and Barbara M Terhal[3]

[1] Centre for Quantum Computation, Clarendon Laboratory, Parks Road, Oxford OX1 3PU, UK
[2] Institute of Theoretical Physics and Astrophysics, University of Gdańsk, 80-952, Gdańsk, Poland
[3] Watson Research Center, PO Box 218, Yorktown Heights, NY 10598, USA

E-mail: patrick.hayden@qubit.org, michalh@iftia.univ.gda.pl and terhal@watson.ibm.com

Abstract
We give a detailed proof of the conjecture that the asymptotic entanglement cost of preparing a state ρ is equal to $\lim_{n\to\infty} E_f(\rho^{\otimes n})/n$ where E_f is the entanglement of formation.

PACS numbers: 03.67.-a, 03.65.Ta

1. Introduction

One of the central issues in quantum entanglement theory is to determine how to optimally convert between different entangled states shared by distant observers Alice and Bob. More precisely, one is interested in the conversion of m pairs of particles, with each pair in a state ρ, into n pairs, each in another state ρ', by means of local quantum operations and classical communication (LOCC) [1], so that the asymptotic ratio m/n is minimal. Of course, the perfect transformation by LOCC

$$\rho^{\otimes m} \to \rho'^{\otimes n}$$

is usually impossible. Therefore, one permits imperfections and requires only asymptotically perfect transformations: the state $\rho^{\otimes m}$ is transformed into some state ρ'_n, which for large n approaches $\rho'^{\otimes n}$ with a chosen distance measure D:

$$\lim_{n\to\infty} D(\rho'_n, \rho'^{\otimes n}) = 0.$$

If the final state ρ is a two-qubit singlet state $\frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$(which we will denote by $|\Psi_-\rangle$), then the process of conversion is called *distillation* [2]. If, instead, it is the initial state that is in the form of singlets, then one refers to *formation* [1]. In this paper we are interested in the latter process. We will call the optimal asymptotic yield the entanglement cost, and denote it

by E_c. In [3] it was shown that for a pure state $\rho = |\psi\rangle\langle\psi|$, E_c is equal to the entropy of either of its reductions, e.g. $\rho_A = \mathrm{Tr}_B(\rho)$. Thus, to produce $|\psi\rangle^{\otimes n}$ one needs $m \approx nS(\rho_A)$ singlets, i.e. the initial state has to be $|\Psi_-\rangle^{\otimes nS(\rho_A)}$.

That result suggested the following stochastic method for the production of ρ out of singlets [1]. One decomposes ρ into an ensemble $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$ of pure states. Then one picks a state $|\psi_i\rangle$ according to the probability distribution $\{p_i\}$, makes the state $|\psi_i\rangle$ from initially shared singlets, and finally forgets the identity of the state. Therefore, one needs on average $\sum_i p_i S(\rho_i^A)$ singlets (where ρ_i^A is the reduction of $|\psi_i\rangle$). One can choose the ensemble which minimizes the above average, with the corresponding minimal cost being called the entanglement of formation of ρ [1], which we will denote by $E_f(\rho)$.

The above scenario can be improved if we realize that it might be more economical to produce the state $\rho^{\otimes n}$ all at once than it is to produce its n constituents one by one (see, e.g., [4]). Thus, the proposed optimal cost should be revised to [5]

$$E_f^\infty(\rho) = \lim_{n\to\infty} \frac{E_f(\rho^{\otimes n})}{n}. \tag{1}$$

(In the following we will argue that the limit on the right-hand side of this equation exists.) This quantity is believed to be equal to the entanglement cost of preparing the state ρ, $E_c(\rho)$. The definition of the entanglement cost, however, refers to a deterministic number of input singlets, while in the stochastic method the number of input singlets is a random variable. Furthermore, it is not clear that the stochastic method is really the optimal way to produce the state ρ. In this paper we resolve these issues by proving that $E_f^\infty(\rho)$ is equal to the entanglement cost $E_c(\rho)$. Our result is in a sense dual to that of Rains [6] concerning the entanglement of distillation. He showed that the entanglement of distillation, if defined in a way analogous to our definition of the entanglement cost, is equal to the expected yield of a stochastic protocol for generating output singlets.

To begin, let us sketch our approach. In our proof, we will first show that $E_f^\infty(\rho) \geqslant E_c(\rho)$, i.e. that there exists a formation protocol that achieves the asymptotic rate $\frac{m}{n} \approx E_f^\infty(\rho)$, by explicitly constructing such a protocol based on the law of large numbers in both its classical and quantum [7] forms. Next, we will show that $E_f^\infty(\rho) \leqslant E_c(\rho)$. The latter inequality will be derived using only some general properties of the entanglement of formation, such as its monotonicity under LOCC operations and its quite strong continuity properties, proved by Nielsen [8]. Such an approach, focusing on abstract features of quantum information-theoretic quantities, has already proven to be fruitful in the domain of quantifying entanglement as well as in the study of quantum channel capacities [9, 10], providing a simple view of intricate topics. To illustrate its power we will now sketch the proof of the inequality $E_f^\infty(\rho) \leqslant E_c(\rho)$.

As mentioned before, the entanglement of formation is monotonically decreasing under LOCC operations [1], i.e.

$$E_f(\rho) \geqslant E_f(\mathcal{L}(\rho)) \tag{2}$$

for any state ρ and LOCC operation $\mathcal{L}$. Moreover, the entanglement of formation is continuous [8] in the sense that for states ρ and ρ' we have

$$\left| E_f(\rho) - E_f(\rho') \right| \leqslant 5D(\rho, \rho') \log_2 \dim \mathcal{H} + 2\eta(D(\rho, \rho')) \tag{3}$$

where ρ and ρ' are supported on the Hilbert space $\mathcal{H}$ and $\eta(x) = -x \log_2 x$, under the assumption that $D(\rho, \rho')$ is sufficiently small. D is the Bures distance, given by $D(\rho, \rho') = 2\sqrt{1 - F(\rho, \rho')}$ with $F(\rho, \rho') = \mathrm{Tr}\sqrt{\rho^{1/2}\rho'\rho^{1/2}}$. (Note that the normalization factor for D is not completely standardized. We have made our choice to agree with [8].) The function F is called the Uhlmann fidelity or square-root-fidelity [11, 12]. Inequality (3) implies that

if two states are close to one another, then so are their *densities* of entanglement. A similar continuity result has been proved in [13] for the relative entropy of entanglement [14].

Now, in order to prove $E_f^\infty \leqslant E_c$, consider the optimal formation protocol, i.e. the sequence of LOCC operations Λ_n producing states $\rho_n \approx \rho^{\otimes n}$ out of $|\Psi_-\rangle^{\otimes m}$, so that m/n tends to E_c. If we take large n, so that $D(\rho^{\otimes n}, \rho_n)$ is small, then by continuity (3) we have

$$\frac{E_f(\rho^{\otimes n})}{n} \approx \frac{E_f(\rho_n)}{n}. \tag{4}$$

Since $\rho_n = \Lambda_n(|\Psi_-\rangle\langle\Psi_-|^{\otimes m})$, from the monotonicity of E_f we obtain

$$\frac{E_f(\rho_n)}{n} \leqslant \frac{E_f(|\Psi_-\rangle\langle\Psi_-|^{\otimes m})}{n} = \frac{m}{n}. \tag{5}$$

Now, by definition, the left-hand side of equation (4) tends to E_f^∞, while the last term of estimate (5) tends to E_c, hence we obtained the required inequality.

Since we used only two properties of the entanglement of formation, we can rephrase the result in a more general setting. Consider any function f, which can be regularized, i.e. for which the limit $f^\infty(\rho) = \lim_{n\to\infty} \frac{f(\rho^{\otimes n})}{n}$ exists. Now, if f is monotonic and continuous in the sense of equation (3), then f^∞ is a lower bound for E_c. This supports the view, according to which the dual measures of entanglement, entanglement of distillation E_D and entanglement cost E_c are in a sense extreme ones [9].

2. Entanglement cost and entanglement of formation

Let us now pass to the fully rigorous part of the paper. Analogous to the definition of asymptotic reducibility in [15] (cf [16]), we define the asymptotic entanglement cost for the preparation of a bipartite state to be

$$E_c(\rho) = \inf \left\{ E \mid \forall \epsilon > 0, \delta > 0, \exists m, n, \mathcal{L}, \left| E - \frac{m}{n} \right| \leqslant \delta \right.$$

$$\left. \text{and } D(\mathcal{L}(|\Psi_-\rangle\langle\Psi_-|^{\otimes m}), \rho^{\otimes n}) \leqslant \epsilon \right\} \tag{6}$$

where $|\Psi_-\rangle$ is the singlet state in $C_2 \otimes C_2$, $\mathcal{L}$ is an LOCC superoperator and D is again the Bures distance.

Our main result is the following:

Theorem 1. *The asymptotic entanglement cost of preparing a state ρ is given by*

$$E_c(\rho) = \lim_{n\to\infty} \frac{E_f(\rho^{\otimes n})}{n} \tag{7}$$

where $E_c(\rho)$ is defined in equation (6) and $E_f(\rho)$ is the entanglement of formation of ρ, defined as

$$E_f(\rho) = \min_{\mathcal{E}=\{p_i,|\psi_i\rangle\}} \sum_i p_i E(|\psi_i\rangle\langle\psi_i|) \tag{8}$$

and $E(|\psi\rangle\langle\psi|)$ is the von Neumann entropy of the reduced density matrix of $|\psi\rangle$.

To make sense of the above claim, we note that the sequence $\left(E_f(\rho^{\otimes n})/n\right)$ has a well defined limit. This is a consequence of the fact that if a sequence (a_n) satisfies $a_n \leqslant cn$ for some constant c and $a_n + a_m \geqslant a_{n+m}$ for any m, n, then (a_n/n) is convergent [17]. It is easy to see that our sequence satisfies the conditions.

We begin by proving that E_c obeys a form of additivity that is highly desirable for an asymptotic cost function, and that will be useful in what follows.

Lemma 1. $E_c(\rho^{\otimes k}) = k E_c(\rho)$ *for all* $k = 1, 2, \dots$.

Proof. One direction is simple: since a protocol to approximate $\rho^{\otimes k}$ is just a protocol to approximate k copies of ρ, $E_c(\rho^{\otimes k}) \geqslant k E_c(\rho)$. The first step in demonstrating that the inequality holds in the opposite direction will be to show that in the definition of E_c, n can be taken to be arbitrarily large. Suppose not, in other words, that for fixed $\epsilon, \delta > 0$, it is impossible to choose $n \geqslant N$, m, and $\mathcal{L}$ such that $|E_c(\rho) - m/n| \leqslant \delta$ and $D(\mathcal{L}(|\Psi_-\rangle\langle\Psi_-|^{\otimes m}), \rho^{\otimes n}) \leqslant \epsilon$. By the definition of $E_c(\rho)$ and the fact that the set of LOCC operations is closed, this implies, however, that there exist $n < N$, m, and $\mathcal{L}$ such that $E_c(\rho) = m/n$ and $\mathcal{L}(|\Psi_-\rangle\langle\Psi_-|^{\otimes m}) = \rho^{\otimes n}$. Setting $r = \lceil N/n \rceil$ and applying $\mathcal{L}^{\otimes r}$ to $|\Psi_-\rangle\langle\Psi_-|^{\otimes mr}$ then violates the assumption that n could not be taken arbitrarily large.

Now, to complete the proof, we will pick $n \gg k$ in order to efficiently approximate some number of copies of $\rho^{\otimes k}$. Formally, we write $n = rk + s$, where r and s are non-negative integers and $s < k$. If we produce an approximation to within ϵ of $\rho^{\otimes n}$ starting from m singlets and then throw away the extra s copies of ρ, we are left with an approximation to within ϵ of $\rho^{\otimes rk}$ that still required m singlets. The waste, however, becomes insignificant since $|E - m/n| \to |E - m/rk|$ as $n \to \infty$. $\square$

Note that full additivity of $E_c(\rho)$, i.e. $E_c(\rho \otimes \sigma) = E_c(\rho) + E_c(\sigma)$ would be a stronger statement, possibly requiring additivity of E_f. With that technical lemma concluded, theorem 1 becomes a consequence of the following two lemmas.

Lemma 2. *Let the entanglement of formation of a (finite-dimensional) density matrix ρ be $E_f(\rho) = \sum_i p_i E_f(|\psi_i\rangle\langle\psi_i|)$ where the optimal ensemble is given by $\{p_i, |\psi_i\rangle\}_{i=1}^k$. We have $E_f(\rho) \geqslant E_c(\rho)$.*

Proof. In the limit of large n, we will approximate $\rho^{\otimes n}$ by

$$\rho_{T_{\delta_1}^{(n)}} = \sum_{s \in T_{\delta_1}^{(n)}} |\psi_s\rangle\langle\psi_s| \tag{9}$$

where $T_{\delta_1}^{(n)}$ is the (strongly) typical set defined as follows. All states $|\psi_s\rangle$ (unnormalized) in this set are tensor products of states $\sqrt{p_i}|\psi_i\rangle$ and they are such that every state $\sqrt{p_i}|\psi_i\rangle$ occurs $p_i n \pm \delta_1 n \log_{p_i}(2)/k$ times in the tensor product. (We assume that $p_i < 1$ for all i. Otherwise, ρ is a pure state.) This implies that the total probability $p_s = \langle\psi_s|\psi_s\rangle$ for a state $|\psi_s\rangle$ is bounded as

$$2^{-n(H(\vec{p})+\delta_1)} \leqslant p_s \leqslant 2^{-n(H(\vec{p})-\delta_1)} \tag{10}$$

where $H(\vec{p})$ is the Shannon entropy of $\vec{p} = (p_i)$. The density matrix $\rho_{T_{\delta_1}^{(n)}}$ is basically the state $\rho^{\otimes n}$ constructed by cutting off the unlikely sequences of states and then renormalizing[4]. In order to bound the fidelity for this approximation, we add these unlikely sequences of states with probability 0 in the sum. For any $\delta_1 > 0$, $\epsilon_1 > 0$ and n sufficiently large we have by the asymptotic equipartition theorem [18] that the total probability $p_{T_{\delta_1}^{(n)}}$ for the typical set is larger than $1 - \epsilon_1$, where $\epsilon_1 \to 0$ as $n \to \infty$. Thus, using the joint concavity of F [11], we can bound

$$F(\rho^{\otimes n}, \rho_{T_{\delta_1}^{(n)}}) \geqslant \sqrt{1 - \epsilon_1}. \tag{11}$$

---

[4] Note that we are using an approximation to $\rho^{\otimes n}$ which is based on the optimal decomposition for the entanglement of formation and not on the eigendecomposition of ρ which is used in Schumacher compression [7]. Both approximations converge to $\rho^{\otimes n}$ in the limit of large n.

Consider the density matrix $\rho_{T_{\delta_1}^{(n)}}$ and its decomposition in terms of the states $|\psi_s\rangle$. Each state $|\psi_s\rangle$ can be obtained from a set of EPR pairs by entanglement dilution [3]. In particular, let $|\psi_s\rangle$ be a state in which every state $|\psi_i\rangle$ occurs $p_i n \pm \delta_1 n \log_{p_i}(2)/k$ times. Starting from a set of maximally entangled EPR pairs, we do entanglement dilution for each state $|\psi_i\rangle^{\otimes p_i n \pm \delta_1 n \log_{p_i}(2)/k}$. For any δ_2 and ϵ_2 greater than zero there exists an n such that starting from $(p_i n + \delta_1 n \log_{p_i}(2)/k)[E_f(|\psi_i\rangle\langle\psi_i|) + \delta_2]$ EPR pairs we can obtain an approximation to $|\psi_i\rangle^{\otimes p_i n \pm \delta_1 n \log_{p_i}(2)/k}$ which has square-root-fidelity larger than $1 - \epsilon_2$. Since there are k states in the optimal ensemble (and k is finite), we can therefore approximate the (normalized) state $|\psi_s\rangle$ with square-root-fidelity $|\langle\psi|\psi'\rangle| \geqslant (1 - \epsilon_2)^k$, starting from $n \sum_i p_i E_f(|\psi_i\rangle\langle\psi_i|) + n(O(\delta_1) + O(\delta_2))$ EPR pairs, with $\epsilon_2 \to 0$, $\delta_1 \to 0$, and $\delta_2 \to 0$ for $n \to \infty$.

The approximation of $\rho_{T_{\delta_1}^{(n)}}$ by $\rho'_{T_{\delta_1}^{(n)}} = \sum_{s \in T_{\delta_1}^{(n)}} |\psi'_s\rangle\langle\psi'_s|$, where $|\psi'_s\rangle$ is the approximation of $|\psi_s\rangle$ which we obtain by entanglement dilution starting from the set of EPR pairs, has the property that

$$F(\rho_{T_{\delta_1}^{(n)}}, \rho'_{T_{\delta_1}^{(n)}}) \geqslant \sum_{s \in T_{\delta_1}^{(n)}} |\langle\psi_s|\psi'_s\rangle| \geqslant (1 - \epsilon_2)^k \equiv 1 - \epsilon_3 \tag{12}$$

where $\epsilon_3 \to 0$ for $n \to \infty$, since k is finite. Furthermore, since we can make every state $|\psi'_s\rangle$ starting from a given set of EPR pairs, we can make any convex combination, for example $\rho'_{T_{\delta_1}^{(n)}}$, of the states $|\psi'_s\rangle$ (see [19] and also [20]), starting from this same set of EPR pairs.

Finally, we can use the triangle inequality for the Bures metric, and equations (11) and (12) to obtain that

$$D(\rho^{\otimes n}, \rho'_{T_{\delta_1}^{(n)}}) \leqslant 2\sqrt{1 - \sqrt{1 - \epsilon_1}} + 2\sqrt{\epsilon_3} \tag{13}$$

which is the desired result.   □

The lemma can be applied to bound $E_c(\rho)$ from above by $E_f(\rho^{\otimes k})/k$ where k is any fixed number of copies of ρ, using lemma 1. Consequently, we have that

$$E_f^\infty(\rho) \geqslant E_c(\rho). \tag{14}$$

Let us now prove the converse of this relation.

Lemma 3. $E_f^\infty(\rho) \leqslant E_c(\rho)$.

Proof. The basic idea of the proof, as sketched in the introduction, is to use the continuity of the entanglement of formation [8] and its monotonicity under LOCC operations. Throughout the proof we will use the notation $A_n = E_f(\rho^{\otimes n})/n$. Now suppose that the lemma does not hold, so that $\lim A_n > E_c(\rho)$. It follows that there exists an integer N, such that for all $k > N$,

$$|A_k - \lim A_n| < \Delta = \frac{\lim A_n - E_c(\rho)}{4} > 0. \tag{15}$$

Let $\rho_k = \mathcal{L}(|\Psi_-\rangle\langle\Psi_-|^{\otimes m})$ be an approximation to $\rho^{\otimes k}$. Now, supposing that ρ is supported on Hilbert space $\mathcal{H}$, fix ϵ such that $5\epsilon \log \dim \mathcal{H} - 2\epsilon \log \epsilon = \Delta$. It then follows from equation (3) and our choice of ϵ that if $D(\rho^{\otimes k}, \rho_k) < \epsilon$ then

$$|E_f(\rho^{\otimes k}) - E_f(\rho_k)| < k\Delta. \tag{16}$$

Now we will apply the definition of $E_c(\rho)$ to fix k. From the definition of E_c, equation (6), and the proof of lemma 1, we have that there exists a $k > N$ as well as m and $\mathcal{L}$ such that

$$\left| E_c(\rho) - \frac{m}{k} \right| \leqslant \Delta \tag{17}$$

and $D(\rho^{\otimes k}, \rho_k) < \epsilon$, where $\rho_k = \mathcal{L}(|\Psi_-\rangle\langle\Psi_-|^{\otimes m})$. Next, we can bound

$$\left| \lim A_n - \frac{E_f(\rho_k)}{k} \right| \leqslant |\lim A_n - A_k| + \left| A_k - \frac{E_f(\rho_k)}{k} \right| < 2\Delta \tag{18}$$

by using equations (15) and (16). This gives $E_f(\rho_k) > k(\lim A_n - 2\Delta)$. On the other hand, by equation (17), we have

$$E_f(|\Psi_-\rangle\langle\Psi_-|^{\otimes m}) = m < k(E_c(\rho) + \Delta) \tag{19}$$

which finally yields

$$E_f(\rho_k) - E_f(|\Psi_-\rangle\langle\Psi_-|^{\otimes m}) > k(\lim A_n - E_c(\rho) - 3\Delta) = k\Delta \tag{20}$$

a contradiction since the entanglement of formation cannot increase under LOCC operations.

□

3. Alternative definitions of the asymptotic cost

While $E_c(\rho)$ is perhaps the most natural function to associate with the asymptotic cost of preparing a bipartite state, other definitions would have been consistent with our discussion in the introduction. An example of a different but perhaps useful definition of the asymptotic entanglement cost is the following:

$$E_{alt}(\rho) = \inf \left\{ E \mid \forall \epsilon > 0, \delta > 0, \exists N \mid \left(\forall n > N \, \exists m, \mathcal{L}, \left| E - \frac{m}{n} \right| \leqslant \delta \right. \right.$$

$$\left. \left. \text{and } D(\mathcal{L}(|\Psi_-\rangle\langle\Psi_-|^{\otimes m}), \rho^{\otimes n}) \leqslant \epsilon \right) \right\}. \tag{21}$$

The difference between E_c and E_{alt} is that for fixed fidelity and cost the former only requires the existence of a single n such that $\rho^{\otimes n}$ can be approximated efficiently, while the second requires that there exist a threshold N such that for all $n > N$ the state $\rho^{\otimes n}$ can be approximated efficiently. One consequence of this difference is that it is immediately clear that E_{alt} is additive in the sense of lemma 1, while some work was required to prove that E_c was. Indeed, *a priori*, one might not expect the two definitions to agree. Suppose, for example, that for fixed ϵ there exists a protocol to make an approximation ρ_n of $\rho^{\otimes n}$ from m singlets, with the approximation satisfying $D(\rho_n, \rho^{\otimes n}) = \epsilon$. In order to produce an approximation to $\rho^{\otimes nk}$ from km singlets, one might think that applying this protocol k times would be a good strategy. One finds, however, that for $\epsilon > 0$,

$$\lim_{k \to \infty} D(\rho^{\otimes nk}, \rho_n^{\otimes k}) = \lim_{k \to \infty} 2\sqrt{1 - (1 - \epsilon^2)^k} = 2. \tag{22}$$

The protocol, therefore, generally fails to produce a good approximation for large k. This example suggests that the existence of a threshold N beyond which approximations to some given fidelity are always possible is a very difficult condition to satisfy. Nonetheless, by applying the results of the previous section, it is actually easy to see that the definitions E_c and E_{alt} are equivalent, so that the extra condition can be met without increasing the

asymptotic unit cost. First, note that the argument of lemma 2 actually also works for E_{alt} so that $E_{alt}(\rho) \leqslant E_f^\infty(\rho)$. Next, since the definition of E_{alt} is more stringent than that of E_c, we have that $E_c(\rho) \leqslant E_{alt}(\rho)$. Combining these inequalities with the result of lemma 3, we obtain

$$E_f^\infty(\rho) \leqslant E_c(\rho) \leqslant E_{alt}(\rho) \leqslant E_f^\infty(\rho) \tag{23}$$

so that these two definitions of the entanglement cost always agree.

A final pair of alternative definitions for the asymptotic entanglement cost would use the trace distance

$$d(\rho, \sigma) = \tfrac{1}{2} \operatorname{Tr} |\rho - \sigma| \tag{24}$$

in place of the Bures distance. Lemma 1 is easily seen to hold for d since its proof uses only axiomatic properties of all metrics and stability with respect to tensor products but no other properties specific to D. The status of lemmas 2 and 3 is slightly more involved. The trace distance and Bures distance are related by the following inequalities [21]:

$$\frac{D}{4}(\rho, \sigma)^2 = 1 - F(\rho, \sigma) \leqslant d(\rho, \sigma) \leqslant \sqrt{1 - F(\rho, \sigma)^2} = \sqrt{\frac{D^2}{2} - \frac{D^4}{16}} \tag{25}$$

which show that d and D are equivalent metrics and bounds them in terms of each other by functions that are independent of the dimension of the underlying Hilbert space, $\mathcal{H}$. In other words, d and D are equivalent metrics even in the asymptotic regime. To be concrete, applying the right-hand inequality to equation (13), for example, is sufficient to show that lemma 2 also holds for the trace distance. In order to prove lemma 3, we used the continuity relation, equation (3), but effectively only required the weaker inequality

$$|E(\rho) - E(\sigma)| \leqslant B\, D(\rho, \sigma)\, \log \dim \mathcal{H} + C \tag{26}$$

where B and C are constants and $\mathcal{H}$ is the supporting Hilbert space. Applying equation (25), however, gives an inequality of the form

$$|E(\rho) - E(\sigma)| \leqslant B\, \sqrt{d(\rho, \sigma)}\, \log \dim \mathcal{H} + C \tag{27}$$

which again is sufficient to carry through the rest of the proof. Therefore, all our conclusions hold even if the entanglement cost is defined using the trace distance. Indeed, they should hold for any metric equivalent to the Bures metric in which the equivalence is given by a function that is independent of the dimension of $\mathcal{H}$.

4. Conclusions

We have given two rigorous definitions of the asymptotic cost of preparing a bipartite mixed state ρ and shown them both to be equal to the regularized entanglement of formation, $E_f^\infty(\rho)$, resolving an important conjecture in the theory of quantum entanglement. Furthermore, we have shown that this asymptotic cost is fairly insensitive to the choice of metric on density operators. In particular, the Bures distance and trace distance result in identical asymptotic costs.

An important problem left open by this work is the question of actually evaluating $E_f^\infty(\rho)$. Even the non-regularized function $E_f(\rho)$ is notoriously difficult to calculate; its value is only known for some very special cases [4, 22]. If it turns out that E_f is not additive for tensor products, then, in spite of the results of this paper, determining the asymptotic cost of preparing a state remains quite a formidable problem.

Acknowledgments

References

[1] Bennett C H, DiVincenzo D P, Smolin J A and Wootters W K 1996 Mixed state entanglement and quantum error correction *Phys. Rev.* A **54** 3824–51
(Bennett C H, DiVincenzo D P, Smolin J A and Wootters W K 1996 *Preprint* quant-ph/9604024)
[2] Bennett C H, Brassard G, Popescu S, Schumacher B, Smolin J and Wootters W K 1996 Purification of noisy entangelment and faithful teleportation via noisy channels *Phys. Rev. Lett.* **76** 722
(Bennett C H, Brassard G, Popescu S, Schumacher B, Smolin J and Wootters W K 1995 *Preprint* quant-ph/9511027)
[3] Bennett C H, Bernstein H J, Popescu S and Schumacher B 1996 Concentrating partial rntanglement by local operations *Phys. Rev.* A **53** 2046
(Bennett C H, Bernstein H J, Popescu S and Schumacher B 1995 *Preprint* quant-ph/9511030)
[4] Wootters W K 1998 Entanglement of formation of an arbitrary state of two qubits *Phys. Rev. Lett.* **80** 2245
(Wootters W K 1997 *Preprint* quant-ph/9709029)
[5] Horodecki P, Horodecki M and Horodecki R 1998 Entanglement and thermodynamical analogies *Acta Phys. Slov.* **48** 141
[6] Rains E M 1999 Rigorous treatment of distillable entanglement *Phys. Rev.* A **60** 173
(Rains E M 1998 *Preprint* quant-ph/9809078)
[7] Schumacher B 1995 Quantum coding *Phys. Rev.* A **51** 2738
[8] Nielsen M A 2000 Continuity bounds for entanglement *Phys. Rev.* A **61** 064301
(Nielsen M A 1999 *Preprint* quant-ph/9908086)
[9] Horodecki M, Horodecki P and Horodecki R 2000 Limits for entanglement measures *Phys. Rev. Lett.* **84** 2014
[10] Horodecki M, Horodecki P and Horodecki R 2000 Unified approach to quantum capacities: towards quantum noisy coding theorem *Phys. Rev. Lett.* **85** 433
[11] Uhlmann A 1976 The transition probability in the state space of a *-algebra *Rep. Math. Phys.* **9** 273
[12] Jozsa R 1994 Fidelity for mixed quantum states *J. Mod. Opt.* **41** 2315
[13] Donald M and Horodecki M 1999 Continuity of relative entropy of entanglement *Phys. Lett.* A **264** 257
[14] Vedral V, Plenio M B, Rippin M A and Knight P L 1997 Quantifying entanglement *Phys. Rev. Lett.* **78** 2275
(Vedral V, Plenio M B, Rippin M A and Knight P L 1997 *Preprint* quant-ph/9702027)
[15] Bennett C H, Popescu S, Rohrlich D, Smolin J A and Thapliyal A V 2001 Exact and asymptotic measures of multipartite pure state entanglement *Phys. Rev.* A **63** 012307
(Bennett C H, Popescu S, Rohrlich D, Smolin J A and Thapliyal A V 1999 *Preprint* quant-ph/9908073)
[16] Keyl M and Werner R F 1999 The rate of the optimal purification procedures *J. Math. Phys.* **40** 3283–99
(Keyl M and Werner R F 1999 *Preprint* quant-ph/9910124)
[17] See e.g. Barnum H, Nielsen M and Schumacher B 1998 Information transmission through a noisy quantum channel *Phys. Rev.* A **57** 4153
(Barnum H, Nielsen M and Schumacher B 1997 *Preprint* quant-ph/9702049)
[18] Cover T M and Thomas J A 1991 *Elements of Information Theory* (New York: Wiley)
[19] Jonathan D and Plenio M 1999 Minimal conditions for local pure-state entanglement manipulation *Phys. Rev. Lett.* **83** 1455
[20] Hayden P, Terhal B M and Uhlmann A 2000 *Preprint* quant-ph/0011095
[21] Fuchs C A and van de Graaf J 1999 Cryptographic distinguishability measures for quantum mechanical states *IEEE T. Inform. Theory* **45** 1216
(Fuchs C A and van de Graaf J 1997 *Preprint* quant-ph/9712042)
[22] Terhal B M and Vollbrecht K G H 2000 The entanglement of formation for isotropic states *Phys. Rev. Lett.* **85** 2625–8
(Terhal B M and Vollbrecht K G H 2000 *Preprint* quant-ph/0005062)