

Communication cost of entanglement transformations

Patrick Hayden*

Institute for Quantum Information, Caltech 107-81, Pasadena, California 91125

Andreas Winter†

Department of Computer Science, University of Bristol, Merchant Venturers Building, Woodland Road, Bristol BS8 1UB, United Kingdom

(Received 16 April 2002; published 31 January 2003)

We study the amount of communication needed for two parties to transform some given joint pure state into another one, either exactly or with some fidelity. Specifically, we present a method to lower bound this communication cost even when the amount of entanglement does not increase. Moreover, the bound applies even if the initial state is supplemented with unlimited entanglement in the form of EPR (Einstein-Podolsky-Rosen) pairs and the communication is allowed to be quantum mechanical. We then apply the method to the determination of the communication cost of asymptotic entanglement concentration and dilution. While concentration is known to require no communication whatsoever, the best known protocol for dilution, discovered by H.-K. Lo and S. Popescu [Phys. Rev. Lett. **83**, 1459 (1999)], requires exchange of a number of bits that is of the order of the square root of the number of EPR pairs. Here we prove a matching lower bound of the same asymptotic order, demonstrating the optimality of the Lo-Popescu protocol up to a constant factor and establishing the existence of a fundamental asymmetry between the concentration and dilution tasks. We also discuss states for which the minimal communication cost is proportional to their entanglement, such as the states recently introduced in the context of “embezzling entanglement” (W. van Dam and P. Hayden, e-print quant-ph/0201041).

DOI: 10.1103/PhysRevA.67.012326

PACS number(s): 03.67.Hk, 03.65.Ta

I. PURE-STATE ENTANGLEMENT TRANSFORMATIONS

The quantification of entanglement began with the study of the following question: Assume that two parties, generically referred to as Alice and Bob, share n copies of a bipartite pure state $|\phi_{AB}\rangle$ which by local operations and classical communication (LOCC) they would like to convert into a state that has high fidelity to k copies of the target state $|\psi_{AB}\rangle$, with k as large as possible. The basic question is then, what is $\lim k/n$ as $n \rightarrow \infty$ and the fidelity goes to 1?

It turns out [1] that this optimal asymptotic ratio is equal to $E(\phi)/E(\psi)$, where

$$E(\phi) = S(\text{Tr}_B |\phi\rangle\langle\phi|) = -\text{Tr}(\text{Tr}_B |\phi\rangle\langle\phi| \log \text{Tr}_B |\phi\rangle\langle\phi|)$$

is the von Neumann entropy of Alice’s reduced state. (Throughout the paper, \log and \exp are understood to be base 2.) For this reason, E is often called the *entropy of entanglement*. One consequence of this result is that pure-state entanglement can be interconverted asymptotically losslessly between its different forms, justifying the introduction of the ebit as a resource quantity, with its ubiquitous “incarnation” the two-qubit Einstein-Podolsky-Rosen (EPR) pair state, which, up to a local change of basis, can be written as

$$|\phi_2^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle).$$

For the sake of quantifying entanglement, however, not only the local actions by Alice and Bob but *classical communication* was considered unlimited. It is precisely these communication requirements that we study in the present paper, in which we follow Lo’s [2] suggestion to study the communication complexity of distributed quantum information processing.

This goal notwithstanding, our point of departure will not be the theory of asymptotically faithful transformations but, rather, its finite (and more refined) variant of transformations from $|\phi_{AB}\rangle$ to $|\psi_{AB}\rangle$ up to fidelity $1 - \epsilon$, as laid out in [3], building on previous work [4–6] for the zero-error case.

Up to local unitaries, pure entangled states are uniquely defined by the spectrum of their reduced states (either at Alice’s or Bob’s side), the eigenvalues known as the *Schmidt coefficients* λ_j . Indeed, it is possible to choose bases in the entangled system such that

$$|\phi_{AB}\rangle = \sum_j \sqrt{\lambda_j} |i\rangle_A \otimes |i\rangle_B.$$

The theory relates the feasibility of such a LOCC transformation to the *majorization order* of the Schmidt coefficients (λ) of $|\phi\rangle$ and (μ) of $|\psi\rangle$, both vectors arranged in nonincreasing order:

$$|\phi\rangle \xrightarrow{\text{LOCC}} |\psi\rangle \quad \text{if and only if } (\lambda) \prec (\mu),$$

where $(\lambda) \prec (\mu)$ is defined to mean

$$\forall k \quad \sum_{j=1}^k \lambda_j \leq \sum_{j=1}^k \mu_j,$$

*Electronic address: patrick@cs.caltech.edu

†Electronic address: winter@cs.bris.ac.uk

which can be shown to be equivalent to the existence of a doubly stochastic matrix M such that $(\lambda) = M(\mu)$. By the results of [5] and [7], any such allowed transformation can always be achieved by one-way communication, say from Alice to Bob, of $2 \log \text{rank} \text{Tr}_B |\phi\rangle\langle\phi|$ classical bits.

The organization of the paper is as follows. In Sec. II we will explain the mathematical model of approximate pure-state transformations and derive our main result, a lower bound on the communication cost of state transformations which holds even if the initial state is supplemented by an unlimited number of EPR pairs, and even if the communication is quantum mechanical. To our knowledge this is the first quantitative statement of its kind. (The need for some nonzero amount of communication in certain transformations was pointed out in [6].) We then apply the result in Sec. III to the asymptotic transformations mentioned in above, proving a lower bound of $\Omega(\sqrt{n})$ on the communication necessary for entanglement dilution, which, up to a constant factor, matches the $O(\sqrt{n})$ construction of Lo and Popescu [8] for this task. In Sec. IV we analyze a class of states that require for their creation from EPR pairs communication of the same order as their entanglement, before ending with a discussion of some open problems.

II. A LOWER BOUND ON THE COMMUNICATION COST

Assume that initially Alice and Bob share the state $|\phi\rangle$, then execute several rounds of local actions and classical communication, and finally end up with some joint state $\tilde{\rho}$ that has high fidelity to $|\psi\rangle$. Allowing the use of quantum bits to communicate, we give Alice and Bob even more power, thereby potentially reducing the communication cost, while at the same time simplifying the appearance of the protocol: Because each of the local actions can be implemented using ancillas and unitary transformations, the whole process can be reduced to a series of exchanges of quantum systems of certain dimensions d_i between Alice and Bob, with a final tracing out (discarding) of part of Alice’s and part of Bob’s system. The total communication cost of such a procedure is just $C = \sum_{i=1}^N \log d_i$ qubits (see Fig. 1).

The idea of the lower bound is very simple, and is explained most straightforwardly for exact state transformations, when $\tilde{\rho} = |\psi\rangle\langle\psi|$. During the process of transformation we monitor a certain quantity Δ associated with Alice’s reduced density operator, showing that, for each qubit communicated, it can only increase by a constant, and then observe that the final partial trace never increases Δ at all. The difference between the initial and the final Δ then provides a lower bound on the communication.

Specifically we shall consider, for $\rho = \text{Tr}_B |\phi\rangle\langle\phi|$,

$$\Delta(\rho) := S_0(\rho) - S_\infty(\rho), \tag{1}$$

where S_α are the Rényi entropies [9] of order α :

$$S_\alpha(\rho) := \frac{1}{1-\alpha} \log \text{Tr}(\rho^\alpha).$$

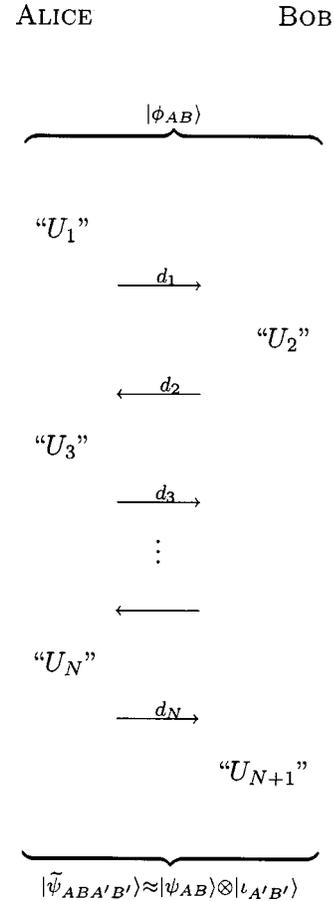


FIG. 1. In round i Alice (Bob) performs some unitary U_i on her (his) system, which separates into a residual system and a d_i -dimensional system that is sent to Bob (Alice). In the last, N th, round, the receiver of the message may perform a unitary on his/her system, and then Alice and Bob trace out subsystems A' and B' .

For $\alpha = 0, 1, \infty$ the Rényi entropies are defined by continuous extension, with the resulting formulas

$$S_0(\rho) = \log \text{rank } \rho,$$

$$S_1(\rho) = -\text{Tr}(\rho \log \rho),$$

$$S_\infty(\rho) = -\log \|\rho\|_\infty,$$

where $\|\cdot\|_\infty$ is the supremum norm: for self-adjoint operators it is the largest absolute value of an eigenvalue. Note that $\Delta(\rho) \geq 0$ since $S_\alpha(\rho)$ is nonincreasing in α [9], or by inspection of the definition. Furthermore, if all the nonzero eigenvalues of ρ are the same then $S_0(\rho) = S_\infty(\rho)$ so that $\Delta(\rho) = 0$. Otherwise, $\Delta(\rho)$ will be strictly greater than zero. Therefore, $\Delta(\rho)$ can be interpreted as a measure of the variation in the eigenvalues of ρ .

The key observation is that, in communication round i , the Rényi entropy of Alice’s reduced state, whose spectrum characterizes the entanglement, cannot change too much. To see this, we assume without loss of generality that it is Alice’s turn to perform a unitary, rotating her reduced state to $\rho_{AA'}$. This step, obviously, does not change the Rényi en-

tropy at all. Next, she gives Bob the d_i -dimensional system A' , leaving her with the new reduced state ρ_A , for which we have the relation (see [10])

$$S_\alpha(\rho_{AA'}) - \log d_i \leq S_\alpha(\rho_A) \leq S_\alpha(\rho_{AA'}) + \log d_i, \quad (2)$$

which implies (inserting $\alpha = 0, \infty$)

$$\Delta(\rho_A) \leq \Delta(\rho_{AA'}) + 2 \log d_i. \quad (3)$$

Thus, the quantity Δ can increase (or decrease) by at most $2 \log d_i$ in step i . After the last round of communication has taken place, the joint state is $|\tilde{\psi}_{ABA'B'}\rangle = |\psi_{AB}\rangle \otimes |\iota_{A'B'}\rangle$. (Note that if this were not a product state, $\tilde{\rho}$ would necessarily not be pure.) Hence, by induction over the number of rounds, summing over Eq. (3) yields

$$\Delta(\text{Tr}_{BB'}|\tilde{\psi}_{ABA'B'}\rangle\langle\tilde{\psi}_{ABA'B'}|) \leq \Delta(\text{Tr}_B|\phi\rangle\langle\phi|) + 2C. \quad (4)$$

The effect on Δ of the final partial trace over the primed system is easy to understand: Because the Rényi entropies are additive under tensor products, i.e.,

$$S_\alpha(\rho \otimes \sigma) = S_\alpha(\rho) + S_\alpha(\sigma),$$

we obtain

$$\Delta(\text{Tr}_B|\psi\rangle\langle\psi| \otimes \text{Tr}_{B'}|\iota\rangle\langle\iota|) = \Delta(\text{Tr}_B|\psi\rangle\langle\psi|) + \Delta(\text{Tr}_{B'}|\iota\rangle\langle\iota|), \quad (5)$$

and the rightmost term is non-negative. This proves the following theorem.

Theorem 1. A (deterministic) pure state transformation of $|\phi_{AB}\rangle$ into $|\psi_{AB}\rangle$ requires at least

$$C \geq \frac{1}{2} [\Delta(\text{Tr}_B|\psi\rangle\langle\psi|) - \Delta(\text{Tr}_B|\phi\rangle\langle\phi|)]$$

bits of communication, even if quantum communication is allowed. ■

We note that in [10] the analogous theorem for the bare Rényi entropies S_α was used to prove bounds on the communication required to perform entanglement transformations in an approximate setting. There, changes in S_α reflected changes in the amount of entanglement present in the system. The advantage of using Δ is precisely that it does not measure entanglement but, rather, variation in the Schmidt coefficients.

Remark 1. Obviously, a similar result holds for

$$\Delta^{\alpha\beta}(\rho) := S_\alpha(\rho) - S_\beta(\rho),$$

with arbitrary $0 \leq \alpha < \beta \leq \infty$. Even though $\Delta^{\alpha\beta}(\rho) \leq \Delta(\rho)$, for some α and β the increase of the former quantity in an entanglement transformation may exceed the increase of the latter.

Remark 2. As an example of a nontrivial consequence of Theorem 1 we may observe that it puts severe restrictions on the entanglement transformations possible without any communication: none of the $\Delta^{\alpha\beta}$ must increase.

For example, from a maximally entangled state only other maximally entangled states (with possibly smaller Schmidt rank) may be obtained. If the Schmidt rank of the target divides that of the initial state this is clearly possible, while inspection of Eq. (5) shows that it is also necessary.

For the case of high-fidelity transformations this approach turns out to be too simple: neither S_0 nor S_∞ can be well controlled if we switch from a state to one close by. For example, for the dilution task, which consists of the creation of $(\alpha|00\rangle + \beta|11\rangle)^{\otimes n}$ from EPR pairs, Theorem 1 implies a lower bound of $\Omega(n)$, while we know from [8] that arbitrarily high fidelity can be achieved with $O(\sqrt{n})$ bits of communication.

Instead, we invent robust versions of S_0 , S_∞ , and Δ : Let the eigenvalues of ρ be denoted r_j and then define, for $0 \leq \epsilon < 1$,

$$S_{0,\epsilon}(\rho) := \log \min \left\{ |J| : \sum_{j \in J} r_j \geq 1 - \epsilon \right\}, \quad (6)$$

$$S_{\infty,\epsilon}(\rho) := -\log \min \left\{ \max_{j \in J} r_j : \sum_{j \in J} r_j \geq 1 - \epsilon \right\}, \quad (7)$$

$$\Delta_\epsilon(\rho) := \log \min \left\{ |J| \left(\max_{j \in J} r_j \right) : \sum_{j \in J} r_j \geq 1 - \epsilon \right\}. \quad (8)$$

All the minimizations are understood to be over subsets J of the eigenvalue indices j . Note that

$$\Delta_\epsilon(\rho) \leq S_{0,\epsilon}(\rho) - S_{\infty,\epsilon}(\rho), \quad (9)$$

with the equality generally only if $\epsilon = 0$, in which case these quantities reduce to the above S_0 , S_∞ , and Δ .

Remark 3. Note that Δ_ϵ has the following “high-fidelity” relation to Δ_0 :

$$\Delta_\epsilon(\rho) = \min \{ \Delta_0(P\rho P) : \text{Tr}(\rho P) \geq 1 - \epsilon \},$$

where the minimization is over all projections P commuting with ρ , extending the definition of Δ_0 to subnormalized density operators. The operators $P\rho P$ can be interpreted as post-measurement states after the event “ P ” has occurred, normalized to the event probability.

More generally, we could allow any $0 \leq B \leq 1$ in the above minimization, such that $\text{Tr}(\rho B) \geq 1 - \epsilon$, and substitute the postmeasurement states $\sqrt{B}\rho\sqrt{B}$. (By a result in [11] this operator has high fidelity to the state ρ .) It is easy to see that the resulting quantity is within a distance of $\log(1 - \epsilon)$ from Δ_ϵ .

We now prove a few lemmas which will together comprise our method of estimating the communication cost, by providing the tools to estimate Δ_ϵ for the appropriate reduced states. We begin with the simple observation that for all states ρ and $\epsilon' < \epsilon < 1$,

$$\Delta_\epsilon(\rho) \geq \log(1 - \epsilon), \quad (10)$$

$$\Delta_\epsilon(\rho) \leq \Delta_{\epsilon'}(\rho). \quad (11)$$

Lemma 1. If for two states ρ and σ $\|\rho - \sigma\|_1 \leq \epsilon$, then

$$\Delta_0(\rho) \geq \Delta_{\sqrt{\epsilon}}(\sigma) + \log(1 - \sqrt{\epsilon})$$

(where $\|\cdot\|_1$ is the trace norm, for self-adjoint operators given by the sum of the absolute values of all eigenvalues, counting multiplicities).

Proof. To begin, denote the eigenvalue lists of ρ and σ by (r) and (s) , respectively, in nonincreasing order. Then, because (see [12])

$$\|(r) - (s)\|_1 \leq \|\rho - \sigma\|_1 \leq \epsilon,$$

we may concentrate on the eigenvalues only. Define, for $\delta = \sqrt{\epsilon}$,

$$J := \{j : (1 - \delta)s_j \leq r_j \leq (1 + \delta)s_j\}.$$

Then, for the complement J^c of J ,

$$\delta s(J^c) = \sum_{j \notin J} \delta s_j \leq \sum_{j \notin J} |r_j - s_j| \leq \epsilon,$$

implying

$$\sum_{j \in J} s_j \geq 1 - \sqrt{\epsilon}.$$

We may clearly assume that s is nonzero on J ; otherwise we can shrink J without affecting the last inequality.

Thus, by the definition of $\Delta_{\sqrt{\epsilon}}$,

$$\log(|J| \max_{j \in J} s_j) \geq \Delta_{\sqrt{\epsilon}}(\sigma).$$

On the other hand, by the definition of J ,

$$j \in J \Rightarrow r_j \neq 0,$$

which implies that $\text{rank } \rho \geq |J|$. Similarly,

$$j \in J \Rightarrow r_j \geq (1 - \sqrt{\epsilon})s_j$$

implies $\max_j r_j \geq (1 - \sqrt{\epsilon}) \max_{j \in J} s_j$. Comparing the last two observations to the definition of $\Delta_0(\rho)$ finishes the proof of the claim. ■

Lemma 2. For any two states τ and ω , and $\epsilon < 1$,

$$\Delta_\epsilon(\tau \otimes \omega) \geq \Delta_{\sqrt{\epsilon}}(\tau) + \log(1 - \sqrt{\epsilon}).$$

Proof. Denote the eigenvalues of τ and ω by t_i and w_k , respectively. Let J be a set of indices i and k such that

$$\Delta_\epsilon(\tau \otimes \omega) = \log(|J| \max_{ik \in J} t_i w_k)$$

and

$$(t \otimes w)(J) = \sum_{ik \in J} t_i w_k \geq 1 - \epsilon. \tag{12}$$

We shall be interested, for certain k , in the sections

$$S_k := \{i : ik \in J\}$$

of J along k , in particular in the set

$$K := \left\{ k : t(S_k) = \sum_{i \in S_k} t_i \geq 1 - \sqrt{\epsilon} \right\}.$$

It follows from the definition of K and the constraint of Eq. (12) that

$$w(K) = \sum_{k \in K} w_k \geq 1 - \sqrt{\epsilon}. \tag{13}$$

The proof is a standard Markov inequality argument: observe that we can rewrite Eq. (12) using the sections

$$1 - \epsilon \leq (t \otimes w)(J) = \sum_k w_k t(S_k).$$

Now the right hand side is a probability average over the values $t(S_k)$, taken with probability w_k . We decompose the sum into two contributions which we estimate separately:

$$\begin{aligned} 1 - \epsilon &\leq \sum_{k \in K} w_k t(S_k) + \sum_{k \notin K} w_k t(S_k) \\ &\leq w(K) + [1 - w(K)](1 - \sqrt{\epsilon}). \end{aligned}$$

Hence $[1 - w(K)]\sqrt{\epsilon} \leq \epsilon$, which is our claim.

Now define

$$J' := \bigcup_{k \in K} S_k \times \{k\},$$

and successively estimate

$$\begin{aligned} |J| \max_{ik \in J} (t_i w_k) &\geq |J'| \max_{ik \in J'} (t_i w_k) \\ &= \sum_{l \in K} |S_l| \max_{ik \in J'} (t_i w_k) \\ &\geq \sum_{k \in K} |S_k| \max_{i \in S_k} (t_i w_k) \\ &= \sum_{k \in K} w_k \left(|S_k| \max_{i \in S_k} t_i \right) \\ &\geq \sum_{k \in K} w_k \exp[\Delta_{\sqrt{\epsilon}}(\tau)] \\ &\geq (1 - \sqrt{\epsilon}) \exp[\Delta_{\sqrt{\epsilon}}(\tau)]; \end{aligned}$$

the second to the last line because of $t(S_k) \geq 1 - \sqrt{\epsilon}$, and the last line by Eq. (13), which proves the lemma. ■

Remark 4. We do not know if a symmetric version of this lemma holds, with an additional term to the right analogous to the one for τ :

$$\Delta_\epsilon(\tau \otimes \omega) \stackrel{?}{\geq} (1 - \epsilon') \Delta_{\epsilon'}(\tau) + (1 - \epsilon') \Delta_{\epsilon'}(\omega) + \epsilon'',$$

with ϵ', ϵ'' functions of ϵ which vanish for $\epsilon \rightarrow 0$.

This would constitute a form of ‘‘quasiadditivity’’ for Δ , since the validity of the analogous reverse inequality

$$\Delta_{2\epsilon}(\tau \otimes \omega) \leq \Delta_{\epsilon}(\tau) + \Delta_{\epsilon}(\omega)$$

is quite easy to see. While it may not be useful to improve on our present results, confirmation of the quasiadditivity would be of conceptual interest.

We are now ready to state our central result, which applies whenever the output state has high Uhlmann fidelity $F(\sigma, \omega) = (\text{Tr} \sqrt{\sigma^{1/2} \omega \sigma^{1/2}})^2$ [13,14] with the desired state, even if the output is mixed.

Theorem 2. Consider a state transformation protocol that takes $|\phi_{AB}\rangle$ to $|\psi_{AB}\rangle$ with fidelity $1 - \epsilon$, exchanging a total of C qubits in the process. Then, with $\delta = \sqrt[8]{4\epsilon}$,

$$2C \geq \Delta_{\delta}(\text{Tr}_B |\psi\rangle\langle\psi|) - \Delta_0(\text{Tr}_B |\phi\rangle\langle\phi|) + 2 \log(1 - \delta).$$

Proof. As in the zero-error case, we follow the increase of Δ_0 over the course of the protocol: After the last communication has taken place, the joint state is $|\tilde{\psi}_{ABA'B'}\rangle$, and we have [compare Eq. (4)]

$$2C \geq \Delta_0(\tilde{\psi}_{AA'}) - \Delta_0(\text{Tr}_B |\phi\rangle\langle\phi|),$$

where $\tilde{\psi}_{AA'} = \text{Tr}_{BB'} |\tilde{\psi}\rangle\langle\tilde{\psi}|$.

Now, since $\text{Tr}_{A'B'} |\tilde{\psi}\rangle\langle\tilde{\psi}|$ has fidelity $1 - \epsilon$ to $|\psi_{AB}\rangle$, we can choose a pure state $|\iota_{A'B'}\rangle$ such that

$$F(|\tilde{\psi}_{ABA'B'}\rangle, |\psi_{AB}\rangle \otimes |\iota_{A'B'}\rangle) \geq 1 - \epsilon.$$

Introducing $\psi_A = \text{Tr}_B |\psi\rangle\langle\psi|$ and $\iota_{A'} = \text{Tr}_{B'} |\iota\rangle\langle\iota|$, we infer, from the monotonicity of the fidelity, that

$$F(\tilde{\psi}_{AA'}, \psi_A \otimes \iota_{A'}) \geq 1 - \epsilon,$$

from which it follows by standard inequalities [12] that

$$\|\tilde{\psi}_{AA'} - \psi_A \otimes \iota_{A'}\|_1 \leq \sqrt{4\epsilon}.$$

Now we can use Lemma 1 to lower bound $\Delta_0(\tilde{\psi}_{AA'})$ in terms of $\Delta_{\sqrt[4]{4\epsilon}}(\psi_A \otimes \iota_{A'})$, which is bounded in turn, using Lemma 2, by $\Delta_{\sqrt[8]{4\epsilon}}(\psi_A)$, which proves the theorem. ■

Using the additivity of the Rényi entropies, and that $S_{\alpha}(1/2) = 1$ for all α , we observe that

$$\Delta_0\left(\rho \otimes \frac{1}{2} \mathbb{1}\right) = \Delta_0(\rho).$$

This implies the following corollary.

Corollary 1. The lower bound on C of Theorem 8 continues to hold even if the starting state $|\phi_{AB}\rangle$ is supplemented by unlimited numbers of EPR pairs. ■

Now suppose that $|\phi_{AB}\rangle$ can be converted into a high-fidelity copy of $|\psi_{AB}\rangle$ using a LOCC protocol in which only C bits are exchanged between Alice and Bob. By consuming EPR pairs for superdense coding [15], this protocol can be converted into a protocol requiring only $C/2$ qubits of communication. Since the lower bound of the corollary applies to

the modified protocol, we conclude that for classical communication our bound can be improved by a factor of 2.

Corollary 2. If the state transformation $|\phi_{AB}\rangle$ to $|\psi_{AB}\rangle$ can be accomplished with fidelity $1 - \epsilon$ by exchanging a total of C classical bits then, with $\delta = \sqrt[8]{4\epsilon}$,

$$C \geq \Delta_{\delta}(\text{Tr}_B |\psi\rangle\langle\psi|) - \Delta_0(\text{Tr}_B |\phi\rangle\langle\phi|) + 2 \log(1 - \delta). \quad \blacksquare$$

Remark 5. Sometimes, direct application of these results can give an overly conservative lower bound because $\Delta_0(\text{Tr}_B |\phi\rangle\langle\phi|)$ can be much larger than the corresponding $\Delta_{\epsilon}(\text{Tr}_B |\phi\rangle\langle\phi|)$.

Here we note that a lower bound on C in terms of Δ_{ϵ} of both the initial and the final state exists: Simply observe that on changing the initial state $|\phi\rangle$ into some state $|\phi'\rangle$ with fidelity $1 - \epsilon_0$, the protocol results in a state ρ' that has fidelity $1 - \epsilon_0$ to ρ (because the fidelity does not decrease under completely positive trace preserving maps), which in turn has fidelity $1 - \epsilon$ to $|\psi\rangle$. By a result of [16] this implies that the transformation from $|\phi'\rangle$ to $|\psi\rangle$ has fidelity $1 - \epsilon'$, with some universal function ϵ' of ϵ and ϵ_0 . We may then apply Theorem 2 to this transformation.

Remark 6. Of course, one can also define a robust version of our previous $\Delta^{\alpha\beta}$ (see Remark 1):

$$\Delta_{\epsilon}^{\alpha\beta}(\rho) := \min \left\{ \frac{\log(\sum_{j \in J} r_j^{\alpha})}{1 - \alpha} - \frac{\log(\sum_{j \in J} r_j^{\beta})}{1 - \beta} \right\},$$

again with minimization over all subsets of indices J such that $\sum_{j \in J} r_j \geq 1 - \epsilon$. Unsurprisingly, a variant of Theorem 2 also holds for this quantity. Consider an entanglement transformation from $|\phi\rangle$ to $|\psi\rangle$ with fidelity $1 - \epsilon$ and a total communication cost of C qubits. Then, for $0 \leq \alpha < 1 < \beta \leq \infty$,

$$2C \geq \Delta_{\delta}^{\alpha\beta}(\text{Tr}_B |\psi\rangle\langle\psi|) - \Delta_0^{\alpha\beta}(\text{Tr}_B |\phi\rangle\langle\phi|) + \delta',$$

with $\delta = \sqrt[8]{4\epsilon}$ and $\delta' = [2\alpha/(1 - \alpha) + 2\beta/(\beta - 1)] \log(1 - \sqrt{\delta})$.

The proof is a slightly more cumbersome version of the proof for the $\Delta_{\delta} = \Delta_{\delta}^{0\infty}$ case.

III. ENTANGLEMENT CONCENTRATION AND DILUTION

In [1] it was shown that, using only local operations, Alice and Bob can convert a state $|\psi_{AB}\rangle^{\otimes n}$ to a high fidelity approximation of $|\phi_2^+\rangle^{\otimes nE(\psi) - O(\sqrt{n})}$. We reproduce the argument here, as the relevant concepts are used again in the dilution protocol and our lower bound.

Diagonalize $\rho_A = \text{Tr}_B |\psi\rangle\langle\psi| = \sum_{i=1}^d r_i |e_i\rangle\langle e_i|$. For a distribution P on $\{1, \dots, d\}$ we can introduce the *type classes* of sequences $i^n = i_1 \cdots i_n$:

$$\mathcal{T}_P^n := \{i^n : \forall i \ N(i|i^n) = nP(i)\},$$

where $N(i|i^n)$ counts the number of occurrences of i in i^n . The number of nonempty type classes is $\binom{n+d-1}{d-1} \leq (n+1)^d$, and the corresponding P are called n types.

For $\delta > 0$ we have the set of typical sequences

$$\mathcal{T}_{r,\delta}^n := \bigcup \left\{ \mathcal{T}_P^n : P \text{ such that } \forall i |P_i - r_i| \leq \frac{\delta \sqrt{r_i(1-r_i)}}{\sqrt{n}} \right\}.$$

Standard facts about these concepts are to be found in [17] (see also [18]):

$$r^{\otimes n}(\mathcal{T}_{r,\delta}^n) \geq 1 - \frac{d}{\delta^2}, \tag{14}$$

$$\forall i^n \in \mathcal{T}_P^n, r^{\otimes n}(i^n) = \exp\{-n[D(P||r) + H(P)]\}, \tag{15}$$

with the relative entropy (or entropy divergence) $D(P||r) = \sum_i P_i \log(P_i/r_i)$. Furthermore,

$$|\mathcal{T}_{r,\delta}^n| \leq \exp[nH(r) + Kd\delta\sqrt{n}], \tag{16}$$

$$|\mathcal{T}_P^n| \leq \exp[nH(P)], \tag{17}$$

$$|\mathcal{T}_P^n| \geq (n+1)^{-d} \exp[nH(P)], \tag{18}$$

$$|\mathcal{T}_P^n| \geq \exp[nH(r) - Kd\delta\sqrt{n}] \text{ if } P \text{ typical}, \tag{19}$$

for an absolute constant $K > 0$. These sets allow for the definition of corresponding projectors $\Pi_P^n := \sum_{i^n \in \mathcal{T}_P^n} |e_{i^n}\rangle\langle e_{i^n}|$, and similarly $\Pi_{r,\delta}^n$, with probability and trace relations identical to Eqs. (14)–(19). Note that $H(r) = S(\rho)$, by definition.

The concentration protocol only requires Alice and Bob to each independently perform the projective measurement $(\Pi_P^n)_{P \text{ n-type}}$. (Without loss of generality $|\psi\rangle$ is in Schmidt diagonal form, and the bases with respect to which the projectors are defined are identical eigenbases of the reduced states.) The result P will be the same for Alice and Bob, and by Eq. (14) it will be typical with probability $\geq 1 - \epsilon$ (choosing δ large enough). Moreover, by Eq. (19) the resulting states $|\phi_P\rangle$ are maximally entangled states of Schmidt rank $\geq \exp[nH(r) - Kd\delta\sqrt{n}]$. Local measurements, corresponding to a partition of \mathcal{T}_P^n into blocks of size 2^m (and a remainder of smaller size), for $m = \lfloor nH(r) - Kd\delta\sqrt{n} + \log \epsilon \rfloor$, project this further down to a state isomorphic to $|\phi_2^+\rangle^{\otimes m}$, with probability $1 - \epsilon$. This shows that $|\psi\rangle^{\otimes n}$ can be converted by local operations into m EPR pairs, with fidelity $1 - 2\epsilon$, establishing that asymptotically $|\psi\rangle$ is worth $E(\psi)$ EPR pairs.

In the same work it was demonstrated that the reverse is true as well: using LOCC, $|\phi_2^+\rangle^{\otimes nE(\psi) + O(\sqrt{n})}$ can be converted to a high fidelity approximation of $|\psi_{AB}\rangle^{\otimes n}$.

Alice simply prepares the state $\Pi_{r,\delta}^n |\psi\rangle^{\otimes n}$ (properly normalized) locally. By Eq. (16) it has Schmidt rank $\leq \exp[nH(r) + Kd\delta\sqrt{n}]$, enabling Alice to teleport [19] the half intended for Bob using $nH(r) + Kd\delta\sqrt{n}$ EPR pairs.

Note that this method requires communication of $2nE(\psi) + O(\sqrt{n})$ classical bits from Alice to Bob, which is of the order of the entanglement manipulated. Whether this amount can be reduced is, therefore, a legitimate and interesting question. In [8] it was shown that, indeed, communication of $O(\sqrt{n})$ classical bits is sufficient, by the following method.

The authors of [8] demonstrated that there exists a state $|\chi\rangle$ entangling $O(\sqrt{n})$ qubits and local unitaries U_A and U_B such that

$$F((U_A \otimes U_B)|\psi\rangle^{\otimes n}, |\phi_2^+\rangle^{\otimes nE - O(\sqrt{n})} \otimes |\chi\rangle) \geq 1 - \epsilon. \tag{20}$$

This state arises naturally by looking at what was done in the concentration procedure above, in a reversible setting. Applying the same dilution procedure as before but to the smaller state $|\chi\rangle$, that is, local preparation by Alice followed by teleportation of Bob’s share, then consumes only $O(\sqrt{n})$ ebits and twice that amount of classical communication (as Lo [2] has shown this factor can be reduced to 1, i.e., a state of Schmidt rank d can be prepared using $\log d$ bits of entanglement and communicating $\log d$ classical bits).

Let us now apply our main result to show that any protocol to create $|\psi\rangle^{\otimes n}$ up to fidelity $1 - \epsilon$ from EPR pairs must use $\Omega(\sqrt{n})$ bits of communication.

Noting first that EPR pairs have $\Delta_0 = 0$, we have only to lower bound $\Delta_\delta(\psi_A^{\otimes n})$ in order to make use of Theorem 2. This we do by using Eq. (9). First, we show that

$$S_{\infty,\delta}(\psi_A^{\otimes n}) \leq nE(\psi) - D(\epsilon)\sqrt{n} + o(\sqrt{n}),$$

with a constant $D(\epsilon) > 0$ (for $\delta = \sqrt[8]{4\epsilon} < 1/2$). Observe that $S_{\infty,\delta}$ is particularly easy to understand; it is the negative logarithm of the largest eigenvalue such that the sum of the eigenvalues exceeding this one is bounded by δ .

Define the independent and identically distributed (i.i.d.) random variables $X_j, j = 1, \dots, n$, by letting

$$\text{Prob}\{X_j = -\log r_i\} = r_i,$$

where (r_i) are the Schmidt coefficients of $|\psi\rangle$. Note that their expectation $\mathbb{E}X_j$ equals $E(\psi)$, and that they are nonconstant, unless $|\psi\rangle$ is maximally entangled, so that the variance σ^2 is nonzero.

Hence we can apply the central limit theorem

$$\text{Prob}\left\{ \sum_{j=1}^n X_j \leq nE(\psi) + x\sigma\sqrt{n} \right\} \rightarrow \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt.$$

This implies that the sum of the largest eigenvalues, from $\exp[-nE(\psi) + D(\epsilon)\sqrt{n} + o(\sqrt{n})]$ up (including multiplicities), is bounded from below by δ , and our claim follows.

Next, we lower bound $S_{0,\delta}(\psi_A^{\otimes n})$. An optimal set J in the definition Eq. (6) must consist of the indices of the $|J|$ largest eigenvalues such that their sum is barely above $1 - \epsilon$.

Once more invoking the central limit theorem, the sum of the smallest eigenvalues (including multiplicities) of $\psi_A^{\otimes n}$ up to $\exp[-nE(\psi) - D(\epsilon)\sqrt{n} + o(\sqrt{n})]$ is at least δ .

We exhibit now a large type class inside the set corresponding to larger eigenvalues, which by the preceding is a subset of J : there exists an n -type P such that $|P_i - r_i| \leq 1/n$, for all i . This entails that for $i^n \in \mathcal{T}_P^n$,

$$\begin{aligned} \log r^{\otimes n}(i^n) &= n \sum_i P_i \log r_i \\ &= n \sum_i \left(r_i \pm \frac{1}{n} \right) \log r_i \\ &= -nH(r) \pm \sum_i |\log r_i| \\ &= -nE(\psi) \pm C. \end{aligned}$$

Thus, $T_P^n \subset J$ as soon as $D(\epsilon) > 0$ and n is large enough.

On the other hand, because $\|P - r\|_1 \leq d/n$, we have (using a well-known estimate for Shannon entropies, see [17]) that

$$|H(P) - H(r)| \leq \frac{d}{n} \log n,$$

and we conclude, by Eq. (18), that

$$\begin{aligned} |J| &\geq |T_P^n| \\ &\geq (n+1)^{-d} \exp[nH(P)] \\ &\geq \exp[nH(r) - d \log n - d \log(n+1)]. \end{aligned}$$

It follows that $S_{0,\delta}(\psi_A^{\otimes n}) \geq nE(\psi) - O(\log n)$.

Combining the estimates of $S_{0,\delta}$ and $S_{\infty,\delta}$, we obtain the following theorem.

Theorem 3. For every bipartite pure state $|\psi_{AB}\rangle$ that is neither separable nor maximally entangled and every sufficiently small ϵ there exists a positive constant $D(\epsilon)$ such that the communication cost of creating $|\psi\rangle^{\otimes n}$ up to fidelity $1 - \epsilon$ from EPR pairs is at least $C \geq D(\epsilon) \sqrt{n} - o(\sqrt{n})$. ■

Remark 7. Recently, secret shared randomness has been proposed as a ‘‘classical analog of entanglement’’ [20], partly to increase intuition on entanglement transformations, and partly to be able to distinguish the quantum effects of entanglement from those that are statistically explainable.

Specifically, pure-state entanglement was considered analogous to classical *perfect* correlation: Alice and Bob share a joint random variable (X, Y) , where X belongs to Alice, Y to Bob, and $X = Y$ with probability 1. Entanglement transformations by LOCC have their analog in transformations of these random variables by local (classical) actions and *public discussion*, which can be listened to by an eavesdropper. The analogs of EPR pairs are shared random bits: $\text{Prob}\{X = Y = 0\} = \text{Prob}\{X = Y = 1\} = 1/2$.

Now it is an easy result of the theory of shared randomness (see [21] for definitions) that in an i.i.d. setting (X, Y) can be asymptotically converted into the Shannon entropy $H(X)$ of X shared secret bits and, inversely, this amount of shared randomness can be used to generate (X, Y) : More precisely, both transformations can be performed with asymptotically vanishing total variational distance of the distributions. These operations are the classical analog of entanglement concentration and dilution.

What is remarkable is that in this setting both the concentration and dilution processes require no public discussion

whatsoever. Thus, our $\Omega(\sqrt{n})$ lower bound is a purely quantum phenomenon that has no counterpart in the ‘‘classical analog.’’

Remark 8. In [25] asymptotic entanglement concentration with exponential fidelity bounds was considered: no classical communication is needed here, either. On the other hand, dilution according to a generalization of the method in [8] is possible with $O(n)$ bits of communication, if the same exponential fidelity bound is imposed. An argument similar to the one that led to Theorem 3 shows that $\Omega(n)$ bits are necessary, though the constant does not appear to be optimal.

IV. STATES WITH LARGE COMMUNICATION COST

In [22], the states

$$|\mu(n)\rangle = \frac{1}{\sqrt{H_n}} \sum_{i=1}^n \frac{1}{\sqrt{i}} |i\rangle |i\rangle,$$

with the harmonic sum $H_n = \sum_{i=1}^n (1/i)$, were introduced to show that the concept of ‘‘approximate pure state transformations with unlimited catalysis’’ allows *any* state transformation (this was dubbed ‘‘embezzling entanglement’’ in [22]). In particular, it was shown that for every pure state $|\phi\rangle$ of Schmidt rank m there are local isometries U_A and U_B such that

$$F(|\mu(n)\rangle \otimes |\phi\rangle, (U_A \otimes U_B)|\mu(n)\rangle) \geq 1 - \frac{\log m}{\log n}.$$

It is straightforward to verify that the entanglement of $|\mu(n)\rangle$ is asymptotically $1/2 \log n$, and we shall demonstrate here that the communication cost to produce it from EPR pairs is of the same order:

Theorem 2 asks us to lower bound Δ_δ of Alice’s reduced state

$$\rho_A = \frac{1}{H_n} \sum_{i=1}^n \frac{1}{i} |i\rangle \langle i|,$$

which we do using Eq. (9):

$$S_{0,\delta}(\rho_A) = \log \min \left\{ k : \sum_{i=k+1}^n \frac{1}{iH_n} \leq \delta \right\}, \quad (21)$$

$$S_{\infty,\delta}(\rho_A) = \log H_n + \log \max \left\{ k : \sum_{i=1}^{k-1} \frac{1}{iH_n} \leq \delta \right\}. \quad (22)$$

Now, asymptotically $(\log n) - 1 \leq H_n \leq \log(n+1)$, and Eqs. (21) and (22) allow us to estimate

$$\Delta_\delta(\rho_A) \geq [(1 - 2\delta) \log n] - 4 - \log \log(n+1), \quad (23)$$

resulting in a lower bound

$$C \geq \left(\frac{1}{2} - \sqrt[8]{4\epsilon} \right) \log n - O(\log \log n)$$

for the communication cost to create $|\mu(n)\rangle$ up to fidelity $1 - \epsilon$ from EPR pairs. In fact, the classical communication cost is, by Corollary 2, lower bounded by $[1 - o(1)] \log n$, asymptotically matching the upper bound $\log n$ from Lo's earlier mentioned state preparation method in [2].

Other states with entanglement of the same order as the communication necessary to create them are the $|\chi\rangle$ of Eq. (20): their entanglement is at most $O(\sqrt{n})$ while Theorem 3 implies a lower bound of $\Omega(\sqrt{n})$ on the communication resources.

V. CONCLUSION

We have exhibited a quantitative lower bound on the communication cost of general entanglement transformations. It is good enough to prove that the Lo-Popescu protocol of entanglement dilution is within a constant factor of being optimal, requiring $\Theta(\sqrt{n})$ bits of communication. Also, it can be used to show that there exist states whose communication cost for creation from EPR pairs is of the same order as their entanglement, making local preparation and teleportation essentially the optimal strategy.

It is unknown to us how tight our lower bound can be made or if there is an upper bound involving similar quantities, so we leave these questions open for future research. On a different note, it has repeatedly been speculated (as in [8]) that the classical communication cost is related to the *loss of entanglement* in a transformation. Observe that this seems to fit perfectly for concentration and dilution, and it might be that in an appropriate model the entanglement loss in a pure

state transformation provides an upper bound on the minimal communication cost required to perform it.

Other applications may include the study of quantum communication complexity, where a technique for lower bounding the communication exists [23,24,10] that requires estimation of the communication cost of certain pure-state entanglement transformations. In the cited works this cost was lower bounded by observing that some measure of entanglement has increased. Our method could be useful as it gives nontrivial lower bounds even when the entanglement remains constant or decreases, and continues to be effective in the presence of unlimited numbers of EPR pairs.

Recently, the independent work of Harrow and Lo [26] came to our attention, which proves the $\Omega(\sqrt{n})$ lower bound on entanglement dilution by a different method (although there are similarities) that simultaneously provides a lower bound on the entanglement loss.

ACKNOWLEDGMENTS

We thank Wim van Dam for his suggestion to also consider general Rényi entropies, and Karol and Michal Horodecki for posing the problem solved in Remark 2. We want to thank Aram Harrow and Hoi-Kwong Lo for making their draft of [26] available to us and for stimulating discussions. P.H. was supported by National Science Foundation Grant No. EIA-0086038 and a grant from the Sherman Fairchild Foundation. A.W. is supported by the U.K. Engineering and Physical Sciences Research Council. This work was carried out during the second author's visit to the Institute of Quantum Information, Caltech.

-
- [1] C. H. Bennett, H. J. Bernstein, S. Popescu, and B. W. Schumacher, Phys. Rev. A **53**, 2046 (1996).
- [2] H.-K. Lo, Phys. Rev. A **62**, 012313 (2000).
- [3] G. Vidal, D. Jonathan, and M. A. Nielsen, Phys. Rev. A **62**, 012304 (2000).
- [4] M. A. Nielsen, Phys. Rev. Lett. **83**, 436 (1999).
- [5] L. Hardy, Phys. Rev. A **60**, 1912 (1999).
- [6] H.-K. Lo and S. Popescu, Phys. Rev. A **63**, 022301 (2001).
- [7] J. G. Jensen and R. Schack, Phys. Rev. A **63**, 062303 (2001).
- [8] H.-K. Lo and S. Popescu, Phys. Rev. Lett. **83**, 1459 (1999).
- [9] A. Rényi, in *Proceedings of the 4th Berkeley Symposium on Mathematics, Statistics and Probability* (University of California Press, Berkeley, CA, 1961), Vol. 1, pp. 547–561.
- [10] W. van Dam and P. Hayden, e-print quant-ph/0204093.
- [11] A. Winter, IEEE Trans. Inf. Theory **45**, 2481 (1999).
- [12] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, U.K., 2000).
- [13] R. Jozsa, J. Mod. Opt. **41**, 2315 (1994).
- [14] A. Uhlmann, Rep. Math. Phys. **9**, 273 (1976).
- [15] C. H. Bennett and S. Wiesner, Phys. Rev. Lett. **69**, 2881 (1992).
- [16] H. Barnum, C. A. Fuchs, R. Jozsa, and B. Schumacher, Phys. Rev. A **54**, 4707 (1996).
- [17] J. Wolfowitz, *Coding Theorems of Information Theory*, 2nd ed. (Springer-Verlag, Berlin, 1964).
- [18] I. Csiszar and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems* (Academic Press, London, 1981).
- [19] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, Phys. Rev. Lett. **70**, 1895 (1993).
- [20] D. Collins and S. Popescu, Phys. Rev. A **65**, 032321 (2002).
- [21] R. Ahlswede and I. Csiszár, IEEE Trans. Inf. Theory **39**, 1121 (1993).
- [22] W. van Dam and P. Hayden, e-print quant-ph/0201041.
- [23] A. Ambainis, L. Schulman, A. Ta-Shma, U. Vazirani, and A. Wigderson, in Proceedings of 39th FOCS (unpublished), pp. 342–351.
- [24] R. Cleve, W. van Dam, M. A. Nielsen, and A. Tapp, in *Proceedings of the 1st NASA International Conference on Quantum Computing and Quantum Communications*, edited by Colin P. Williams, (Springer Verlag, Berlin, 1998), pp. 61–74.
- [25] M. Hayashi, M. Koashi, K. Matsumoto, F. Morikoshi, and A. Winter, J. Phys. A: Math. Gen. **36**, 527 (2003).
- [26] A. Harrow and H.-K. Lo, e-print quant-ph/0204096.