

Entanglement monotones for W -type states

Eric Chitambar,^{*} Wei Cui,[†] and Hoi-Kwong Lo[‡]

Center for Quantum Information and Quantum Control (CQIQC), Department of Physics and Department of Electrical & Computer Engineering, University of Toronto, Toronto, Ontario M5S 3G4, Canada

(Received 8 July 2011; revised manuscript received 10 April 2012; published 19 June 2012)

In this article, we extend recent results concerning random-pair Einstein-Podolsky-Rosen distillation and the operational gap between separable operations (SEPs) and local operations with classical communication (LOCC). In particular, we consider the problem of obtaining bipartite maximal entanglement from an N -qubit W -class state (i.e., that of the form $\sqrt{x_0}|00\dots 0\rangle + \sqrt{x_1}|10\dots 0\rangle + \dots + \sqrt{x_n}|00\dots 1\rangle$) when the target pairs are *a priori* unspecified. We show that when $x_0 = 0$, the optimal probabilities for SEPs can be computed using semidefinite programming. On the other hand, to bound the optimal probabilities achievable by LOCC, we introduce entanglement monotones defined on the N -qubit W class of states. The LOCC monotones we construct can be increased by SEPs, and in terms of transformation success probability, we are able to quantify a gap as large as 37% between the two classes. Additionally, we demonstrate transformations $\rho^{\otimes n} \rightarrow \sigma^{\otimes n}$ that are feasible by SEP for any n but impossible by LOCC.

DOI: [10.1103/PhysRevA.85.062316](https://doi.org/10.1103/PhysRevA.85.062316)

PACS number(s): 03.67.Mn, 03.67.Hk, 03.65.Ud

I. INTRODUCTION

Quantum entanglement is a celebrated aspect of quantum theory and represents one of the sharpest departures from the classical world. From a practical perspective, entanglement provides a key tool for novel technologies such as quantum teleportation [1], dense coding [2], and entanglement-based quantum cryptography [3]. Formally treating entanglement as a physical resource involves specifying a quantitative measure so that it makes sense to discuss “how much” entanglement a certain quantum system possesses. For bipartite pure states, the von Neumann entropy serves as the unequivocal measure of entanglement [4]. However, for multipartite pure states and even mixed bipartite states, there does not appear to exist one unifying entanglement measure [5,6]. Instead, it seems more appropriate to quantify the amount of entanglement in a given system relative to some particular task or physical characteristic.

A necessary (and arguably sufficient) property that every entanglement measure must satisfy is the so-called *local operations and classical communication (LOCC) constraint* [7–12]. In a realistic multipartite setting, each party will possess a laboratory in which he or she performs quantum measurements on only one part of the whole system. At the same time, the parties may wish to coordinate their measurement strategies by using a classical communication channel to share their measurement outcomes. This paradigm is known as LOCC, and it describes the basic setting for nearly all practical quantum communication schemes. The LOCC constraint means that entanglement cannot be increased on average by LOCC. Therefore, a function μ fulfills the LOCC constraint if, for any LOCC process that converts ρ into σ_i with probability p_i , the following inequality holds: $\mu(\rho) \geq \sum_i p_i \mu(\sigma_i)$.

While it is very easy to describe the idea of LOCC operations, giving a precise mathematical description is notoriously difficult [13–15]. For many purposes—such as upper bounding the success probability of some LOCC task—a finely tuned description is not necessary. Instead, one can turn to a more general (but not too general) class of quantum operations and see what is possible under this relaxation. The most natural approximation to LOCC is the class of separable operations (SEPs). For an N -partite quantum system, an operation is called separable if it admits a Kraus operator representation $\mathcal{E}(\cdot) = \sum_\lambda A_\lambda(\cdot)A_\lambda^\dagger$, where $A_\lambda = M_{1,\lambda} \otimes M_{2,\lambda} \otimes \dots \otimes M_{N,\lambda}$. As every LOCC is built by a successive composition of local maps $\mathcal{E}^{(k)} \otimes \mathbb{I}^{(\bar{k})}$, it follows that every LOCC map is separable. Compared to a LOCC, the structure of a SEP is easier to analyze, and studying it has been useful for proving LOCC impossibility results [7,16–20].

A somewhat unexpected finding is the existence of SEPs that cannot be implemented by LOCC [13]. A dramatic example of this is the phenomenon of “nonlocality without entanglement,” which refers to certain sets of product states that can be distinguished by SEP but not by LOCC [13,21]. Following the initial finding that LOCC \subsetneq SEP, additional examples were constructed that demonstrated this fact [22–26]. Like LOCC, SEPs have the property that they cannot generate entanglement. Indeed, if a separable map is applied to a general separable state $\sum_i p_i \rho_i^{(1)} \otimes \dots \otimes \rho_i^{(N)}$, the resultant state will likewise be separable. The fact that LOCC \neq SEP then implies a certain irreversibility to non-LOCC separable maps since these operations are unable to create entanglement, but nevertheless, they require some preshared entanglement to be performed in the multipartite setting. Thus, such maps may be interpreted as the operational analog to “bound entanglement” [19], where the latter refers to multipartite states that cannot be converted into pure entanglement but, nevertheless, require some initial entanglement to be created. Consequently, studying the gap between LOCC and SEP is crucial to understanding the nature of quantum entanglement.

Unfortunately, very little quantitative research has been conducted into the difference between LOCC and SEP. Thus

^{*}e.chitambar@utoronto.ca

[†]cuiwei@physics.utoronto.ca

[‡]hklo@comm.utoronto.ca

it becomes difficult to say just how much more powerful a SEP is than LOCC. Previous numerical results that compared SEP versus LOCC for the task of distinguishing certain quantum states was very small in scale. For instance, Ref. [13] demonstrated a minimum of $O(10^{-6})$ between the two classes (in terms of the attainable mutual information), while in Ref. [23], optimal success probabilities in distinguishability were shown to diverge by at most 0.8%. Recently, however, we were able to provide the first appreciable gap between SEP and LOCC in terms of a 12.5% difference in the probability of successfully performing a particular state transformation [27]. In this article we vastly improve on our previous result and demonstrate a difference of 37% between LOCC and SEPs. The key step in proving this result is the construction of new entanglement monotones for a particular subset of N -qubit states that can be increased by SEP.

Specifically, we turn to the problem of randomly distilling an EPR pair from one copy of a multipartite W -class state, as first initiated by Fortescue and Lo [28,29]. An EPR random distillation refers to a transformation of multipartite entanglement into bipartite maximal pure entanglement in which the two parties sharing the final entanglement are allowed to vary among the different outcomes. We denote such a transformation

$$|\varphi\rangle_{1\dots N} \rightarrow \{p_{ij}, |\Phi^{(ij)}\rangle\}, \quad (\star)$$

where $|\Phi^{(ij)}\rangle$ is a maximally entangled two-qubit state shared between party i and party j obtained from $|\varphi\rangle_{1\dots N}$ with probability p_{ij} .

It is often more convenient to represent transformation (\star) by a *distillation configuration graph* $\mathcal{G} = (V, E \subset V \times V)$ in which each party i is assigned to a vertex $v_i \in V$, and an edge $(i, j) \in E$ is drawn between v_i and v_j if and only if p_{ij} is nonzero (see Fig. 1). Let $E_k \subset E$ denote the set of edges connected to vertex v_k .

In terms of overall success probability, often the random distillation of some state can be more efficient than if entanglement is distilled to a fixed pair. Perhaps the most impressive demonstration of this effect is the *Fortescue-Lo protocol*, which performs transformations (\star) on the three-qubit W state $|W_3\rangle = \sqrt{1/3}(|100\rangle + |010\rangle + |001\rangle)$ for any value of $p_{12} + p_{23} + p_{13}$ less than 1; this should be compared to

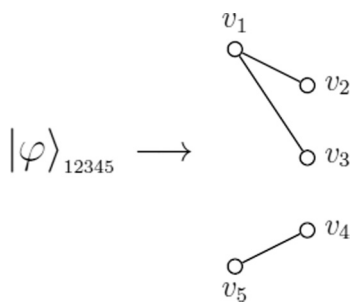


FIG. 1. Graph representation \mathcal{G} of one particular EPR random distillation configuration for state $|\varphi\rangle_{12345}$. Each edge represents a possible outcome EPR state shared between the two parties corresponding to the connected nodes. The probability of obtaining a given edge is $p_{ij} > 0$, and E_k is the set of all edges connected to vertex v_k .

the maximum probability for transformation $|W_3\rangle \rightarrow |\Phi^{(ij)}\rangle$, which is $2/3$ for any $i \neq j \in \{1, 2, 3\}$ [28]. The Fortescue-Lo protocol also extends to distilling an EPR pair from $|W_N\rangle$ with a probability arbitrarily close to 1. Such a finding demonstrates the importance of considering random distillations in the multipartite setting.

1. Summary of main results and article outline

This article compares the LOCC versus SEP feasible probabilities of transformation (\star) when $|\varphi\rangle_{1\dots N}$ belongs to the N -qubit W class of states, i.e., any state reversibly obtainable from $|W_N\rangle = \sqrt{1/N}(|10\dots 0\rangle + |01\dots 0\rangle + \dots + |00\dots 1\rangle)$ by LOCC with a nonzero probability. We begin our investigation in Sec. II with a review of results by Kintaş and Turgut on the subject of W -class transformations [30]. There we also define the notation used throughout the paper.

In Sec. III we show that for states of the form $\sqrt{x_1}|10\dots 0\rangle + \dots + \sqrt{x_N}|00\dots 1\rangle$, the possibility of transformation (\star) by SEP can be phrased as a semidefinite programming feasibility question. Thus, numerically it has an efficient solution. When the initial state is $|W_N\rangle$, we are able to obtain simple necessary and sufficient criteria for transformation feasibility by studying the dual problem, as carried out in the Appendix. Note that the results of this section also provide LOCC upper bounds.

Next, in Sec. IV, we turn to the LOCC setting specifically, and we introduce two new types of entanglement monotones defined on the N -qubit W -class of states. To prove that these functions are monotonic under LOCC, we decompose a general LOCC transformation into a sequence of local weak measurements. However, these functions are not monotonic under SEPs. While a general separable measurement can also be decomposed into a sequence of weak measurements, these measurements need not be local, and our functions are sensitive precisely to this relaxation in constraint. We prove that the monotones have operational meanings as the supremum success probabilities for the distillation of EPR states for certain distillation configuration graphs. Moreover, these monotones can be saturated by an “equal or vanish” measurement scheme, which we describe further in Sec. IV. Thus we are able to prove LOCC optimal rates for certain configuration graphs of transformation (\star) . In particular, we solve the one-shot analog of “entanglement combing” studied by Yang and Eisert [31], in which one particular party is selected to be a shareholder of the bipartite entanglement for each of the possible outcomes. Formal comparisons between SEP and LOCC for transformation (\star) are made in Sec. V.

Finally, in Sec. VI we move beyond the single-copy case and investigate a particular n -copy random distillation problem. Interestingly, we are able to show the existence of a state transformation $|\psi\rangle\langle\psi|^{\otimes n} \rightarrow \rho^{\otimes n}$ that, for any n , is impossible by LOCC but always possible by SEP. This result is the first of its kind. Brief concluding remarks are then given in Sec. VII.

2. Relationship to previous work

This article complements recent work we conducted on the random distillation problem and its connection to the structure of LOCC [26,27]. In particular, Ref. [26] presents a general LOCC procedure for completing transformation (\star) on W -class

states and also computes tight bounds for four-qubit systems. The distinguishing feature of this article is a solution to (\star) by SEPs for a wide class of states and the construction of N -partite entanglement monotones that generalize those presented in [27]. Additionally, we consider here the many-copy variant of the random distillation problem, which has previously only been investigated in Ref. [29].

II. NOTATION AND THE KINTAŞ AND TURGUT MONOTONES

Throughout the paper, we deal exclusively with pure states $|\varphi\rangle_{1\dots N}$. If we wish to express the state as the rank 1 density operator $|\varphi\rangle\langle\varphi|_{1\dots N}$, we will denote it $\varphi^{(1\dots N)}$. For some operator A acting on a multipartite state space, we let A^{Γ_i} denote its partial transpose in the computational basis with respect to a party (or parties) i .

It is often useful to consider two states equivalent if they can be reversibly converted from one to the other by LOCC with some nonzero probability. Such a transformation is known as stochastic LOCC (SLOCC), and the well-known criterion for $|\varphi\rangle_{1\dots N} \xrightarrow{\text{SLOCC}} |\varphi'\rangle_{1\dots N}$ is the existence of invertible $M^{(k)}$ such that $\bigotimes_{k=1}^N M^{(k)}|\varphi\rangle_{1\dots N} = |\varphi'\rangle_{1\dots N}$ [32]. In this way, multipartite state space can then be partitioned into SLOCC equivalence classes.

The N -party W class is the set of states SLOCC equivalent to $|W_N\rangle = \sqrt{1/N}(|10\dots 0\rangle + |01\dots 0\rangle + |00\dots 1\rangle)$, and such states take the form $\sqrt{x_0}|00\dots 0\rangle + \sqrt{x_1}|10\dots 0\rangle + \dots + \sqrt{x_N}|00\dots 1\rangle$. More importantly, even after a local unitary (LU) transformation, $|0\rangle \rightarrow |0'\rangle$ and $|1\rangle \rightarrow |1'\rangle$, the component values $\sqrt{x_i}$ always remain unchanged for $N \geq 3$ [30]. Therefore, we can uniquely characterize any W -class state by the N -component vector,

$$\vec{x} = (x_1, x_2, \dots, x_N) \updownarrow \sqrt{x_0}|00\dots 0\rangle + \sqrt{x_1}|10\dots 0\rangle + \dots + \sqrt{x_N}|00\dots 1\rangle, \quad (1)$$

and $x_0 = 1 - \sum_{i=1}^N x_i$. When $N = 2$, uniqueness can be ensured by demanding that $x_0 = 0$ and $x_1 \geq x_2$.

The order in value of these components is highly important to our investigation. Thus, we often use the indices $\{n_1, n_2, \dots, n_N\} = \{1, 2, \dots, N\}$ such that $x_{n_1} \geq x_{n_2} \geq \dots \geq x_{n_N}$. We let $n_1(\vec{x})$ denote the largest component in state $\vec{x} = (x_1, \dots, x_N)$.

A main result of Kintaş and Turgut’s work was proving that the component values, $-x_0$ and x_i for $1 \leq i \leq N$, are entanglement monotones [30]. In other words, for an LOCC transformation converting $\vec{x} \rightarrow \vec{x}_\lambda$ with probability p_λ , the relations

$$x_0 \leq \sum_\lambda p_\lambda x_{\lambda,0}, \quad x_i \geq \sum_\lambda p_\lambda x_{\lambda,i} \quad (2)$$

hold for all $1 \leq i \leq N$. We refer to these as the K - T monotones and they place an upper bound of $\min\{x_i/y_i\}_{i=1\dots N}$ on the probability for any W -class transformation $\vec{x} \rightarrow \vec{y}$. Recently, necessary and sufficient conditions were obtained for when this upper bound can be achieved [33].

To study the effects of measurement on a W -class state, first note that any measurement operator A is a 2×2 matrix expressible in the form $A = U \cdot \begin{pmatrix} \sqrt{a} & b \\ 0 & \sqrt{c} \end{pmatrix}$, where U

is a unitary matrix. Thus, up to a final LU operation, any local measurement corresponds to a set of upper triangular matrices $\{M_\lambda\}_\lambda$ with $\sum_\lambda M_\lambda^\dagger M_\lambda = \mathbb{I}$. When it is party k who performs the measurement, we denote the measurement operators $M_\lambda^{(k)}$. It is easy to see that this measurement on state $\sqrt{x_0}|00\dots 0\rangle + \sqrt{x_1}|10\dots 0\rangle + \dots + \sqrt{x_N}|00\dots 1\rangle$ will transform the components as

$$x_k \rightarrow \frac{c_\lambda}{p_\lambda} x_k, \quad x_j \rightarrow \frac{a_\lambda}{p_\lambda} x_j, \quad 1 \leq j \neq k \leq N, \quad (3)$$

where p_λ is the probability that outcome λ occurs. We can simplify matters even further by noting that any transformation possible by LOCC can always be achieved by a protocol in which each party only performs two-outcome measurements [34]. Since our chief concern is the possibility of transformations, we can assume without loss of generality that each local measurement consists of two upper triangular matrices $\{M_1^{(k)}, M_2^{(k)}\}$ whose entries are

$$M_1^{(k)} = \begin{pmatrix} \sqrt{a_1} & b_1 \\ 0 & \sqrt{c_1} \end{pmatrix}, \quad M_2^{(k)} = \begin{pmatrix} \sqrt{a_2} & b_2 \\ 0 & \sqrt{c_2} \end{pmatrix}, \quad (4)$$

with $a_1 + a_2 = 1$ and $c_1 + c_2 \leq 1$, in which equality is achieved by the latter if and only if $M_1^{(k)}$ and $M_2^{(k)}$ are both diagonal.

III. SEPARABLE TRANSFORMATIONS

In this section we derive the conditions for which transformation (\star) is possible by SEPs when the initial state is a W -class state with $x_0 = 0$. As shown in the following lemma, the unique structure of such states allows for a major simplification in the analysis.

Lemma 1. Suppose that $\{\Pi_\lambda := M_\lambda^{(1)} \otimes \dots \otimes M_\lambda^{(N)}\}_{\lambda=1\dots t}$ corresponds to a complete measurement that achieves transformation (\star) with probabilities p_{12}, \dots, p_{N-1N} when $|\varphi\rangle_{1\dots N} = \sqrt{x_1}|10\dots 0\rangle + \dots + \sqrt{x_N}|00\dots 1\rangle$. Then up to LU operations, there exists a measurement $\{\hat{M}_\lambda^{(1)} \otimes \dots \otimes \hat{M}_\lambda^{(N)}\}_{\lambda=1\dots 2t}$ that achieves transformation (\star) with the same probabilities and with each $\hat{M}_\lambda^{(k)}$ being diagonal.

Proof. Up to an LU operation, each $M_\lambda^{(k)}$ takes the form $M_\lambda^{(k)} = \begin{pmatrix} \sqrt{a_{\lambda k}} & b_{\lambda k} \\ 0 & \sqrt{c_{\lambda k}} \end{pmatrix}$ so that

$$\sum_\lambda \Pi_\lambda^\dagger \Pi_\lambda = \sum_\lambda \bigotimes_{k=1}^N \begin{pmatrix} a_{\lambda k} & \sqrt{a_{\lambda k}} b_{\lambda k} \\ \sqrt{a_{\lambda k}} b_{\lambda k}^* & |b_{\lambda k}|^2 + c_{\lambda k} \end{pmatrix} = \mathbb{I}. \quad (5)$$

Let $\hat{M}_\lambda^{(k)} := \begin{pmatrix} \sqrt{a_{\lambda k}} & 0 \\ 0 & \sqrt{c_{\lambda k}} \end{pmatrix}$. It is straightforward to see that the operators $\{\hat{\Pi}_\lambda := \bigotimes_{k=1}^N \hat{M}_\lambda^{(k)}\}_{\lambda=1\dots t}$ correspond to an incomplete measurement that achieves transformation (\star) with the same probabilities as the $\{\Pi_\lambda\}_{\lambda=1\dots t}$. From Eq. (5), the collection of separable operators $\{\bigotimes_{k=1}^N \begin{pmatrix} 0 & 0 \\ 0 & |b_{\lambda k}| \end{pmatrix}\}_{\lambda=1\dots t}$ can be combined with $\{\hat{\Pi}_\lambda\}_{\lambda=1\dots t}$ to form a set which corresponds to a complete measurement.

One immediate consequence of this lemma is that for any incomplete separable transformation of the form (\star) with $\sum_\lambda \Pi_\lambda^\dagger \Pi_\lambda \ll \mathbb{I}$, we can always assume that $\mathbb{I} - \sum_\lambda \Pi_\lambda^\dagger \Pi_\lambda$ has a diagonal representation and is therefore separable. As a result, when $|\varphi\rangle_{1\dots N}$ is a W -class state, it is sufficient to

consider the feasible probabilities of transformation (\star) under incomplete separable transformations.

Now for measurement $\{\Pi_\lambda := M_\lambda^{(1)} \otimes \cdots \otimes M_\lambda^{(N)}\}_{\lambda=1\dots t}$, if we let S_{ij} denote the set of all outcomes λ such that $\Pi_\lambda|\varphi\rangle_{1\dots N} \propto |\Psi^{(ij)}\rangle$, we can form a Choi matrix Ω_{ij} for each edge $(i, j) \in E$ of graph \mathcal{G} [35]:

$$\Omega_{ij} = \sum_{\lambda \in S_{ij}} \Pi_\lambda \otimes \mathbb{I} \left(\bigotimes_{i=1}^N \Phi^{(i')} \right) (\Pi_\lambda^\dagger) \otimes \mathbb{I}. \quad (6)$$

Here, Π_λ acts on systems $1, 2, \dots, N$, while \mathbb{I} is the identity acting on their copies $1', 2', \dots, N'$. By Lemma 1, the Π_λ can be taken as diagonal matrices so that Ω_{ij} has support only on the span of $\{|i_1 i_1\rangle_{11'} |i_2 i_2\rangle_{22'} \cdots |i_N i_N\rangle_{NN'}\}_{i_1, i_2, \dots, i_N \in \{0, 1\}}$. Furthermore, since all parties besides i and j hold pure states in the end, $M_\lambda^{(k)}$ must be a rank 1 matrix for $k \neq i, j$ and $\lambda \in S_{ij}$. Thus, up to LUs and a permutation of spaces, Ω_{ij} has the form

$$\Omega^{(ij)} = \chi^{(i'j'j')} \otimes |0\rangle\langle 0|^{\overline{(i'j'j')}}},$$

where $\chi^{(i'j'j')}$ is effectively a separable $2 \otimes 2$ density matrix having support on $\{|mm\rangle_{i'j'} |nn\rangle_{j'j'}\}_{m, n \in \{0, 1\}}$; equivalently, $\chi^{(i'j'j')}$ has a positive partial transpose [36]. In terms of the Choi matrix, the condition of obtaining $|\Phi^{(ij)}\rangle$ with probability p_{ij} is given by

$$\text{tr}_{1'\dots N'}(\Omega_{ij} \varphi^{(1'\dots N')}) = p_{ij} \Phi^{(ij)} \otimes |0\rangle\langle 0|^{\overline{(ij)}}. \quad (7)$$

Here we use the fact that $\varphi^{(1'\dots N')}$ is taken to have only real components. Finally, the constraint that $\sum_{(i,j) \in E} \sum_{\lambda \in S_{ij}} \Pi_\lambda^\dagger \Pi_\lambda \leq \mathbb{I}$ is captured by

$$\sum_{(i,j) \in E} \text{tr}_{1\dots N}(\Omega_{ij}) \leq \mathbb{I}. \quad (8)$$

This construction is completely reversible such that, given matrices satisfying the above conditions, we can always construct a separable measurement facilitating transformation (\star) [37]. Thus the necessary and sufficient conditions for a feasible separable map are 4×4 complex matrices $\chi^{(i'j'j')}$ for all $(i, j) \in E$ which satisfy

$$\chi^{(i'j'j')} \geq 0, \quad [\chi^{(i'j'j')}]_{\Gamma_{i'j'}} \geq 0, \quad (9)$$

as well as Eqs. (7) and (8). This is a semidefinite feasibility problem which can be efficiently solved using a variety of numerical tools [38]. Furthermore, duality theory can be used to analytically prove instances of infeasibility. We perform such an analysis in the Appendix for the initial state $|\varphi_{1\dots N}\rangle = |W_N\rangle$. The result is given by the following theorem, which also provides an LOCC upper bound.

Theorem 1. For $|\varphi_{1\dots N}\rangle = |W_N\rangle$, transformation (\star) with graph representation \mathcal{G} is possible by SEPs if and only if

$$\frac{N^2}{4} \sum_{(i,j) \in E} p_{ij}^2 \leq 1, \quad \frac{N}{2} \sum_{(i,j) \in E_k} p_{ij}, \leq 1, \quad 1 \leq k \leq N. \quad (10)$$

Remark. In practice, it may be helpful to use the inequality $\sum_{i=1}^n x_i^2 \geq \frac{1}{n} (\sum_{i=1}^n x_i)^2$ so that the first constraint in Eq. (10)

becomes

$$\frac{N^2}{4|E|} \left(\sum_{(i,j) \in E} p_{ij} \right)^2 \leq 1. \quad (11)$$

IV. ENTANGLEMENT MONOTONES

In this section, we introduce new entanglement monotones on the N -qubit W class of states. An important property of quantum measurements is the universality of weak measurements. This means that any general measurement can be replaced by a sequence of measurements that obtains the same overall outcomes but only changes the state by an arbitrarily small increment with each individual measurement [13,39]. Consequently, to prove LOCC monotonicity of a given function, it is sufficient to prove it nonincreasing, on average, under two-outcome infinitesimal measurements by a single party. The full generality of the latter consideration was explored in Ref. [40]. Here, a weak measurement of $\{M_1^{(k)}, M_2^{(k)}\}$ corresponds to (a_1, c_1, a_2, c_2) lying in a small neighborhood of $(1/2, 1/2, 1/2, 1/2)$, and the relatively simple structure of the W class eases analysis in this infinitesimal setting.

We define our monotones as follows. For an N -party W state (x_1, x_2, \dots, x_N) , set $\{n_1, n_2, \dots, n_N\} = \{1, 2, \dots, N\}$ such that $x_{n_1} \geq x_{n_2} \geq \dots \geq x_{n_N}$ and consider the continuous functions:

$$\begin{aligned} \eta(\vec{x}) &= x_{n_1} - \left(\frac{1}{x_{n_1}} \right)^{N-2} \prod_{i=2}^N (x_{n_1} - x_{n_i}), \\ \kappa(\vec{x}) &= \sum_{i=2}^N x_{n_i} + \eta(\vec{x}). \end{aligned} \quad (12)$$

Theorem 2.

(I) η is nonincreasing on average for any single local measurement in which n_1 is the same value for the initial and all possible final states,

(II) κ is an entanglement monotone. It is strictly decreasing on average for any nontrivial measurement by party n_1 .

The three-qubit form of this theorem has been proven in Ref. [27]. Here, in the general case, our proof technique will be very similar.

Proof. (I) We consider case-by-case measurements of each party under the conditions of I. The function η transforms as $\eta \rightarrow \eta_\lambda$ for $\lambda = 1, 2$, and we are interested in the average change, $\overline{\eta_\lambda} = p_1 \eta_1 + p_2 \eta_2$, under infinitesimal measurements. First, suppose that party n_1 measures. According to Eq. (3), the average change in η is

$$\begin{aligned} \overline{\eta(\vec{x}_\lambda)} &= c_1 x_{n_1} \left(1 - \prod_{i=2}^N \left(1 - \frac{a_1 x_{n_i}}{c_1 x_{n_1}} \right) \right) \\ &\quad + c_2 x_{n_1} \left(1 - \prod_{i=2}^N \left(1 - \frac{a_2 x_{n_i}}{c_2 x_{n_1}} \right) \right). \end{aligned} \quad (13)$$

We demonstrate that in the weak-measurement setting, this quantity is maximized by equality of the upper bound:

$c_1 + c_2 = 1$. Indeed, we have

$$\left. \frac{\partial \overline{\eta}_\lambda}{\partial c_\lambda} \right|_{\substack{a_1 = a_2 = 1/2 \\ c_1 = c_2 = 1/2}} = x_{n_1} \left\{ \left(1 - \prod_{i=2}^N \left(1 - \frac{x_{n_i}}{x_{n_1}} \right) \right) - \sum_{i=2}^N \frac{x_{n_i}}{x_{n_1}} \prod_{j \neq i} \left(1 - \frac{x_{n_j}}{x_{n_1}} \right) \right\} \quad (14)$$

for $\lambda = 1, 2$, and it suffices to show that this expression is strictly positive. Now if we differentiate Eq. (14) with respect to any x_{n_k} , we obtain

$$\begin{aligned} & \prod_{i \neq k}^N \left(1 - \frac{x_{n_i}}{x_{n_1}} \right) - \prod_{i \neq k}^N \left(1 - \frac{x_{n_i}}{x_{n_1}} \right) + \sum_{i \neq k} \frac{x_{n_i}}{x_{n_1}} \prod_{j \neq i, k}^N \left(1 - \frac{x_{n_j}}{x_{n_1}} \right) \\ & = \sum_{i \neq k} \frac{x_{n_i}}{x_{n_1}} \prod_{j \neq i, k}^N \left(1 - \frac{x_{n_j}}{x_{n_1}} \right) \geq 0 \quad (2 \leq k \leq N), \end{aligned}$$

and since Eq. (14) vanishes when $x_{n_k} = 0$ for all n_k , it follows that for nonzero values of x_{n_k} , Eq. (14) is strictly positive. Thus, the maximal change in η occurs when $c_1 + c_2 = 1$. As we are interested in this upper bound, we assume the measurement is characterized by $a \equiv a_1$, $1 - a = a_2$, $c \equiv c_1$, and $1 - c = c_2$. We then have

$$\begin{aligned} \eta - \overline{\eta(\vec{x}_\lambda)} & = -x_{n_1} \prod_{i=2}^N \left(1 - \frac{x_{n_i}}{x_{n_1}} \right) + cx_{n_1} \prod_{i=2}^N \left(1 - \frac{ax_{n_i}}{cx_{n_1}} \right) \\ & + (1 - c)x_{n_1} \prod_{i=2}^N \left(1 - \frac{(1-a)x_{n_i}}{(1-c)x_{n_1}} \right). \quad (15) \end{aligned}$$

Expanding this to second order about the point $(a, c) = (1/2, 1/2)$ yields

$$\eta - \overline{\eta(\vec{x}_\lambda)} \approx 4(a - c)^2 \sum_{i,j} \frac{x_{n_i} x_{n_j}}{x_{n_1}} \prod_{l \neq i, j}^N \left(1 - \frac{x_{n_l}}{x_{n_1}} \right) \geq 0. \quad (16)$$

And this expression will be positive whenever $a \neq c$, which is whenever party n_1 performs a nontrivial measurement. In the case in which party n_i performs a measurement for some $i > 1$, η changes as

$$\begin{aligned} \overline{\eta(\vec{x}_\lambda)} & = x_{n_1} - (a_1 x_{n_1} - c_1 x_{n_i}) \prod_{j \neq i}^N \left(1 - \frac{x_{n_j}}{x_{n_1}} \right) \\ & - (a_2 x_{n_1} - c_2 x_{n_i}) \prod_{j \neq i}^N \left(1 - \frac{x_{n_j}}{x_{n_1}} \right) \leq \eta(\vec{x}). \quad (17) \end{aligned}$$

(II) We can always decompose a general transformation into a sequence of weak measurements for which either each measurement satisfies the conditions of I or its pre-measurement state \vec{y} satisfies $y_{n_1} = y_{n_2}$. In the first case, κ is monotonic by part I and the fact that $\sum_{i=2}^N x_{n_i}$ is nonincreasing on average by the K-T monotones. In the second case, we have $\kappa(\vec{y}) = 1 - y_0$. Since $1 - y_{\lambda,0}$ is an upper bound on $\kappa(\vec{y}_\lambda)$ for each of the postmeasurement states \vec{y}_λ , and $1 - y_0$ is nonincreasing on average by the K-T monotones, it follows that

$\kappa(\vec{y}) \geq \sum_\lambda p_\lambda \kappa(\vec{y}_\lambda)$. Thus, κ is an entanglement monotone in general.

Theorem 2 also applies to any fixed collection of subsystems. Indeed for N -qubit systems, let S denote some subset of parties, and consider the unnormalized state \vec{s} which has $|S|$ components, each belonging to a different party in S . Then Theorem 2 also holds for the functions $\eta(\vec{s})$ and $\kappa(\vec{s})$. The proof of this is exactly the same as above, with the added note that whenever a measurement is performed by a party not in S , $\eta(\vec{s})$ and $\kappa(\vec{s})$ remain invariant on average, which follows from Eq. (3).

For example, in a four-party system, let S be parties 1, 2, and 3. Now for any four-qubit state \vec{x} , take $\{x_{\max}, x_{\text{mid}}, x_{\min}\} = \{x_1, x_2, x_3\}$ such that $x_{\max} \geq x_{\text{mid}} \geq x_{\min}$. Then the function

$$2x_{\text{mid}} + 2x_{\min} - \frac{x_{\text{mid}} x_{\min}}{x_{\max}} \quad (18)$$

is an entanglement monotone.

The condition in part I of Theorem 2 can be extended beyond single measurements.

Corollary 1. Suppose the transformation $\vec{x} \rightarrow \{p_i, \vec{y}_i\}$ is possible by LOCC, where $n_1(\vec{x}) = n_1(\vec{y}_i)$ for all i . Then $\eta(\vec{x}) \geq \sum_i p_i \eta(\vec{y}_i)$.

Proof. We can partition any transformation into sections where $n_1(\vec{x})$ is the largest component and where it is not. By weak measurement theory, we can assume that when passing from one section to the other, we always first obtain a state \vec{s} on the border such that $\eta(\vec{s}) = s_{n_1(\vec{x})}$. Therefore, since the $n_1(\vec{x})$ component is always monotonic by the K-T monotones, (2), we have that η will not have increased on average within any region for which $n_1(\vec{x})$ is not the largest component. For sections when $n_1(\vec{x})$ is the largest, we know that η is monotonic by part I of the previous theorem.

3. Interpretation of monotones

A natural question is whether the functions η and κ possess any physical interpretation. Here we show that for states \vec{x} having $x_0 = 0$, $2\eta(\vec{x})$ gives the optimal probability for transformation (\star) when the configuration graph \mathcal{G} consists of all edges connected to node $v_{n_1(\vec{x})}$. We refer to this as a ‘‘combing transformation’’ since it represents a single-copy version of the entanglement combing procedure described in Ref. [31]. On the other hand, $\kappa(\vec{x})$ gives the optimal probability when \mathcal{G} is complete, i.e., each vertex is connected to every other one (see Fig. 2). The following theorem gives a precise statement of this result.

Theorem 3. For an N -party W state $\vec{x} = (x_1, x_2, \dots, x_N)$, let P_{tot} be the optimal total probability of obtaining an EPR pair by LOCC, and P_k the optimal total probability of party k becoming EPR entangled. Then

- (I) $P_{\text{tot}} < \kappa(\vec{x})$, and
- (II)

$$P_k \leq \begin{cases} 2x_k & \text{if } x_k < x_l \text{ for some } l, \\ 2\eta(\vec{x}) & \text{if } x_k \geq x_l \text{ for all } l. \end{cases}$$

When $x_0 = 0$, the upper bound in I can be approached arbitrarily closely, while in II it can be achieved exactly.

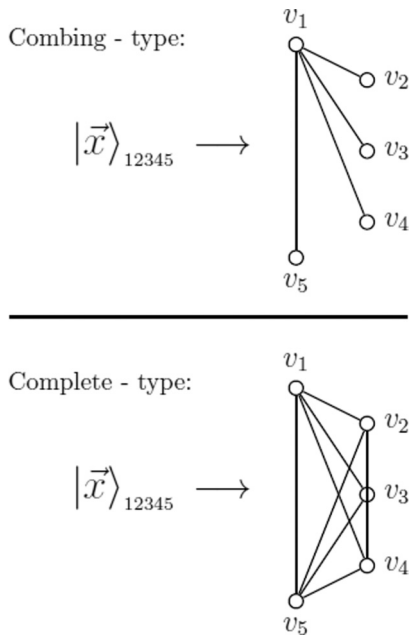


FIG. 2. Distillation configurations for η vs κ . Top: A “combing-type” distillation: when $x_0 = 0$, $2\eta(\vec{x})$ is the optimal probability for a random distillation in which party n_1 shares one-half of each EPR pair. Bottom: A “complete-type” distillation: when $x_0 = 0$, $\kappa(\vec{x})$ gives the optimal probability for a random distillation in which the target pairs are any two of the parties.

Proof. First recall that $\kappa(\Phi^{(ij)}) = 1$. Then the upper bounds follow from Theorem 2 and the K-T monotones. Assume now that $x_0 = 0$. To show that the upper bounds are effectively tight, we construct a specific protocol based on an “equal or vanish” (e/v) measuring scheme [28]. On its own, an e/v measuring scheme is just one way in which a W -class state $|\varphi\rangle_{1\dots N}$ can be converted into either EPR pairs or W states $|W_m\rangle$ for $3 \leq m \leq N$. Each party k performs a two-outcome measurement for which outcome 1 is a state whose k th component equals the maximum component, and outcome 2 is a state whose k th component is 0. The specific measurement operators are given by $M_1 = \text{diag}[\sqrt{\frac{x_k}{x_{n_1}}}, 1]$ and $M_2 = \text{diag}[\sqrt{1 - \frac{x_k}{x_{n_1}}}, 0]$. When each party does this, the possible resultant states are $|\Phi^{(n_1 k)}\rangle$ for $n_2 \leq k \leq N$, $|W_m\rangle$ for $3 \leq m \leq N$, or a product state (see Fig. 3).

For a complete-type distillation, the parties first perform e/v measurements and then implement the Fortescue-Lo Protocol on the resultant $|W_m\rangle$ states. When $x_{n_1} = x_{n_2}$ for an initial state \vec{x} , a product state is never obtained by the e/v measurements, and the total success probability is therefore arbitrarily close to 1. When $x_{n_1} > x_{n_2}$, we prove the success rate by induction on the number of parties. For $N = 2$, the rate of $\kappa(\vec{x}) = 2x_{n_2}$ can be achieved [41]. Suppose now that probability κ is obtained arbitrarily close with $N - 1$ parties, and consider the N -party case. If party n_2 is the first to perform an e/v measurement, then with probability q this measurement will raise his component to equal the largest; i.e., the resultant state \vec{y} has $y_{n_1} = y_{n_2}$. Thus, random EPR distillation can be accomplished deterministically on \vec{y} . For the “vanish” outcome occurring with probability $1 - q$, the resultant state \vec{z} is shared

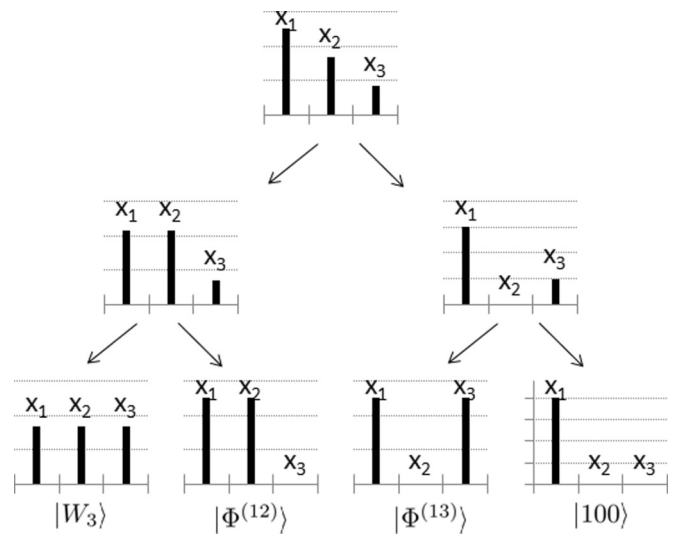


FIG. 3. A three-qubit “equal or vanish” measurement scheme. The initial state is (x_1, x_2, x_3) with $x_1 > x_2 > x_3$ and $x_0 = 0$. Bob (party 2) measures first and either obtains a state in which his component is a maximum or becomes entangled in the other two. In the next round Charlie (party 3) performs the same type of measurement. The possible outcome states are $|W_3\rangle$, $|\Phi^{(12)}\rangle$, $|\Phi^{(13)}\rangle$, and a product state. A “complete-type” distillation begins with this measurement scheme and then the Fortescue-Lo protocol is performed on the $|W_3\rangle$ outcome. A “combing-type” distillation is exactly this measurement scheme except that the premeasurement state of $|W_3\rangle$ is converted into either $|\Phi^{(12)}\rangle$ or a product state (and not $|W_3\rangle$).

among $N - 1$ qubits with $z_{n_i} = x_{n_i} \frac{x_{n_1} - x_{n_2}}{x_{n_1}(1 - q)}$ for $n_i \neq 2$. By the inductive hypothesis, we then have

$$\begin{aligned}
 p_{\text{tot}}(\vec{x}) &= q + (1 - q) \left[1 - \left(\frac{1}{z_{n_1}} \right)^{N-3} \prod_{i=3}^N (z_{n_1} - z_{n_i}) \right] \\
 &= 1 - \frac{x_{n_1} - x_{n_2}}{x_{n_1}} \left(\frac{1}{x_{n_1}} \right)^{N-3} \prod_{i=3}^N (x_{n_1} - x_{n_i}) \\
 &= 1 - \left(\frac{1}{x_{n_1}} \right)^{N-2} \prod_{i=2}^N (x_{n_1} - x_{n_i}). \tag{19}
 \end{aligned}$$

For a combing-type distillation, when $x_k \leq x_l$ for some party l , $2x_k$ is known to be an achievable rate [42,43]. When $x_k > x_l$ for all parties l , the procedure is for each party to perform an e/v measurement (in any order), except that when the first party l obtains an “equal” outcome, a nonrandom EPR distillation is made between party k and party l . This occurs with total probability $2x_l$, and a completely analogous inductive argument to the one given above shows that this full measurement scheme succeeds with probability exactly equal to $\eta(\vec{x})$. ■

Remark. We make two remarks here. First, for three-qubit systems, combing- and complete-type transformations represent the only two types of random distillations. Thus, for three-qubit states with $x_0 = 0$, Theorem 3 gives a complete solution to transformation (\star) . Second, a natural question is whether the “equal or vanish” scheme is always optimal for distilling EPR pairs. In other words, for some random

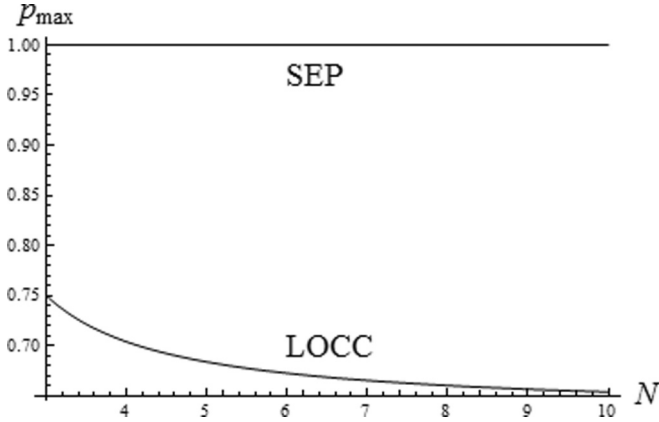


FIG. 4. LOCC vs SEP for the maximum probability of party 1 becoming EPR entangled as a function of N when the initial state is $\sqrt{\frac{1}{2}}|10\dots 0\rangle + \sqrt{\frac{1}{2(1-N)}}(|01\dots 0\rangle + \dots + |00\dots 1\rangle)$. The LOCC probability is $1 - (1 - \frac{1}{N-1})^{N-1}$. A gap of 37% exists between SEP and LOCC.

distillation configuration graph \mathcal{G} , is it always best to first perform e/v measurements and then implement the Fortescue-Lo protocol? We have found that this is not the case, and we describe specific counterexamples in Ref. [26].

V. SEP VS LOCC

In this section we use results from Sec. III and Theorem 3 to compare the distillation performances of SEP and LOCC. In particular, we consider an N -qubit combing-type distillation.

The state we consider is $|\psi_{1/2}\rangle_{1\dots N} = \sqrt{\frac{1}{2}}|10\dots 0\rangle + \sqrt{\frac{1}{2(1-N)}}(|01\dots 0\rangle + \dots + |00\dots 1\rangle)$. By LOCC, the optimal probability for a combing-type transformation is

$$2\eta(\psi_{1/2}) = 1 - \left(1 - \frac{1}{N-1}\right)^{N-1} \rightarrow 1 - e^{-1}, \quad (20)$$

where we have taken the limit for large N . However, it is easy to see that the following SEPs (defined up to a reordering of spaces) represent a complete measurement which, with total probability 1, will obtain an EPR pair shared by the first party:

$$M_k = \mathbb{I}^{(1)} \otimes \sqrt{\frac{1}{N-1}}|0\rangle\langle 0|^{(k)} + |1\rangle\langle 1|^{(k)} \bigotimes_{j \neq 1, k}^N |0\rangle\langle 0|^{(j)}$$

for $1 < k \leq N$,

$$M_0 = \sqrt{\mathbb{I} - \sum_{i=1}^N M_i^\dagger M_i}. \quad (21)$$

We plot this separation between LOCC and SEP as a function of N in Fig. 4.

VI. MULTICOPY DISTILLATIONS

So far we have only considered transformations of a single W -class state. However, in this section, we consider a particular n -copy variant of transformation (\star) . While the following discussion pertains to the tripartite case, its generalization to more parties is straightforward.

Suppose the trio starts with n copies of the state $|\psi_{1/2}\rangle = \sqrt{1/2}|100\rangle + 1/2(|010\rangle + |001\rangle)$, and they wish to distill n EPR pairs such that Alice is always one of the shareholders (actually Bob and Charlie need not have the same components in the following argument). The problem can be phrased as follows:

$$|\psi_{1/2}\rangle^{\otimes n} \rightarrow |\Psi^{(AB)}\rangle^{\otimes k} |\Psi^{(AC)}\rangle^{\otimes n-k} \quad (22)$$

with probability $p_k = \binom{n}{k} / 2^n$ for $k = 0, \dots, n$.

This is a combing-type transformation, and by the previous section we know that for any n , the transformation can always be completed with probability 1 by SEP. On the other hand, even if the parties are allowed to act coherently on the n copies of their local state, the following theorem still gives a no-go result.

Theorem 4. The transformation given by Eq. (22) is not possible by LOCC for any n . Nor is it possible for any other distribution of the specified target states.

We give the proof below. The only technical component needed is Lemma 2, which relies heavily on the special form of state $|\psi_{1/2}\rangle$. The main idea is that when viewed as a bipartite transformation with respect to A:BC, the reduced-state entropies are the same for the initial and all the final states. Consequently, the reduced-state entropy must remain invariant for each measurement outcome in the LOCC protocol, and following the lines of Theorem 1 in Ref. [44], this implies that Alice is restricted to performing only LUs.

However, due to the form of $|\psi_{1/2}\rangle$, invariance of the reduced state entropy also implies that Bob and Charlie can perform only LUs, as we now show. Without loss of generality, suppose that Bob acts first, before Charlie, in the protocol. Since Alice can only have performed an LU up to this point, Bob and Charlie's reduced state is

$$\left(\frac{1}{2}|00\rangle\langle 00| + \frac{1}{2}|\Psi\rangle\langle\Psi|\right)^{\otimes n} = \frac{1}{2^n} \sum_{x \in \{0,1\}^n} |\tilde{x}\rangle\langle\tilde{x}|, \quad (23)$$

where we introduce the notation that for a binary vector $x \in \{0,1\}^n$ with components $x_i \in \{0,1\}$, the corresponding string \tilde{x} has symbolic components $\tilde{x}_i = 00$ if $x_i = 0$ and $\tilde{x}_i = \Psi$ if $x_i = 1$. For example,

$$x = 010 \Rightarrow |\tilde{x}\rangle = |00\rangle|\Psi\rangle|00\rangle.$$

The reason for introducing this notation becomes evident in the following.

Lemma 2. (i) For $x, y \in \{0,1\}^n$, let $S \subset \{0,1\}^n$ be the set such that $b \in S$ if $b_i = 0$ whenever $x_i \cdot y_i \neq 1$. Then for any operator A acting on Bob's system,

$$\langle \tilde{x} | A \otimes \mathbb{I} | \tilde{y} \rangle \propto \sum_{b \in S} \langle x + b | A | y + b \rangle. \quad (24)$$

(ii) If $\langle \tilde{x} | A \otimes \mathbb{I} | \tilde{y} \rangle = 0$ for all $x \neq y \in \{0,1\}^n$, then $\langle x | A | y \rangle = 0$ for all $x \neq y \in \{0,1\}^n$. (iii) If $\langle \tilde{x} | A \otimes \mathbb{I} | \tilde{x} \rangle = k$ for all $x \in \{0,1\}^n$, then $\langle x | A | x \rangle = k$ for all $x \in \{0,1\}^n$.

Proof. Part i can be verified from the relations $\langle 00 | T \otimes \mathbb{I} | 00 \rangle \propto \langle 0 | T | 0 \rangle$, $\langle 00 | T \otimes \mathbb{I} | \Psi \rangle \propto \langle 0 | T | 1 \rangle$, $\langle \Psi | T \otimes \mathbb{I} | 00 \rangle \propto \langle 1 | T | 0 \rangle$, and $\langle \Psi | T \otimes \mathbb{I} | \Psi \rangle \propto \langle 1 | T | 1 \rangle + \langle 0 | T | 0 \rangle$. For ii, we use induction on $\log_2 |S|$, i.e., on the number of coordinates simultaneously equal to 1 in both x and y . By part i, when

$\log_2 |S| = 0$, then the statement is easily seen to be true from Eq. (24) since the only $b \in S$ is the all-zero vector $\vec{0}$. Now suppose the claim is true when $\log_2 |S| = m$, and consider two vectors x, y such that $\log_2 |S| = m + 1$. Again, by part i,

$$0 = \langle \vec{x}|A \otimes \mathbb{I}|\vec{y} \rangle \propto \sum_{\vec{0} \neq b \in S} \langle x + b|A|y + b \rangle + \langle x|A|y \rangle.$$

But for $\vec{0} \neq b \in S$, the strings $x + b$ and $y + b$ will have no more than m coordinates that are both equal to 1. Therefore, by the inductive assumption, each term in the sum vanishes, and so $\langle x|A|y \rangle = 0$. Part iii can be proven by using a similar inductive argument and noting that, for $\langle \vec{x}|A \otimes \mathbb{I}|\vec{x} \rangle$, the proportionality factor in part i is $1/|S|$.

Now let M be one of Bob's measurement operators. By invariance of the von Neumann entropy, we must have [45]

$$\begin{aligned} n &= S \left(\frac{1}{2^n p_M} \sum_{x \in \{0,1\}^n} M \otimes \mathbb{I}|\vec{x} \rangle \langle \vec{x}|M^\dagger \otimes \mathbb{I} \right) \\ &\leq H \left\{ \frac{\langle \vec{x}|M^\dagger M \otimes \mathbb{I}|\vec{x} \rangle}{2^n p_M} \right\}_{x \in \{0,1\}^n} \leq n, \end{aligned} \quad (25)$$

which requires that $\langle \vec{x}|M^\dagger M \otimes \mathbb{I}|\vec{x} \rangle$ is some positive constant for all $x \in \{0,1\}^n$ and the $M \otimes \mathbb{I}|\vec{x} \rangle$ are orthogonal. By ii and iii of Lemma 2, this is only possible if $M^\dagger M$ is proportional to the identity. In other words, M is of the form $\sqrt{p}U$ for some unitary U .

In the next round of measurement it will be Charlie's turn. However, the above argument will apply for Charlie's measurement even after Bob performs an LU rotation. Thus, in all rounds the parties can perform LUs and therefore transformation (22) cannot be accomplished by any LOCC protocol.

Up to a conditional LU transformation, transformation (22) can be phrased as the mixed-state transformation $|\psi_{1/2}\rangle\langle\psi_{1/2}|^{\otimes n} \rightarrow \sigma^{\otimes n}$, where

$$\sigma = 1/2(|\Psi^{(AB)}\rangle\langle\Psi^{(AB)}| \otimes |0\rangle\langle 0| + |\Psi^{(AC)}\rangle\langle\Psi^{(AC)}| \otimes |1\rangle\langle 1|).$$

Here, the $|0\rangle$ and $|1\rangle$ is classical information accessible to all parties, and it encodes which particular duo holds the EPR state. Thus, LOCC impossibility of transformation (22) means that the transformation $|\psi_{1/2}\rangle\langle\psi_{1/2}|^{\otimes n} \rightarrow \sigma^{\otimes n}$ is LOCC infeasible.

Finally, we can consider the asymptotic setting and when the trio wishes to distill maximal entanglement with unit efficiency such that the entanglement is distributed equally to the pairs Alice-Bob and Alice-Charlie. More precisely, we seek for every n an LOCC map Γ^n such that

$$\text{tr}[\Gamma^n(\psi_{1/2}^{\otimes n}) \cdot \Psi^{(AB)\otimes n/2} \Psi^{(AC)\otimes n/2}] \rightarrow 1.$$

In fact, as given by the entanglement combing protocol in Ref. [31], this transformation is asymptotically feasible. Moreover, their protocol holds for various distributions of final entanglement and not just equal shares between Alice-Bob and Alice-Charlie. Consequently, we have shown that for particular state transformations, $\text{SEP} > \text{LOCC}$ regardless of the number of copies considered. However, when the same transformations are considered in asymptotic form, we have that $\text{SEP} = \text{LOCC}$.

VII. CONCLUSION

In this article, we have studied the random distillation of W -class states by SEPs and LOCC. Based on the transformation results of bipartite pure states [46], one may suspect that SEP and LOCC have equivalent transformation capabilities. However, here we have shown that SEP is strictly more powerful.

For SEPs, the general solution to transformation (\star) can be solved by semidefinite programming optimization when $x_0 = 0$. This then places an upper bound on the problem for LOCC. Tightening the LOCC bound requires analyzing each configuration graph on a case-by-case basis. Two particular transformations we have considered are combing- and complete-type transformations (Fig. 2). Theorem 3 provides an upper bound for the success probabilities of these transformations. For states with $x_0 = 0$, the upper bounds can be approached arbitrarily close.

To obtain these results, our general strategy has been (i) to start with a general W -class state and compute the combing- or complete-type transformation probability using an e/v' protocol and (ii) to prove that the general probability expression (as a function of components x_i) is an entanglement monotone. This strategy isolates essential properties of LOCC beyond the tensor product structure of its measurement operators as it has generated entanglement monotones that can be increased under SEPs.

When $x_0 \neq 0$, we know these upper bounds are not tight, a prime example being the state $\sqrt{1-3s}|000\rangle + \sqrt{s}(|100\rangle + |010\rangle + |001\rangle)$ with $s > 0$. For a combing-type transformation of this state with Alice always being a shareholder in the outcome entanglement, the probability of success is upper bounded by $2\eta = 2s$. However, it is known that such a rate cannot be achieved [30,33]. We leave it as an open problem to determine the optimal random distillation rates when $x_0 \neq 0$.

In terms of success probability, Fig. 4 shows a maximum percentage difference of roughly 37% for the combing-type distillation. We conjecture that much larger gaps between SEP and LOCC exist than the ones shown in this article. Even for state $|W_N\rangle$, we predict that different distillation configuration graphs \mathcal{G} restrict the feasible probabilities for LOCC much stronger than the separable upper bounds of Theorem 3 (see Ref. [26] for more details).

Finally, we observe that for particular random distillations, the advantage of SEP over LOCC does not appear in the asymptotic setting, while it does when only finite resources are considered, regardless of the amount. While we have shown this specifically for transformation (22), the result holds true for more general combing-type transformations. This suggests that the intriguing conjecture that SEP and LOCC are operationally equivalent in the many-copy limit. It is our hope that this article will lead to a deeper understanding of multipartite entanglement and the structure of LOCC.

ACKNOWLEDGMENTS

We thank Andreas Winter, Debbie Leung, Graeme Smith, and John Smolin for providing helpful suggestions and discussing related material. We also thank Jonathan Oppenheim,

Ben Fortescue, Sandu Popescu, and Runyao Duan for offering insightful comments on the subject. We thank the funding

agencies including the NSERC, QuantumWorks, the CRC program, and CIFAR for financial support.

APPENDIX: DUAL SOLUTION TO $|W_N\rangle$ DISTILLATION BY SEP

We begin by writing Eqs. (8) and (9) in standard semidefinite programming form. Fix some encoding function $\phi : E \rightarrow |E|$ and define the matrices:

$$\begin{aligned}
 F_1 &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \oplus \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, & F_2 &= \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \oplus \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \\
 F_3 &= \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \oplus \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, & F_4 &= \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \oplus \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \\
 F_5 &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \oplus \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, & F_6 &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \oplus \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \\
 F_7 &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \oplus \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, & & (A1)
 \end{aligned}$$

$$\begin{aligned}
 G_1^{(ij)} &= [-1] \bigoplus_{k=1}^{|E|} [0] \bigoplus_{k=1}^{\phi(i,j)-1} [0]_{4 \times 4} \oplus F_1 \bigoplus_{k=\phi(i,j)+1}^{|E|} [0]_{4 \times 4}, \\
 G_2^{(ij)} &= [0] \bigoplus_{k=1}^{|E|} [0] \bigoplus_{k=1}^{\phi(i,j)-1} [0]_{4 \times 4} \oplus F_2 \bigoplus_{k=\phi(i,j)+1}^{|E|} [0]_{4 \times 4}, \\
 G_3^{(ij)} &= [0] \bigoplus_{k=1}^{|E|} [0] \bigoplus_{k=1}^{\phi(i,j)-1} [0]_{4 \times 4} \oplus F_3 \bigoplus_{k=\phi(i,j)+1}^{|E|} [0]_{4 \times 4}, \\
 &\dots \\
 G_7^{(ij)} &= [0] \bigoplus_{k=1}^{\phi(i,j)-1} [0] \oplus [-1] \bigoplus_{k=\phi(i,j)+1}^{|E|} [0] \bigoplus_{k=1}^{\phi(i,j)-1} [0]_{4 \times 4} \oplus F_7 \bigoplus_{k=\phi(i,j)+1}^{|E|} [0]_{4 \times 4}, \\
 G_0 &= [1] \bigoplus_{k=1}^{|E|} [1] \bigoplus_{\phi(i,j)=1}^{|E|} \left[\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{Np_{ij}}{2} & \frac{Np_{ij}}{2} & 0 \\ 0 & \frac{Np_{ij}}{2} & \frac{Np_{ij}}{2} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \oplus \begin{pmatrix} 0 & 0 & 0 & \frac{Np_{ij}}{2} \\ 0 & \frac{Np_{ij}}{2} & 0 & 0 \\ 0 & 0 & \frac{Np_{ij}}{2} & 0 \\ \frac{Np_{ij}}{2} & 0 & 0 & 0 \end{pmatrix} \right]. & (A2)
 \end{aligned}$$

Then Eqs. (9) and (8) are captured by the existence of $x_k^{(ij)} \in \mathbb{C}$ such that

$$G_0 + \sum_{(i,j) \in E} \sum_{m=1}^7 x_m^{(ij)} G_m^{(ij)} \geq 0, \quad (A3)$$

with the additional constraints that

$$\sum_{(i,j) \in E_k} \frac{Np_{ij}}{2} \leq 1 \quad \text{for } 1 \leq k \leq N. \quad (A4)$$

The dual problem to this asks

$$\max -\text{tr}(ZG_0), \quad \text{s.t. } 0 = \text{tr}(ZG_m^{(ij)}) \quad \text{for all } G_m^{(ij)}, \quad Z \geq 0. \quad (A5)$$

A critical relationship between the dual and the primal formulations is that if (A3) can be satisfied for some $x_k^{(ij)}$, then for any Z satisfying the constraints of (A5), we must have $\text{tr}(ZG_0) \geq 0$. Thus infeasibility is proven by the existence of some $Z \geq 0$ such that $\text{tr}(ZG_m^{(ij)}) = 0$ for all $G_m^{(ij)}$ and $\text{tr}(ZG_0) < 0$. We construct a certificate for infeasibility as follows. For each $(i, j) \in E$, define the matrix

$$Z^{(ij)} = \begin{bmatrix} 1 \\ |E| \end{bmatrix} \bigoplus_{k=1}^{\phi(i,j)-1} [0] \oplus \begin{bmatrix} N^2 p_{ij}^2 \\ 4 \end{bmatrix} \bigoplus_{k=\phi(i,j)+1}^{|E|} [0] \bigoplus_{k=1}^{\phi(i,j)-1} [0]_{8 \times 8} \oplus [0]_{4 \times 4} \oplus \begin{pmatrix} 1 & 0 & 0 & \frac{-Np_{ij}}{2} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \frac{-Np_{ij}}{2} & 0 & 0 & \frac{N^2 p_{ij}^2}{4} \end{pmatrix} \bigoplus_{k=\phi(i,j)+1}^{|E|} [0]_{8 \times 8}. \quad (\text{A6})$$

The claim is that the matrix

$$Z := \sum_{(i,j) \in E} Z^{(ij)}$$

is dual feasible with $\text{tr}(ZG_0) < 0$ whenever $\frac{N^2}{4} \sum_{(i,j) \in E} p_{ij}^2 > 1$. Indeed, it can easily be seen that $Z \geq 0$ and $\text{tr}[ZG_m^{(ij)}] = 0$ for $1 \leq m \leq 7$ and $(i, j) \in E$. And finally,

$$\text{tr}[ZG_0] = 1 + \frac{N^2}{4} \sum_{(i,j) \in E} p_{ij}^2 - \frac{N^2}{2} \sum_{(i,j) \in E} p_{ij}^2 < 0.$$

We have thus proven Theorem 1.

-
- [1] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, *Phys. Rev. Lett.* **70**, 1895 (1993).
[2] C. H. Bennett and S. J. Wiesner, *Phys. Rev. Lett.* **69**, 2881 (1992).
[3] A. K. Ekert, *Phys. Rev. Lett.* **67**, 661 (1991).
[4] S. Popescu and D. Rohrlich, *Phys. Rev. A* **56**, R3319 (1997).
[5] N. Linden, S. Popescu, B. Schumacher, and M. Westmoreland, *Quant. Inf. Proc.* **4**, 241 (2005).
[6] A. Acin, J. Vidal, and J. Cirac, *Quant. Inf. Comp.* **3**, 55 (2003).
[7] V. Vedral and M. B. Plenio, *Phys. Rev. A* **57**, 1619 (1998).
[8] G. Vidal, *J. Mod. Opt.* **47**, 355 (2000).
[9] M. Horodecki, P. Horodecki, and R. Horodecki, *Phys. Rev. Lett.* **84**, 2014 (2000).
[10] M. Horodecki, *Quant. Inf. Comp.* **1**, 3 (2001).
[11] M. B. Plenio, *Phys. Rev. Lett.* **95**, 090503 (2005).
[12] M. B. Plenio and S. Virmani, *Quant. Inf. Comp.* **7**, 1 (2007).
[13] C. H. Bennett, D. P. DiVincenzo, C. A. Fuchs, T. Mor, E. Rains, P. W. Shor, J. A. Smolin, and W. K. Wootters, *Phys. Rev. A* **59**, 1070 (1999).
[14] E. M. Rains, *Phys. Rev. A* **60**, 173 (1999).
[15] M. Donald, M. Horodecki, and O. Rudolph, *J. Math. Phys.* **43**, 4252 (2002).
[16] E. M. Rains, [arXiv:quant-ph/9707002v3](https://arxiv.org/abs/quant-ph/9707002v3) (1997).
[17] A. Kent, *Phys. Rev. Lett.* **81**, 2839 (1998).
[18] A. Chefles, *Phys. Rev. A* **69**, 050307 (2004).
[19] M. Horodecki, P. Horodecki, and R. Horodecki, *Phys. Rev. Lett.* **80**, 5239 (1998).
[20] D. Stahlke and R. B. Griffiths, *Phys. Rev. A* **84**, 032316 (2011).
[21] C. H. Bennett, D. P. DiVincenzo, T. Mor, P. W. Shor, J. A. Smolin, and B. M. Terhal, *Phys. Rev. Lett.* **82**, 5385 (1999).
[22] D. P. DiVincenzo, T. Mor, P. W. Shor, J. A. Smolin, and B. M. Terhal, *Commun. Math. Phys.* **238**, 379 (2003).
[23] M. Koashi, F. Takenaga, T. Yamamoto, and N. Imoto, [arXiv:0709.3196v1](https://arxiv.org/abs/0709.3196v1).
[24] R. Duan, Y. Feng, Y. Xin, and M. Ying, *IEEE Trans. Inf. Theory* **55**, 1320 (2009).
[25] E. Chitambar and R. Duan, *Phys. Rev. Lett.* **103**, 110502 (2009).
[26] W. Cui, E. Chitambar, and H.-K. Lo, *Phys. Rev. A* **84**, 052301 (2011).
[27] E. Chitambar, W. Cui, and H.-K. Lo, *Phys. Rev. Lett.* **108**, 240504 (2012).
[28] B. Fortescue and H.-K. Lo, *Phys. Rev. Lett.* **98**, 260501 (2007).
[29] B. Fortescue and H.-K. Lo, *Phys. Rev. A* **78**, 012348 (2008).
[30] S. Kintaş and S. Turgut, *J. Math. Phys.* **51**, 092202 (2010).
[31] D. Yang and J. Eisert, *Phys. Rev. Lett.* **103**, 220501 (2009).
[32] W. Dür, G. Vidal, and J. I. Cirac, *Phys. Rev. A* **62**, 062314 (2000).
[33] W. Cui, E. Chitambar, and H.-K. Lo, *Phys. Rev. A* **82**, 062314 (2010).
[34] E. Andersson and D. K. L. Oi, *Phys. Rev. A* **77**, 052104 (2008).
[35] A. Jamiolkowski, *Rep. Math. Phys.* **3**, 275 (1972).
[36] M. Horodecki, P. Horodecki, and R. Horodecki, *Phys. Lett. A* **223**, 1 (1996).
[37] J. I. Cirac, W. Dür, B. Kraus, and M. Lewenstein, *Phys. Rev. Lett.* **86**, 544 (2001).
[38] L. Vandenberghe and S. Boyd, *SIAM Rev.* **38**, 49 (1994).
[39] O. Oreshkov and T. A. Brun, *Phys. Rev. Lett.* **95**, 110409 (2005).
[40] O. Oreshkov and T. Brun, *Phys. Lett. A* **73**, 042314 (2006).
[41] H.-K. Lo and S. Popescu, *Phys. Rev. A* **63**, 022301 (2001).
[42] S. Turgut, Y. Gül, and N. K. Pak, *Phys. Rev. A* **81**, 012317 (2010).
[43] W. Cui, W. Helwig, and H.-K. Lo, *Phys. Rev. A* **81**, 012111 (2010).
[44] C. H. Bennett, S. Popescu, D. Rohrlich, J. A. Smolin, and A. V. Thapliyal, *Phys. Rev. A* **63**, 012307 (2000).
[45] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).
[46] V. Gheorghiu and R. B. Griffiths, *Phys. Rev. A* **78**, 020304(R) (2008).