arXiv:2204.00781v3 [quant-ph] 30 Mar 2023

The Round Complexity of Local Operations and Classical Communication (LOCC) in Random-Party Entanglement Distillation

Guangkuo Liu[1,2,3], Ian George[4,5], and Eric Chitambar[4,5]

[1]Department of Physics, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801,USA

[2]JILA, University of Colorado/NIST, Boulder, CO, 80309, USA

[3]Department of Physics, University of Colorado, Boulder CO 80309, USA

[4]Department of Electrical and Computer Engineering, Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA

[5]Illinois Quantum Information Science and Technology (IQUIST) Center, University of Illinois Urbana-Champaign, Urbana, IL 61801

March 28, 2023

A powerful operational paradigm for distributed quantum information processing involves manipulating pre-shared entanglement by local operations and classical communication (LOCC). The LOCC round complexity of a given task describes how many rounds of classical communication are needed to complete the task. Despite some results separating one-round versus two-round protocols, very little is known about higher round complexities. In this paper, we revisit the task of one-shot random-party entanglement distillation as a way to highlight some interesting features of LOCC round complexity. We first show that for random-party distillation in three qubits, the number of communication rounds needed in an optimal protocol depends on the entanglement measure used; for the same fixed state some entanglement measures need only two rounds to maximize whereas others need an unbounded number of rounds. In doing so, we construct a family of LOCC instruments that require an unbounded number of rounds to implement. We then prove explicit tight lower bounds on the LOCC round number as a function of distillation success probability. Our calculations show that the original W-state random distillation protocol by Fortescue and Lo is essentially optimal in terms of round complexity.

1 Introduction

Quantum entanglement is an essential ingredient for realizing the full capabilities of distributed quantum information processing. To understand the research character of entanglement, one typically considers the "distant-lab" paradigm in which spatially separated laboratories have access to different parts of a globally-entangled states, but they can only process it using local operations and classical communication (LOCC) [40, 25]. In this scenario, parties take turns performing a local quantum measurement, announcing the outcome, and choosing a new local measurement for the next round based on the global history of previous results. Every LOCC protocol then has a tree-like structure with each branch in the tree representing a different sequence of measurement outcomes. Unfortunately, for many tasks such as entanglement distillation [4, 20] or state discrimination [39, 3, 28, 5, 10], precise bounds on LOCC capabilities are hard to prove. One reason for this difficulty is that a potentially unbounded number of classical communication exchanges are allowed in an LOCC protocol.

The largely unexplored topic of LOCC round complexity studies how many rounds of classical communication are needed to perform some distributed quantum information task. We will say that an LOCC protocol has r rounds if there is at least one branch having a sequence of local measurements alternating r times between different parties (see Section 5.1 for a slightly more general definition). Note that each local measurement might involve a combination of measurements all performed in the same laboratory. Most results on LOCC round complexity involve quantum state discrimination problems and focus just on separating one-round versus two-round success rates [38, 34, 10, 11, 17, 43, 50, 49]. Much more challenging is proving that a given task requires more than two rounds of LOCC to complete, and only a few results are known of this form. Wakakuwa *et al.* have proven separation results between two and three rounds for the task of entanglement-assisted nonlocal gate implementation [45, 46]. For higher rounds, Wang and Duan have constructed families of bipartite quantum states hav-

ing increasing dimension but requiring more rounds to perfectly discriminate the states as the dimension grows [48]. A similar finding has been shown for the convertibility of certain quantum (resp. classical) mixed states [12]. On the other hand, some round compression results are known. For example, in the task of LOCC state conversion, any bipartite protocol can always be reduced to one-way LOCC provided the initial state is pure, regardless of the system's dimensions [32, 35, 44]. In another direction, Cohen has shown that certain LOCC tasks of high round complexity can be approximated arbitrarily well using one-way LOCC [15]. Complementing the work on round or "depth" compression, it has recently been shown that the "width" of an arbitrary LOCC state discrimination protocol can be compressed to a standard form [30]. One could in general study the interplay between depth and width complexities in different LOCC tasks.

The purpose of this work is to further advance the study of LOCC round complexity. To exemplify certain new facts about round complexity, we focus on the specific task of random-party entanglement distillation in three-qubit W-class states [22], which has already proven to be a fruitful problem for demonstrating interesting properties of LOCC [18, 7, 9, 8]. A W-class state $|\Psi_W\rangle$ is one that can be obtained from the canonical W state $|W\rangle = \frac{1}{\sqrt{3}}(|100\rangle + |010\rangle + |001\rangle)$ by stochastic LOCC (SLOCC) [21]. A random-party EPR distillation protocol transforms a tripartite W-class state $|\Psi_W\rangle$ into a bipartite EPR state $|\Psi^+\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle)$ with the target pair unspecified; any branch in the protocol is deemed a success branch provided $|\Psi^+\rangle$ is obtained between some pair of parties at the end of the branch. Fortescue and Lo devised a family LOCC protocols with increasing round number (details given below) that completes this task with success probabilities approaching one. The limit of such protocols is then some map that distills $|W\rangle$ into random-party EPR pairs with probability one. However, it has been shown [9] that this map or any map achieving unit success probability is not implementable by LOCC. Consequently, LOCC constitutes a class of quantum operations that is not closed. This argument was later extended in Ref. [13] to show that the set of bipartite LOCC maps is likewise not topologically closed. Related topological properties have been reported in Ref. [6, 15, 16], and we also note that an experimental demonstration of random distillation has recently been performed using a three-photon W-state [31].

Variations to the random-party EPR distillation task can also be considered. For example, one can demand that the final states obtained in each success branch be just some entangled bipartite state $|\phi\rangle$, and not necessarily a maximally entangled one [23, 7]. If E is some measure of bipartite entanglement, the goal then is to maximize the average final value E across all success branches. More precisely, we define

$$E^{(\mathrm{rnd})}(\Psi_W) := \sup_{\Psi_W \to \{p_i, \varphi_i\}} \sum_i p_i E(\varphi_i), \qquad (1)$$

where the supremum is taken over all finite-round LOCC transformations that convert $|\Psi_W\rangle$ into some bipartite state $|\varphi_i\rangle$ with probability p_i. A convenient entanglement measure for bipartite systems is the concurrence [47], which for a pure state $|\varphi\rangle^{XY}$ is defined as $C(\varphi) = 2\left(\det[\mathrm{Tr}_X|\varphi\rangle\langle\varphi|]\right)^{1/2}$. Thus, $C^{(\mathrm{rnd})}(\Psi_W)$ is the largest average pure-state concurrence obtainable from $|\Psi_W\rangle$. We will refer to the task of optimizing the expected bipartite concurrence as random-party concurrence distillation. Another entanglement measure is given $E_2(\varphi) = 2\lambda_{\min}[\mathrm{Tr}_X|\varphi\rangle\langle\varphi|]$, which is twice the smaller eigenvalue of the reduced density matrix $\mathrm{Tr}_X|\varphi\rangle\langle\varphi|$. Operationally, it is known that $E_2(\varphi)$ corresponds to the supremum probability of transforming the state $|\varphi\rangle$ to an EPR state by LOCC [32]. From this it immediately follows that $E_2^{(\mathrm{rnd})}(\Psi_W)$ is equal to the supremum probability of obtaining an EPR pair between any two parties starting from $|\Psi_W\rangle$, and so the optimization in Eq. (1) can be restricted to LOCC transformations with only EPR states being the target, i.e. the original task of random-party EPR distillation.

A second variation to the random-party EPR distillation task is to fix one of the parties ($\star$) and only deem the transformation a success if party $\star$ is entangled in the end. We refer to this as $\star$-random-party distillation, as opposed to *total* random-party distillation described above. For a fixed party $\star \in \{A, B, C\}$, the optimal average bipartite entanglement for party $\star$ is

$$E^{(\star\text{-rnd})}(\Psi_W) := \sup_{\Psi_W \to \{p_i, \varphi_i\}} \sum_i p_i E(\varphi_i), \qquad (2)$$

where the supremum is taken over all LOCC transformations in which each state $|\varphi_i\rangle$ is bipartite entangled between party $\star$ and some other party in $\{A, B, C\} \setminus \{\star\}$. Restricted random-party distillation has been partially studied for the concurrence and E_2 entanglement measures in Refs. [13] and [8], respectively.

In terms of round complexity, one of our main results is that $E_2^{(\star\text{-rnd})}(\Psi_W)$ is achievable in finite rounds of LOCC for any tripartite W-state, whereas $C^{(\star\text{-rnd})}(\Psi_W)$ is not. Hence, the amount of rounds needed in an optimal entanglement distillation protocol depends on the type of entanglement measure considered. What makes this result particularly surprising is that concurrence and E_2 measures are in one-to-one correspondence: $E_2(|\varphi\rangle) = 1 + \sqrt{1 - C(|\varphi\rangle)^2}$. Our second main result is establishing a tight lower bound on the number of LOCC rounds needed to achieve a random-party EPR distillation of $|W\rangle$ with probability $> 1 - \delta$, for any $\delta > 0$.

To our knowledge, this is the first time any type of trade-off has been obtained between round complexity and success probability of a given distributed quantum information processing task. This result also shows that the original Fortescue-Lo protocol is essentially optimal for random-party distillation of $|W\rangle$, and we prove optimality of more general types of transformations by enlarging the class of LOCC to encompass all operations that completely preserve positivity of the partial transpose (PPT) [41]. As a corollary of our work, we introduce a new family of quantum instruments that lie on the boundary of LOCC but not inside LOCC itself [13].

2 The Structure of W-class States

In this paper, we refer to the W-class as the collection of three-qubit states that can be obtained from the canonical W state $|W\rangle$ by SLOCC. Up to a local change in basis, every W-class state can be written as

$$|\mathbf{x}\rangle := \sqrt{x_0}|000\rangle + \sqrt{x_A}|100\rangle + \sqrt{x_B}|010\rangle + \sqrt{x_C}|001\rangle,$$

with the x_i being non-negative. Normalization requires that $x_0 = 1 - (x_A + x_B + x_C)$, and therefore the state $|\mathbf{x}\rangle$ only depends on three non-negative parameter $\mathbf{x} = (x_A, x_B, x_C)$. If all three of these numbers are strictly positive, then the state is genuinely three-way entangled (i.e. it is not a product state with respect to some bi-partition of the three parties). Moreover, in this case it can be shown that the vector $\mathbf{x}$ *uniquely* identifies the state [27]. More precisely, if we express $|\mathbf{x}\rangle$ in some other basis as

$$\sqrt{x'_0}|0'0'0'\rangle + \sqrt{x'_A}|1'0'0'\rangle + \sqrt{x'_B}|0'1'0'\rangle + \sqrt{x'_C}|0'0'1'\rangle$$

then necessarily $x_k = x'_k$ for all components. Equivalently stated, the components x_k are invariant under local unitaries (LU). This LU invariance, combined with the fact that the W-class is closed under LOCC, makes the W-class very attractive for studying properties of LOCC. Indeed, if Alice, Bob, and Charlie start out sharing a W-class state $|\mathbf{x}\rangle$, then any multi-outcome LOCC protocol on $|\mathbf{x}\rangle$ can be described compactly by a probabilistic transformation of vectors,

$$\mathbf{x} \mapsto \mathbf{y}_i \quad \text{with probability } p_i, \tag{3}$$

in which the $|\mathbf{y}_i\rangle$ are different W-class states obtained along different branches of the protocol.

Every bipartite entangled state belongs to the W-class, and in canonical form, they have one and only one of the coordinates $\{x_A, x_B, x_C\}$ equaling zero. For example, $|\varphi\rangle = \sqrt{x_0}|000\rangle + \sqrt{x_A}|100\rangle + \sqrt{x_B}|010\rangle$ is a bipartite entangled state shared between Alice and Bob whenever $x_A, x_B > 0$. Its concurrence is readily computed to be (see Appendix A)

$$C(\varphi) = 2\sqrt{x_A x_B}. \tag{4}$$

From this we see that the concurrence is homogeneous under multiplication by a non-negative scalar. That is, $C(\alpha\varphi) = \alpha C(\varphi)$ for any $\alpha \geq 0$. We also observe that $|\varphi\rangle$ is maximally entangled iff $x_A = x_B = \frac{1}{2}$ and $x_0 = 0$.

In an LOCC protocol, each party takes turns measuring their local system and announcing the result. Since each party holds a qubit system, every local measurement can be described by a set of 2×2 Kraus operators $\{M_i\}_i$. Without loss of generality, we can assume that each M_i has been brought into upper triangular form by a local unitary. Indeed, if $U_i M_i$ is upper triangular, then we can always append $U_i^\dagger$ to the start of the next round of local measurement to recover the action of the original Kraus operator M_i. Hence we will write the Kraus operators of every local measurement as

$$M_i = \begin{pmatrix} \sqrt{a_i} & b_i \\ 0 & \sqrt{c_i} \end{pmatrix}, \tag{5}$$

where $a_i, c_i \geq 0$. The completion relation $\sum_i M_i^\dagger M_i = \mathbb{I}$ requires that $\sum_i a_i = 1$ and $\sum_i c_i \leq 1$. When party k performs a local measurement $\{M_i\}_i$ on the W-state $|x\rangle$ and obtains outcome i, then the post-measurement state $|\mathbf{y}_i\rangle = \frac{1}{\sqrt{p(i)}} M_i|\mathbf{x}\rangle$ has undergone a coordinate transformation

$$x_k \mapsto y_{i,k} = \frac{c_i}{p_i} x_k, \tag{6}$$

$$x_j \mapsto y_{i,j} = \frac{a_i}{p_i} x_j \qquad \text{for } j \neq k \text{ or } 0, \tag{7}$$

where $p_i = \langle \mathbf{x}|M_i^\dagger M_i|\mathbf{x}\rangle$. Observe the monotonicity conditions [27]

$$\sum_i p_i y_{i,k} = \sum_i c_i x_k \leq x_k \tag{8}$$

$$\sum_i p_i y_{i,j} = \sum_i a_i x_j = x_j \qquad \text{for } j \neq k \text{ or } 0. \tag{9}$$

In other words, the component of the measuring party is non-increasing on average where as the components of the non-measuring parties remain unchanged on average.

An important property of any local measurement is that it can be decomposed into a sequence of weak measurements [36]. Hence, we can envision a given LOCC W-state conversion as a tree of smooth trajectories in the positive quadrant of $\mathbb{R}^3$ taking the initial state $\mathbf{x}$ to its possible target states $\{\mathbf{y}_i\}_i$. Consequently, if $x_k \mapsto y_{i,k}$ is the component transformation for party k along one branch of some protocol, then there exists an equivalent sequence of protocols that passes through any desired coordinates between x_k and $y_{i,k}$.

Finally, we introduce three important functions of three-qubit W-class states that play an important role in this work. For coordinates (x_A, x_B, x_C) of state $|\mathbf{x}\rangle$, let $n_1, n_2, n_3 \in \{A, B, C\}$ be distinct party

labels such that $x_{n_1} \geq x_{n_2} \geq x_{n_3}$ and let $n_i(\mathbf{x}) := x_{n_i}$. Suppose also that $x_{n_2} > 0$ so that $|\mathbf{x}\rangle$ is not a product state. Then define

$$\begin{aligned}\eta(\mathbf{x}) &:= x_{n_2} + x_{n_3} - \frac{x_{n_2}x_{n_3}}{x_{n_1}} \\ \kappa(\mathbf{x}) &:= 2(x_{n_2} + x_{n_3}) - \frac{x_{n_2}x_{n_3}}{x_{n_1}} \\ \zeta(\mathbf{x}) &:= 2\sqrt{x_{n_1}x_{n_2}} + \frac{2}{3}x_{n_3}\sqrt{\frac{x_{n_1}}{x_{n_2}}} + \frac{1}{3}\frac{x_{n_2}x_{n_3}}{x_{n_1}}. \end{aligned} \quad (10)$$

With a slight abuse of the notation, we introduce one additional function. Let $\star \in \{1,2,3\}$ be any fixed party, and take $n_1, n_2 \in \{A,B,C\} \setminus \{\star\}$ as distinct party labels such that $x_{n_1} \geq x_{n_2} > 0$. Then define

$$\zeta^\star(\mathbf{x}) := 2\sqrt{x_\star x_{n_1}} + \frac{2}{3}x_{n_2}\sqrt{\frac{x_\star}{x_{n_1}}}. \quad (11)$$

It has been shown that $\eta(\mathbf{x})$, $\kappa(\mathbf{x})$, and $\zeta^\star(\mathbf{x})$ are entanglement monotones when manipulating W-class states under LOCC [9, 8, 13]; i.e.

$$\eta(\mathbf{x}) \geq \sum_i p_i \eta(\mathbf{y}_i) \quad (12)$$

for any LOCC transformation $|\mathbf{x}\rangle \to \{p_i, |\mathbf{y}_i\rangle\}$, and likewise for $\kappa(\mathbf{x})$ and $\zeta^\star(\mathbf{x})$. Notice that since the W-class is closed under LOCC [21], $|\mathbf{x}\rangle \to \{p_i, |\mathbf{y}_i\rangle\}$ describes the most general type of pure-state transformations possible by LOCC and hence we have monotonicity under arbitrary pure state transformations. Below we show that the same is true for the $\zeta(\mathbf{x})$, and we also reveal its operational meaning in terms of random-party concurrence distillation.

3 Total Random-Party Distillation

For the task of total random-party EPR distillation, the original Fortescue-Lo (F-L) protocol gives $E_2^{(\mathrm{rnd})}(W) = 1$ for the W state $|W\rangle$ [22]. The F-L protocol consists of each party performing the same binary-outcome measurement with Kraus operators

$$M_0^\epsilon = \sqrt{1-\epsilon}|0\rangle\langle 0| + |1\rangle\langle 1|, \quad M_1^\epsilon = \sqrt{\epsilon}|0\rangle\langle 0|. \quad (13)$$

Consider one iteration of the protocol in which each party performs this measurement. If all three parties obtain outcome M_0, then the post-measurement state is $M_0 \otimes M_0 \otimes M_0 |W\rangle \propto |W\rangle$; the W-state is left on changed, and they all repeat the same measurement in the next iteration of the protocol. On the other hand if one party obtains outcome M_1 while the other two obtain M_0, then the latter are left in an EPR state. The final possibility is that at least two parties obtain outcome M_1 which breaks all the entanglement in the state. For any $\delta > 0$, the value ϵ can be chosen sufficiently small and the number of iterations can be chosen sufficiently large such that an EPR state is obtained with probability $> 1 - \delta$. In Section 5.1 we identify a precise trade-off between the value δ and the number of rounds needed to obtain a success probability $> 1 - \delta$. The F-L protocol was later generalized to other W-class states. It was shown that

$$E_2^{(\mathrm{rnd})}(\mathbf{x}) = \kappa(\mathbf{x}) \quad (14)$$

for any W-state $|\mathbf{x}\rangle$ such that $x_0 = 0$ [9].

Our goal now is to establish a similar result for concurrence distillation.

Lemma 1. Let $\zeta(\mathbf{x}) := 2\sqrt{x_{n_1}x_{n_2}} + \frac{2}{3}x_{n_3}\sqrt{\frac{x_{n_1}}{x_{n_2}}} + \frac{1}{3}\frac{x_{n_2}x_{n_3}}{x_{n_1}}$, as introduced above. Then

$$C^{(\mathrm{rnd})}(\mathbf{x}) = \zeta(\mathbf{x}) \quad (15)$$

for any W-class state $|\mathbf{x}\rangle$. Furthermore, $\zeta(\mathbf{x})$ is an entanglement monotone that strictly decreases whenever party n_2 or n_3 performs a non-trivial measurement.

Remark. The equality in Eq. (15) holds for all W-class states, even those for which $x_0 \neq 0$. This is in contrast to Eq. (14) which holds only under the assumption $x_0 = 0$. The value $E_2^{(\mathrm{rnd})}(\mathbf{x})$ for general W-class states remains an open problem.

Proof. The monotonicity of $\zeta(\mathbf{x})$ and it strictly decreasing under measurement of party n_2 and n_3 is proven in Appendix B. Here we prove the achievability. Consider an arbitrary W-class state $|\mathbf{x}\rangle = \sqrt{x_0}|000\rangle + \sqrt{x_A}|100\rangle + \sqrt{x_B}|010\rangle + \sqrt{x_C}|001\rangle$. Since we are trying to optimize the total average bipartite concurrence for any two parties, without loss of generality we can relabel the parties and assume that $x_A \geq x_B \geq x_C$. As depicted in Fig. 1, the following three-step protocol achieves a total average bipartite concurrence of $\overline{C} = \zeta(\mathbf{x}) - \delta$ for arbitrary $\delta > 0$.

Step 1. Let Charlie perform a measurement described by the following Kraus Operators,

$$M_0 = \sqrt{\frac{x_C}{x_B}}|0\rangle\langle 0| + |1\rangle\langle 1|, \quad M_1 = \sqrt{1 - \frac{x_C}{x_B}}|0\rangle\langle 0|. \quad (16)$$

There will be two measurement outcomes. Using the transformation rule of Eq. (9), the post-measurement state of outcome 1 is the bipartite pure state $|\varphi\rangle^{AB} = \frac{1}{\sqrt{p_1}}\sqrt{1 - \frac{x_C}{x_B}}(\sqrt{x_0}|00\rangle + \sqrt{x_A}|10\rangle + \sqrt{x_B}|01\rangle)$. Hence from Eq. (4), the concurrence of this state weighted by its probability is $C_1 := p_1 C(\varphi) = 2(1 - \frac{x_C}{x_B})\sqrt{x_A x_B}$. The post-measurement state of outcome 0 is still a tripartite W-class state, which, when weighted by the probability of outcome 0 has components $(\frac{x_A x_C}{x_B}, x_C, x_C)$.

Step 2. Continuing with the post-measurement state of outcome 0, Bob and Charlie then perform n iterations of a *restricted* Fortescue-Lo protocol, which involves just Bob and Charlie performing n

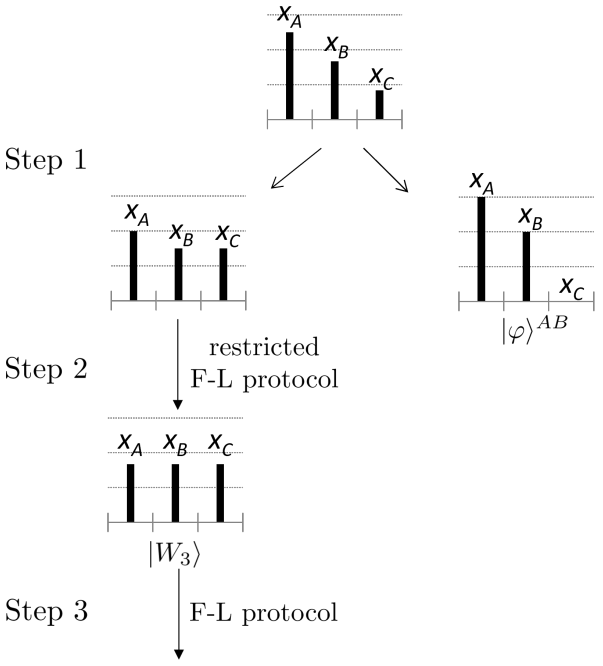


Figure 1: A depiction of the protocol described in Lemma 1.

iterations of the the weak measurement specified by Kraus operators in Eq. (13). Here, n is a parameter that we are free to choose, and for a given choice, we take $\epsilon > 0$ such that $(1-\epsilon)^n = x_B/x_A$. So when $n \to \infty$ we have $\epsilon \to 0$.

For each iteration of Bob and Charlie measuring $\{M_0^\epsilon, M_1^\epsilon\}$, they halt if at least one of them obtains outcome 1. If the both obtain outcome 0 then they proceed with another iteration of measurement. This is done for n total iterations. One can verify that if the protocol has not halted by the j^{th} iterations, then outcome 00 in the j^{th} iteration yields the unnormalized state with coordinates $((1-\epsilon)^{2j}\frac{x_A x_C}{x_B}, (1-\epsilon)^j x_C, (1-\epsilon)^j x_C)$. On the other hand, if either outcome 01 or 10 is obtained, then the concurrence of the post-measurement state $|\phi_{2,j}\rangle$, weighted by its probability, is

$$C_2^j := (1-\epsilon)^{(j-1)3/2} \cdot 4\epsilon\sqrt{1-\epsilon}\cdot x_C\sqrt{\frac{x_A}{x_B}}. \quad (17)$$

Should the protocol not halt for all n iterations, we find that all all three coefficients end up being equal, $(1-\epsilon)^{2n}\frac{x_A x_C}{x_B} = (1-\epsilon)^n x_C = \frac{x_B x_C}{x_A}$, since $(1-\epsilon)^n = x_B/x_A$. After normalization, this state is the W-state $|W\rangle$.

Step 3. The original F-L protocol is performed for n' iterations on $|W\rangle$ with all three parties using measurement $\{M_0^{\epsilon'}, M_1^{\epsilon'}\}$ and ϵ' a freely chosen parameter. If a bipartite entangled state is obtained in the j^{th} iteration of the F-L protocol, then the probability-weighted concurrence is

$$C_3^j = (1-\epsilon')^{2(j-1)}\cdot 6\epsilon'(1-\epsilon')\frac{x_B x_C}{x_A}. \quad (18)$$

As we let $n, n \to \infty$ and $\epsilon' \to 0$, the expected concurrence is given by

$$\begin{aligned}\overline{C} :=& C_1 + \lim_{n\to\infty}\sum_{j=1}^{n} C_2^j + \lim_{\epsilon'\to 0}\lim_{n'\to\infty}\sum_{j=1}^{n'} C_3^j \\ =& 2\sqrt{x_A x_B} + \frac{2}{3}x_C\sqrt{\frac{x_A}{x_B}} + \frac{1}{3}\frac{x_B x_C}{x_A} \\ =& \zeta(\mathbf{x}). \end{aligned} \quad (19)$$

□

Notice that this protocol uses two unbounded-round subroutines: the restricted F-L protocol (step 2) and the total F-L protocol (step 3). In Section 5.2.2 we will explore finite-round approximations and how the expected concurrence changes as we trade round numbers between these two subroutines.

4 ★-Random-Party Distillation

The problem of $\star$-random-party EPR distillation in W-class states has been partially solved in Ref. [13]. Namely, it was shown that

$$E_2^{(\star\text{-rnd})}(\mathbf{x}) = \begin{cases} 2x_\star & \text{if} \quad x_\star \neq \max\{x_A, x_B, x_C\} \\ 2\eta(\mathbf{x}) & \text{if} \quad x_\star = \max\{x_A, x_B, x_C\} \end{cases} \quad (20)$$

for any W-state $|\mathbf{x}\rangle$ such that $x_0 = 0$. Similar to the scenario of total random-party distillation, the value of $E_2^{(\star\text{-rnd})}(\mathbf{x})$ is unknown when $x_0 \neq 0$. However, the RHS of Eq. (20) still serves as an upper bound due to the monotonicity of $x_\star$ and $\eta(\mathbf{x})$. Furthermore, the protocol achieving these distillation probabilities requires no more than three rounds of LOCC.

Our contribution here is solving the $\star$-Random-Party distillation problem for concurrence.

Lemma 2. For arbitrary $\star \in \{A, B, C\}$, let $\zeta^\star(\mathbf{x}) = 2\sqrt{x_\star x_{n_1}} + \frac{2}{3}x_{n_2}\sqrt{\frac{x_\star}{x_{n_1}}}$ with $n_1, n_2 \in \{A, B, C\} \setminus \{\star\}$ satisfying $x_{n_1} \geq x_{n_2} > 0$. Then

$$C^{(\star\text{-rnd})}(\mathbf{x}) = \zeta^\star(\mathbf{x}). \quad (21)$$

Proof. In Ref. [13] it was shown that $\zeta^\star(\mathbf{x})$ is an entanglement monotone that that strictly decreases on average whenever party $\star$ or n_1 performs a non-trivial measurement. What remains to be demonstrated is that $\zeta^\star(\mathbf{x})$ is indeed an achievable concurrence distillation average for pairs $(\star, x_{n_1})$ and $(\star, x_{n_2})$.

Without loss of generality we assume $x_A > 0$ since otherwise the lemma is trivially true. The distillation protocol is very similar to the one given in Lemma 1. Formally, we replace $x_\star \leftrightarrow x_A$, $x_{n_1} \leftrightarrow x_B$, and $x_{n_2} \leftrightarrow x_C$. However, while it was assumed that $x_A \geq x_B \geq x_C$ in Lemma 1, we only assume that $x_B \geq x_C$ in this protocol.

Step 1 of this protocol is then the same Step 1 as Lemma 1, with Charlie performing the measurement given in Eq. (16). The weighted concurrence for outcome 1 is $C_1 := p_1 C(\varphi) = 2(1 - \frac{x_C}{x_B})\sqrt{x_A x_B}$. In step 2, Bob and Charlie again perform the restricted F-L protocol; however now there is no halting round. The weighted concurrence of a bipartite state $|\phi_{2,j}\rangle$ obtained in round j is C_2^j, as given in Eq. (17). Note that Alice then never performs a measurement in this protocol. After replacing our system labels by $\{\star, n_1, n_2\}$, the total expected concurrence as the restricted F-L protocol in step 2 continues indefinitely is

$$\begin{aligned}\overline{C} &= C_1 + \lim_{n\to\infty} \sum_{j=1}^{n} C_2^n = 2\sqrt{x_\star x_{n_1}} + \frac{2}{3} x_{n_2} \sqrt{\frac{x_\star}{x_{n_1}}} \\ &= \zeta^\star(\mathbf{x}). \end{aligned} \tag{22}$$

□

5 Round Complexity in Random-Party Distillation

We now turn to the question of round complexity, which is the motivating topic of this work.

5.1 Lower Bounds in EPR Distillation

Suppose the parties are only afforded a finite number of LOCC rounds. What is the largest achievable probability of obtaining an EPR state starting from $|W\rangle$? Here we prove a lower bound on this probability as a function of round number.

Consider any finite-round random EPR distillation protocol $\mathcal{P}$. We can assume without loss of generality that every branch in the protocol ends with either an EPR state $|\Psi^+\rangle$ or a product state; this is because every weakly entangled state $|\phi\rangle$ can be transformed into $|\Psi^+\rangle$ with some probability without extending the number of rounds. Any branch that terminates with an EPR state will be called a *success branch* (many different success branches will overlap).

Our first step will be to argue that we can always transform the protocol $\mathcal{P}$ into one in which Alice and Bob just perform diagonal measurements. As described in the introduction, we can characterize each local measurement by upper-triangular Kraus operators, $M_i = \begin{pmatrix} \sqrt{a_i} & b_i \\ 0 & \sqrt{c_i} \end{pmatrix}$. Notice that the composition of two such matrices has form

$$M_i M_i' = \begin{pmatrix} \sqrt{a_i a_i'} & b_i\sqrt{c_i'} + b_i'\sqrt{a_i} \\ 0 & \sqrt{c_i c_i'} \end{pmatrix}. \tag{23}$$

If we then consider the full set of Kraus operators constituting protocol $\mathcal{P}$, each success branch λ will be characterized by a product Kraus operator of the form

$$T_\lambda = \begin{pmatrix} \sqrt{a_{1,\lambda}} & b_{1,\lambda} \\ 0 & \sqrt{c_{1,\lambda}} \end{pmatrix} \otimes \begin{pmatrix} \sqrt{a_{2,\lambda}} & b_{2,\lambda} \\ 0 & \sqrt{c_{2,\lambda}} \end{pmatrix} \otimes \begin{pmatrix} \sqrt{a_{3,\lambda}} & b_{3,\lambda} \\ 0 & \sqrt{c_{3,\lambda}} \end{pmatrix},$$

where each matrix in the tensor product is obtained by concatenating all the local operators along branch λ performed by Alice, Bob, and Charlie, respectively. Since the only maximally entangled two-qubit state of the form $|\mathbf{x}\rangle$ is $|\Psi^+\rangle$, we must have $T_\lambda|W\rangle \propto |\Psi^+\rangle|0\rangle$, with $|\Psi^+\rangle$ being held by some pair of parties. The branch probability is then given by $|\langle\Psi^+|\langle 0|T_\lambda|W\rangle|^2$, which is independent of the off-diagonal terms $b_{k,\lambda}$ in T_λ. Furthermore, by repeatedly applying the composition rule of Eq. (23), the diagonal terms in T_λ only depend on the diagonal terms of the individual local measurements in each round. Hence, we would obtain the same success branch probabilites $|\langle\Psi^+|\langle 0|T_\lambda|W\rangle|^2$ if Alice and Bob only performed the diagonal parts of their local Kraus operators. In general, the diagonal parts of the Kraus operators will not form a complete measurement themselves, but they can easily be completed without affecting the overall success of the protocol. Specifically, if the $M_i = \begin{pmatrix} \sqrt{a_i} & b_i \\ 0 & \sqrt{c_i} \end{pmatrix}$ are the Kraus operators for some local measurement, then it is replaced by the Kraus operators $M_i' = \begin{pmatrix} \sqrt{a_i} & 0 \\ 0 & \sqrt{c_i} \end{pmatrix}$ with an additional outcome $M_0' = \begin{pmatrix} 0 & 0 \\ 0 & \sqrt{\sum_i |b_i|^2} \end{pmatrix}$. Hence we have established the following simplification.

Proposition 1. The optimal success probability for random-party EPR distillation of $|W\rangle$ in N rounds can always be obtained by an N-round protocol in which all parties perform local measurements with diagonal Kraus operators.

With this proposition, let us consider an arbitrary distillation protocol $\mathcal{P}$ with all measurements in diagonal form. As described in Section 2, using weak measurement theory, we can depict the protocol as a tree that starts with the state $|W\rangle$ and evolves continuously along different branches. Crucially, the conversion of a general protocol to one in which the coordinates (x_A, x_B, x_C) transform smoothly can be done without changing the number of LOCC rounds. This is because each local measurement is partitioned into a sequence of weak measurements with no communication needed between each measurement in the sequence.

To proceed with the analysis, we will need to introduce some concepts that characterize a given protocol $\mathcal{P}$. We begin by defining an important class of states that appear in $\mathcal{P}$.

Definition 1. A state $|\mathbf{x}\rangle$ with ordered components $x_{n_1} \geq x_{n_2} \geq x_{n_3}$ obtained in $\mathcal{P}$ is called a *block state* if (i) $x_{n_1} = x_{n_2}$ and $x_{n_3} \neq 0$, (ii) a party whose component value is maximal is the next to perform a local measurement on $|\mathbf{x}\rangle$, and (iii) an EPR state is obtained with some nonzero probability from $|\mathbf{x}\rangle$.

Using the concept of a block state, we can analyze an arbitrary random-party distillation protocol in a systematic way. First, with continuous evolution of the states in the protocol, it follows that the party having the largest component value cannot change along a success branch without first passing through a block state. We can then partition the protocol in blocks determined by when there is a change in the party having the largest component value. More precisely, define a sub-block of $\mathcal{P}$ as any sub-tree in the overall protocol tree that begins with a block state $|\mathbf{x}\rangle$ and terminates as soon as it reaches some latter block state $|\mathbf{y}\rangle$ (which may be equivalent to $|\mathbf{x}\rangle$), an EPR state, or a product state (see Fig. 2). Each terminal block state in one sub-block is then the initial block state of a subsequent sub-block (unless the protocol halts), and the entire LOCC protocol can be divided into a disjoint union of sub-blocks. We say that an N-block protocol is one that traverses N different sub-blocks along the longest branch of the protocol. We further organize the protocol into N layers such that all sub-blocks belonging to layer k have an initial block state laying on a branch that has previously traveled through $k-1$ sub-blocks. Finally, we introduce the notion of a canonical sub-block.

Definition 2. A sub-block is called *canonical* if all its terminal block states are $|W\rangle$ (see Fig. 3 (a)). Hence a canonical sub-block always ends with $|W\rangle$, an EPR state, or a product state. A canonical sub-block is called *W-canonical* if it begins with $|W\rangle$ and returns to $|W\rangle$ along only one branch within the sub-block (see Fig. 3 (b)).

Note that all sub-blocks in the N^{th} layer are canonical since they obtain only EPR and product states.

Our first goal will be to reduce any N-block protocol $\mathcal{P}$ into another $\mathcal{P}'$ whose sub-blocks are W-canonical. The following technical lemma and its corollary provide the main ingredients for this procedure.

Lemma 3. Suppose that an arbitrary state $|\mathbf{x}\rangle = (x_{n_1}, x_{n_2}, x_{n_3})$ is transformed into either $|W\rangle$, an EPR pair, or a product state with respective probabilities P_W, P_{EPR}, and P_F. Further suppose that party $n_1(\mathbf{x})$ maintains the largest component value in each of the final states (if more than one party has maximal value, fix $n_1(\mathbf{x})$ to be one them). Then

$$P_{\text{EPR}} \leq 2(x_{n_2} + x_{n_3}) - 4\frac{x_{n_2}x_{n_3}}{x_{n_1}} \tag{24}$$

$$P_W = \frac{3}{2}(x_{n_2} + x_{n_3}) - \frac{3}{4}P_{\text{EPR}}. \tag{25}$$

Moreover, the upper bound on P_{EPR} is achievable.

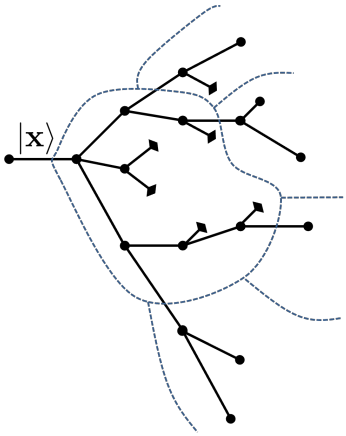


Figure 2: Any random-distillation protocol $\mathcal{P}$ can partitioned into disjoint sub-blocks (enclosed by dashed line). Each sub-block starts with a block state and terminates with either another block state, an EPR state, or a product state. EPR and product states are indicated by diamond end points. Block states are obtained along any edge intersecting a dashed line. A terminal block state of one sub-block is the initial block state of a subsequent sub-block unless the protocol halts.

Proof. Since we are considering only protocols with diagonal local Kraus operators, the average component value for each party remains unchanged. Hence, $x_{n_1} = \frac{1}{2}P_{\text{EPR}} + \frac{1}{3}P_W + P_F$. Since $P_F = 1 - P_{\text{EPR}} - P_W$ and $x_{n_1} + x_{n_2} + x_{n_3} = 1$, we have $x_{n_2} + x_{n_3} = \frac{1}{2}P_{\text{EPR}} + \frac{2}{3}P_W$, which is Eq. (25). On the other hand, the η monotone says that $\eta(\mathbf{x}) = x_{n_2} + x_{n_3} - \frac{x_{n_2}x_{n_3}}{x_{n_1}} \geq \frac{1}{2}P_{\text{EPR}} + \frac{1}{3}P_W$. Combining this with Eq. (25) yields Eq. (24).

The upper bound on P_{EPR} is achievable using an "equal or vanish" protocol [8]. This involves party $n_k(\mathbf{x})$, for $k = 2, 3$, performing a binary-measurement with Kraus operators $M_1 = \text{Diag}[\sqrt{x_{n_k}/x_{n_1}}, 1]$ and $M_2 = \text{Diag}[\sqrt{1 - x_{n_k}/x_{n_1}}, 0]$. If they both obtain outcome 1 the resulting state is $|W\rangle$, and this occurs with probability $P_W = 3\frac{x_{n_2}x_{n_3}}{x_{n_1}}$. On the other hand, if one and only one party obtains outcome 1, then an EPR pair is obtained. This occurs with total probability

$$\begin{aligned} P_{\text{EPR}} &= 2x_{n_2}(1 - \frac{x_{n_3}}{x_{n_1}}) + 2x_{n_3}(1 - \frac{x_{n_2}}{x_{n_1}}) \\ &= 2(x_{n_2} + x_{n_3}) - 4\frac{x_{n_2}x_{n_3}}{x_{n_1}}. \end{aligned}$$

□

Lemma 3 requires that party $n_1(\mathbf{x})$ have the largest component value in all terminal states of the sub-block. We next relax this condition and derive similar bounds on P_{EPR} and P_W when there is no party with maximal value in all the outcomes. To conduct this analysis, let a_i and c_i be the numbers characterizing the measurement on an initial block state $|\mathbf{x}\rangle$. Let $\mathcal{A}$ (resp. $\mathcal{C}$) be the set of indices

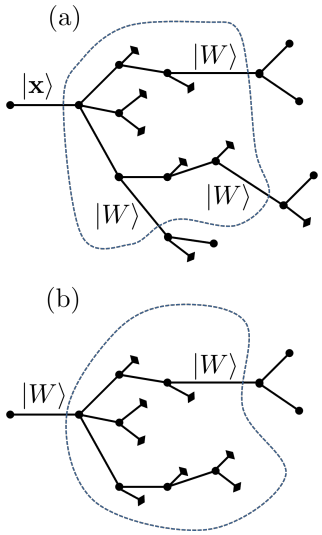


Figure 3: (a) A canonical sub-block is one in which every terminating block state is $|W\rangle$. (b) A W-canonical sub-block is a canonical sub-block whose initial block state is $|W\rangle$ and that has at most one terminating block state.

such that $a_i > c_i$ (resp. $c_i > a_i$). Note that we can always assume that $a_i \neq c_i$ for every branch i. Indeed, suppose that $a_i = c_i$, which means the state is left unchanged along branch i. If all other branches lead to a higher expected EPR probability than the EPR probability along branch i, then one can just remove branch i and scale the probability of the other branches up to unity. On the other hand, if the subsequent sub-protocol along branch i has a higher expected EPR probability than that along the other branches, one can just perform this sub-protocol on the original state $|\mathbf{x}\rangle$. Therefore, all branches with $a_i = c_i$ have been eliminated without decreasing the overall EPR probability or the number of rounds. Since $|\mathbf{x}\rangle$ is a block state, we have $x_{n_1} = x_{n_2} \geq x_{n_3}$. The post-measurement states of $|\mathbf{x}\rangle$ will thus split into two categories: for outcomes $i \in \mathcal{A}$ we will have $a_i x_{n_1} \geq a_i x_{n_3} > c_i x_{n_1}$ and for outcomes $i \in \mathcal{C}$ we will have $c_i x_{n_1} > a_i x_{n_1} \geq a_i x_{n_3}$. In both these cases, the conditions of Lemma 3 apply since the party having maximal component cannot switch without passing through another block state, and by assumption, the only block state obtained within the given sub-block is $|W\rangle$ (for which all parties have maximal component value). We can then prove the following corollary.

Corollary 1. Suppose that $|\mathbf{x}\rangle$ is the initial block state of a canonical sub-block and a local measurement is performed with Kraus operator parameters a_i and c_i. Let $a = \sum_{i\in\mathcal{A}} a_i$, $c = \sum_{i\in\mathcal{A}} c_i$, and $\tilde{a} = a - c$. Then

$$P_{\text{EPR}} \leq 2x_{n_1}(1-\tilde{a}) + 2x_{n_3} - 4x_{n_3}(1-\tilde{a})^2 \quad (26)$$

$$P_W = \frac{3}{2}x_{n_1}(1-\tilde{a}) + \frac{3}{2}x_{n_3} - \frac{3}{4}P_{\text{EPR}}. \quad (27)$$

Furthermore, the upper bound of P_{EPR} can be saturated using binary-outcome measurements that obtain $|W\rangle$ along just one branch.

Proof. Group outcomes i into sets $\mathcal{A}$ and $\mathcal{C}$ as described above. Applying Lemma 3 to each of the post-measurement states yields

$$\begin{aligned}
P_{\text{EPR}} \leq& \sum_{i\in\mathcal{A}} (2x_{n_1}c_i + 2x_{n_3}a_i - 4x_{n_3}c_i) \\
&+ \sum_{i\in\mathcal{C}} \left(2x_{n_1}a_i + 2x_{n_3}a_i - 4x_{n_3}\frac{a_i^2}{c_i}\right) \\
\leq& 2x_{n_1}(1-\tilde{a}) + 2x_{n_3} - 4x_{n_3}\left(c + \frac{(1-a)^2}{1-c}\right) \\
\leq& 2x_{n_1}(1-\tilde{a}) + 2x_{n_3} - 4x_{n_3}(1-(a-c))^2 \\
=& 2x_{n_1}(1-\tilde{a}) + 2x_{n_3} - 4x_{n_3}(1-\tilde{a})^2, \quad (28)
\end{aligned}$$

where the second inequality follows from the Cauchy-Schwarz Inequality:

$$\left(\sum_{i\in\mathcal{C}} \sqrt{c_i}\cdot\sqrt{c_i}\right)\left(\sum_{i\in\mathcal{C}} \frac{a_i}{\sqrt{c_i}}\cdot\frac{a_i}{\sqrt{c_i}}\right) \geq \left(\sum_{i\in\mathcal{C}} a_i\right)^2,$$

and the third comes from the fact that $c + \frac{(1-a)^2}{1-c} \geq (1-(a-c))^2$. The equality of P_W again follows from Lemma 3 as

$$\begin{aligned}
P_W &= \frac{3}{2}x_{n_1}\left(\sum_{i\in\mathcal{A}} c_i + \sum_{i\in\mathcal{C}} a_i\right) + \frac{3}{2}x_{n_3} - \frac{3}{4}P_{\text{EPR}} \\
&= \frac{3}{2}x_{n_1}(1-\tilde{a}) + \frac{3}{2}x_{n_3} - \frac{3}{4}P_{\text{EPR}}. \quad (29)
\end{aligned}$$

The upper bound on P_{EPR} can be achieved by having party $n_1(\mathbf{x})$ perform a binary outcome measurement with Kraus operators $M_1 = \text{Diag}[\sqrt{1-\tilde{a}}, 1]$ and $M_2 = \text{Diag}[\sqrt{\tilde{a}}, 0]$. For outcome 2, an EPR pair can be subsequently obtained between parties $n_2(\mathbf{x})$ and $n_3(\mathbf{x})$ with probability $2\tilde{a}x_{n_3}$. For outcome 1, the "equal or vanish" protocol is subsequently performed, as described in Lemma 3. Since $x_{n_1} = x_{n_2}$, the total EPR probbility is then

$$P_{\text{EPR}} = 2\tilde{a}x_{n_3} + 2(1-\tilde{a})(x_{n_1} + x_{n_3}) - 4x_{n_3}(1-\tilde{a})^2.$$

From the "equal or vanish" protocol we also obtain a W-state with probability $P_W = 3x_{n_3}(1-\tilde{a})^2$. □

We next convert an arbitrary canonical sub-block into a W-canonical sub-block.

Lemma 4. Any canonical sub-block of $\mathcal{P}$ can be replaced by a W-canonical sub-block without decreasing the total EPR probability or increasing the total number of sub-blocks in the protocol.

Proof. Let S be any canonical sub-block starting with block state $|\mathbf{x}\rangle$. Let a_i and c_i be the Kraus operator parameters of the initial measurement as before. We transition S into a W-canonical sub-block in two steps. First, we limit the the total number of branches within the sub-block that obtain $|W\rangle$ to at most one. In general, there may be multiple branches within the sub-block that obtain state $|W\rangle$. For each of these, the protocol $\mathcal{P}$ specifies a

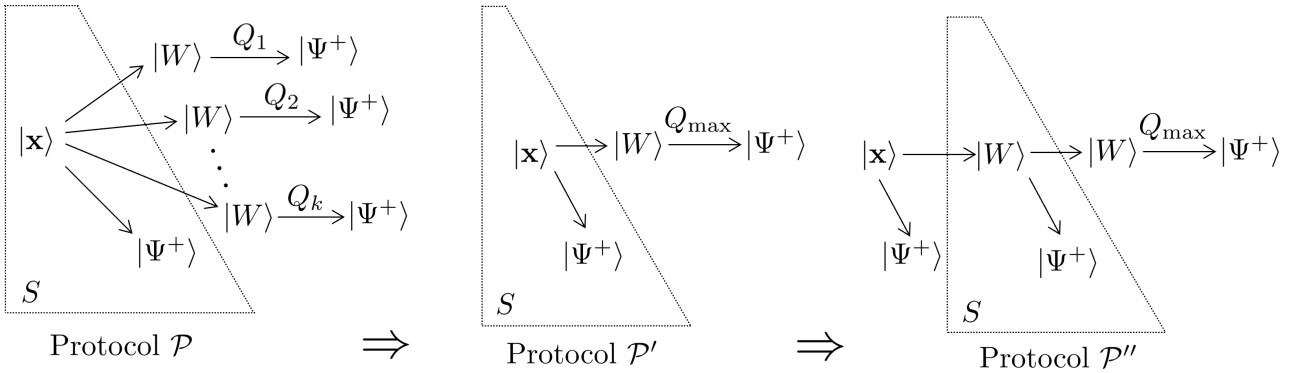


Figure 4: Lemma 4 describes the reduction of a canonical sub-block S to a W-canonical sub-block in two steps. The first is to reduce the number of terminal $|W\rangle$ states to just one. The second replaces the initial block state $|\mathbf{x}\rangle$ with $|W\rangle$ by performing a prior measurement on $|\mathbf{x}\rangle$ in the preceding sub-block.

subsequent distillation sub-protocol to be performed on $|W\rangle$. Let $Q_{\max}$ denote the maximum probability of obtaining EPR pairs from $|W\rangle$ among all these sub-protocols. Therefore, if P_S is the total probability of obtaining EPR pairs when starting from the block state $|\mathbf{x}\rangle$, we can use Corollary 1 to obtain the bound

$$\begin{aligned} P_S &\leq P_{\text{EPR}} + P_W \cdot Q_{\max} \\ &= P_{\text{EPR}} + \left(\frac{3}{2}x_{n_1}(1-\tilde{a}) + \frac{3}{2}x_{n_3} - \frac{3}{4}P_{\text{EPR}}\right) Q_{\max} \\ &= \frac{3}{2}x_{n_1}(1-\tilde{a}) + \frac{3}{2}x_{n_3} + \left(1 - \tfrac{3}{4}Q_{\max}\right) P_{\text{EPR}}. \end{aligned} \tag{30}$$

Since the coefficient of P_{EPR} is non-negative on the RHS, we can further bound P_S using Eq. (26). This bound can be saturated by performing the optimal binary-outcome measurement scheme of Corollary 1 on the initial block state $|\mathbf{x}\rangle$. For the single $|W\rangle$ outcome, the original sub-protocol that attains EPR probability $Q_{\max}$ is then performed. Hence, if P'_S denotes the total EPR probability starting from $|\mathbf{x}\rangle$ in our modified protocol $\mathcal{P}'$, then

$$\begin{aligned} P'_S &= 2x_{n_1}(1-\tilde{a}) + 2x_{n_3} - 4x_{n_3}(1-\tilde{a})^2 \\ &\quad + 3x_{n_3}(1-\tilde{a})^2 \cdot Q_{\max} \\ &\geq P_S. \end{aligned} \tag{31}$$

Next, we modify the initial block state $|\mathbf{x}\rangle$ to be $|W\rangle$ (if $|\mathbf{x}\rangle \neq |W\rangle$). This is accomplished by stochastically transforming $|\mathbf{x}\rangle$ into $|W\rangle$ or $|\Psi^+\rangle$ in the sub-block prior to S by having party $n_3(\mathbf{x})$ measure with Kraus operators $M_1 = \text{Diag}[\sqrt{x_{n_3}/x_{n_1}}, 1]$ and $M_2 = \text{Diag}[\sqrt{1 - x_{n_3}/x_{n_1}}, 0]$. The probability of obtaining $|W\rangle$ (outcome 1) is $3x_{n_3}$ while the probability of $|\Psi^+\rangle$ (outcome 2) is $2(x_{n_1} - x_{n_3})$. Now $|W\rangle$ is the initial block state of S. The optimal transformation of Corollary 1 is performed on $|W\rangle$ with the measurement parameter $\tilde{a}$ being the same as in $\mathcal{P}'$. This yields EPR states within sub-block S with probability $P_{\text{EPR}} = \frac{2}{3}(1-\tilde{a}) + \frac{2}{3} - \frac{4}{3}(1-\tilde{a})^2 = \frac{2}{3}(3-2\tilde{a})\tilde{a}$, while the W-state is obtained with probability $P_W = (1-\tilde{a})^2$. For the $|W\rangle$ outcome, the $Q_{\max}$ sub-protocol is again performed in subsequent sub-blocks. Letting P''_S denote the total EPR probability starting from $|\mathbf{x}\rangle$ in this new protocol $\mathcal{P}''$, we have

$$\begin{aligned} P''_S =& 2(x_{n_1} - x_{n_3}) + 2x_{n_3}(3-2\tilde{a})\tilde{a} \\ &+ 3x_{n_3}(1-\tilde{a})^2 \cdot Q_{\max}. \end{aligned} \tag{32}$$

One can verify that

$$P''_S - P_S \geq P''_S - P'_S = 2a(x_{n_1} - x_{n_3}) \geq 0. \tag{33}$$

Hence, we have transformed an arbitrary canonical sub-block S into a W-canonical sub-block without decreasing the total EPR probability or increasing the number of sub-blocks in the protocol.

□

Corollary 2. Any N-block LOCC protocol $\mathcal{P}$ performing random-party EPR distillation can be converted into an N-block protocol consisting only of W-canonical sub-blocks and obtaining EPR pairs with at least as large as a probability as $\mathcal{P}$.

Proof. All sub-blocks at layer N are canonical, and we perform Lemma 4 to convert them to W-canonical sub-blocks. We then move backward to layer $N-1$. All of these sub-blocks must be canonical since their terminal block states have been converted to $|W\rangle$ in the previous step. We can thus transform all these canonical sub-blocks to W-canonical sub-blocks. Proceeding iteratively layer-by-layer back to the original state $|W\rangle$, we obtain the stated structure of Corollary 2. □

We now possess the necessary components to prove a general lower bound.

Theorem 1. Let P_{tot} be the total probability of obtaining EPR pairs from $|W\rangle$ in some LOCC random-party distillation protocol $\mathcal{P}$. Then

$$\text{Rounds in } \mathcal{P} \geq \frac{1}{1 - P_{tot}} - 2. \tag{34}$$

Moreover, this bound is tight.

Proof. Suppose that $\mathcal{P}$ is a finite-round protocol; otherwise Eq. (34) holds trivially. In light of Corollary 2, we can assume without loss of generality that $\mathcal{P}$ consists only of W-canonical sub-blocks. For a non-negative number a_k, the probability of distilling EPR pairs in the k^{th} sub-block is $\frac{2}{3}(2-a_k)-\frac{4}{3}B_k$, where $B_k=(1-a_k)^2$ is the probability of obtaining $|W\rangle$ and therefore carrying the protocol onto sub-block $k+1$. Consequently, the total EPR probability for an N-block protocol is

$$\begin{aligned}P_{\text{tot}} =&\frac{4}{3}-\frac{2}{3}a_1-\frac{4}{3}B_1+B_1(\frac{4}{3}-\frac{2}{3}a_2-\frac{4}{3}B_2)\\ &+...+\frac{2}{3}(B_1B_2\cdots B_{N-1})\\ =&\frac{4}{3}-\frac{2}{3}a_1-\frac{2}{3}B_1a_2\\ &-...-\frac{2}{3}(B_1\cdots B_{N-2})a_{N-1}-\frac{2}{3}(B_1\cdots B_{N-1})\\ =&\frac{4}{3}-\frac{2}{3}a_1-\frac{2}{3}\sum_{i=1}^{N-1}\prod_{j=1}^{i}B_ja_{i+1}, \end{aligned}\tag{35}$$

where $a_N=1$. To optimize, we compute partial derivatives, which ultimately leads to a system of equations:

$$\begin{aligned}0&=1-2(1-a_{N-1})\\ 0&=1-2(1-a_{N-2})(a_{N-1}+B_{N-1})\\ 0&=1-2(1-a_{N-3})(a_{N-2}+B_{N-2}(a_{N-1}+B_{N-1}))\\ &\vdots\end{aligned}$$

The solutions are given by the recursive formula $a_k=\frac{a_{k+1}}{1+a_{k+1}}$ for $k=1,...,N-1$. Thus, $a_k=\frac{1}{N-(k-1)}$ and $B_k=\left(\frac{N-k}{N-(k-1)}\right)^2$. Substituting back into P_{tot}, we find that for any N-block LOCC distillation of $|W\rangle$,

$$\begin{aligned}P_{\text{tot}} &\leq \frac{4}{3}-\frac{2}{3}\frac{1}{N}-\frac{2}{3}\sum_{i=1}^{N-1}\prod_{j=1}^{i}\left(\frac{N-j}{N-(j-1)}\right)^2\left(\frac{1}{N-i}\right)\\ &=\frac{4}{3}-\frac{2}{3}\frac{1}{N}-\frac{2}{3}\sum_{i=1}^{N-1}\frac{N-i}{N^2}\\ &=1-\frac{1}{3N}.\end{aligned}\tag{36}$$

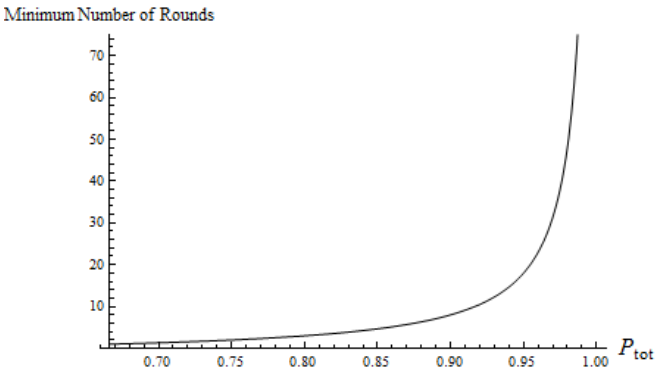


Figure 5: Minimum number of LOCC rounds to obtain EPR pairs from $|W\rangle$ with probability P_{tot}. This lower bound is tight.

This upper bound can be achieved by using measurements with the optimal values of a_k. Finally, in terms of LOCC round number, we have that each sub-block contains at least three rounds unless it is the final sub-block, which only requires one sub-block since it involves fixed-party distillation. Therefore, we obtain Eq. (34) with a corresponding plot in Fig. 5. □

Remark. Throughout this section we have been modeling an LOCC protocol as a discrete sequence time steps in which only a single party measures at each step. In this model, the round number of the protocol is the largest number of times the measuring party switches in some branch. By only allowing one party to measure at a time, the protocol is easier to analyze. However, a more general model allows all parties to measure at each time step. This has been referred to as *broadcast* LOCC (BLOCC) in Ref. [24], and this model has also been consider in Ref. [42]. In the BLOCC setting, a protocol's round number is then just the largest number of broadcasts made along some branch. In Section 5.2.2, we adopt the BLOCC model to count the number of iterations in the F-L protocol for convenience.

Notice that the BLOCC round complexity of a task can be reduced by no more than a factor of N, compared to the "one at a time" round complexity. In terms of random-party EPR distillation of $|W\rangle$, the lower bound of Theorem 1 becomes

$$\text{Number of classical broadcasts in } \mathcal{P} \geq \frac{1}{3(1-P_{tot})}.$$

This is also tight since each W-canonical sub-block requires only one round of broadcast communication to implement.

Remark. The lower bound of Theorem 1 matches in form the original F-L distillation probability [22] as well as the slightly optimized distillation probability by Li *et al.* [31].

5.2 Lower Bounds in Concurrence Distillation

5.2.1 Infinite Rounds is Optimal

We now turn to the task of random-party concurrence distillation. We first establish that every finite-round protocol is sub-optimal.

Theorem 2. Suppose $\mathcal{P}$ is an N-round total random-party concurrence distillation protocol on a tripartite W-state $|\mathbf{x}\rangle$. Then there exists a protocol $\mathcal{P}'$ having more than N rounds but distills a larger average concurrence than $\mathcal{P}$. The same statement holds for $\star$-random-party concurrence distillation.

Proof. Consider the last instance in the protocol for which some local measurement is performed on a tripartite entangled W-state. Call this state $|\mathbf{x}'\rangle$. Since this is the final time such a state appears in the protocol, the local measurement must be a hard

measurement that completely decouples the measuring party from the other two. Moreover, the protocol halts after this measurement since any further bipartite processing can never increase the concurrence. The greatest average post-measurement concurrence obtained from a hard measurement on $|\mathbf{x}'\rangle$ is called the concurrence of assistance (CoA) [29], and if (x'_A, x'_B, x'_C) are the coordinates of state $|\mathbf{x}'\rangle$, then

$$\text{CoA of } |\mathbf{x}'\rangle \leq 2\sqrt{x'_i x'_j}, \tag{37}$$

where $i, j \in \{A, B, C\}$ are the two non-measuring parties [29]. It is obvious that both $\zeta(\mathbf{x}')$ and $\zeta^\star(\mathbf{x}')$ are strictly larger than this value. Hence if instead of performing the hard measurement and terminating the protocol the parties were to continue with the protocols of Lemma 1 and 2, respectively, they would obtain a greater average concurrence. □

This result shows that infinite-round LOCC (i.e. protocols with unbounded round number) is needed to optimally distill concurrence from W-class states.

Remark. As noted above Lemma 2, for the task of $\star$-random-party EPR distillation, an optimal protocol can be found that consumes only three rounds of LOCC. Thus, we see that finite rounds of LOCC suffice to optimize the entanglement measure $E_2^{(\star-\mathrm{rnd})}$, whereas infinite rounds of LOCC are needed to optimize the measure $C^{(\star-\mathrm{rnd})}$, even though the tasks share the same goal of breaking tripartite entanglement into bipartite form.

Remark. The proof of Theorem 2 uses the fact that both ζ and $\zeta^\star$ are defined on the entire class of W-states, including those states (x_A, x_B, x_C) for which $x_0 \neq 0$. In contrast, the proof will not directly work for $E_2^{(\mathrm{rnd})}$ since the equality $E_2^{(\mathrm{rnd})}(\mathbf{x}) = \kappa(\mathbf{x})$ only holds for states with $x_0 = 0$, and a general LOCC protocol might attain states for which the latter condition does not hold. However, by using Proposition 1 and its generalization to arbitrary W-class states, one can assume without loss of generality that only diagonal Kraus operators are used on the protocol. Hence if the initial states $\mathbf{x}$ has $x_0 = 0$, then it will remain zero throughout. Consequently, we can repeat the same argument as Theorem 2 and likewise conclude that $E_2^{(\mathrm{rnd})}(\mathbf{x})$ is not achievable in finite round LOCC whenever $x_0 = 0$.

5.2.2 Lower Bounds for F-L Protocols

We now consider the problem of finite round concurrence distillation. Specifically, we restrict the protocol to be the ones in Section 3 and 4, but allow the ϵ parameters in the F-L measurements to vary depending on the round number. We find the optimal choices of these ϵ's given the total number of BLOCC rounds.

$\star$-Random-Party Concurrence Distillation. We first find the optimal protocol for the $\star$-random-party concurrence distillation task. We consider just the type of protocol specified in Section 4.

The first step is always a hard measurement on party n_2 if $x_{n_2} \neq x_{n_1}$, but can be omitted if they are equal. We thus exclude this step in our round counts for either case, assuming an additional round if $x_{n_2} \neq x_{n_1}$. The rest of the protocol consists of Fortescue-Lo measurements but with the ϵ_k now able to vary for each round $k = 1, \ldots, N-1$. In the last round, we perform a hard measurement on party n_2 to retrieve the concurrence in the bipartite state between party $\star$ and n_1. Then the concurrence distilled for round k is

$$C_k = \begin{cases} 4x_{n_2}\sqrt{\frac{x_\star}{x_{n_1}}}\prod_{j=1}^{k-1}(1-\epsilon_j)^{\frac{3}{2}}\epsilon_k\sqrt{1-\epsilon_k}, & k < N, \\ 2x_{n_2}\sqrt{\frac{x_\star}{x_{n_1}}}\prod_{j=1}^{N-1}(1-\epsilon_j)^{3/2}, & k = N. \end{cases} \tag{38}$$

Our goal is to maximize the total concurrence $\sum_{k=1}^{N} C_k$ by varying ϵ_k, $k = 1, \ldots, N$. By requiring the partial derivatives to be zero, the ϵ_k's are calculated to be

$$\epsilon_k = \begin{cases} 1/3 & \text{for } k = N-1, \\ \frac{1}{1+1/(2-3A_k-\frac{3}{2}B_k)} & \text{for } k = 1, \ldots, N-2, \end{cases} \tag{39}$$

where

$$A_k = \sum_{j=k+1}^{N-1}\left[\prod_{i=k+1}^{j-1}(1-\epsilon_i)^{3/2}\right]\epsilon_j\sqrt{1-\epsilon_j}, \tag{40}$$

$$B_k = \prod_{i=k+1}^{N-1}(1-\epsilon_i)^{3/2}. \tag{41}$$

Note that each ϵ_k only depends on ϵ_j's with $j > k$. So these parameters can be calculated one by one in a descending manner, starting from $\epsilon_{N-1} = 1/3$.

For a $|W\rangle$ state, the relation between the number of rounds and the optimal average concurrence is plotted in Fig. 6. Note that as the number of rounds increases, the concurrence distilled approaches the asymptotic bound $\zeta^\star$.

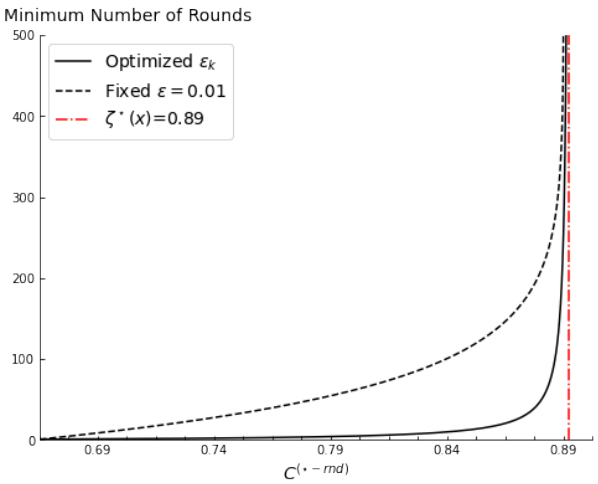


Figure 6: Minimum number of BLOCC rounds to obtain average bipartite concurrence $C^{(\star\text{-rnd})}$ from $|W\rangle$. Comparison is made between choosing a uniform ϵ and optimizing ϵ_k for each round.

Total Random-Party Concurrence Distillation. We next find the optimal finite-round parameters for total random party concurrence distillation protocol described in Section 3. Notice that in this case there are two separate steps (Step 2 and 3) that require infinitesimal measurements to achieve the asymptotic bound. Therefore, for a fixed finite number of rounds N, we distribute it so that we spend N_2 in step 2 and N_3 in step 3 with $N = N_2 + N_3$.

In step 2, the expected concurrence distilled for round k is

$$C_{\text{Step }2,k} = 4x_C\sqrt{\frac{x_A}{x_B}}\prod_{j=1}^{k-1}(1-\epsilon_j)^{3/2}\epsilon_k\sqrt{1-\epsilon_k}, \tag{42}$$

where $k = 1, \ldots, N_2$. As in the asymptotic protocol, we require that $\prod_{k=1}^{N_2}(1-\epsilon_k) = x_B/x_A = \gamma$. Given a fixed N_2, we wish to maximize $C_{\text{Step }2,N_2} := \sum_{k=1}^{N_2} C_{\text{Step }2,k}$ over ϵ_k's subject to the constraint above. Using the Lagrange multiplier method, the maximizing condition is

$$\nabla_{\lambda,\epsilon_1,\ldots,\epsilon_{N_2}}\left[\sum_{k=1}^{N_2} C_{\text{Step }2,k} + \lambda\left(\prod_{k=1}^{N_2}(1-\epsilon_k)^{3/2} - \gamma^{3/2}\right)\right] = 0. \tag{43}$$

Solving this we get

$$(1-\epsilon_k) = \begin{cases} \frac{1}{3(1-\lambda)}, & \text{for } k = N_2, \\ \frac{1}{3-2\sqrt{1-\epsilon_{k+1}}}, & \text{for } k = 1, 2, \ldots, N_2 - 1. \end{cases} \tag{44}$$

Therefore, every ϵ_k can be expressed in terms of λ, and the exact number can be calculated by plugging into the constraint that $\prod_{k=1}^{N_2}(1-\epsilon_k) = x_B/x_A = \gamma$.

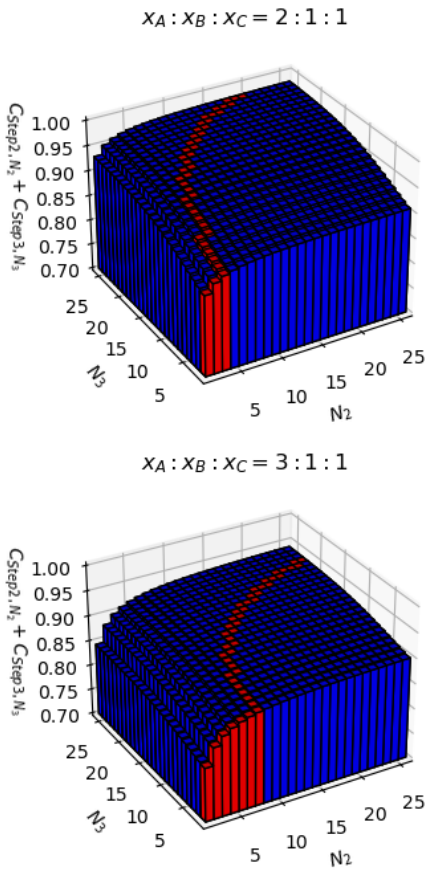


Figure 7: Maximum average bipartite concurrence that can be distilled using N_2 number of BLOCC rounds in Step 2 and N_3 in Step 3. Each red bar represents the best value in the diagonal line where $N_2 + N_3$ is constant. Notice that for different states there are different ways to optimally distribute the number of rounds into each step. For states with a large ratio between x_A and x_C, more rounds are needed in Step 2 to gradually lower the ratio.

In step 3, the expected concurrence distilled for round k is

$$C_{\text{Step }3,k} = \begin{cases} 6\frac{x_B x_C}{x_A}\prod_{j=1}^{k-1}(1-\epsilon_j)^2\epsilon_k(1-\epsilon_k), & k < N_2, \\ 2\frac{x_B x_C}{x_A}\prod_{j=1}^{k-1}(1-\epsilon_j)^2, & k = N_2, \end{cases} \tag{45}$$

with a hard measurement on party C in the end. (With an abuse of notation we use ϵ_k's for the parameters in both step 2 and step 3, but they are different variables.) With a fixed N_3, we wish to maximize $C_{\text{Step }3,N_3} := \sum_{k=1}^{N_3} C_{\text{Step }3,k}$ over ϵ_k's. By requiring the partial derivatives with respect to ϵ_k's to be zero, we get

$$\epsilon_k = \begin{cases} 1/4 & \text{for } k = N_3 - 1, \\ 1 - \frac{3}{2(3-3A_k-B_k)} & \text{for } k = 1, \ldots, N_3 - 2, \end{cases} \tag{46}$$

where

$$A_k = \sum_{j=k+1}^{N_3-1} \left[\prod_{i=k+1}^{j-1} (1-\epsilon_i)^2 \right] \epsilon_j \sqrt{1-\epsilon_j}, \quad (47)$$

$$B_k = \prod_{i=k+1}^{N_3-1} (1-\epsilon_i)^2. \quad (48)$$

Again, the ϵ_k's can be calculated from the last one to the first, and the total concurrence distilled can be calculated by plugging the values into (45).

The final goal is to determine the best way to distribute the total number of rounds N into N_2 and N_3, i.e.

$$\max_{N_2,N_3\in\mathbb{N} \text{ s.t. } N=N_2+N_3} C_{\text{Step 2},N_2} + C_{\text{Step 3},N_3}. \quad (49)$$

This can be done by brute-force search over the $N-1$ ways of distributing round numbers to Step 2 and 3. Results are shown in Fig. 7 and Fig. 8.

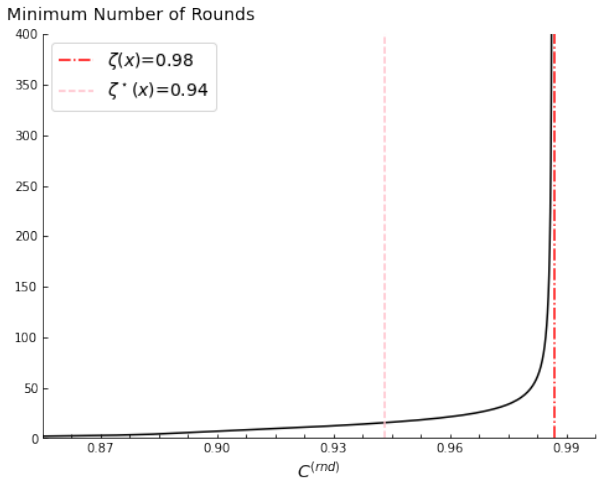


Figure 8: Minimum number of BLOCC rounds to obtain average bipartite concurrence $C^{(\text{rnd})}$ from a state with $x_A : x_B : x_C = 2:1:1$. Notice the gap between total random-party distillation bound $\zeta(\mathbf{x})$ and $\star$-random-party distillation bound $\zeta^\star(\mathbf{x})$.

6 A Family of Maps on $\overline{\text{LOCC}}$

The restricted Fortescue-Lo protocol described in Section 3 served as an example of a map on the boundary of LOCC but not in LOCC [14], thus proving that LOCC is not closed. Now we show that the operations in the step 2 of the above protocol indicate that there exist a family of maps $\{\mathfrak{J}_\gamma : 0<\gamma<1\}$ on the LOCC boundary but not in LOCC, ie. $\overline{\text{LOCC}} \ni \mathfrak{J}_\gamma \notin \text{LOCC}$ for $0<\gamma<1$.

We define $\mathfrak{J}_\gamma$ as follows. First fix a $0<\gamma<1$. Let Bob and Charlie both perform the measurement

$$M_0 = \sqrt{1-\epsilon}|0\rangle\langle 0| + |1\rangle\langle 1|, \quad M_1 = \sqrt{\epsilon}|0\rangle\langle 0|. \quad (50)$$

If the joint outcome is one of 01, 10, or 11, they stop. If the joint outcome is 00, then they repeat the same measurement. After a maximum number of n iterations they stop. The ϵ and n are deliberately chosen so that $(1-\epsilon)^n = \gamma$. We coarse-grain the instrument to obtain $\mathfrak{J}_{\gamma,n} = (\mathcal{E}^\gamma_{n00}, \mathcal{E}^\gamma_{n01}, \mathcal{E}^\gamma_{n10}, \mathcal{E}^\gamma_{n11})$ where $\mathcal{E}^\gamma_{nij}$ include all the cases when Bob and Charlie stop upon obtaining the outcome ij. Thus the four maps are respectively generated by the following sets of Kraus operators: $\{M_0^n \otimes M_1^n\}, \{M_0^j \otimes M_1 M_0^{j-1} : j = 1,2,\ldots,n\}, \{M_1 M_0^{j-1} \otimes M_0^j : j = 1,2,\ldots,n\}$, and $\{M_1 M_0^{j-1} \otimes M_1 M_0^{j-1} : j = 1,2,\ldots,n\}$.

The map $\mathfrak{J}_\gamma$ is obtained by taking $\lim_{n\to\infty} \mathfrak{J}_{\gamma,n}$. The Choi matrices $\{\Omega^\gamma_{ij}\}$ of $\mathfrak{J}_\gamma$ can be obtained from a similar calculation as in [14]. We have

$$\Omega^\gamma_{00} = \begin{pmatrix} \gamma & \sqrt{\gamma} \\ \sqrt{\gamma} & 1 \end{pmatrix}^{A'A} \otimes \begin{pmatrix} \gamma & \sqrt{\gamma} \\ \sqrt{\gamma} & 1 \end{pmatrix}^{B'B}, \quad (51)$$

$$\Omega^\gamma_{01} = \begin{pmatrix} \frac{1}{2}(1-\gamma^2) & \frac{2}{3}(1-\gamma^{3/2}) \\ \frac{2}{3}(1-\gamma^{3/2}) & 1-\gamma \end{pmatrix}^{A'A} \otimes |00\rangle\langle 00|^{B'B}, \quad (52)$$

$$\Omega^\gamma_{01} = |00\rangle\langle 00|^{A'A} \otimes \begin{pmatrix} \frac{1}{2}(1-\gamma^2) & \frac{2}{3}(1-\gamma^{3/2}) \\ \frac{2}{3}(1-\gamma^{3/2}) & 1-\gamma \end{pmatrix}^{B'B}, \quad (53)$$

$$\Omega^\gamma_{11} = 0. \quad (54)$$

Thus the corresponding instruments are

$$\mathcal{E}^\gamma_{ij} : \rho^{AB} \mapsto \text{Tr}_{A'B'} \left[\Omega^\gamma_{ij} \left(\mathbb{I}^{AB} \otimes (\rho^T)^{A'B'} \right) \right] \quad (55)$$

It is evident that this map is on the boundary of LOCC. We argue that this map indeed does not belong to LOCC. For any $0<\gamma<1$, we consider the W-class state represented by $(x, \gamma x, \gamma x)$. The discussion in the step 2 of the previous section showed that performing the map $\mathfrak{J}_\gamma$ will optimally distill expected concurrence, thus not decreasing the value of $\zeta(\mathbf{x})$. In the end the state transforms to $(\gamma^2 x, \gamma^2 x, \gamma^2 x)$. The same can never be achieved by any LOCC protocol, because we have shown that any nontrivial measurement performed by Alice or Bob will strictly decrease the expected distillable concurrence.

6.1 EPR distillation with PPT and SEP Instruments

So far we have discussed converting $|\vec{x}\rangle$ to an EPR state between two of the three parties using LOCC. A natural question is how this bound relates to the same probability when one only restricts the quantum instrument to being positive partial transpose (PPT) with respect to all parties. If we consider only the cases where two of the three parties end up with an EPR state or declare the outcome a failure (i.e. there is no outcome of distilling a W state), then we can coarse-grain the instrument to having a single round and four branches— one for each pair of parties and one for failure. We can then express the failure probability of this instrument on an arbitrary tripartite state ρ^{ABC} as a semidefinite program (SDP).

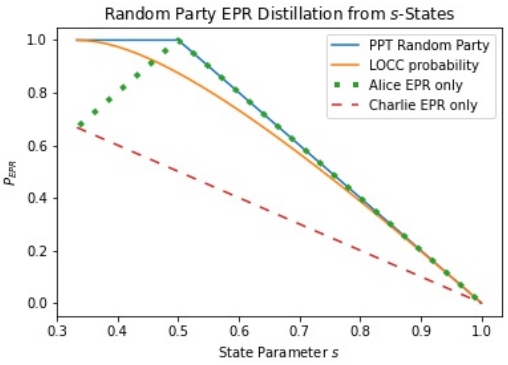


Figure 9: This depicts the gap between random party EPR distillation using a PPT (and SEP) instrument and an LOCC instrument. It also depicts the asymmetry in PPT distillation of EPR states shared between Alice and Bob or Charlie versus Charlie (resp. Bob) with Bob (resp. Charlie) or Alice.

For conciseness, let the input (resp. output) space be denoted in (resp. out) standing for the space ABC (resp. $A'B'C'$). For A, B, C, let $\Phi^{AB|C} := |\Phi^+\rangle\langle\Phi^+|^{AB} \otimes |0\rangle\langle 0|^C$. For any space B, let $\Gamma^B(X^{AB}) := (\mathrm{id}^A \otimes T)(X^{AB})$ where T is the transpose on the space B. Let $\mathcal{O} := \{0,1,2,3\}$ and $\mathcal{P} = \{AA', BB', CC'\}$. Then the SDP can be expressed as

$$\begin{aligned}
\text{min. } & \mathrm{Tr}[J_0(\rho^{\mathrm{T}} \otimes \mathbb{I}^{out})] \\
\text{s.t. } & \mathrm{Tr}_{out}\left[\sum_{i=0}^{3} J_i\right] = \mathbb{I}_{in} \\
& \mathrm{Tr}[\mathrm{Tr}_{in}[J_1(\rho^{\mathrm{T}} \otimes \mathbb{I}^{out})](\mathbb{I}^{out} - \Phi^{AB|C})] = 0 \\
& \mathrm{Tr}[\mathrm{Tr}_{in}[J_2(\rho^{\mathrm{T}} \otimes \mathbb{I}^{out})](\mathbb{I}^{out} - \Phi^{AC|B})] = 0 \\
& \mathrm{Tr}[\mathrm{Tr}_{in}[J_3(\rho^{\mathrm{T}} \otimes \mathbb{I}^{out})](\mathbb{I}^{out} - \Phi^{BC|A})] = 0 \\
& \Gamma^j(J_i) \geq 0 \quad \forall i \in \mathcal{O}, \forall j \in \mathcal{P} \\
& J_i \geq 0 \quad \forall i \in \mathcal{O}\,.
\end{aligned} \tag{56}$$

Recalling that for a linear map $\mathcal{N}_{A\to B}$ and linear operator X^A, $\mathcal{N}(X) := \mathrm{Tr}_A[J_{\mathcal{N}}(X^{\mathrm{T}} \otimes \mathbb{I}^B)]$ where $J_{\mathcal{N}}$ is the Choi operator, it's clear that J_0 represents the failure outcome, so the objective function is minimizing the failure probability. The next constraint guarantees coarse-graining the instrument results in a quantum channel as required. The following three constraints guarantee each success outcome is an instrument outcome that only maps the input state to an EPR state shared between two of the parties. The next line is the PPT constraints on the elements of the instruments, and the final line is simply positivity constraints on the instrument elements. Therefore we can see we have constructed an SDP for the relevant scenario.

We consider the set of s-states

$$|\psi_s\rangle := \sqrt{s}|100\rangle + \sqrt{\frac{1-s}{2}}(|010\rangle + |001\rangle)\,,$$

which are a specific case of $|\vec{x}\rangle$ states. Implementing the SDP (56) on the s-states numerically using CVX [19] with the solver MOSEK [2] along with QETLAB [26] in MATLAB [33], we find the gap between PPT and LOCC P_{EPR} (See Fig. 9), where the LOCC curve is given by $2(1-s) - (1-s)^2/4s$ [9]. Fig. 9 shows a gap between LOCC and PPT success probability, as well as the optimal strategy for $s \in [1/2, 1]$ establishes EPR pairs only between Alice and the other two parties. A natural further question would be if a separation between LOCC and separable operations also exists in this regime. In [9], it was shown that a SEP operations could achieve a success probability of one for $s \in [1/3, 1/2]$, which means it agrees with the PPT curve in Fig. 9 in this regime. The remaining question is where a separation between LOCC and SEP exists for $s \in (1/2, 1]$, as well as the possibility of a gap between SEP and PPT. We now show the above PPT curve is the same as the SEP curve. Note that Fig. 9, up to numerical error, shows that for PPT operations $P_{EPR}(s) = 2(1-s)$ for $s \in [1/2, 1]$. Consider $|\psi_s\rangle$ where $s \in [1/2, 1]$. Consider Alice applying a quantum instrument to her local portion where the first outcome uses a Kraus operator $M_0 = |0\rangle\langle 0| + \sqrt{(1-s)/s}\,|1\rangle\langle 1|$. One can verify that the state conditioned on that outcome is $|\psi_{1/2}\rangle$. Moreover, this calculation shows the probability of this outcome is $2(1-s)$. By [9], we know $|\psi_{1/2}\rangle$ can be distilled to an EPR with success probability one with SEP operations. Therefore, for $|\psi_s\rangle$ with $s \in [1/2, 1]$, $P^{\mathrm{Sep}}_{EPR} \geq 2(1-s) = P^{\mathrm{PPT}}_{EPR} \geq P^{\mathrm{Sep}}_{EPR}$. Thus, the PPT curve in Fig. 9 is the same as the SEP curve, and there exists a gap between random party EPR distillation from $|\psi_s\rangle$ using SEP and LOCC operations for all $s \in (1/3, 1)$.

7 Conclusion

In this paper, we have studied how more rounds of classical communication can increase the expected bipartite entanglement yield in random-party distillation protocols. Interestingly, the optimal number of rounds needed is highly sensitive to the type of entanglement measure considered. For certain distillation tasks, an unbounded number of rounds is needed if the bipartite concurrent measures is optimized, whereas a finite number is needed if the E_2 measure is optimized. Tight lower bounds were derived on the number of LOCC rounds required to attain a certain entanglement value.

Our work focused on W-class states since they possess a form highly amenable for analysis in the LOCC setting. In the future, it would be interesting to continue the study of round complexity for other types of multipartite entangled states. We conjecture that high round complexity is a general feature of random-party entanglement distillation. Intuitively, sequences of alternating weak local measurements enable a more "gentle" and less lossy decoupling of certain parties than when performing a hard

decoupling measurement in one just one iteration of local measurements. This leads to a greater entanglement yield when using more rounds of LOCC. We leave it as future work to further test and develop this intuition.

8 Acknowledgements

E.C. thanks Hoi-Kwong Lo and Laura Mančhinska for helpful discussions in the development of this paper. This work was supported by NSF Award No. 2016136.

Appendices

A Concurrence of Bipartite W-class States

In all concurrence distillation protocols that we discuss, the bipartite outcome of any LOCC branch must be a bipartite W-class state of the form

$$\sqrt{x_0}|00\rangle + \sqrt{x_1}|01\rangle + \sqrt{x_2}|10\rangle. \tag{57}$$

This is due to the fact that the W-class is closed under LOCC. The concurrence of such state is given by $2|\sigma_1\sigma_2|$, where σ_1 and σ_2 are the Schmidt coefficients of this state. The Schmidt coefficients are the singular values of the matrix

$$\begin{pmatrix} \sqrt{x_0} & \sqrt{x_1} \\ \sqrt{x_2} & 0 \end{pmatrix},$$

which are the square roots of the eigenvalues of the matrix

$$\begin{pmatrix} \sqrt{x_0} & \sqrt{x_1} \\ \sqrt{x_2} & 0 \end{pmatrix} \begin{pmatrix} \sqrt{x_0} & \sqrt{x_2} \\ \sqrt{x_1} & 0 \end{pmatrix} = \begin{pmatrix} x_0 + x_1 & \sqrt{x_0 x_2} \\ \sqrt{x_0 x_2} & x_2 \end{pmatrix}.$$

The eigenvalues are the roots of the polynomial $\lambda^2 - (x_0+x_1+x_2)\lambda + x_1x_2 = 0$. From Vieta's formulas, we have $\lambda_1\lambda_2 = x_1x_2$. So the concurrence is $2|\sigma_1\sigma_2| = 2\sqrt{x_1x_2}$.

B Monotonicity of ζ

Lemma 2 (Monotonicity part)**.** The function $\zeta(\mathbf{x})$ is non-increasing on average under LOCC. Furthermore, it is strictly decreasing on average when either $n_1(\mathbf{x})$ or $n_2(\mathbf{x})$ performs a non-trivial measurement.

Proof. The proof is analogous to the one given in [14] which proved that $\zeta^\star$ is LOCC monotone. We can decompose any local measurement into a sequence of binary measurements [1][36]. Thus each measurement can be specified by two Kraus operators $\{M_\lambda : \lambda = 1, 2\}$,

$$M_\lambda := U_\lambda \cdot \begin{pmatrix} \sqrt{a_\lambda} & b_\lambda \\ 0 & \sqrt{c_\lambda} \end{pmatrix}, \tag{58}$$

where a_λ, $c_\lambda \geq 0$, $b_\lambda \in \mathbb{C}$ and U_λ is unitary (which does not affect the representation $\vec{x}$.) By the completeness relation $\sum_{\lambda=0}^{1} M_\lambda^\dagger M_\lambda = I$, we have

$$\sum_{\lambda=0}^{1} \begin{pmatrix} a_\lambda & \sqrt{a_\lambda} b_\lambda \\ \sqrt{a_\lambda} b_\lambda^* & |b_\lambda|^2 + c_\lambda \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \tag{59}$$

thus $a_1 + a_2 = 1$ and $c_1 + c_2 \leq 1$ where equality holds if and only if $b_1 = b_2 = 0$.

To see how the state is changed by such measurement, consider when the first party Alice measures, then

$$\begin{aligned} &\sqrt{x_0}|000\rangle + \sqrt{x_A}|100\rangle + \sqrt{x_B}|010\rangle + \sqrt{x_C}|001\rangle \\ \mapsto &\frac{1}{\sqrt{p_\lambda}}[(\sqrt{a_\lambda x_0} + b_\lambda\sqrt{x_A})|000\rangle + \sqrt{c_\lambda x_A}|100\rangle \\ &+ \sqrt{a_\lambda x_B}|010\rangle + \sqrt{a_\lambda x_C}|001\rangle], \end{aligned} \tag{60}$$

where p_λ is the normalization constant that also indicates the probability of obtaining outcome λ. The post-measurement states can be represented by $\vec{x}_\lambda = (\frac{c_\lambda}{p_\lambda}x_A, \frac{a_\lambda}{p_\lambda}x_B, \frac{a_\lambda}{p_\lambda}x_C)$ if $c_\lambda x_A \geq a_\lambda x_B$, otherwise the ordering should be changed such that $\vec{x}_\lambda = (\frac{a_\lambda}{p_\lambda}x_B, \frac{c_\lambda}{p_\lambda}x_A, \frac{a_\lambda}{p_\lambda}x_C)$. More generally, when party $K \in \{A, B, C\}$ measures and outcome λ occurs, the components of the representing vector $\vec{x}_\lambda$ are $\{\frac{c_\lambda}{p_\lambda}x_K, \frac{a_\lambda}{p_\lambda}x_I, \frac{a_\lambda}{p_\lambda}x_J\}$ after sorting in decreasing order, where $I, J \in \{A, B, C\} \setminus \{K\}$.

Define the average change in ζ incurred by the measurement as

$$\overline{\Delta\zeta} := p_1\zeta(\vec{x}_1) + p_2\zeta(\vec{x}_2) - \zeta(\vec{x}). \tag{61}$$

We will show that $\overline{\Delta\zeta} \leq 0$ for all six possible cases.

Case (i) Alice measures when $x_A > x_B \geq x_C$.

Since any measurement can be decomposed as a sequence of weak measurements (ie. ones where $a_1 \approx c_1$ and $a_2 \approx c_2$) [36][37], without loss of generality we can assume the measurement is sufficiently weak so that $c_\lambda x_A \geq a_\lambda x_B$ is guaranteed. Thus the post-measurement states can be represented by $p_\lambda\vec{x}_\lambda = (c_\lambda x_A, a_\lambda x_B, a_\lambda x_C)$, and $\overline{\Delta\zeta}$ is calculated to be

$$\begin{aligned} \overline{\Delta\zeta} = &\left(2\sqrt{x_A x_B} + \frac{2}{3}x_C\sqrt{\frac{x_A}{x_B}}\right)(\sqrt{a_1c_1} + \sqrt{a_2c_2} - 1) \\ &+ \frac{1}{3}\frac{x_B x_C}{x_A}\left(\frac{a_1^2}{c_1} + \frac{a_2^2}{c_2} - 1\right). \end{aligned} \tag{62}$$

A necessary condition for $\overline{\Delta\zeta}$ to be maximized is that it is maximized over allowed values of c_2 when a_1, a_2 and c_1 are fixed. Taking the partial derivative, we have

$$\begin{aligned} \frac{\partial\overline{\Delta\zeta}}{\partial c_2} =&\sqrt{\frac{a_2}{c_2}}\left(\sqrt{x_A x_B} + \frac{x_C}{3}\sqrt{\frac{x_A}{x_B}} - \frac{x_C}{3}\frac{a_2 x_B}{c_2 x_A}\sqrt{\frac{a_2}{c_2}}\right) \\ \geq&\sqrt{\frac{a_2}{c_2}}\left[\sqrt{x_A x_B} + \frac{x_C}{3}\left(\sqrt{\frac{x_A}{x_B}} - \sqrt{\frac{a_2}{c_2}}\right)\right] \geq 0, \end{aligned} \tag{63}$$

where inequalities follow from the weakness constraint. Thus, for fixed a_1, a_2 and c_1, the value of $\overline{\Delta\zeta}$ is maximized when $c_2 = 1 - c_1$. Substitute this into (62) and evaluate around the point $(a_1, c_1) = (1/2, 1/2)$, we expand to the lowest order of $(a_1 - 1/2)$ and $(c_1 - 1/2)$ and obtain

$$\overline{\Delta\zeta} \approx -\left(\sqrt{x_A x_B} + \frac{x_C}{3}\sqrt{\frac{x_A}{x_B}} - \frac{4}{3}\frac{x_B x_C}{x_A}\right)(a_1 - c_1)^2. \tag{64}$$

Let $\gamma = \sqrt{\frac{x_B}{x_A}}$, then $\gamma \leq 1$, and

$$\overline{\Delta\zeta} \approx -\left[\sqrt{x_A x_B} + \frac{x_C}{3}\sqrt{\frac{x_A}{x_B}}\left(1 - 4\gamma^3\right)\right](a_1 - c_1)^2. \tag{65}$$

If $1 - 4\gamma^3 \geq 0$, then $\overline{\Delta\zeta} \leq 0$ holds.
If $1 - 4\gamma^3 < 0$, then from $x_C \leq x_B$ and $\gamma \leq 1$,

$$\overline{\Delta\zeta} \leq -\sqrt{x_A x_B} \cdot \frac{4}{3} \cdot (1 - \gamma^3)(a_1 - c_1)^2 \leq 0, \tag{66}$$

with equality obtained only by the trivial measurement where $a_1 = c_1$ and $a_2 = c_2$.
Case (ii) Bob measures when $x_A > x_B > x_C$.

WLOG, we assume that the measurement is weak enough so that $a_\lambda x_A \geq c_\lambda x_B \geq a_\lambda x_C$. Then the post-measurement states can be represented by $p_\lambda \vec{x}_\lambda = (a_\lambda x_A, c_\lambda x_B, a_\lambda x_C)$, and $\overline{\Delta\zeta}$ is calculated to be

$$\overline{\Delta\zeta} = \overline{\Delta\zeta^\star} + \frac{1}{3}\frac{x_B x_C}{x_A}(c_1 + c_2 - 1), \tag{67}$$

where $\overline{\Delta\zeta^\star} := p_1\zeta^\star(\vec{x}_1) + p_2\zeta^\star(\vec{x}_2) - \zeta^\star(\vec{x})$ with Alice being the ★ party. Since $\zeta^\star$ have been shown to be LOCC monotone [14], we have $\overline{\Delta\zeta^\star} \leq 0$. And since $c_1 + c_2 \leq 1$, $\overline{\Delta\zeta} \leq 0$ holds in this case with equality obtained only by the trivial measurement where $a_1 = c_1$ and $a_2 = c_2$.
Case (iii) Charlie measures when $x_A \geq x_B > x_C$.

WLOG, we assume that the measurement is weak enough so that $a_\lambda x_B \geq c_\lambda x_C$. Then the post-measurement states can be represented by $p_\lambda \vec{x}_\lambda = (a_\lambda x_A, a_\lambda x_B, c_\lambda x_C)$, and $\overline{\Delta\zeta}$ is calculated to be

$$\overline{\Delta\zeta} = \overline{\Delta\zeta^\star} + \frac{1}{3}\frac{x_B x_C}{x_A}(c_1 + c_2 - 1), \tag{68}$$

Again, $\overline{\Delta\zeta} \leq 0$ holds in this case. But since Charlie correspond to the party n_2 when calculating $\overline{\Delta\zeta^\star}$, non-trivial measurement performed by Charlie does not necessarily decrease $\zeta^\star$ [14], and equality holds whenever $c_1 + c_2 - 1 = 0$.
Case (iv) Bob measures when $x_A > x_B = x_C$.

WLOG, we assume that $a_1 \leq c_1$ (thus $a_2 \geq c_2$.) We also assume that the measurement is weak enough so that $a_1 x_A \geq c_1 x_B$. Then the post-measurement states can be represented by

$$p_1\vec{x}_1 = (a_1 x_A, c_1 x_B, a_1 x_B), \tag{69}$$
$$p_2\vec{x}_2 = (a_2 x_A, a_2 x_B, c_2 x_B), \tag{70}$$

and $\overline{\Delta\zeta}$ is calculated to be

$$\overline{\Delta\zeta} = \overline{\Delta\zeta^\star} + \frac{1}{3}\frac{x_B x_C}{x_A}(c_1 + c_2 - 1), \tag{71}$$

Again, $\overline{\Delta\zeta} \leq 0$ holds in this case with equality obtained only by the trivial measurement where $a_1 = c_1$ and $a_2 = c_2$.
Case (v) Alice measures when $x_A = x_B > x_C$.

WLOG, we assume that $a_1 \leq c_1$ (thus $a_2 \geq c_2$.) We also assume that the measurement is weak enough so that $c_2 x_A \geq a_2 x_B$. Then the post-measurement states can be represented by

$$p_1\vec{x}_1 = (c_1 x_A, a_1 x_A, a_1 x_C), \tag{72}$$
$$p_2\vec{x}_2 = (a_2 x_A, c_2 x_A, a_2 x_C), \tag{73}$$

and $\overline{\Delta\zeta}$ is calculated to be

$$\overline{\Delta\zeta} = \overline{\Delta\zeta^\star} + \frac{1}{3}\frac{x_C}{c_1}(a_1^2 - c_1^2), \tag{74}$$

Again, $\overline{\Delta\zeta} \leq 0$ holds in this case with equality obtained only by the trivial measurement where $a_1 = c_1$ and $a_2 = c_2$.
Case (vi) Alice measures when $x_A = x_B = x_C = x$.

WLOG, we assume that $a_1 \leq c_1$ (thus $a_2 \geq c_2$.) Then the post-measurement states can be represented by

$$p_1\vec{x}_1 = (c_1 x, a_1 x, a_1 x), \tag{75}$$
$$p_2\vec{x}_2 = (a_2 x, a_2 x, c_2 x), \tag{76}$$

and $\overline{\Delta\zeta}$ is calculated to be

$$\overline{\Delta\zeta} = \overline{\Delta\zeta^\star} + \frac{1}{3}\frac{x_C}{c_1}(a_1^2 - c_1^2), \tag{77}$$

Again, $\overline{\Delta\zeta} \leq 0$ holds in this case with equality obtained only by the trivial measurement where $a_1 = c_1$ and $a_2 = c_2$.

□

References

[1] Erika Andersson and Daniel K. L. Oi. Binary search trees for generalized measurements. *Phys. Rev. A*, 77:052104, May 2008. doi:10.1103/PhysRevA.77.052104.

[2] MOSEK ApS. *The MOSEK optimization toolbox for MATLAB manual. Version 9.0.*, 2019. URL: http://docs.mosek.com/9.0/toolbox/index.html.

[3] Charles H. Bennett, David P. DiVincenzo, Christopher A. Fuchs, Tal Mor, Eric Rains, Peter W. Shor, John A. Smolin, and William K. Wootters. Quantum nonlocality without entanglement. *Phys. Rev. A*, 59:1070–1091, Feb 1999. doi:10.1103/PhysRevA.59.1070.

[4] Charles H. Bennett, David P. DiVincenzo, John A. Smolin, and William K. Wootters. Mixed-state entanglement and quantum error correction. *Physical Review A*, 54(5):3824–3851, November 1996. doi:10.1103/physreva.54.3824.

[5] Andrew M. Childs, Debbie Leung, Laura Mančinska, and Maris Ozols. A framework for bounding nonlocality of state discrimination. *Communications in Mathematical Physics*, 323(3):1121–1153, September 2013. doi:10.1007/s00220-013-1784-0.

[6] Andrew M. Childs, Debbie Leung, Laura Mančinska, and Maris Ozols. Interpolatability distinguishes LOCC from separable von neumann measurements. *Journal of Mathematical Physics*, 54(11):112204, November 2013. doi:10.1063/1.4830335.

[7] Eric Chitambar. Local quantum transformations requiring infinite rounds of classical communication. *Physical Review Letters*, 107(19), November 2011. doi:10.1103/physrevlett.107.190502.

[8] Eric Chitambar, Wei Cui, and Hoi-Kwong Lo. Entanglement monotones forW-type states. *Physical Review A*, 85(6), June 2012. doi:10.1103/physreva.85.062316.

[9] Eric Chitambar, Wei Cui, and Hoi-Kwong Lo. Increasing entanglement monotones by separable operations. *Physical Review Letters*, 108(24), June 2012. doi:10.1103/physrevlett.108.240504.

[10] Eric Chitambar and Min-Hsiu Hsieh. Revisiting the optimal detection of quantum information. *Phys. Rev. A*, 88:020302, Aug 2013. doi:10.1103/PhysRevA.88.020302.

[11] Eric Chitambar and Min-Hsiu Hsieh. Asymptotic state discrimination and a strict hierarchy in distinguishability norms. *Journal of Mathematical Physics*, 55(11):112204, November 2014. doi:10.1063/1.4902027.

[12] Eric Chitambar and Min-Hsiu Hsieh. Round complexity in the local transformations of quantum and classical states. *Nature Communications*, 8(1), December 2017. doi:10.1038/s41467-017-01887-5.

[13] Eric Chitambar, Debbie Leung, Laura Mančinska, Maris Ozols, and Andreas Winter. Everything you always wanted to know about LOCC (but were afraid to ask). *Communications in Mathematical Physics*, 328(1):303–326, March 2014. doi:10.1007/s00220-014-1953-9.

[14] Eric Chitambar, Debbie Leung, Laura Mančinska, Maris Ozols, and Andreas Winter. Everything you always wanted to know about locc (but were afraid to ask). *Communications in Mathematical Physics*, 328(1):303–326, Mar 2014. doi:10.1007/s00220-014-1953-9.

[15] Scott M. Cohen. Structure of local quantum operations and classical communication: Finite versus infinite rounds. *Phys. Rev. A*, 91:042106, Apr 2015. doi:10.1103/PhysRevA.91.042106.

[16] Scott M. Cohen. General approach to quantum channel impossibility by local operations and classical communication. *Phys. Rev. Lett.*, 118:020501, Jan 2017. doi:10.1103/PhysRevLett.118.020501.

[17] Sarah Croke and Stephen M. Barnett. Difficulty of distinguishing product states locally. *Phys. Rev. A*, 95:012337, Jan 2017. doi:10.1103/PhysRevA.95.012337.

[18] Wei Cui, Eric Chitambar, and Hoi-Kwong Lo. Randomly distillingW-class states into general configurations of two-party entanglement. *Physical Review A*, 84(5), November 2011. doi:10.1103/physreva.84.052301.

[19] CVX Research, Inc. CVX: Matlab software for disciplined convex programming, version 2.0. http://cvxr.com/cvx, August 2012.

[20] Igor Devetak and Andreas Winter. Distillation of secret key and entanglement from quantum states. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 461(2053):207–235, January 2005. doi:10.1098/rspa.2004.1372.

[21] W. Dür, G. Vidal, and J. I. Cirac. Three qubits can be entangled in two inequivalent ways. *Phys. Rev. A*, 62:062314, Nov 2000. doi:10.1103/PhysRevA.62.062314.

[22] Ben Fortescue and Hoi-Kwong Lo. Random bipartite entanglement fromWandW-like states. *Physical Review Letters*, 98(26), June 2007. doi:10.1103/physrevlett.98.260501.

[23] Ben Fortescue and Hoi-Kwong Lo. Random-party entanglement distillation in multiparty states. *Physical Review A*, 78(1), July 2008. doi:10.1103/physreva.78.012348.

[24] Alvin Gonzales and Eric Chitambar. Bounds on instantaneous nonlocal quantum computation. *IEEE Transactions on Information Theory*, 66(5):2951–2963, 2020. doi:10.1109/TIT.2019.2950190.

[25] Ryszard Horodecki, Paweł Horodecki, Michał Horodecki, and Karol Horodecki. Quantum entanglement. *Rev. Mod. Phys.*, 81:865–942, Jun 2009. doi:10.1103/RevModPhys.81.865.

[26] Nathaniel Johnston. QETLAB: A MATLAB toolbox for quantum entanglement, version 0.9.

http://qetlab.com, Jan 2016. doi:10.5281/zenodo.44637.

[27] S. Kıntaş and S. Turgut. Transformations of w-type entangled states. *Journal of Mathematical Physics*, 51(9):092202, September 2010. doi:10.1063/1.3481573.

[28] M. Kleinmann, H. Kampermann, and D. Bruß. Asymptotically perfect discrimination in the local-operation-and-classical-communication paradigm. *Phys. Rev. A*, 84:042326, Oct 2011. doi:10.1103/PhysRevA.84.042326.

[29] T. Laustsen, F. Verstraete, and Steven J. van Enk. Local vs. joint measurements for the entanglement of assistance. *Quantum Inf. Comput.*, 3:64–83, 2003.

[30] Debbie Leung, Andreas Winter, and Nengkun Yu. LOCC protocols with bounded width per round optimize convex functions. *Reviews in Mathematical Physics*, 33(05):2150013, January 2021. doi:10.1142/s0129055x21500136.

[31] Zheng-Da Li, Xiao Yuan, Xu-Fei Yin, Li-Zheng Liu, Rui Zhang, Yue-Yang Fei, Li Li, Nai-Le Liu, Xiongfeng Ma, He Lu, Yu-Ao Chen, and Jian-Wei Pan. Experimental random-party entanglement distillation via weak measurement. *Phys. Rev. Research*, 2:023047, Apr 2020. doi:10.1103/PhysRevResearch.2.023047.

[32] Hoi-Kwong Lo and Sandu Popescu. Concentrating entanglement by local actions: Beyond mean values. *Phys. Rev. A*, 63:022301, Jan 2001. doi:10.1103/PhysRevA.63.022301.

[33] MATLAB. *version R2021a*. The MathWorks Inc., Natick, Massachusetts, 2021.

[34] Michael Nathanson. Three maximally entangled states can require two-way local operations and classical communication for local discrimination. *Phys. Rev. A*, 88:062316, Dec 2013. doi:10.1103/PhysRevA.88.062316.

[35] M. A. Nielsen. Conditions for a class of entanglement transformations. *Phys. Rev. Lett.*, 83:436–439, Jul 1999. doi:10.1103/PhysRevLett.83.436.

[36] Ognyan Oreshkov and Todd A. Brun. Weak measurements are universal. *Physical Review Letters*, 95(11), Sep 2005. doi:10.1103/physrevlett.95.110409.

[37] Ognyan Oreshkov and Todd A. Brun. Infinitesimal local operations and differential conditions for entanglement monotones. *Phys. Rev. A*, 73:042314, Apr 2006. doi:10.1103/PhysRevA.73.042314.

[38] Masaki Owari and Masahito Hayashi. Two-way classical communication remarkably improves local distinguishability. *New Journal of Physics*, 10(1):013006, January 2008. doi:10.1088/1367-2630/10/1/013006.

[39] Asher Peres and William K. Wootters. Optimal detection of quantum information. *Phys. Rev. Lett.*, 66:1119–1122, Mar 1991. doi:10.1103/PhysRevLett.66.1119.

[40] Martin B. Plenio and Shashank Virmani. An introduction to entanglement measures. *Quantum Info. Comput.*, 7(1):1–51, jan 2007. doi:10.48550/arXiv.quant-ph/0504163.

[41] E. M. Rains. Bound on distillable entanglement. *Phys. Rev. A*, 60:179–184, Jul 1999. doi:10.1103/PhysRevA.60.179.

[42] Filip Rozpiędek, Thomas Schiet, Le Phuc Thinh, David Elkouss, Andrew C. Doherty, and Stephanie Wehner. Optimizing practical entanglement distillation. *Phys. Rev. A*, 97:062333, Jun 2018. doi:10.1103/PhysRevA.97.062333.

[43] Guojing Tian, Xia Wu, Ya Cao, Fei Gao, and Qiaoyan Wen. General existence of locally distinguishable maximally entangled states only with two-way classical communication. *Scientific Reports*, 6(1), July 2016. doi:10.1038/srep30181.

[44] Guifré Vidal. Entanglement of pure states for a single copy. *Phys. Rev. Lett.*, 83:1046–1049, Aug 1999. doi:10.1103/PhysRevLett.83.1046.

[45] Eyuri Wakakuwa, Akihito Soeda, and Mio Murao. A coding theorem for bipartite unitaries in distributed quantum computation. *IEEE Transactions on Information Theory*, 63(8):5372–5403, August 2017. doi:10.1109/tit.2017.2709754.

[46] Eyuri Wakakuwa, Akihito Soeda, and Mio Murao. Complexity of causal order structure in distributed quantum information processing: More rounds of classical communication reduce entanglement cost. *Phys. Rev. Lett.*, 122:190502, May 2019. doi:10.1103/PhysRevLett.122.190502.

[47] William K. Wootters. Entanglement of formation of an arbitrary state of two qubits. *Phys. Rev. Lett.*, 80:2245–2248, Mar 1998. doi:10.1103/PhysRevLett.80.2245.

[48] Yu Xin and Runyao Duan. Local distinguishability of orthogonal $2 \otimes 3$ pure states. *Phys. Rev. A*, 77:012315, Jan 2008. doi:10.1103/PhysRevA.77.012315.

[49] Ying-Hui Yang, Jiang-Tao Yuan, Cai-Hong Wang, and Shi-Jiao Geng. Locally distinguishable maximally entangled states by two-way LOCC. *Quantum Information Processing*, 20(1), January 2021. doi:10.1007/s11128-020-02957-2.

[50] Jiang-Tao Yuan, Ying-Hui Yang, and Cai-Hong Wang. Constructions of locally distinguishable sets of maximally entangled states which require two-way LOCC. *Journal of Physics A: Mathematical and Theoretical*, 53(50):505304, November 2020. doi:10.1088/1751-8121/abc43b.