

Tripartite Haar Random State Has No Bipartite Entanglement

Zhi Li,^{1,2,*} Takato Mori^{3,2,4,†} and Beni Yoshida^{2,§}

¹*National Research Council Canada, Waterloo, Ontario N2L 3W8, Canada*

²*Perimeter Institute for Theoretical Physics, Waterloo, Ontario N2L 3W8, Canada*

³*Department of Physics, Rikkyo University, 3-34-1 Nishi-Ikebukuro, Toshima-ku, Tokyo 171-8501, Japan*

⁴*Yukawa Institute for Theoretical Physics, Kyoto University, Kitashirakawa Oiwakecho, Sakyo-ku, Kyoto 606-8502, Japan*



(Received 11 September 2025; revised 20 May 2026; accepted 28 July 2026; published 13 October 2026)

We show that no Einstein-Podolsky-Rosen (EPR)-like bipartite entanglement can be distilled from a tripartite Haar random state $|\Psi\rangle_{ABC}$ by local unitaries or local operations when each subsystem A , B , or C has fewer than half of the total qubits. Specifically, we derive an upper bound on the probability of sampling a state with EPR-like entanglement at a given EPR fidelity tolerance, showing a doubly exponential suppression in the number of qubits. Our proof relies on a simple volume argument supplemented by an ϵ -net argument and concentration of measure. Viewing $|\Psi\rangle_{ABC}$ as a bipartite quantum error-correcting code $C \rightarrow AB$, this implies that neither output subsystem A nor B supports any nontrivial logical operator. We also establish general constraints on the structure of tripartite entanglement in Haar random states, showing that W- or Greenberger-Horne-Zeilinger-like entanglement cannot be distilled and that nontrivial global symmetries are absent. Finally, we discuss a physical interpretation in the AdS/CFT correspondence, indicating that a connected entanglement wedge does not necessarily imply bipartite entanglement, contrary to a previous belief.

DOI: [10.1103/d7b9-ctw2](https://doi.org/10.1103/d7b9-ctw2)

Subject Areas: Quantum Physics,
Quantum Information, String Theory

I. INTRODUCTION

Quantum entanglement lies at the heart of many fundamental questions in quantum physics. However, the study of strongly entangled quantum systems poses significant challenges, as analytical and numerical approaches are often limited. Haar random states provide useful insights into the physical properties of certain many-body quantum systems. These states, characterized by their entanglement properties averaged over random ensembles, provide a powerful tool for understanding complex quantum systems with some degree of analytical tractability. In condensed matter physics, Haar random states (and unitaries) have proven instrumental in understanding dynamical properties, as in the eigenstate thermalization hypothesis [1], scrambling dynamics [2], and random matrix theory [3]. In

the AdS/CFT correspondence, Haar random states and unitaries serve as minimal toy models of a quantum black hole [4–6], and Haar random tensor networks serve as toy models obeying the Ryu-Takayanagi formula at the AdS scale at the leading order [7,8]. Also, the properties of Haar random states have played a central role in quantum information theory, as many protocols and fundamental questions rely on these states or on approximations that mimic their behavior [9]. These examples represent only a fraction of the widespread applications of Haar randomness, which has become an essential concept across diverse areas of modern physics.

Bipartite entanglement of a Haar random state $|\Psi\rangle_{AB}$, when the total system is bipartitioned into two complementary subsystems A and B , is well understood [4,10,11]. In particular, $|\Psi\rangle_{AB}$ contains approximately $\min(n_A, n_B)$ copies of approximate Einstein-Podolsky-Rosen (EPR) pairs shared between A and B up to some local unitary (LU) transformations $U_A \otimes U_B$. However, the nature of *tripartite entanglement* in a Haar random state $|\Psi\rangle_{ABC}$, where the system is divided into three complementary subsystems A , B , and C , remains much less understood. While various entanglement measures and properties of tripartite Haar random states have been studied extensively, the most fundamental question remains open: whether two subsystems A and B in $|\Psi\rangle_{ABC}$ exhibit bipartite EPR-like entanglement or not (Fig. 1).

*Present address: IBM Quantum.

†Contact author: zli@ibm.com

‡Contact author: takato.mori@yukawa.kyoto-u.ac.jp

§Contact author: byoshida@perimeterinstitute.ca

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/). Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

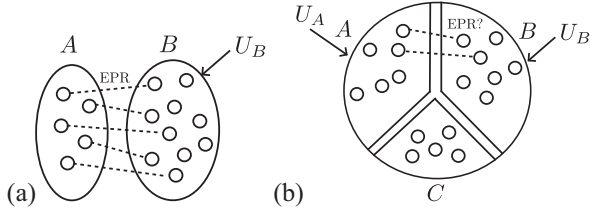


FIG. 1. (a) Bipartite Haar random states with $|A| < |B|$. EPR pairs can be distilled by applying a local unitary U_B . (b) Tripartite Haar random states. Can EPR pairs be distilled by local unitary rotations or local operations?

In this paper, we prove that no EPR-like entanglement can be distilled between two subsystems by LU transformations or local operations (LO) when each subsystem A, B, C has fewer than half of the total qubits. Specifically, we derive an upper bound with doubly exponential suppression (in terms of n) on the probability of sampling a quantum state with bipartite entanglement. Hence, we show that quantum entanglement in a tripartite Haar random state $|\Psi\rangle_{ABC}$ is nonbipartite, despite the fact that the reduced mixed state ρ_{AB} possesses a large number of quantum (nonclassical) correlations [e.g., the mutual information $I(A:B) \sim O(n)$ and the logarithmic negativity $E_N(A:B) \sim O(n)$]. We also discuss an application of our results in the context of quantum error-correcting codes. Viewing $|\Psi\rangle_{ABC}$ as a random encoding isometry $C \rightarrow AB$ with input C and output AB , our results imply that each output subsystem A or B supports no logical operator of the code if $n_C < n_A + n_B$ and $|n_A - n_B| < n_C$.

We also discuss the implications of our results in the context of the AdS/CFT correspondence. Namely, our results on LU and LO distillability suggest that a connected entanglement wedge does not necessarily imply the presence of EPR-like entanglement, contrary to previous beliefs. Furthermore, our results on logical operators lead to a surprising prediction concerning entanglement wedge reconstruction: They suggest the possible existence of extensive bulk regions whose degrees of freedom cannot be reconstructed on A or $B = A^c$ when the boundary is bipartitioned into A and B .

A. Entanglement in Haar random states

Consider an n -qubit Haar random state $|\Psi_{AB}\rangle$ where qubits are bipartitioned into two complementary subsystems A and B with n_A and $n_B = n - n_A$ qubits, respectively. For $n_A < n_B$, we have

$$\mathbb{E} \left\| \rho_A - \frac{1}{2^{n_A}} I_A \right\|_1 \lesssim 2^{(n_A - n_B)/2}, \quad (1)$$

where $\mathbb{E}$ represents the Haar average [12]. This result is often referred to as Page's theorem [4, 10, 11], suggesting that A is nearly maximally entangled with a 2^{n_A} -

dimensional subspace in B . Namely, there exists a local unitary $I_A \otimes U_B$ acting exclusively on B such that

$$(I_A \otimes U_B) |\Psi\rangle_{AB} \approx |\text{EPR}\rangle^{\otimes n_A} \otimes |\text{something}\rangle, \quad (2)$$

$$|\text{EPR}\rangle = \frac{1}{\sqrt{2}} (|00\rangle + |11\rangle),$$

where $A' \subseteq B$ and $|A| = |A'|$ with $|R|$ representing the number of qubits in a subsystem R . Hence, n_A approximate EPR pairs can be distilled between A and B by applying some local unitary transformations without using measurements or classical communications. Such unitary transformations can be explicitly constructed by unitarily approximating the Petz recovery map [13]. It is worth noting that EPR pairs can be distilled from a single copy of $|\Psi_{AB}\rangle$ without considering the asymptotic (many-copy) scenario.

Next, consider a tripartite n -qubit Haar random state $|\Psi_{ABC}\rangle$ on A, B , and C . We focus on regimes where each subsystem occupies less than half of the system and, thus, satisfies

$$S_R \approx n_R \quad \left(0 < n_R < \frac{n}{2} \right) \quad (3)$$

for $R = A, B, C$. In particular, we are interested in the asymptotic limit of large n . Two subsystems, say, A and B , have a large number of correlations as seen in the mutual information

$$S_{AB} \approx n_C, \quad I(A:B) \equiv S_A + S_B - S_{AB} \approx n_A + n_B - n_C \sim O(n). \quad (4)$$

While the mutual information does not distinguish classical and quantum correlations, in general, it can be verified that these are nonclassical by computing the logarithmic negativity [14, 15]:

$$E_N(A:B) \approx \frac{1}{2} I(A:B), \quad E_N(A:B) \equiv \log_2 \left(\sum_j |\lambda_j| \right), \quad (5)$$

where λ_j are eigenvalues of the partial-transposed density matrix $\rho_{AB}^{T_A}$.

A naturally arising question concerns the nature of quantum entanglement in ρ_{AB} . Namely, we are interested in whether EPR pairs can be distilled in a single copy of ρ_{AB} by applying some local unitary transformation $U_A \otimes U_B$ or some local operation $\Phi_A \otimes \Phi_B$.

When $|\Psi_{ABC}\rangle$ is randomly sampled from (qubit) stabilizer states, the mixed state ρ_{AB} contains approximately $\frac{1}{2} I(A:B)$ copies of unitarily (Clifford) rotated EPR pairs. This is essentially due to the fact that tripartite qubit

stabilizer states have only two types of entanglement: bipartite entanglement (i.e., Clifford rotated EPR pairs) or Greenberger-Horne-Zeilinger (GHZ)-like entanglement [16,17], with the latter being rare in random stabilizer states [18].

Do tripartite Haar random states also consist mostly of bipartite entanglement? There has been a significant body of previous works addressing this question in a variety of contexts, but no definite or quantitative answer has been provided. In particular, in the AdS/CFT correspondence, there have been extensive studies based on the bit thread formalism (see Refs. [19–23], for instance) which, implicitly or explicitly, assume that entanglement in ρ_{AB} is mostly bipartite with approximately $\frac{1}{2}I(A:B)$ EPR pairs. However, another line of research [24–27] has presented evidence for the presence of genuinely tripartite entanglement in a Haar random state.

B. Main results: Tripartite Haar random state

Consider a Haar random pure state $|\Psi_{ABC}\rangle$ supported on a d -dimensional Hilbert space with n qubits ($d = 2^n$). In this paper, we prove that no EPR pairs can be distilled from $\rho_{AB} = \text{Tr}_C(|\Psi_{ABC}\rangle\langle\Psi_{ABC}|)$ via local unitary transformations or local operations when $n_R < (n/2)$ for $R = A, B, C$ in the large- n limit.

1. Local unitaries

For a non-negative constant $0 < h \leq 1$, the (one-shot) LU-distillable entanglement is defined as

$$\text{ED}_h^{[\text{LU}]}(A:B) \equiv \sup_{m \in \mathbb{N}} \sup_{\Lambda \in \text{LU}} \left\{ m \left| \text{Tr}(\Lambda(\rho_{AB}) \Pi_{R_A R_B}^{[\text{EPR}]}) \right| \geq h^2 \right\}, \quad (6)$$

where $\Lambda = U_A \otimes U_B \in \text{LU}$ represents a local unitary acting on $A \otimes B$ and $\Pi_{R_A R_B}^{[\text{EPR}]}$ is a projection operator onto m EPR pairs supported on R_A and R_B . Here, $R_A \subseteq A$ and $R_B \subseteq B$, and $|R_A| = |R_B| = m$ with $|R_A|$ and $|R_B|$ denoting the number of qubits in subsystems R_A and R_B , respectively. The parameter h controls the fidelity of EPR pairs, where $h = 1$ corresponds to perfect EPR pairs while $h \approx 0$ corresponds to low-fidelity EPR pairs.

Theorem 1. If $\delta \stackrel{\text{def}}{=} h^2 - 2^{-2m} > 0$, then, for an arbitrary constant $0 < c < 2$, we have

$$\begin{aligned} & \log \mathbb{P} \left(\text{ED}_h^{[\text{LU}]}(A:B) \geq m \right) \\ & \leq -c\delta^2 d + O \left((d_A^2 + d_B^2) \log \frac{1}{\delta} \right). \end{aligned} \quad (7)$$

Note that the assumption $h^2 > 2^{-2m}$ is necessary. In fact, even without applying any unitaries, the region $R_A R_B$

already contains m EPR pairs with fidelity approximately 2^{-2m} , since $R_A R_B$ is nearly maximally mixed.

Theorem 1 provides a meaningful bound when the second term of Eq. (7) is subleading, ensuring that the right-hand side is negative. The bound then implies that EPR pairs cannot be LU distilled from ρ_{AB} with a fidelity better than that from a maximally mixed state. In other words, any attempt to enhance the EPR fidelity by applying LU transformations $U_A \otimes U_B$ will be useless.

Specifically, if we require δ to be a constant, then it suffices to assume $n_A, n_B = (n/2) - \omega(1)$ [here, $\omega(1)$ means superconstant: A function $f(n)$ is $\omega(1)$ if and only if $\lim_{n \rightarrow +\infty} f(n) = +\infty$], which ensures that $d_A^2, d_B^2 = o(d)$. Namely,

- (i) if $n_A, n_B = (n/2) - \omega(1)$, then $\mathbb{P}(\text{ED}_h^{[\text{LU}]}(A:B) \geq m) \leq \exp(-\Theta(d))$ whenever $h^2 - 2^{-2m} = \Theta(1)$.

Noting that $d = 2^n$, we see that the probability of A and B containing EPR-like entanglement is *doubly* exponentially small in the number of qubits.

More generally, an even lower EPR fidelity, allowing δ to vanish, is also permissible. In fact, it suffices to require $\delta > c'(d_A + d_B) \sqrt{[(\log d)/d]}$ for a sufficiently large constant c' , which ensures that the second term remains subleading and the first term diverges to negative infinity. As long as $n_A, n_B = [(n - \log_2 n)/2] - \omega(1)$, we have $(d_A + d_B) \sqrt{[(\log d)/d]} = o(1)$, ensuring our choice of δ can be satisfied. Therefore,

- (i) if $n_A, n_B = [(n - \log_2 n)/2] - \omega(1)$, then $\mathbb{P}(\text{ED}_h^{[\text{LU}]}(A:B) \geq m) = o(1)$ whenever $h^2 - 2^{-2m} > c' 2^{\max(n_A, n_B) - [(n - \log_2 n)/2]}$ for a sufficiently large constant c' .

2. Local operations

Analogous to Eq. (6), the (one-shot) LO-distillable entanglement with fidelity h^2 is defined as

$$\text{ED}_h^{[\text{LO}]}(A:B) \equiv \sup_{m \in \mathbb{N}} \sup_{\Lambda \in \text{LO}} \left\{ m \left| \text{Tr}(\Lambda(\rho_{AB}) \Pi_{R_A R_B}^{[\text{EPR}]}) \right| \geq h^2 \right\}, \quad (8)$$

where $\Lambda = \Phi_A \otimes \Phi_B$ represents a local operation (channel) acting on $A \otimes B$ and $\Pi_{R_A R_B}^{[\text{EPR}]}$ is a projection operator onto m EPR pairs supported on R_A, R_B , where $|R_A| = |R_B| = m$.

Theorem 2. If $\delta \stackrel{\text{def}}{=} h^2 - 2^{-m} > 0$, then, for an arbitrary constant $0 < c < 1$, we have

$$\begin{aligned} & \log \mathbb{P} \left(\text{ED}_h^{[\text{LO}]}(A:B) \geq m \right) \\ & \leq -c\delta^2 d + O \left(2^{2m} (d_A^2 + d_B^2) \log \frac{1}{\delta} \right). \end{aligned} \quad (9)$$

The threshold value 2^{-m} for h^2 is also optimal. In fact, we can consider a simple quantum channel on both sides:

attaching m ancillas in $|0\rangle^{\otimes m}$ and doing nothing. Making Bell measurement on the ancilla, the probability of getting m EPR pairs is already 2^{-m} . Our bound essentially suggests that any attempt to enhance the EPR fidelity cannot outperform this simple quantum channel.

This result is similar to Theorem 1, with the threshold value 2^{-2m} replaced by 2^{-m} and n_A (and n_B) replaced by $n_A + m$ (and $n_B + m$) in the second term. Following the discussions above, we conclude that

- (i) if $n_A, n_B = (n/2) - m - \omega(1)$, then $\mathbb{P}(\text{ED}_h^{\text{LU}}(A:B) \geq m) \leq \exp(-\Theta(d))$ whenever $h^2 - 2^{-m} = \Theta(1)$;
- (ii) if $n_A, n_B = [(n - \log_2 n)/2] - m - \omega(1)$, then $\mathbb{P}(\text{ED}_h^{\text{LU}}(A:B) \geq m) = o(1)$ whenever $h^2 - 2^{-m} > c' 2^{\max(n_A, n_B) - [(n - \log_2 n)/2] + m}$ for a sufficiently large constant c' .

3. Logical operators

Since the subsystem C is nearly maximally entangled with AB (assuming $n_C < n_A + n_B$), one can view a Haar random state $|\Psi_{ABC}\rangle$ as an approximate isometry $V: C \rightarrow AB$ that encodes $k = n_C$ logical qubits

$$V : \begin{array}{c} \begin{array}{cc} A & B \end{array} \\ \begin{array}{c} \text{output} \\ \text{input} \end{array} \end{array} \quad (10)$$

via the Choi isomorphism, where C (AB) corresponds to the input (output) Hilbert space.

We are interested in whether the encoded quantum information is recoverable from a single subsystem A or not while the complementary subsystem B is traced out (e.g., under erasure errors). This question can be addressed by asking whether a *logical operator* can be supported on A or not. Loosely speaking, $\bar{U}$ is said to be a logical unitary of U when $\bar{U}$ implements an action of U in the encoded code-word subspace. If a logical operator $\bar{U}_A$ can be supported on A , then a piece of information about U_C with respect to the initial state can be deduced from ρ_A .

The relation to the LU-distillation problem becomes evident by considering a pair of anticommuting Pauli logical operators. Namely, if Pauli logical operators $\bar{X}$ and $\bar{Z}$ could be supported on A , it would imply that an EPR pair could be LU distilled between A and C . Hence, one can deduce that logical Pauli operators $\bar{X}$ and $\bar{Z}$ cannot be supported inside A . This observation enables us to establish the following no-go result on random encoding. (A rigorous and quantitative bound will be presented in Theorem 4).

Theorem 3 (informal). Consider a random encoding $C \rightarrow AB$. If $n_C < n_A + n_B$ and $n_A < n_B + n_C$, then A contains no quantum information about C . Namely, A does not support any nontrivial logical unitary operator [28].

One interesting corollary of this result is that the so-called cleaning lemma does not necessarily hold for nonstabilizer codes. To recap, the cleaning lemma for a stabilizer code asserts that, if a subsystem A supports no nontrivial logical operators, then the complementary subsystem $B = A^c$ supports all the logical operators of the code [29]. This fundamental result is central in establishing the fault tolerance of topological stabilizer codes (those with geometrically local generators), as it ensures that logical operators can be supported on regions that avoid damaged qubits. While the original formulation is restricted to stabilizer codes, analogous properties (e.g., deformability of string logical operators) are known to hold in various models of topological phases beyond the stabilizer formalism. Despite these examples, our result suggests that the cleaning lemma, in its original formulation given by Ref. [29], does not extend to general nonstabilizer quantum error-correcting codes.

4. Generalizations

We also generalize our main theorems in three different directions in Sec. V.

- (i) *Tripartite entanglement*—Essentially no nontrivial tripartite pure state of smaller local dimension, such as GHZ, or W-like states, can be LU distilled from a Haar random tripartite state.
- (ii) *Chirality*—Tripartite Haar random states are intrinsically chiral. Namely, there is no tripartite LU which converts $|\psi\rangle_{ABC}$ into $|\psi^*\rangle_{ABC}$.
- (iii) *Global symmetry*—Tripartite Haar random states do not possess any (approximate) global symmetry. Namely, there is no tripartite LU satisfying $(U_A \otimes U_B \otimes U_C)|\psi\rangle_{ABC} \approx |\psi\rangle_{ABC}$.

C. Miscellaneous comments

Let us mention another relevant question concerning local operation and classical communication (LOCC)-distillable entanglement. Recall the *hashing lower bound* for LOCC-distillable entanglement in the asymptotic scenario [30,31]:

$$\begin{aligned} \text{hash}(A:B) &\leq E_D(A:B), \\ \text{hash}(A:B) &\equiv \max(S_A - S_{AB}, S_B - S_{AB}, 0). \end{aligned} \quad (11)$$

In Ref. [32], we considered one-shot one-way LOCC-distillable entanglement in ρ_{AB} and showed (see Ref. [33] also)

$$\text{ED}_D^{\text{one-shot one-way}}(A:B) \approx \text{hash}(A:B). \quad (12)$$

Hence, the hashing lower bound is tight for a tripartite Haar random state under the one-shot one-way scenario.

D. Physical and proof intuition

In this paper, we provide separate proofs for Theorems 1 and 2. The first proof follows an elementary approach with

volume counting; the second proof requires a few additional prerequisites, but the core idea remains similar. In fact, the proof of Theorem 2 can be simplified to establish Theorem 1. Since the proofs are somewhat technically involved, it will be helpful to first provide some physical intuition and highlight the key technical ingredients and challenges.

1. Counting argument

We begin with the intuition behind the first proof (Sec. II). Consider a tripartite state $|\psi_{ABC}\rangle$ where each subsystem is smaller than half the total size, $n_A, n_B, n_C < (n/2)$. Suppose that, by acting only with local unitaries on A and B , we can distill m EPR pairs between them. After removing these distilled EPR pairs, the remaining state lives on $n' = n - 2m$ qubits and can be arbitrary:

$$(U_A \otimes U_B)|\psi\rangle_{ABC} = |\text{EPR}\rangle_{AB}^{\otimes m} \otimes |\text{something}\rangle. \quad (13)$$

Our goal is to argue that states admitting such a structure form an extremely small subset of the whole Hilbert space.

In an n -qubit system, there exist 2^n mutually orthogonal pure states. However, if we relax the requirement of exact orthogonality and demand only that distinct states have small mutual overlaps, one can show that there are, in fact, doubly exponentially many such *nearly* orthogonal states. Namely, letting $\Phi_{\text{state}}(n)$ denote the number of nearly orthogonal n -qubit states within some small inner-product cutoff, we have

$$\Phi_{\text{state}}(n) \sim e^{O(2^n)}. \quad (14)$$

Drawing a Haar random state is effectively equivalent to picking a state from this enormous set of quantum states. Likewise, letting $\Phi_{\text{unitary}}(n)$ denote the number of distinguishable n -qubit unitaries, $\Phi_{\text{unitary}}(n)$ is also doubly exponential:

$$\Phi_{\text{unitary}}(n) \sim e^{O(4^n)}. \quad (15)$$

Here, the difference in the exponents can be understood via the Choi-Jamiołkowski isomorphism where an n -qubit unitary can be viewed as a $2n$ -qubit pure state.

Now we can estimate the total number of states with LU-distillable EPR pairs. The first step is to pick local unitaries $U_A \otimes U_B$ that distill m EPR pairs. Since Φ_{unitary} grows doubly exponentially, the number of choices for such unitary operators is upper bounded by

$$\Phi_{\text{unitary}}(n_A)\Phi_{\text{unitary}}(n_B) \sim e^{O(4^{n_A})+O(4^{n_B})} < \Phi_{\text{unitary}}(n_R + c), \quad (16)$$

$$n_R = \max(n_A, n_B),$$

where $c > 0$ is an $O(1)$ constant. Once the EPR pairs are removed, the remaining $n' = n - 2m$ qubits can be in an

arbitrary state, giving $\Phi_{\text{state}}(n')$ possibilities. Therefore, the total number of states with LU-distillable EPR pairs is upper bounded by

$$\Phi_{\text{unitary}}(n_R + c)\Phi_{\text{state}}(n') \sim e^{O(2^{2n_R+2c})+O(2^{n-2m})}. \quad (17)$$

Importantly, if $n_R < (n/2)$ [more precisely, if $n_R = (n/2) - \omega(1)$], then

$$e^{O(2^{2n_R+2c})+O(2^{n-2m})}/e^{O(2^n)} < e^{-c'O(2^n)}, \quad (18)$$

which is doubly exponentially small. [Here, $c' > 0$ is another $O(1)$ constant.] In other words, among the doubly exponentially large number of nearly orthogonal states in an n -qubit Hilbert space, only a doubly exponentially tiny fraction of states admit LU-distillable entanglement between A and B . This intuition also clarifies when the counting argument applies: When $2n_R \gtrsim n$ (i.e., one subsystem contains half of the system), the counting argument no longer suppresses such states, and LU distillation of EPR pairs becomes possible.

2. ϵ -net

While the above argument is intuitively appealing, promoting it to a rigorous proof requires overcoming two technical hurdles. First, the argument's scope is limited to exact EPR pairs [i.e., Eq. (13) holds perfectly], but we also hope to consider states that contain approximate EPR pairs as measured by fidelity. Second, although Haar measure is defined over a continuous space, the above intuitive argument is based on a discrete set of states. To address these challenges, we employ the idea of an ϵ -net, which allows us to discretize the continuous Hilbert space. Specifically, an ϵ -net is a finite set of reference states, $\mathcal{M}_\epsilon$, with the property that any state in the entire space is guaranteed to be within a distance ϵ of some state in $\mathcal{M}_\epsilon$ [34]. Elements of an ϵ -net serve as reference points that allow us to characterize the Hilbert space.

A generic lower bound on the number of states (the cardinality) of an ϵ -net $\mathcal{M}_\epsilon$ can be obtained by a simple volume argument. However, known constructions contain significantly more states than this lower bound [34]. In this paper, we prove the existence of $\mathcal{M}_\epsilon$ whose cardinality is essentially tight with respect to this bound (Lemma 1 and Corollary 1). This construction is crucial and enables us to obtain an essentially tight threshold for LU-distillation results. The construction may also be of independent interest.

3. Concentration of measure

While the above counting argument provides physically transparent intuition, it unfortunately does not apply to LO-distillation problems. The main reason is that LO allows ancilla qubits that are initially unentangled, which spoils the volume-counting argument. Moreover, LO modifies the

relevant fidelity threshold significantly, as explained below in Theorem 2. As such, we need to use a different strategy.

The key idea is to use concentration of measure. Many physical quantities computed for Haar random states exhibit concentration of measure, meaning that their probability distributions sharply concentrate around their expectation values. This phenomenon is most precisely captured by Levy's lemma, which establishes exponential concentration for observables satisfying a continuity (more precisely, Lipschitz) property. It turns out that the EPR fidelity satisfies this condition.

We are interested in the distribution of the EPR fidelity. For given LOs (characterized by isometries V_A and V_B), the expectation value of the EPR fidelity depends on the choice of isometries. Thus, we cannot directly identify the location of concentration. The key observation, however, is that we can uniformly upper bound the average fidelity over all possible isometries. This allows us to conclude that measure concentration occurs at or below the threshold, and, therefore, the probability of LO distillation is exponentially suppressed.

To extend the argument from fixed isometries to arbitrary local operations, we apply the concentration bound uniformly over an ϵ -net of isometries and then take a union bound. Although the number of candidate local operations grows exponentially in subsystem dimensions, the concentration estimate is exponentially strong in the total Hilbert space dimension. Consequently, the union bound does not overwhelm the concentration effect, and the probability that any local operation achieves fidelity above the threshold remains exponentially suppressed.

II. LOCAL UNITARY: PROOF OF THEOREM 1

A. Spherical cap

Recall that a pure state can be written as

$$|\Psi\rangle = \sum_{j=1}^d (a_j + ib_j)|j\rangle, \quad \sum_{j=1}^d a_j^2 + b_j^2 = 1, \quad (19)$$

where $a_j, b_j \in \mathbb{R}$. Hence, a pure state corresponds to a point on a unit sphere in $\mathbb{R}^{2d}$. Sampling a Haar random state

corresponds to choosing a point uniformly at random from a unit sphere [35]. Throughout this paper, a unit sphere in $\mathbb{R}^p$ is called a unit $(p-1)$ -sphere and is denoted by S^{p-1} while its interior is called a unit p -ball. Their area and volume are given by

$$S_{p-1} = \frac{2\pi^{p/2}}{\Gamma(\frac{p}{2})}, \quad V_p = \frac{\pi^{p/2}}{\Gamma(\frac{p}{2} + 1)}, \quad (20)$$

where $\Gamma(z)$ is the Gamma function.

Consider a unit $(p-1)$ -sphere in $\mathbb{R}^p$. Letting $(x_1, \dots, x_p)$ be the coordinates, a spherical cap $\text{Cap}(h)$ is a portion of the sphere cut off by a hyperplane at height $x_1 = h$, where $0 < h \leq 1$, namely,

$$\text{Cap}(h) = \left\{ (x_1, \dots, x_p) \in \mathbb{R}^p \mid h \leq x_1 \leq 1, \sum_{j=1}^p x_j^2 = 1 \right\}, \quad (21)$$

where the radius of the cap is $r = \sqrt{1-h^2}$. See Fig. 2(a) for an illustration. We also consider a generalized spherical cap defined as

$$\begin{aligned} \text{Cap}^{(q)}(h) = & \left\{ (x_1, \dots, x_p) \in \mathbb{R}^p \mid h^2 \right. \\ & \left. \leq \sum_{j=1}^q x_j^2 \leq 1, \sum_{j=1}^p x_j^2 = 1 \right\} \end{aligned} \quad (22)$$

for $1 \leq q \leq p-1$. Note that $\text{Cap}^{(1)}(h)$ consists of two copies of $\text{Cap}(h)$ for $x_1 > 0$ and $x_1 < 0$.

B. ϵ -net

For $0 < \epsilon < 1$, a set of states $\mathcal{M}_\epsilon = \{|\tilde{\Psi}_j\rangle\}$ is said to be an ϵ -net in the 2-norm [34] when, for every pure state $|\Psi\rangle$ in the Hilbert space, there exists $|\tilde{\Psi}_i\rangle \in \mathcal{M}_\epsilon$ such that

$$\| |\Psi\rangle - |\tilde{\Psi}_i\rangle \|_2 \leq \epsilon. \quad (23)$$

See Fig. 3 for an illustration. For a d -dimensional Hilbert space, the 2-norm distance between a pair of pure states

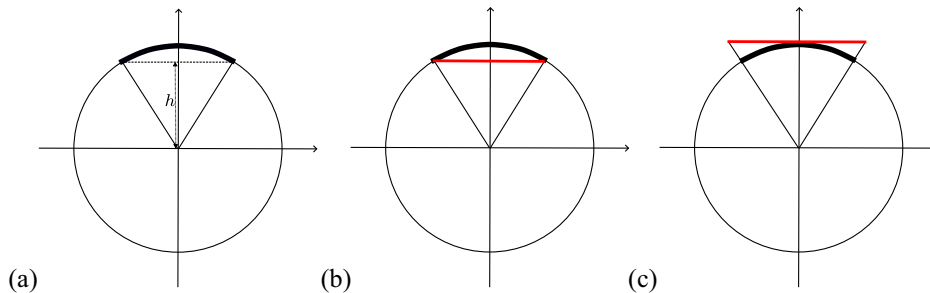


FIG. 2. A spherical cap and upper and lower bounds on its surface area on a unit sphere S^{p-1} . Figures are depicted for $p = 2$; x_1 is the vertical axis. (a) A spherical cap Cap in $\mathbb{R}^p$. (b) A lower bound. (c) An upper bound.

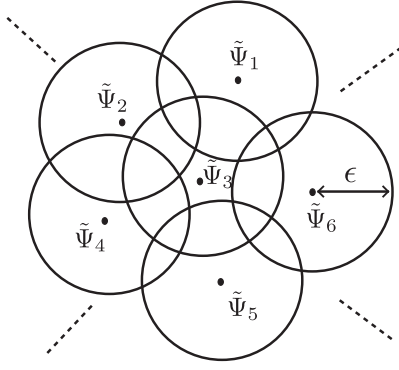


FIG. 3. A schematic picture of an ϵ -net. Any point in the Hilbert space has some ϵ -neighbor point in the ϵ -net.

corresponds to the Euclidean distance in $\mathbb{R}^{2d}$. An ϵ -net is often discussed in terms of the 1-norm in the literature, but an ϵ -net in the 2-norm suffices to establish the desired result. Note that an ϵ -net in the 2-norm is a 2ϵ -net in the 1-norm [34], but the converse is not true.

We begin by proving that there exists an ϵ -net $\mathcal{M}_\epsilon$ with $|\mathcal{M}_\epsilon| \lesssim (1/\epsilon)^{2d}$ up to subleading factors. This construction yields an ϵ -net with nearly optimal cardinality, a result that is exponentially tighter (in dimension) than the one in [34]. For this, we use the following lemma regarding ϵ -nets on a unit sphere S^{p-1} .

Lemma 1. Assume $p \geq 3$ and $0 < \epsilon < \sqrt{2}$. For any region $R \subseteq S^{p-1}$, there exists an absolute constant $\alpha > 0$ and an ϵ -net $\mathcal{M}_\epsilon(R)$ of R such that

$$|\mathcal{M}_\epsilon(R)| \leq \alpha p \log p \frac{\text{Area}(R^{+2\epsilon})}{\text{Area}(B_\epsilon)}. \quad (24)$$

Here, $B_\epsilon \subseteq S^{p-1}$ is the ϵ spherical ball (in the Euclidean distance), $R^{+2\epsilon} = \{x \in S^{p-1} | \text{dist}(x, R) \leq 2\epsilon\}$, and $\mathcal{M}_\epsilon(R) \subseteq S^{p-1}$ does not need to be contained in R .

Proof. It is proven in Ref. [36] that S^{p-1} can be covered by ϵ spherical balls such that every point on a unit sphere S^{p-1} is covered by less than $400p \ln p$ times. We pick such a covering. For any $R \subseteq S^{p-1}$, we collect ϵ spherical balls that intersect with R . Denote $\mathcal{M}_\epsilon(R)$ as the set of centers of these balls. Then, $\mathcal{M}_\epsilon(R)$ is an ϵ -net of R , since the union of these balls covers R . On the other hand, the union of those balls is contained in $R^{+2\epsilon}$ due to the triangle inequality, and each point of $R^{+2\epsilon}$ is counted by less than $400p \ln p$ times; hence,

$$|\mathcal{M}_\epsilon(R)| \text{Area}(B_\epsilon) \leq 400p \ln p \text{Area}(R^{+2\epsilon}). \quad (25)$$

This completes the proof. $\blacksquare$

Observing that B_ϵ is a spherical cap $\text{Cap}(\sqrt{1-r^2})$ with radius $r = \epsilon\sqrt{1 - (\epsilon^2/4)}$, its area can be lower bounded by

$$\text{Area}(B_\epsilon) > \left(\epsilon \sqrt{1 - \frac{\epsilon^2}{4}} \right)^{p-1} V_{p-1}; \quad (26)$$

see Fig. 2(b) for an illustration. Hence, by choosing $R = S^{2d-1} = R^{+2\epsilon}$, we obtain the following corollary.

Corollary 1. For pure states in a d -dimensional Hilbert space, there exists an ϵ -net $\mathcal{M}_\epsilon$ of S^{2d-1} satisfying

$$|\mathcal{M}_\epsilon| \leq 2\alpha d \log(2d) \left(\frac{1}{f(\epsilon)} \right)^{2d-1} \frac{S_{2d-1}}{V_{2d-1}} \quad (27)$$

for $0 < \epsilon < \sqrt{2}$, where $f(\epsilon) \equiv \epsilon\sqrt{1 - (\epsilon^2/4)}$.

C. State with bare EPR pairs

Consider a set of states where m EPR pairs are already prepared approximately on smaller subsystems $R_A \subseteq A$ and $R_B \subseteq B$ without the need to apply local unitary transformations. Namely, we define

$$\mathbf{N}^{(m,h)} \equiv \left\{ |\Psi\rangle \in \mathcal{H}_{ABC} \left| \langle \Psi | \Pi_{R_A R_B}^{[\text{EPR}]} | \Psi \rangle \geq h^2 \right. \right\} \quad (28)$$

for some fixed R_A, R_B with $|R_A| = |R_B| = m$.

Lemma 2. For $0 < 2\epsilon \leq h \leq 1$, there exists an ϵ -net $\mathcal{N}_\epsilon^{(m,h)}$ for $\mathbf{N}^{(m,h)}$ such that

$$|\mathcal{N}_\epsilon^{(m,h)}| \leq 2\alpha d \log(2d) \left(\frac{1}{f(\epsilon)} \right)^{2d-1} \frac{\text{Area}(\text{Cap}^{(2\tilde{d})}(h-2\epsilon))}{V_{2d-1}}, \quad (29)$$

where $\tilde{d} = 2^{n-2m}$.

Proof. Let $R = R_A \cup R_B$ with $|R| = 2m$. Expand $|\Psi\rangle$ as

$$|\Psi\rangle = \sum_{j=1}^{\tilde{d}} \alpha_j |\tilde{1}\rangle_R \otimes |j\rangle_{R^c} + \sum_{i=2}^{2^{2m}} \sum_{j=1}^{\tilde{d}} \beta_{ij} |\tilde{i}\rangle_R \otimes |j\rangle_{R^c} \quad (30)$$

with orthonormal basis on R and R^c while choosing $|\tilde{1}\rangle_R = |\text{EPR}\rangle_{R_A R_B}$. The state $|\Psi\rangle$ is in $\mathbf{N}^{(m,h)}$ if and only if

$$\sum_{j=1}^{\tilde{d}} |\alpha_j|^2 \geq h^2, \quad \sum_{i=2}^{2^{2m}} \sum_{j=1}^{\tilde{d}} |\beta_{ij}|^2 \leq 1 - h^2. \quad (31)$$

Hence, the total surface region occupied by states in $\mathbf{N}^{(m,h)}$ is $\text{Cap}^{(2\tilde{d})}(h)$ in $\mathbb{R}^{2d}$. Using Lemma 1 with $R = \text{Cap}^{(2\tilde{d})}(h)$ and observing $R^{+2\epsilon} \subseteq \text{Cap}^{(2\tilde{d})}(h-2\epsilon)$, we obtain the desired result. $\blacksquare$

The area of $\text{Cap}^{(2\tilde{d})}(h)$ has a simple asymptotic expression for large d , namely,

$$\frac{\text{Area}(\text{Cap}^{(2\tilde{d})}(h))}{S_{2d-1}} \approx u(2^{-2m} - h^2), \quad (32)$$

where $u(x)$ is a “step function”: For $h^2 > 2^{-2m}$, the area of $\text{Cap}^{(2\tilde{d})}(h)$ will be negligibly small, while for $h^2 < 2^{-2m}$,

the area will be almost as large as that of the unit sphere. For our purpose, we need the estimation only when $h^2 > 2^{-2m}$, which is formalized in the following lemma.

Lemma 3. Let $\delta \stackrel{\text{def}}{=} h^2 - 2^{-2m}$. For $\delta > 0$,

$$\frac{\text{Area}(\text{Cap}^{(2\tilde{d})}(h))}{S_{2d-1}} \leq \exp(-2(d+1)\delta^2). \quad (33)$$

Proof. The area of $\text{Cap}^{(2\tilde{d})}(h)$ has an exact expression. Using the polar coordinates on S^{2d-1} , we have

$$\frac{\text{Area}(\text{Cap}^{(2\tilde{d})}(h))}{S_{2d-1}} = \frac{1}{B(\tilde{d}, d - \tilde{d})} \int_{h^2}^1 x^{\tilde{d}-1} (1-x)^{d-\tilde{d}-1} dx, \quad (34)$$

where $B(\cdot, \cdot)$ is the Beta function. Namely, it equals $\mathbb{P}(X > h^2)$ where the random variable X follows the Beta distribution $\text{Beta}(\tilde{d}, d - \tilde{d})$. For an $X \sim \text{Beta}(\tilde{d}, d - \tilde{d})$, we have

$$\mathbb{E}(X) = \frac{\tilde{d}}{d} = 2^{-2m}, \quad \text{Var}(X) = \frac{\tilde{d}(d - \tilde{d})}{d^2(d+1)}. \quad (35)$$

Therefore, for large d , the distribution is narrowly peaked at the mean value 2^{-2m} . In particular, Ref. [37] shows that $\text{Beta}(\tilde{d}, d - \tilde{d})$ is $[1/4(d+1)]$ -sub-Gaussian, implying that

$$\frac{\text{Area}(\text{Cap}^{(2\tilde{d})}(h))}{S_{2d-1}} \leq \exp(-2(d+1)\delta^2), \quad (36)$$

where $\delta = h^2 - 2^{-2m} > 0$. ■

D. State with distillable EPR pairs

Let us define a set of states with LU-distillable EPR pairs

$$\hat{\mathbf{N}}^{(m,h)} \equiv \left\{ |\Psi\rangle \in \mathcal{H}_{ABC} \mid \text{ED}_h^{\text{LU}}(A:B) \geq m \right\}. \quad (37)$$

Note that $\hat{\mathbf{N}}^{(m,h)}$ can be constructed by applying unitaries with the form of $U_A \otimes U_B$ to states in $\mathbf{N}^{(m,h)}$.

Lemma 4. Let $\mathcal{N}_\epsilon^{(m,h)}$ be an ϵ -net for $\mathbf{N}^{(m,h)}$. Let $\hat{\epsilon} \equiv \epsilon + 2\epsilon'$, where $\epsilon' < \sqrt{2}$. Then, there exists an $\hat{\epsilon}$ -net $\hat{\mathcal{N}}_\epsilon^{(m,h)}$ for $\hat{\mathbf{N}}^{(m,h)}$ such that

$$|\hat{\mathcal{N}}_\epsilon^{(m,h)}| \leq \alpha_1 d_A^2 d_B^2 \log(d_A) \log(d_B) \left(\frac{1}{f(\epsilon')} \right)^{2d_A^2 + 2d_B^2 - 2} \times \frac{S_{2d_A^2-1} S_{2d_B^2-1}}{V_{2d_A^2-1} V_{2d_B^2-1}} |\mathcal{N}_\epsilon^{(m,h)}| \quad (38)$$

for some absolute constant $\alpha_1 > 0$.

Proof. For each $|\Psi\rangle \in \mathcal{N}_\epsilon^{(m,h)}$, we can Schmidt decompose it as

$$|\Psi\rangle = \sum_{i=1}^K \lambda_i |\psi_i\rangle_A \otimes |\phi_i\rangle_{BC}, \quad (39)$$

where the Schmidt rank $K \leq d_A$ since $|A| < |B| + |C|$. Therefore, we can represent $U_A \otimes I_{BC} |\Psi\rangle$ as $(\lambda_1 U_A |\psi_1\rangle, \lambda_2 U_A |\psi_2\rangle, \dots)$, a normalized complex vector with complex dimension $K d_A$. Such representation is an isometry from the space $\{U_A \otimes I_{BC} |\Psi\rangle \mid U_A \in \text{U}(\mathcal{H}_A)\}$ to $S^{2K d_A - 1}$, where $\text{U}(\mathcal{H}_R)$ denotes the group of unitaries acting on a subsystem R . Conversely, any point on $S^{2K d_A - 1}$ also corresponds to a normalized pure state on ABC . Therefore, we can use an ϵ' -net on $S^{2K d_A - 1}$ to construct an ϵ' -net of pure states for the space $\{U_A \otimes I_{AC} |\Psi\rangle\}$. The cardinality of this net is upper bounded by $2\alpha d_A^2 \log(2d_A^2) \{1/f(\epsilon')\}^{2d_A^2-1} (S_{2d_A^2-1}/V_{2d_A^2-1})$ due to Corollary 1 and $K \leq d_A$.

Repeating the same argument for U_B and using the triangle inequality for the 2-norm, we find a $2\epsilon'$ -net for the space $\{U_A \otimes U_B |\Psi\rangle \mid U_A \in \text{U}(\mathcal{H}_A), U_B \in \text{U}(\mathcal{H}_B)\}$. Repeating this procedure for all states $|\Psi\rangle$ in the ϵ -net $\mathcal{N}_\epsilon^{(m,h)}$, we conclude that there exists an $\hat{\epsilon}$ -net $\hat{\mathcal{N}}_\epsilon^{(m,h)}$ for $\hat{\mathbf{N}}^{(m,h)}$ whose cardinality is upper bounded as specified in the lemma. ■

Combining Lemmas 2 and 4, we arrive at the following result.

Lemma 5. For $0 < 2\epsilon \leq h \leq 1$, $\epsilon' < \sqrt{2}$, there exists an $\hat{\epsilon}$ -net with $\hat{\epsilon} = \epsilon + 2\epsilon'$ such that

$$|\hat{\mathcal{N}}_\epsilon^{(m,h)}| \leq \alpha_2 d_A^2 d_B^2 d \log(d_A) \log(d_B) \log(d) \times \left(\frac{1}{f(\epsilon')} \right)^{2d_A^2 + 2d_B^2 - 2} \left(\frac{1}{f(\epsilon)} \right)^{2d-1} \cdot \frac{\text{Area}(\text{Cap}^{(2\tilde{d})}(h - 2\epsilon))}{V_{2d-1}} \frac{S_{2d_A^2-1} S_{2d_B^2-1}}{V_{2d_A^2-1} V_{2d_B^2-1}} \quad (40)$$

for some absolute constant $\alpha_2 > 0$.

E. Putting together

Recall

$$|\Psi\rangle \in \hat{\mathbf{N}}^{(m,h)} \Rightarrow \exists |\tilde{\Psi}\rangle \in \hat{\mathcal{N}}_\epsilon^{(m,h)} \text{ such that } \|\Psi - \tilde{\Psi}\|_2 \leq \hat{\epsilon}. \quad (41)$$

Hence, we have

$$\mathbb{P}(|\Psi\rangle \in \hat{\mathbf{N}}^{(m,h)}) \leq \mathbb{P}(\exists |\tilde{\Psi}\rangle \in \hat{\mathcal{N}}_\epsilon^{(m,h)} \text{ such that } \|\Psi - \tilde{\Psi}\|_2 \leq \hat{\epsilon}). \quad (42)$$

Let $\{|\tilde{\Psi}_j\rangle\}$ be elements of $\hat{\mathcal{N}}_{\hat{\epsilon}}^{(m,h)}$. Using the union bound, we have

$$\mathbb{P}\left(\exists |\tilde{\Psi}\rangle \in \hat{\mathcal{N}}_{\hat{\epsilon}}^{(m,h)} \text{ such that } \|\Psi\rangle - |\tilde{\Psi}\rangle\|_2 \leq \hat{\epsilon}\right) \leq \sum_j \mathbb{P}\left(\|\Psi\rangle - |\tilde{\Psi}_j\rangle\|_2 \leq \hat{\epsilon}\right). \quad (43)$$

Lemma 6. Let $|\Psi\rangle$ be chosen according to the Haar measure and $|\Psi_0\rangle$ be an arbitrary, fixed state. Let $\epsilon < 1$. Then,

$$\mathbb{P}(\|\Psi\rangle - |\Psi_0\rangle\|_2 \leq \epsilon) < \frac{g(\epsilon)^{2d-1} V_{2d-1}}{S_{2d-1}}, \quad (44)$$

where $g(\epsilon) \equiv (\epsilon/\sqrt{1-\epsilon^2})$.

Proof. The ϵ -ball around $|\Psi_0\rangle$ is given by $\text{Cap}(\sqrt{1-r^2})$, where $r(\epsilon) = \epsilon\sqrt{1-(\epsilon^2/4)}$. This area can be upper bounded by $(2d-1)$ -ball of radius

$$\frac{r(\epsilon)}{\sqrt{1-r(\epsilon)^2}} < \frac{\epsilon}{\sqrt{1-\epsilon^2}} = g(\epsilon). \quad (45)$$

See Fig. 2(c) for an illustration. This completes the proof. ■

Using Lemmas 5 and 6 and Eq. (43), we have

$$\begin{aligned} \mathbb{P}\left(|\Psi\rangle \in \hat{\mathcal{N}}^{(m,h)}\right) &< \alpha_2 d_A^2 d_B^2 d \log(d_A) \log(d_B) \log(d) \left(\frac{1}{f(\epsilon')}\right)^{2d_A^2+2d_B^2-2} \left(\frac{1}{f(\epsilon)}\right)^{2d-1} g(\hat{\epsilon})^{2d-1} \\ &\cdot \frac{S_{2d_A^2-1} S_{2d_B^2-1}}{V_{2d_A^2-1} V_{2d_B^2-1}} \frac{\text{Area}(\text{Cap}^{(2\tilde{d})}(h-2\epsilon))}{S_{2d-1}}. \end{aligned} \quad (46)$$

Taking the logarithm with base 2 leads to

$$\begin{aligned} \log \left[\mathbb{P}\left(|\Psi\rangle \in \hat{\mathcal{N}}^{(m,h)}\right) \right] &< \log \left(\frac{S_{2d_A^2-1} S_{2d_B^2-1}}{V_{2d_A^2-1} V_{2d_B^2-1}} \right) + \log \left(\frac{\text{Area}(\text{Cap}^{(2\tilde{d})}(h-2\epsilon))}{S_{2d-1}} \right) \\ &- (2d_A^2 + 2d_B^2 - 2) \log f(\epsilon') - (2d-1) \log f(\epsilon) + (2d-1) \log g(\hat{\epsilon}) + O(\log d), \end{aligned} \quad (47)$$

where the multiplicative factor $\alpha_2 d_A^2 d_B^2 d \log(d_A) \log(d_B) \log(d)$ leads to a contribution of $O(\log d)$. Let us split terms into three groups as follows:

$$\begin{aligned} A_1 &\equiv \log \left(\frac{S_{2d_A^2-1} S_{2d_B^2-1}}{V_{2d_A^2-1} V_{2d_B^2-1}} \right), \\ A_2 &\equiv \log \left(\frac{\text{Area}(\text{Cap}^{(2\tilde{d})}(h-2\epsilon))}{S_{2d-1}} \right), \\ A_3 &\equiv -(2d_A^2 + 2d_B^2 - 2) \log f(\epsilon') - (2d-1) \log f(\epsilon) + (2d-1) \log g(\hat{\epsilon}). \end{aligned} \quad (48)$$

We then have

$$\log \left[\mathbb{P}\left(|\Psi\rangle \in \hat{\mathcal{N}}^{(m,h)}\right) \right] < A_1 + A_2 + A_3 + O(\log d). \quad (49)$$

Note that the bound is valid for arbitrary ϵ, ϵ' (as long as $\hat{\epsilon} = \epsilon + 2\epsilon' < 1$).

Let us evaluate each term. As for A_1 , we have

$$A_1 = \log \left(\frac{2\pi^{d_A^2}}{\Gamma(d_A^2)} \frac{2\pi^{d_B^2}}{\Gamma(d_B^2)} \frac{\Gamma(d_A^2 + \frac{1}{2})}{\pi^{d_A^2 - (1/2)}} \frac{\Gamma(d_B^2 + \frac{1}{2})}{\pi^{d_B^2 - (1/2)}} \right) = \log \left(4\pi \frac{\Gamma(d_A^2 + \frac{1}{2})\Gamma(d_B^2 + \frac{1}{2})}{\Gamma(d_A^2)\Gamma(d_B^2)} \right) < \log(4\pi d_A d_B), \quad (50)$$

where we have used $\Gamma(n^2 + \frac{1}{2}) < n\Gamma(n^2)$. Hence, we have

$$A_1 < O(\log d). \quad (51)$$

As for A_2 , Lemma 3 implies

$$A_2 < -2(d+1) \left((h-2\epsilon)^2 - 2^{-2m} \right)^2 < -2d(h^2 - 4\epsilon - 2^{-2m})^2 < -2d\delta^2 + O(d\delta\epsilon), \quad (52)$$

where the second inequality follows from $(h-2\epsilon)^2 > h^2 - 4\epsilon$ and assumes $\delta = h^2 - 2^{-2m} > 4\epsilon$.

As for A_3 , recall that

$$f(\epsilon) \equiv \epsilon \sqrt{1 - \frac{\epsilon^2}{16}} = \epsilon[1 + O(\epsilon^2)], \quad g(\epsilon) \equiv \frac{\epsilon}{\sqrt{1 - \epsilon^2}} = \epsilon[1 + O(\epsilon^2)]. \quad (53)$$

Hence, we have

$$\begin{aligned} A_3 &= 2d \log \frac{g(\epsilon + 2\epsilon')}{f(\epsilon)} + O\left(d_A^2 \log \frac{1}{\epsilon'}, d_B^2 \log \frac{1}{\epsilon'}\right) \\ &= O\left(d \frac{\epsilon'}{\epsilon}\right) + O(d\epsilon^2) + O(d\epsilon'^2) + O\left(d_A^2 \log \frac{1}{\epsilon'}, d_B^2 \log \frac{1}{\epsilon'}\right). \end{aligned} \quad (54)$$

Putting these together, for $\delta = h^2 - (\tilde{d}/d) > 0$, we have

$$\log[\mathbb{P}(|\Psi\rangle \in \hat{\mathbf{N}}^{(m,h)})] < -2d\delta^2 + O\left(\log d, d\epsilon^2, d\epsilon'^2, d\epsilon\delta, d\frac{\epsilon'}{\epsilon}, d_A^2 \log \frac{1}{\epsilon'}, d_B^2 \log \frac{1}{\epsilon'}\right). \quad (55)$$

By choosing $\epsilon = \tilde{c}\delta$ and $\epsilon' = \epsilon^2$ where $\tilde{c}$ is a small enough constant, we have [38]

$$\log[\mathbb{P}(|\Psi\rangle \in \hat{\mathbf{N}}^{(m,h)})] < -[2 - O(1)]d\delta^2 + O\left(d_A^2 \log \frac{1}{\delta}, d_B^2 \log \frac{1}{\delta}\right). \quad (56)$$

This completes the proof of Theorem 1.

III. LOCAL OPERATION: PROOF OF THEOREM 2

In this section, we extend our bound on distillable entanglement to local operations. For convenience, we work within the Stinespring dilation picture, where a local operation is described as first applying an isometry (or attaching ancillas and applying a unitary) and then tracing out a subsystem.

Fact 1—Every quantum channel $\Phi: \mathcal{S}(\mathbb{C}^{d_1}) \rightarrow \mathcal{S}(\mathbb{C}^{d_2})$ can be expressed as

$$\Phi(\rho) = \text{Tr}_{\mathbb{C}^{d_3}}(V\rho V^\dagger), \quad (57)$$

where $V: \mathbb{C}^{d_1} \rightarrow \mathbb{C}^{d_2} \otimes \mathbb{C}^{d_3}$ is an isometry. Furthermore, since the Kraus rank of Φ is at most $d_1 d_2$ [39], we can always choose $d_3 \leq d_1 d_2$.

Applying it to Φ_A and Φ_B in our setting, we obtain two isometries $V_A: \mathbb{C}^{d_A} \rightarrow \mathbb{C}^{d'_A}$ and $V_B: \mathbb{C}^{d_B} \rightarrow \mathbb{C}^{d'_B}$, where [40]

$$d'_A \leq d_A 2^{2m}, \quad d'_B \leq d_B 2^{2m}. \quad (58)$$

In other words, crucially, we do not need to attach a large number of ancilla, as the relevant output system on each side contains only m qubits.

Let us first consider the case where the isometries V_A and V_B are fixed.

Lemma 7. Suppose the quantum channel Φ_A (and Φ_B) is defined via Eq. (57) from a given isometry $V_A: \mathbb{C}^{d_A} \rightarrow \mathbb{C}^{d'_A}$ (and $V_B: \mathbb{C}^{d_B} \rightarrow \mathbb{C}^{d'_B}$, respectively), $\Lambda = \Phi_A \otimes \Phi_B$, then

$$\mathbb{P}\left(\text{Tr}(\Lambda(\rho_{AB})\Pi^{[\text{EPR}]}) > 2^{-m} + \eta\right) \leq \exp(-d\eta^2). \quad (59)$$

Here, the probability is taken over the Haar measure on states $|\psi\rangle$ on ABC , $d = 2^{n_{ABC}}$, and $\eta > 0$ is arbitrary.

Proof. We define a function $f_{V_A V_B}(\psi)$ as

$$f_{V_A V_B}(\psi) = \text{Tr}(\Lambda(\rho_{AB})\Pi^{[\text{EPR}]}) = \langle \psi | V_A^\dagger V_B^\dagger \Pi^{[\text{EPR}]} V_A V_B | \psi \rangle, \quad (60)$$

where we slightly abuse the notation of $\Pi^{[\text{EPR}]}$: In the third term, it now acts as an operator on $\mathbb{C}^{d'_A} \otimes \mathbb{C}^{d'_B}$. We claim f is 1-Lipschitz. In fact,

$$|f(\psi) - f(\phi)| = \left| \text{Tr}\left(\tilde{\Pi}(|\psi\rangle\langle\psi| - |\phi\rangle\langle\phi|)\right) \right| \leq \|\tilde{\Pi}\|_\infty \sin\theta(\psi, \phi) \leq \|\psi - \phi\|_2 \leq \text{dist}(\psi, \phi). \quad (61)$$

Here, $\tilde{\Pi} = V_A^\dagger V_B^\dagger \Pi^{[\text{EPR}]} V_A V_B$, $\theta(\psi, \phi)$ is the (Hermitian) angle between $|\psi\rangle$ and $|\phi\rangle$ defined as $\cos\theta(\psi, \phi) = |\langle\psi|\phi\rangle|$, and $\text{dist}(\psi, \phi)$ is the (geodesic) distance when regarding $|\psi\rangle$ and $|\phi\rangle$ as points on S^{2d-1} . The first inequality is due to the fact that the eigenvalues of $|\psi\rangle\langle\psi| - |\phi\rangle\langle\phi|$ are $\{\sin\theta, -\sin\theta, 0, \dots, 0\}$.

Let us compute the average of $f_{V_A V_B}(\psi)$ over Haar random $|\psi\rangle$. By standard Haar average computation, we have

$$\mathbb{E}(f_{V_A V_B}(\psi)) = \frac{1}{d} \text{Tr}\left(V_A^\dagger V_B^\dagger \Pi^{[\text{EPR}]} V_A V_B\right). \quad (62)$$

While it may depend on V_A and V_B , the following always holds:

$$\mathbb{E}(f_{V_A V_B}(\psi)) = \frac{1}{d} 2^{-m} \left(\text{Diagram 1} \right) \leq \frac{1}{d} 2^{-m} \left(\text{Diagram 2} \right) = 2^{-m}, \quad (63)$$

where the inequality comes from the fact that $\text{Tr}(XY) \leq \text{Tr}(X)\text{Tr}(Y)$ for two positive semidefinite matrices. Therefore, $f_{V_A V_B}(\psi) > 2^{-m} + \eta$ implies $f_{V_A V_B}(\psi) > \mathbb{E}(f_{V_A V_B}(\psi)) + \eta$.

Now, applying Levy's lemma (see below) for f on S^{2d-1} , we obtain the desired bound:

$$\mathbb{P}(f_{V_A V_B}(\psi) > 2^{-m} + \eta) \leq \exp\left(-\frac{2d\eta^2}{2}\right) = \exp(-d\eta^2). \quad (64)$$

In the above proof, we used the following fact about the concentration of measure on high-dimensional spheres [41,42].

Fact 2 (Levy's lemma)—If a function $f: S^{p-1} \rightarrow \mathbb{R}$ is K -Lipschitz in the sense that $|f(x) - f(y)| \leq K \text{dist}(x, y)$ (geodesic distance), then

$$\Pr[f > \mathbb{E}(f) + \eta] \leq \exp\left(-\frac{p\eta^2}{2K^2}\right). \quad (65)$$

Next, to allow arbitrary isometries on A and B , we consider ϵ -nets on the space of isometries. We denote the

space of isometries $\mathbb{C}^d \rightarrow \mathbb{C}^{d'}$ as $\mathcal{V}_d(\mathbb{C}^{d'})$. It is a Stiefel manifold $U(d')/U(d'-d)$, and $\dim \mathcal{V}_d(\mathbb{C}^{d'}) = 2dd' - d^2$. We quote the following result from Ref. [43,44].

Fact 3—There exists an absolute number c such that $\mathcal{V}_d(\mathbb{C}^{d'})$ has an ϵ -net $\mathcal{M}_\epsilon$ in operator norm $\|\cdot\|_\infty$ such that

$$|\mathcal{M}_\epsilon| \leq \left(\frac{c}{\epsilon}\right)^{2dd'-d^2}. \quad (66)$$

Now we prove Theorem 2.

Proof of Theorem 2. We pick ϵ -nets for isometries according to Fact 3 and denote them as $\mathcal{M}_\epsilon^A$ and $\mathcal{M}_\epsilon^B$, respectively. We choose ϵ later. For any $|\psi\rangle$, V_A and V_B , we can always choose $V'_A \in \mathcal{M}_\epsilon^A$ and $V'_B \in \mathcal{M}_\epsilon^B$ such that $\|V_A - V'_A\|_\infty \leq \epsilon$ and $\|V_B - V'_B\|_\infty \leq \epsilon$ and, hence, $\|V_A V_B |\psi\rangle - V'_A V'_B |\psi\rangle\|_\infty \leq 2\epsilon$. It follows from Eq. (61) that

$$f_{V'_A V'_B}(\psi) \geq f_{V_A V_B}(\psi) - 2\epsilon. \quad (67)$$

We denote $\delta = h^2 - 2^{-m}$. By definition, for any $|\psi\rangle$,

$$\text{ED}_h^{[\text{LO}]}(A:B) \geq m \iff \exists V_A, V_B \text{ such that } f_{V_A V_B}(\psi) > 2^{-m} + \delta. \quad (68)$$

Therefore, due to the union bound, we have

$$\mathbb{P}(\text{ED}_h^{[\text{LO}]}(A:B) \geq m) \leq \sum_{V_A \in \mathcal{M}_\epsilon^A} \sum_{V_B \in \mathcal{M}_\epsilon^B} \mathbb{P}(f_{V_A V_B}(\psi) > 2^{-m} + \delta - 2\epsilon). \quad (69)$$

Applying Fact 3 and Lemma 7, we get

$$\mathbb{P}(\text{ED}_h^{[\text{LO}]}(A:B) \geq m) \leq |\mathcal{M}_\epsilon^A| |\mathcal{M}_\epsilon^B| \exp(-d(\delta - 2\epsilon)^2) \leq \left(\frac{c}{\epsilon}\right)^{2d_A d'_A + 2d_B d'_B} \exp(-d(\delta - 2\epsilon)^2), \quad (70)$$

as long as $\delta > 2\epsilon$.

Now we pick $\epsilon = c'\delta$, where the proportional constant may be taken arbitrarily small. Taking the logarithm of Eq. (70), we obtain

$$\log \mathbb{P}(\text{ED}_h^{[\text{LO}]}(A:B) \geq m) \leq -[1 - O(1)]\delta^2 d + O\left(2^{2m}(d_A^2 + d_B^2) \log \frac{1}{\delta}\right). \quad (71)$$

■

IV. LOGICAL OPERATORS IN RANDOM ENCODING

A. No logical operators in bipartite subsystems

Given two parameters $0 \leq h, w \leq 1$, we say an isometric encoding $V: C \rightarrow AB$ admits a logical unitary operator U_{AB} if there exists a unitary U_C such that

$$\left| \frac{\text{Tr}(U_C^\dagger V^\dagger U_{AB} V)}{d_C} \right| \geq h^2 \quad \text{and} \quad \frac{|\text{Tr}(U_C)|}{d_C} \leq w. \quad (72)$$

The first inequality imposes a fidelity h on a logical unitary operator U_{AB} . Namely, it compares two isometries $U_{AB}V$ and VU_C via the fidelity between their Choi states:

$$\begin{aligned} \frac{\text{Tr}(U_C^\dagger V^\dagger U_{AB} V)}{d_C} &= \langle \text{EPR} | U_C^\dagger V^\dagger U_{AB} V | \text{EPR} \rangle \\ &= \langle \Psi | U_{AB} \otimes U_C^* | \Psi \rangle, \end{aligned} \quad (73)$$

where $|\Psi\rangle = (V \otimes I)|\text{EPR}\rangle$ is the Choi state for V . Note that

$$1 - |\langle \text{EPR} | U_C^\dagger V^\dagger U_{AB} V | \text{EPR} \rangle| \leq \frac{1}{2} \|U_{AB} V - V U_C\|_\infty^2; \quad (74)$$

thus, a large h is also a necessary condition for $U_{AB}V$ and VU_C to be close in the operator norm.

The second inequality ensures the nontriviality of the logical operator. Namely, $w < 1$ is required to ensure that U_C acts nontrivially on the input state. Otherwise, $w = 1$ would imply that U_C is a phase $e^{i\theta} I_C$, and $e^{i\theta} I_{AB}$ is always a trivial logical operator for $e^{i\theta} I_C$ with $h = 1$. On the other hand, if U_C is a unitary conjugation of a Pauli operator, then $w = 0$.

Theorem 4. Assuming $\delta \stackrel{\text{def}}{=} h^2 - w > 0$, there exists an absolute constant $c' > 0$ such that, for a Haar random isometry V ,

$$\begin{aligned} &\log \mathbb{P}(V \text{ admits a logical operator on } A) \\ &\leq -c'\delta^2 d + O\left((d_A^2 + d_C^2) \log \frac{1}{\delta}\right). \end{aligned} \quad (75)$$

Here, $d = 2^{n_{ABC}}$.

To prove this theorem, we need the following result (Ref. [41], Sec. II.1) concerning the concentration of measurement on $\mathcal{V}_q(\mathbb{C}^p)$, the space of isometries $\mathbb{C}^q \rightarrow \mathbb{C}^p$.

Fact 4—If $f: \mathcal{V}_q(\mathbb{C}^p) \rightarrow \mathbb{R}$ is K -Lipschitz: $|f(V_1) - f(V_2)| \leq K \|V_1 - V_2\|_F$, where the norm is defined as $\|V_1 - V_2\|_F = \sqrt{\text{Tr}[(V_1 - V_2)^\dagger (V_1 - V_2)]}$, then there exist absolute constants $c_1, c_2 > 0$, independent of p and q , such that

$$\Pr[|f - \mathbb{E}(f)| > \eta] \leq c_1 \exp\left(-\frac{c_2 p \eta^2}{K^2}\right). \quad (76)$$

Proof of Theorem 4. Fixing unitaries U_A and U_C , define a function $f_{U_A U_C}: \mathcal{V}_q(\mathbb{C}^p) \rightarrow \mathbb{C}$ as follows:

$$(V: \mathcal{H}_C \rightarrow \mathcal{H}_{AB}) \mapsto f_{U_A U_C}(V) = \langle \text{EPR} | (V \otimes I)^\dagger U_A \otimes U_C^* (V \otimes I) | \text{EPR} \rangle. \quad (77)$$

Here, $q = 2^{n_C}$, and $p = 2^{n_{AB}}$. In the following, we sometimes omit the subscript. By standard Haar average computation, we have

$$\mathbb{E}(f(V)) = \frac{\text{Tr}(U_A) \text{Tr}(U_C^*)}{d_A d_C}. \quad (78)$$

We now establish the Lipschitz property of f . Denote $|\psi_i\rangle = (V_i \otimes I) | \text{EPR} \rangle$ ($i = 1, 2$). We have

$$\langle \psi_1 | \psi_2 \rangle = \frac{1}{2^{n_C}} \text{Tr}(V_1^\dagger V_2); \quad (79)$$

hence,

$$\begin{aligned} \|V_1 - V_2\|_F^2 &= \text{Tr}(2I_C - V_1^\dagger V_2 - V_2^\dagger V_1) \\ &= 2^{n_C+1}(1 - \text{Re}\langle \psi_1 | \psi_2 \rangle) \\ &\geq 2^{n_C+1}(1 - |\langle \psi_1 | \psi_2 \rangle|) \geq 2^{n_C} \sin^2 \theta(\psi_1, \psi_2). \end{aligned} \quad (80)$$

In the last step, we used an elementary inequality $1 - \cos \theta \geq \frac{1}{2} \sin^2 \theta$. Therefore, similar to Eq. (61), we have

$$|f(V_1) - f(V_2)| \leq \sin \theta(\psi_1, \psi_2) \leq 2^{-n_C/2} \|V_1 - V_2\|_F. \quad (81)$$

Therefore, f is $2^{-n_C/2}$ -Lipschitz with respect to the Frobenius norm. Applying Fact 4 to the real part of $f(V)$, we obtain

$$\begin{aligned} \mathbb{P}(\text{Re } f(V) - \text{Re } \mathbb{E}(f(V)) > \eta) &\leq c_1 \exp\left(-\frac{c_2 2^{n_{AB}} \eta^2}{(2^{-n_C/2})^2}\right) \\ &= c_1 \exp(-c_2 d \eta^2), \end{aligned} \quad (82)$$

where $d = 2^{n_{ABC}}$.

Now we pick ϵ -nets for $\text{U}(\mathcal{H}_A)$ [and $\text{U}(\mathcal{H}_C)$, respectively] using Fact 3, such that the cardinality is less than $(c/\epsilon)^{d_A}$ [and $(c/\epsilon)^{d_C}$, respectively]. Note that if $\|U_A - U'_A\|_\infty < \epsilon$ and $\|U_C - U'_C\|_\infty < \epsilon$, then

$$|f_{U_A U_C}(V) - f_{U'_A U'_C}(V)| < 2\epsilon, \quad \forall V \in \mathcal{V}_q(\mathbb{C}^p). \quad (83)$$

Taking the expectation over V , it also follows that

$$|\mathbb{E}(f_{U_A U_C}(V)) - \mathbb{E}(f_{U'_A U'_C}(V))| < 2\epsilon. \quad (84)$$

We now assume a given isometry V admits a logical operator that is fully supported on region A : There exist U_A and U_C such that $|\text{Tr}(U_C)|/d_C \leq w$ and $|f_{U_A U_C}(V)| \geq h^2$ as in Eq. (72). We can always assume without loss of generality that $f_{U_A U_C}(V) \geq h^2$ (otherwise, we can multiply U_A or U_C with a phase), which implies that

$$\begin{aligned} \text{Re } f_{U_A U_C}(V) - \text{Re } \mathbb{E}(f_{U_A U_C}(V)) \\ \geq h^2 - \left| \frac{\text{Tr}(U_A) \text{Tr}(U_C^*)}{d_A d_C} \right| \geq h^2 - w. \end{aligned} \quad (85)$$

This inequality, together with Eqs. (83) and (84), implies that there exist U'_A and U'_C in two ϵ -nets respectively such that

$$\text{Re } \Re f_{U'_A U'_C}(V) - \text{Re } \mathbb{E}(f_{U'_A U'_C}(V)) > h^2 - 4\epsilon - w. \quad (86)$$

Substituting it into Eq. (82) and applying the union bound over the ϵ -nets, we obtain

$$\begin{aligned} \mathbb{P}(V \text{ admits a logical operator on } A) &\leq 2c_1 \left(\frac{c}{\epsilon}\right)^{d_A^2 + d_C^2} \\ &\times \exp(-c_2 d (h^2 - 4\epsilon - w)^2). \end{aligned} \quad (87)$$

Then Theorem 4 is proved by choosing $\epsilon \propto h^2 - w$ with a small enough proportional constant, similar to the proof of Eq. (71). ■

B. Miscellaneous comments

We have established that, when $|\Psi_{ABC}\rangle$ is a Haar random state (or when $V: C \rightarrow AB$ is a random isometry), encoded logical qubits cannot be recovered from subsystems A or B assuming $n_A, n_B, n_C < \frac{1}{2}(n_A + n_B + n_C)$. On the contrary, when $|\Psi_{ABC}\rangle$ is a random stabilizer state (or when V is a random Clifford isometry), part of the encoded logical qubits can be recovered from subsystems A and B . Namely, let g_R be the number of independent nontrivial logical operators supported on a subsystem R ; the following relations are well known [16]:

$$\begin{aligned} g_A &= I(A: C) \approx n_A + n_C - n_B, \\ g_B &= I(B: C) \approx n_B + n_C - n_A, \quad g_A + g_B = 2k, \end{aligned} \quad (88)$$

where $k = n_C$ in our setting and the approximations hold when $n_R < (n/2)$ for $R = A, B, C$. This suggests that $I(A: C) \sim O(n)$ logical operators can be supported on A . In fact, when V is sampled randomly, it is likely that a given logical operator ℓ_A on A can find some other logical operator r_A on A that anticommutes with ℓ_A . As such, one can choose $(g_A/2) - o(1)$ pairs of mutually anticommuting

basis logical operators on A , suggesting that $(g_A/2) - o(1)$ logical qubits can be recovered from A .

In the above discussion, it is crucial to consider unitary logical operators $\overline{U}_C$. In fact, some *nonunitary* logical operators can be constructed on A or B . For instance, let us split C further into two subsystems $C = C_0 C_1$, where C_0 consists only of one qubit while C_1 consists of $n_C - 1$ qubits. Consider the following operator:

$$V_C = X_{C_0} \otimes |0\rangle\langle 0|_{C_1}, \quad (89)$$

where $|0\rangle\langle 0|_{C_1}$ is a projection operator acting on C_1 . Observe that $|0\rangle\langle 0|_{C_1}$ acting on a Haar random state $|\Psi_{ABC}\rangle$ effectively creates another Haar random state $|\Phi_{ABC_0}\rangle \propto |0\rangle\langle 0|_{C_1} |\Psi_{ABC}\rangle$ after an appropriate normalization. Then, finding a logical operator $\overline{V}_C$ for $|\Psi_{ABC}\rangle$ is equivalent to finding a logical operator $\overline{X}_{C_0}$ for $|\Phi_{ABC_0}\rangle$, reducing the problem to an EPR distillation in the projected state $|\Phi_{ABC_0}\rangle$. Then, if A contains more than half of ABC_0 , namely, $n_A > n_B + 1$, $\overline{X}_{C_0}$ can be supported on A even when no logical *unitary* operator can be supported on A .

We note that, in fact, Theorem 4 also holds for nonunitary operators with bounded operator norm. The inability to exclude nonunitary logical operators stems from the fidelity threshold h : For V_C in Eq. (89), we have $h^2 = O(1/d_C)$, rendering the bound in Theorem 4 vacuous when $n_A > n_B + 1$. Referring to Eq. (89) as a nonunitary logical operator implicitly assumes postselection on the $|0\rangle\langle 0|$ outcome, which effectively rescales the fidelity threshold by the success probability of the measurement.

It is worth noting that the reconstruction of a nonunitary logical operator $\overline{V}_C$ on A can be interpreted as (one-shot) LOCC entanglement distillation where one (i) performs projective measurements $\{|i\rangle\langle i|\}_{C_1}$ on C_1 , (ii) sends the measurement outcome i to A , and (iii) applies an appropriate LU on A to prepare an EPR pair between A and C_0 . See Ref. [32] for details.

V. GENERALIZATIONS

The method developed in this paper is general and admits broader applicability. In this section, we establish three theorems highlighting the tripartite entanglement structure of Haar random tripartite states. We omit some technical details in this section, as the proofs are similar to those in Secs. III and IV.

A. Nothing is LU-distillable

We have shown that bipartite EPR-like entanglement is essentially absent in tripartite Haar random states. This naturally raises a broader question: Can any other tripartite pure state be LU-distilled? For instance, could one distill GHZ-like or W-like entanglement?

We now show a much stronger no-go statement: Essentially no nontrivial tripartite pure state of smaller local dimension can be LU-distilled from a tripartite Haar random state. Let the tripartition be A, B, C , and decompose each subsystem as

$$A = A_1 \otimes A_2, \quad B = B_1 \otimes B_2, \quad C = C_1 \otimes C_2. \quad (90)$$

We ask whether there exists a local unitary $U = U_A \otimes U_B \otimes U_C$ and a pure state $|\phi\rangle_{A_1 B_1 C_1}$ such that

$$U|\psi\rangle_{ABC} = |\phi\rangle_{A_1 B_1 C_1} \otimes |\text{something}\rangle_{A_2 B_2 C_2}. \quad (91)$$

Without loss of generality, we may assume $1 \leq n_1 \leq n_2$ (here, $n_i = n_{A_i} + n_{B_i} + n_{C_i}$). If $n_1 \geq n_2$, we can consider an equivalent problem of outputting a pure state $|\phi'\rangle$ on $A_2 B_2 C_2$.

Theorem 5 (nothing is LU-distillable). Let $|\psi\rangle_{ABC}$ be a Haar random state with $n_A, n_B, n_C < n/2$. Then, the probability that there exists a tripartite pure state $|\phi\rangle_{A_1 B_1 C_1}$ which can be LU-distilled from $|\psi\rangle_{ABC}$ is $o(1)$ in the large- n limit.

Proof sketch. Consider the entanglement entropy of $|\psi\rangle$ with respect to the bipartition into $A_1 B_1 C_1$ and $A_2 B_2 C_2$, denoted as $S(\psi)$. If there exists a LU-distillable tripartite pure state $|\phi\rangle$ on $A_1 B_1 C_1$ [namely, Eq. (91) holds], then $S(U\psi) = 0$. Here, this entropy is across the $(A_1 B_1 C_1) : (A_2 B_2 C_2)$ cut. In the following, we relax this condition to $S(U\psi) < \delta$ and upper bound its probability.

Page's theorem shows that $\mathbb{E}_\psi[S(U\psi)] = \mathbb{E}_\psi[S(\psi)] \approx n_1$ for any U . Moreover, we have the following lemma regarding the Lipschitz continuity of the entanglement entropy.

Lemma 8. $S(\psi)$ is $O(n_1)$ -Lipschitz as a function of ψ .

Proof. Given $|\psi\rangle$ and $|\phi\rangle$, we consider the two-dimensional subspace spanned by them and define H as a rotation that rotates $|\psi\rangle$ to $|\phi\rangle$ in unit time. We have $\|H\|_\infty = \theta(\psi, \phi)$, where $\cos \theta(\psi, \phi) = |\langle \psi | \phi \rangle|$. Reference [45] shows that

$$\frac{dS(e^{-iHt}\psi)}{dt} \leq c \|H\|_\infty \log_2 \min\{d_1, d_2\}, \quad (92)$$

where $c < 1.9123$. Hence, $|S(\psi) - S(\phi)| \leq c n_1 \theta(\psi, \phi)$.

Levy's lemma then implies that $\mathbb{P}(S(\psi) < \delta)$ is doubly exponentially small when δ is bounded well below n_1 (say, $\delta \leq n_1 - c\sqrt{n_1}$ with some $c > 0$):

$$\mathbb{P}(S(\psi) < \delta) \lesssim \exp\left(-\frac{d}{n_1}\right) \lesssim \exp\left(-\frac{d}{n}\right), \quad d = 2^n. \quad (93)$$

Covering the space of LUs by an ϵ -net with size $\sim (1/\epsilon)^{d_A^2 + d_B^2 + d_C^2}$ and applying the union bound, we can

establish that, with overwhelming probability, no smaller tripartite pure state $|\phi\rangle_{A_1 B_1 C_1}$ can be LU-distilled. ■

This result shows that as n_A , n_B , and n_C increase (while remaining below $n/2$), a Haar random state represents a genuinely new form of tripartite entanglement that cannot be reduced, by local unitaries, to any smaller tripartite pure state. In this sense, tripartite Haar random states are maximally irreducible under LU distillation [46].

B. Chirality of wave function

Another interesting application of our method concerns the intrinsic chirality of wave functions. Conventionally, chirality has been characterized dynamically through responses to external perturbations, such as robust edge currents or quantized transport coefficients. Recently, there have been significant developments toward characterizing chirality directly from a single many-body wave function [47]. In this subsection, we focus on the characterization based on multipartite local unitaries [48]. More precisely, an m -partite pure state $|\psi\rangle_{R_1 \dots R_m}$ is said to be *m-partite nonchiral* if there exists a local unitary of the form

$$U = U_{R_1} \otimes \dots \otimes U_{R_m} \quad \text{such that } U|\psi\rangle = |\psi^*\rangle, \quad (94)$$

where $|\psi^*\rangle$ denotes the complex conjugate of $|\psi\rangle$ in the computational basis. Each component U_{R_j} may be arbitrary. If no such local unitary exists, the state is called *m-partite chiral*.

We now ask whether Haar random states are chiral in this sense. It is instructive to first consider the bipartite case. A Haar random state $|\psi\rangle_{AB}$ is not bipartite chiral. Indeed, by Page's theorem [Eq. (2)], it can be converted by bipartite local unitaries into a tensor product of approximate EPR pairs and a product state. Since each EPR pair may be chosen to be real, the resulting state can be mapped to its complex conjugate by further local unitaries. Thus, bipartite Haar random states are generically nonchiral.

The situation changes dramatically in the tripartite setting.

Theorem 6 (tripartite chirality). Let $|\psi\rangle_{ABC}$ be a Haar random state with $n_A, n_B, n_C < n/2$. Then, $|\psi\rangle_{ABC}$ is tripartite chiral with probability $1 - o(1)$ in the large- n limit.

Proof sketch. Suppose there exists a tripartite local unitary $U = U_A \otimes U_B \otimes U_C$ such that $|\psi^*\rangle = U|\psi\rangle$ or, equivalently,

$$\langle \psi^* | U | \psi \rangle = 1. \quad (95)$$

Define

$$f_U(\psi) := \langle \psi^* | U | \psi \rangle. \quad (96)$$

One verifies that f_U is $O(1)$ -Lipschitz on the unit sphere S^{2d-1} .

We next evaluate the Haar average of $f_U(\psi)$. The Haar measure is invariant under global phase rotations $|\psi\rangle \mapsto e^{i\theta}|\psi\rangle$. Under this transformation, we have

$$f_U(\psi) \mapsto e^{i2\theta} f_U(\psi). \quad (97)$$

Hence,

$$\mathbb{E}_\psi[f_U(\psi)] = e^{i2\theta} \mathbb{E}_\psi[f_U(\psi)] \quad (98)$$

for all θ , which implies

$$\mathbb{E}_\psi[f_U(\psi)] = 0. \quad (99)$$

Applying Levy's lemma, we find, for any fixed U , $\mathbb{P}(f_U(\psi) \approx 1)$ is doubly exponentially small in n . Therefore, using the ϵ -net and the union bound argument, we can establish that, with overwhelming probability, no tripartite local unitary maps $|\psi\rangle$ to $|\psi^*\rangle$. ■

C. Absence of global symmetry

The final application our method concerning the absence of global symmetry. Conventionally, a global symmetry is characterized as a tensor-product representation of a group acting on local degrees of freedom. A prototypical example is a global $\mathbb{Z}_2$ symmetry on qubits generated by $\bigotimes_j X_j$, which implements a simultaneous spin flip on every site and corresponds to conservation of total $\mathbb{Z}_2$ charge.

More generally, one may characterize a global symmetry of a pure state as invariance under a tensor-product unitary. Namely, given an m -partite state $|\psi\rangle_{R_1 \dots R_m}$, we say that it possesses a (possibly approximate) global symmetry if there exists a tensor-product unitary

$$U = U_{R_1} \otimes \dots \otimes U_{R_m} \quad \text{such that } U|\psi\rangle \approx |\psi\rangle, \quad (100)$$

where at least one U_{R_j} is nontrivial (not proportional to the identity) in a sense of Eq. (72). This is a broad characterization, as we do not require the operators to form a finite group, nor do we require exact equality. Nevertheless, this formulation contains the conventional notion of (exact or approximate) global symmetry under finite groups as a special case.

A natural question is whether Haar random states admit any such (approximate) tensor-product symmetry. The bipartite case again provides useful intuition. A Haar random state $|\psi\rangle_{AB}$ can be brought, by local unitaries, into a tensor product of approximate EPR pairs and a product state. Since each EPR pair is invariant under nontrivial operators such as $X \otimes X$ and $Z \otimes Z$, conjugating these symmetries back by the inverse local unitaries produces nontrivial tensor-product operators $U_A \otimes U_B$ that leave $|\psi\rangle_{AB}$ invariant. Hence, bipartite Haar random states generically possess bipartite tensor-product symmetries.

In contrast, this structure disappears in the tripartite setting.

Theorem 7 (absence of global symmetry). Let $|\psi\rangle_{ABC}$ be a Haar random state with $n_A, n_B, n_C < n/2$. Then, with probability $1 - o(1)$ in the large- n limit, there does not exist a tensor-product unitary

$$U = U_A \otimes U_B \otimes U_C, \quad \text{such that } U|\psi\rangle = |\psi\rangle, \quad (101)$$

with at least one U_R being nontrivial.

Here and in Theorem 8, nontriviality is defined as in the second inequality of Eq. (72).

Proof sketch. Suppose such a tensor-product unitary exists. Then, $\langle\psi|U|\psi\rangle = 1$. Define

$$g_U(\psi) := \langle\psi|U|\psi\rangle, \quad (102)$$

where g_U is $O(1)$ -Lipschitz on the unit sphere. We first compute its Haar average:

$$\mathbb{E}_\psi[g_U(\psi)] = \frac{\text{Tr}(U)}{d}, \quad (103)$$

where $d = 2^n$. If at least one of U_A , U_B , and U_C is nontrivial, then $|\text{Tr}(U)| \leq wd$ and, hence,

$$\mathbb{E}_\psi[g_U(\psi)] = w < 1. \quad (104)$$

Applying Levy's lemma, we find, for fixed U , $\mathbb{P}(g_U(\psi) \approx 1)$ is doubly exponentially small. Covering the space of tensor-product unitaries by an ϵ -net and applying the union bound, we can establish that, with overwhelming probability, $|\psi\rangle$ does not have nontrivial tripartite global symmetry. ■

An analogous statement holds for Haar random isometries, which extends Theorem 4.

Theorem 8 (absence of tensor-product logical operator). Let $V: C \rightarrow AB$ be a Haar random isometry with $n_C < n_A + n_B$ and $|n_A - n_B| < n_C$. Then, with overwhelming probability, V does not admit a nontrivial logical operator of the form $U_A \otimes U_B$.

The proof proceeds in the same manner as Theorem 4, by considering the concentration of

$$\text{Tr}[V^\dagger(U_A \otimes U_B)VU_C^\dagger] \quad (105)$$

over the Haar measure and applying an ϵ -net argument.

VI. HOLOGRAPHY

In this section, we illustrate two particular applications of our results in the AdS/CFT correspondence. The first question concerns the presence or absence of bipartite entanglement in holographic mixed states. The second question concerns whether the converse of entanglement wedge reconstruction holds or not. Both questions have remained unresolved and led to important conceptual puzzles. While we keep the presentation of this section minimal, we encourage readers to refer to Ref. [32] for

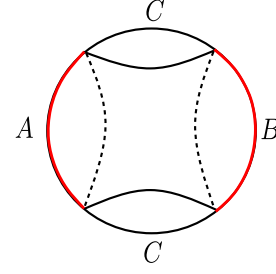


FIG. 4. Connected entanglement wedge. Here, the minimal area surface of AB is given by geodesics shown in solid lines, leading to a large mutual information $I(A:B)$. Do two subsystems A and B contain EPR-like entanglement?

detailed discussions and background of these problems in the AdS/CFT correspondence.

A. Entanglement distillation

In the AdS/CFT correspondence, entanglement entropy S_A of a boundary subsystem A is given by the Ryu-Takayanagi (RT) formula

$$S_A = \frac{1}{4G_N} \min_{\gamma_A} \text{Area}(\gamma_A) + \dots \quad (106)$$

at the leading order in $1/G_N$ for static geometries, where γ_A is a bulk surface homologous to A in the asymptotically AdS spacetime and G_N is the Newton's constant. (Recall that G_N is a very small constant.) This formula predicts that two boundary subsystems A and B can have large mutual information even when they are spatially disconnected on the boundary with a separating subsystem C :

$$I(A:B) \equiv S_A + S_B - S_{AB} = O(1/G_N) \quad (107)$$

when the minimal surface γ_{AB} extends into the bulk and connects A and B with a *connected entanglement wedge*. [49] A prototypical example is depicted in Fig. 4 for the AdS₃/CFT₂.

The entanglement structure in ρ_{AB} remains mysterious. For one thing, the mutual information is sensitive to classical correlations such as those in the GHZ state. Evidence from quantum gravity thought experiments and toy models [17,50,51] suggests that correlations in ρ_{AB} in holography are not of classical nature at the leading order in $1/G_N$. Furthermore, a previous work [24] showed that $|\Psi_{ABC}\rangle$ in holography must contain some tripartite entanglement at the leading order in $1/G_N$.

In Ref. [32], we proposed that a holographic mixed state ρ_{AB} does not contain bipartite entanglement when two individual minimal surfaces γ_A and γ_B do not overlap in the bulk. Namely, a connected entanglement wedge does not necessarily imply EPR-like bipartite entanglement. The present paper provides supporting evidence for our proposal. Recall that a Haar random state serves as a minimal toy model of holography. We have $S_A \approx \min(n_A, n - n_A)$ at

the leading order in n , which can be interpreted as the RT-like formula with the area (equals the total number of qubits across the cut) minimization by employing tensor diagrams:

$$S_A \approx \min \left(A \text{---} \boxed{\psi} \text{---} B, A \text{---} \boxed{\psi} \text{---} B \right). \quad (108)$$

$$S_{AB} \approx A \text{---} \boxed{\psi} \text{---} B = n_C,$$

Our result shows that no EPR-like entanglement is contained between A and B even when they have a “connected entanglement wedge.”

While our work provides insights into entanglement properties of holographic mixed states, we hasten to emphasize that whether ρ_{AB} in the real holography possesses EPR-like entanglement or not remains open. It is important to recall that Haar random states (and their networks) reproduce entanglement properties of fixed area states, where quantum fluctuations of area operators are strongly suppressed due to the flat spectrum. Real holographic states have subleading fluctuations that may potentially lead to bipartite entanglement. Also, bipartite entanglement in bulk matter fields can contribute to subleading bipartite entanglement in the boundary.

Our result on tripartite Haar random states is formulated for a finite-dimensional Hilbert space with a tensor product structure $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$. However, many physically relevant systems, particularly those arising in quantum field theory (QFT), are infinite-dimensional. Furthermore, in relativistic QFT, the Hilbert space does not literally factorize into a tensor product for spatially adjacent regions due to the Reeh-Schlieder property and the type-III nature of local von Neumann algebras. This naturally leads to the question of how our results should be interpreted in the context of continuum or “nonqubit” systems.

In continuum theories, bipartite entanglement can be unbounded or even divergent. For example, the entanglement entropy of a finite spatial region A in relativistic QFT diverges, and two complementary regions A and $B = A^c$ may share an infinite amount of entanglement. Such divergences are typically regulated using a UV cutoff (e.g., a lattice spacing or energy cutoff). The resulting entanglement is dominated by short-distance modes near

When $n_A, n_B, n_C < (n/2)$, we have

$$S_R \approx n_R \quad (R = A, B, C), \quad (109)$$

where the minimal surface γ_R does not contain the tensor at the center. This mimics the situation with a connected entanglement wedge as in Fig. 4. Namely, by splitting C into two subsystems, the minimal surface of AB can be schematically depicted as follows:

$$I(A : B) \approx n_A + n_B - n_C \sim O(n). \quad (110)$$

the entangling surface and is accordingly referred to as UV entanglement. This contribution indeed contains large amounts of bipartite, EPR-like entanglement between neighboring subregions.

At first sight, this conventional picture of QFT entanglement seems to be at odds with our result, which asserts the absence of EPR-like pairs. A natural interpretation is that the Haar random state result concerns the IR, long-distance structure of entanglement, after coarse-graining the UV correlations. In this sense, the Haar random state may be viewed as a toy model for the entanglement structure that remains after all UV, short-range bipartite correlations have been removed.

This can be seen clearly in the holographic entanglement structure illustrated in Fig. 4. In that geometry, neighboring regions A and C , or B and C , share large amounts of UV bipartite entanglement. By contrast, spatially separated regions A and B do not exhibit UV-divergent entanglement, although their total entanglement remains parametrically large, of the order of $1/G_N$. The absence of UV contributions is reflected in the fact that the mutual information $I(A : B)$ is independent of the cutoff scale. Equivalently, the minimal surfaces associated with A and B are spatially separated in the bulk, ensuring that short-distance modes near the boundary do not contribute to the A - B correlation. The resulting entanglement is, therefore, purely IR in nature. This IR contribution is encoded geometrically in the central bulk region bounded by the minimal surfaces of A , B , and C .

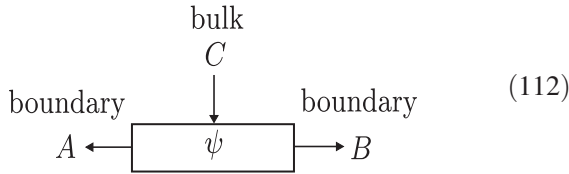
B. Entanglement wedge reconstruction (and its converse)

In the AdS/CFT correspondence, the physics of bulk quantum gravity is holographically encoded into boundary

quantum systems like a quantum error-correcting code. The conceptual pillar behind this interpretation is *entanglement wedge reconstruction* [52] which asserts that, if a bulk operator ϕ lies inside the entanglement wedge $\mathcal{E}_A$ of a boundary subsystem A , then it can be expressed as some boundary operator O_A which is supported exclusively on A :

$$\phi \text{ can be reconstructed on } A \Leftarrow \phi \text{ is inside } \mathcal{E}_A. \quad (111)$$

While the microscopic mechanisms of the bulk reconstruction still remain somewhat mysterious, random tensor network toy models can provide crucial insights on how the bulk operators may be reconstructed on the boundary subsystems. Let us illustrate the idea by using a Haar random state $|\Psi\rangle$ as a minimal toy model. Let us first assume that the bulk consists only of one qubit, which is encoded into $n - 1$ boundary qubits by viewing an n -qubit Haar random state $|\Psi\rangle$ as an encoding isometry $1 \rightarrow n - 1$ as schematically shown below:



where the bulk qubit is denoted by C and the boundary qubits are partitioned into AB .

In this toy model, the question of whether the bulk unitary operator U_C can be reconstructed on a subsystem A is equivalent to whether a logical unitary operator $\overline{U}_C$ can be supported on A in the $C \rightarrow AB$ quantum error-correcting code. It is well known that A supports a nontrivial logical operator $\overline{U}_C$ when $n_A > (n/2)$ (and does not support one when $n_A < (n/2)$) [53].

This standard result on Haar encoding can be understood as entanglement wedge reconstruction. Recall that, for static cases in the AdS/CFT correspondence, the entanglement wedge is computed by minimizing the generalized entropy

$$S_A = \min_{\gamma_A} \frac{\text{Area}(\gamma_A)}{4G_N} + S_{\text{bulk}}, \quad (113)$$

where S_{bulk} is a bulk entropy on a subregion surrounded by γ_A . When A occupies more than half of the total system, we have

$$S_A = \begin{array}{c} n_B + 1 \\ C \\ \downarrow \\ A \leftarrow \boxed{\psi} \rightarrow B \end{array} \quad (114)$$

where the bulk qubit C is inside the $\mathcal{E}_A$, suggesting its recoverability on A . Here, $+1$ comes from $S_{\text{bulk}} = S_C = 1$.

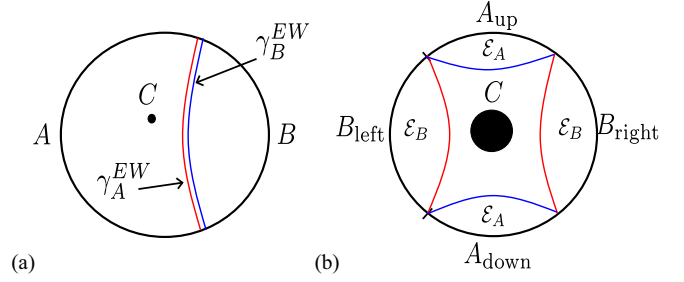


FIG. 5. Entanglement wedge reconstruction. (a) When C carries subleading entropy, the minimal surfaces γ_A^{EW} and γ_B^{EW} coincide. Hence, C can be reconstructed on one and only one subsystem A or B . (b) Here, $A = A_{\text{up}} \cup A_{\text{down}}$ and $B = B_{\text{left}} \cup B_{\text{right}}$. When C carries leading-order entropy, the minimal surfaces γ_A^{EW} and γ_B^{EW} do not necessarily coincide, and C may not be contained inside $\mathcal{E}_A$ or $\mathcal{E}_B$. Can A or B reconstruct C ?

On the other hand, when A occupies less than half of the total system, the minimal surface is given by

$$S_A = \begin{array}{c} n_A \\ C \\ \downarrow \\ A \leftarrow \boxed{\psi} \rightarrow B \end{array} \quad (115)$$

where the bulk qubit is outside the $\mathcal{E}_A$. Instead, in this case, the bulk qubit is inside $\mathcal{E}_B$, suggesting the recoverability on B . Hence, the bulk information about C can be recovered from either A or B (unless the sizes of A and B match exactly).

An analogous setup can be considered in the AdS₃/CFT₂ correspondence as depicted in Fig. 5(a). Here, observe that when the bulk degree of freedom C carries a subleading entropies $[O(1)]$ or, more generally, $O(1/G_N^a)$ with $0 < a < 1$, the minimal surface γ_A^{EW} for defining the entanglement wedge $\mathcal{E}_A$ matches with the minimal area (Ryu-Takayanagi) surface γ_A^{RT} :

$$\gamma_A^{\text{EW}} \approx \gamma_A^{\text{RT}} \text{ at leading order in } 1/G_N. \quad (116)$$

Hence, if the bulk d.o.f. C is enclosed by γ_A^{RT} , an operator acting on C can be reconstructed on A . On the other hand, if C lies outside γ_A^{RT} , it will be enclosed by γ_B^{RT} since $\gamma_A^{\text{RT}} = \gamma_B^{\text{RT}}$ (unless we fine-tune the sizes of A and B so that there are multiple minimal area surfaces). Thus, an operator acting on C can be reconstructed on B . Recalling the no-cloning theorem, this also implies that an operator on C cannot be reconstructed on A if C lies outside γ_A^{RT} . (Otherwise, quantum information encoded on C could be reconstructed on both A and B , creating two copies of C .) Hence, unless the bulk d.o.f. C lies exactly at the minimal surface γ_A^{EW} , it can always be reconstructed on *one and only one* subsystem, A or B .

In the above arguments for a Haar random state as well as holography, we have observed that, when the bulk C carries a subleading entropy, entanglement wedge reconstruction is an *if and only if* statement at the leading order in $1/G_N$ (or n):

$$\phi \text{ can be reconstructed on } A \Leftrightarrow \phi \text{ is inside } \mathcal{E}_A \quad (\text{when } S_{\text{bulk}} \text{ is subleading}). \quad (117)$$

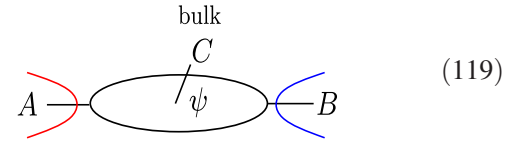
A naturally arising question concerns whether this remains the case when C carries a leading-order entropy. Indeed, it is worth emphasizing that the original entanglement wedge reconstruction Eq. (111) is an *if* statement, implying that a bulk operator can be reconstructed on A *if* it is contained inside $\mathcal{E}_A$. Some of previous works shows that the converse statement also holds if the complementary recovery is possible [55]. However, the complementary recovery generally fails when the bulk region C carries leading-order entropies. In such cases, physical or analytical evidence supporting the converse statement has thus far been lacking. In other words, whether the converse statement holds or not remains unclear:

$$\phi \text{ can be reconstructed on } A \stackrel{?}{\Rightarrow} \phi \text{ is inside } \mathcal{E}_A. \quad (118)$$

When the bulk C is subleading, we were able to promote it to an *if and only if* statement since $\gamma_A^{\text{EW}} = \gamma_B^{\text{EW}}$ at the leading order. However, this may fail when C is not subleading.

One particular holographic setup, highlighting this subtlety, was considered in Ref. [56] [Fig. 5(b)]. Here, the boundary is divided into four segments of roughly equal sizes and organized into A and B . In the absence of the bulk d.o.f., the minimal surfaces of A and B would be the same and given by the smaller of the two geodesic lines, colored in red and blue in Fig. 5(b). However, when $S_C = O(1/G_N)$, the minimal surface locations may change at the leading order as $S_{\text{bulk}} = S_C$ needs to be included in the evaluation of the generalized entropy. In particular, we can observe that $\gamma_A^{\text{EW}} \neq \gamma_B^{\text{EW}}$ at the leading order when the length difference for the red and blue geodesics, multiplied by $1/4G_N$, is smaller than S_C [Fig. 5(b)]. In this case, the bulk C lies outside $\mathcal{E}_A$ or $\mathcal{E}_B$ [57].

The key question is whether the bulk C is recoverable from A or B . The converse of entanglement wedge reconstruction would say no, since C lies outside $\mathcal{E}_A$ or $\mathcal{E}_B$. However, we can observe that $I(C:A), I(C:B) = O(1/G_N)$ in the Choi state, which suggests the presence of leading-order correlations between the input C and the output subsystems A and B . Our result on logical operators supports the converse of entanglement wedge reconstruction as a Haar random encoding with $n_A, n_B, n_C < (n/2)$ mimics the situation considered in Fig. 5(b):



where $\mathcal{E}_A$ and $\mathcal{E}_B$ do not contain C . Theorem 4 from Sec. IV then suggests that a nontrivial logical unitary operator U_C cannot be reconstructed on either A or B .

In summary, our result suggests that the converse of entanglement wedge reconstruction is true even when the bulk d.o.f. carries a leading-order entropy. One interesting consequence is that there can be a bulk region whose information cannot be reconstructed on either A or B . In Ref. [32], such a bulk region was referred to as shadow of entanglement wedge. This is in strong contrast with random stabilizer states, where the bulk information C can be recovered from either A or B at the leading order. Note that it is essential to restrict to unitary logical operators in the above argument, as discussed in Sec. IV.

VII. OUTLOOK

In this paper, we showed that a Haar random state does not contain bipartite entanglement under nontrivial tripartition satisfying $n_A, n_B, n_C < (n/2)$ at the limit of large n . In the quantum error-correction picture ($C \rightarrow AB$), this implies that neither subsystem A nor B can support any logical operator if $n_C < n_A + n_B$ and $|n_A - n_B| < n_C$. We also discussed two particular applications of our results in the AdS/CFT correspondence: one about the presence or absence of bipartite entanglement in holographic mixed states and the other about the converse of entanglement wedge reconstruction and the possible extensive bulk region that cannot be reconstructed from bipartite subsystems.

A natural next step is to consider more structured toy models of holography, such as Haar random tensor networks. While Haar random states serve as minimal toy models, tensor networks incorporate geometric locality and are, therefore, closer in spirit to AdS/CFT. An important question is whether the absence of bipartite distillable entanglement persists once such geometric structures are introduced. Technically, this would require understanding whether concentration-of-measure arguments, underlying our results, remain valid in random tensor network ensembles.

The possible breakdown of complementary recovery in holography may also have important implications. Our result predicts the existence of a semiclassical region that is not contained in the entanglement wedge of any bipartite boundary subsystem. As an example, in an accompanying work we proposed that the breakdown of complementary recovery for Haar random isometries provides a possible mechanism for the emergence of semiclassical closed baby universes [58]. We expect that this perspective may prove

useful in understanding the black hole interior and the role of black hole singularities, where Haar random (or pseudorandom) ideas have already played an important role.

Another important direction is to determine how much randomness is actually required for our conclusions to hold. Haar randomness is an extreme assumption, and physically relevant states may only approximate it. It would be interesting to investigate whether similar suppression results hold for states drawn from less random ensembles, such as approximate unitary k designs or T -doped Clifford circuits. Technically, this amounts to asking whether the concentration phenomena underlying our proofs persist under weaker pseudorandomness assumptions and to identifying the minimal degree of scrambling necessary to eliminate bipartite distillable entanglement.

Furthermore, our analysis places no restriction on the computational complexity of the local operations used in the distillation protocol. In practice, however, implementable operations are constrained by circuit depth and gate count. It would, therefore, be interesting to study entanglement distillation under bounded-complexity restrictions and to understand how computational constraints modify the suppression behavior.

ACKNOWLEDGMENTS

We thank the anonymous referees for their stimulating comments. Research at Perimeter Institute is supported in part by the Government of Canada through the Department of Innovation, Science and Economic Development and by the Province of Ontario through the Ministry of Colleges and Universities. This work is supported by JSPS KAKENHI Grants No. 23KJ1154 and No. 24K17047. This work is supported by the Applied Quantum Computing Challenge Program at the National Research Council of Canada.

DATA AVAILABILITY

There are no publicly available research data or software supporting this manuscript. Requests for further information or data should be sent to the authors.

-
- [1] M. Srednicki, *Chaos and quantum thermalization*, *Phys. Rev. E* **50**, 888 (1994).
 - [2] P. Hayden and J. Preskill, *Black holes as mirrors: Quantum information in random subsystems*, *J. High Energy Phys.* **09** (2007) 120.
 - [3] M. L. Mehta, *Random Matrices* (Elsevier, New York, 2004), Vol. 142.
 - [4] D. N. Page, *Average entropy of a subsystem*, *Phys. Rev. Lett.* **71**, 1291 (1993).
 - [5] P. Hosur, X.-L. Qi, D. A. Roberts, and B. Yoshida, *Chaos in quantum channels*, *J. High Energy Phys.* **02** (2016) 004.
 - [6] J. S. Cotler, G. Gur-Ari, M. Hanada, J. Polchinski, P. Saad, S. H. Shenker *et al.*, *Black holes and random matrices*, *J. High Energy Phys.* **05** (2017) 118.
 - [7] F. Pastawski, B. Yoshida, D. Harlow, and J. Preskill, *Holographic quantum error-correcting codes: Toy models for the bulk/boundary correspondence*, *J. High Energy Phys.* **06** (2015) 149.
 - [8] P. Hayden, S. Nezami, X.-L. Qi, N. Thomas, M. Walter, and Z. Yang, *Holographic duality from random tensor networks*, *J. High Energy Phys.* **11** (2016) 009.
 - [9] A. Ambainis and J. Emerson, *Quantum t -designs: t -wise independence in the quantum world*, in *Proceedings of the Twenty-Second Annual IEEE Conference on Computational Complexity (CCC'07)* (IEEE, New York, 2007), p. 129.
 - [10] E. Lubkin, *Entropy of an n -system from its correlation with a k -reservoir*, *J. Math. Phys. (N.Y.)* **19**, 1028 (1978).
 - [11] S. Lloyd and H. Pagels, *Complexity as thermodynamic depth*, *Ann. Phys. (N.Y.)* **188**, 186 (1988).
 - [12] Choosing a Haar random state means that one picks a quantum state uniformly at random from the set of all the n -qubit pure states. In particular, a Haar measure can be characterized as a unique probability distribution that is left and right invariant under any unitary operators.
 - [13] B. Yoshida and A. Kitaev, *Efficient decoding for the Hayden-Preskill protocol*, [arXiv:1710.03363](https://arxiv.org/abs/1710.03363).
 - [14] T.-C. Lu and T. Grover, *Entanglement transitions as a probe of quasiparticles and quantum thermalization*, *Phys. Rev. B* **102**, 235110 (2020).
 - [15] H. Shapourian, S. Liu, J. Kudler-Flam, and A. Vishwanath, *Entanglement negativity spectrum of random mixed states: A diagrammatic approach*, *PRX Quantum* **2**, 030347 (2021).
 - [16] B. Yoshida and I. L. Chuang, *Framework for classifying logical operators in stabilizer codes*, *Phys. Rev. A* **81**, 052302 (2010).
 - [17] S. Nezami and M. Walter, *Multipartite entanglement in stabilizer tensor networks*, *Phys. Rev. Lett.* **125**, 241602 (2020).
 - [18] G. Smith and D. Leung, *Typical entanglement of stabilizer states*, *Phys. Rev. A* **74**, 062314 (2006).
 - [19] M. Freedman and M. Headrick, *Bit threads and holographic entanglement*, *Commun. Math. Phys.* **352**, 407 (2017).
 - [20] S. X. Cui, P. Hayden, T. He, M. Headrick, B. Stoica, and M. Walter, *Bit threads and holographic monogamy*, *Commun. Math. Phys.* **376**, 609 (2019).
 - [21] C. A. Agón, J. De Boer, and J. F. Pedraza, *Geometric aspects of holographic bit threads*, *J. High Energy Phys.* **05** (2019) 075.
 - [22] J. Harper and M. Headrick, *Bit threads and holographic entanglement of purification*, *J. High Energy Phys.* **08** (2019) 101.
 - [23] N. Bao and G. Suer, *Holographic entanglement distillation from the surface state correspondence*, *J. High Energy Phys.* **01** (2024) 091.
 - [24] C. Akers and P. Rath, *Entanglement wedge cross sections require tripartite entanglement*, *J. High Energy Phys.* **04** (2020) 208.
 - [25] C. Akers, T. Faulkner, S. Lin, and P. Rath, *The page curve for reflected entropy*, *J. High Energy Phys.* **06** (2022) 089.

- [26] C. Akers, T. Faulkner, S. Lin, and P. Rath, *Reflected entropy in random tensor networks. Part II. A topological index from canonical purification*, *J. High Energy Phys.* **01** (2023) 067.
- [27] C. Akers, T. Faulkner, S. Lin, and P. Rath, *Reflected entropy in random tensor networks. Part III. Triway cuts*, *J. High Energy Phys.* **12** (2024) 209.
- [28] The inequalities here are informal, serving as the counterpart of the $n_R < n/2$ ($R = A, B, C$) condition in the results for tripartite Haar random states. Precise conditions may be obtained via Theorem 4 following the discussions below Theorems 1 and 2.
- [29] S. Bravyi and B. Terhal, *A no-go theorem for a two-dimensional self-correcting quantum memory based on stabilizer codes*, *New J. Phys.* **11**, 043029 (2009).
- [30] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, *Mixed state entanglement and quantum error correction*, *Phys. Rev. A* **54**, 3824 (1996).
- [31] I. Devetak and A. Winter, *Distillation of secret key and entanglement from quantum states*, *Proc. R. Soc. A* **461**, 207 (2005).
- [32] T. Mori and B. Yoshida, *Does connected wedge imply distillable entanglement?*, *J. High Energy Phys.* **01** (2026) 125.
- [33] P. Hayden, D. W. Leung, and A. Winter, *Aspects of generic entanglement*, *Commun. Math. Phys.* **265**, 95 (2006).
- [34] P. Hayden, D. Leung, P. W. Shor, and A. Winter, *Randomizing quantum states: Constructions and applications*, *Commun. Math. Phys.* **250**, 371 (2004).
- [35] Here and throughout this paper, we consider pure states with phases, so that we work with the sphere S^{2d-1} rather than the complex projective space $\mathbb{C}P^{d-1}$. Two formulations are equivalent due to the global phase-rotation symmetry.
- [36] K. Böröczky and G. Wintsche, *Covering the sphere by equal spherical balls*, in *Discrete and Computational Geometry: The Goodman-Pollack Festschrift*, edited by B. Aronov, S. Basu, J. Pach, and M. Sharir (Springer, Berlin, Heidelberg, 2003), pp. 235–251.
- [37] O. Marchal and J. Arbel, *On the sub-Gaussianity of the beta and dirichlet distributions*, *Electron. Commun. Probab.* **22** (2017).
- [38] Here, we omit the $O(\log d)$ term, which is unfortunately unavoidable in this proof due to the $p \log p$ term in Sec. II B. However, it can be avoided if we prove Theorem 1 via the argument used for Theorem 2. Moreover, in the regimes of interest—when $\delta = \Theta(1)$ and/or $\delta \gg (d_A + d_B) \times \sqrt{[(\log d)/d]}$ —the additional $O(\log d)$ term does not affect the asymptotic behavior.
- [39] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, England, 2010).
- [40] When constructing V_A , we have $d_1 = d_A$ and $d_2 = 2^m$, and we denote $d'_A = d_2 d_3$. Similar comments apply to B .
- [41] M. Ledoux, *The Concentration of Measure Phenomenon*, Mathematical Surveys and Monographs Vol. 89 (American Mathematical Society, Providence, 2001).
- [42] G. Aubrun, J. Jenkinson, and S. J. Szarek, *Optimal constants in concentration inequalities on the sphere and in the Gauss space*, in *High Dimensional Probability X* (Birkhäuser, Cham, 2026).
- [43] S. J. Szarek, *Metric entropy of homogeneous spaces*, [arXiv:math/9701213](https://arxiv.org/abs/math/9701213).
- [44] We may simply consider ϵ -nets for $U(d')$, since an isometry can be (nonuniquely) extended to a unitary. Then, Theorem 2 still holds with 2^{2m} replaced by 2^{4m} in the subleading term. To obtain a tighter bound on the subleading term, we may also quotient out another $U(\sqrt{dd'})$, where dd' is the dimension of the subsystem being traced out. This would give us the dimension $dd' - d^2$, which merely improves a constant factor before 2^{2m} .
- [45] S. Bravyi, *Upper bounds on entangling rates of bipartite Hamiltonians*, *Phys. Rev. A* **76**, 052319 (2007).
- [46] While this result concerns LU distillation of pure states supported on smaller tripartite subsystems, it also has implications for the LU distillation of certain target states supported on the original tripartite systems. For example, it implies that the probability for a tripartite state to be LU equivalent to a GHZ state vanishes, since a GHZ state over ABC is the tensor product of two GHZ states over $A_1 B_1 C_1$ and $A_2 B_2 C_2$ (the same conclusion also follows from a direct counting argument).
- [47] I. H. Kim, B. Shi, K. Kato, and V. V. Albert, *Chiral central charge from a single bulk wave function*, *Phys. Rev. Lett.* **128**, 176402 (2022).
- [48] S. Vardhan, B. Shi, I. H. Kim, and Y. Zou, *Chirality, magic, and quantum correlations in multipartite quantum states*, *SciPost Phys.* **20**, 066 (2026).
- [49] At the leading order in G_N , the entanglement wedge of a boundary subregion R is defined as a bulk subregion enclosed by the minimal surface γ_R together with the AdS boundary. See Fig. 5(a) for its illustration.
- [50] L. Susskind, *ER=EPR, GHZ, and the consistency of quantum measurements*, *Fortschr. Phys.* **64**, 72 (2016).
- [51] X. Dong, X.-L. Qi, and M. Walter, *Holographic entanglement negativity and replica symmetry breaking*, *J. High Energy Phys.* **06** (2021) 024.
- [52] A. Almheiri, X. Dong, and D. Harlow, *Bulk locality and quantum error correction in AdS/CFT*, *J. High Energy Phys.* **04** (2015) 163.
- [53] This can be done by applying the Petz recovery map [54].
- [54] H. Barnum and E. Knill, *Reversing quantum dynamics with near-optimal quantum and classical fidelity*, *J. Math. Phys.* (N.Y.) **43**, 2097 (2002).
- [55] C. Akers and G. Penington, *Quantum minimal surfaces from quantum error correction*, *SciPost Phys.* **12**, 157 (2022).
- [56] C. Akers and G. Penington, *Leading order corrections to the quantum extremal surface prescription*, *J. High Energy Phys.* **04** (2021) 062.
- [57] The reason why we consider a partition into four segments is that it leaves a sufficiently large bulk region (AdS size) where a bulk d.o.f. with $1/G_N$ entropy can be placed without worrying about backreaction. One may replace C with a small (sub-AdS size) black hole or a conical singularity in order to explicitly account for backreaction. We also emphasize that this setup does not require fine-tuning of the sizes of A and B as long as the condition on the two geodesic lengths is met.
- [58] T. Mori and B. Yoshida, *Baby universe as logical qubits: Information recovery in random encoding*, [arXiv:2511.20747](https://arxiv.org/abs/2511.20747).