Detecting Topological Order at Finite Temperature Using Entanglement Negativity

Tsung-Cheng Lu[1], Timothy H. Hsieh,[2] and Tarun Grover[1]

[1]*Department of Physics, University of California at San Diego, La Jolla, California 92093, USA*
[2]*Perimeter Institute for Theoretical Physics, Waterloo, Ontario N2L 2Y5, Canada*

We propose a diagnostic for finite temperature topological order using "topological entanglement negativity," the long-range component of a mixed-state entanglement measure. As a demonstration, we study the toric code model in d spatial dimensions for $d = 2,3,4$, and find that when topological order survives thermal fluctuations, it possesses a nonzero topological entanglement negativity, whose value is equal to the topological entanglement entropy at zero temperature. Furthermore, we show that the Gibbs state of 2D and 3D toric code at any nonzero temperature, and that of 4D toric code above a certain critical temperature, can be expressed as a convex combination of short-range entangled pure states, consistent with the absence of topological order.

Strongly interacting quantum many-body systems at zero temperature can exhibit exotic order beyond Laudau-Ginzburg paradigm, dubbed topological order, whose defining property is that the ground state degeneracy depends on the topology of the space [1–3]. While the theory of topological order in ground states (i.e., zero temperature) is well developed, our understanding for topological order at finite temperature is less clear. In particular, the pursuit of a model supporting a stable topological order at finite temperature has been difficult since typically topological order is fragile against thermal fluctuations [4–8]. The most well-known model exhibiting finite-T topological order is the toric code model in four spatial dimensions [9,10], and it remains unclear whether such a model exists below four dimensions. Apart from being a fundamental question in many-body physics, a stable finite-T topological order also has profound implications for quantum computing since it serves as a stable self-correcting quantum memory (encoded information is protected against thermal decoherence) [9,11].

Even for models supporting finite-T topological order, it is not obvious how to define an appropriate nonlocal order parameter. Hastings defined topological order at finite T by the requirement that the corresponding thermal density matrix cannot be connected to a separable mixed state via a finite-depth quantum channel [12]. While this provides a precise operational definition, it remains desirable to have a computable order parameter for finite-T topological order, analogous to the characterization of ground state topological order using topological entanglement entropy [13–15].

Previous works have studied the subleading term S_{topo} of the von Neumann entropy $S = -\text{tr}\rho \log \rho$ at finite temperature, in models that support topological order at $T = 0$ [16–22]. Nevertheless, S_{topo} cannot distinguish quantum correlations from the classical ones: even a purely classical

$\mathbb{Z}_2$ gauge theory in three dimensions has a nonzero S_{topo} consistent with the fact that it exhibits a self-correcting classical memory [11,23].

In this Letter, we propose an entanglement-based diagnostic for finite-T topological order that is sensitive only to quantum correlations. Specifically, we employ entanglement negativity E_N, a mixed-state entanglement measure, to quantify nonlocal quantum correlations resulting from finite-T topological order. The intuition behind our approach is that if a mixed state possesses long-range entanglement, then it is nonseparable over a length scale proportional to the system size, and therefore, such entanglement cannot be undone via any finite-depth quantum channel.

Given a Gibbs state corresponding to a local model, for a smooth entangling boundary, one can express E_N as a sum of local and nonlocal terms [24], analogous to the entanglement entropy for gapped ground states [25]: $E_N = E_{N,\text{local}} - E_{N,\text{topo}}$, where $E_{N,\text{local}} = \alpha_{d-1}L_A^{d-1} + \alpha_{d-3}L_A^{d-3} + \cdots$ characterizes the short-range entanglement, while $E_{N,\text{topo}}$ denotes the nonlocal entanglement, which is not expressible as a functional of local curvature along the entangling boundary. We will denote the nonlocal term as "topological entanglement negativity" and use it as a diagnostic for finite-T topological order.

We will primarily focus on toric code models at finite T in d spatial dimensions for $d = 2, 3, 4$: $H = -\lambda_A \sum_s A_s - \lambda_B \sum_p B_p$ where A_s/B_p are products of Pauli-X/Z operators (their precise forms depend on the dimensionality). This model has two critical temperatures T_A and T_B above which the excitations corresponding to A_s and B_p operators proliferate. Intuitively, a stable topological order at finite temperature can protect the encoded qubits against the thermal decoherence without the need of active error
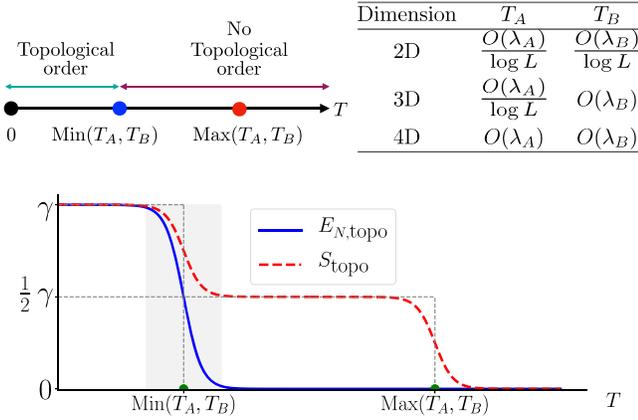
Dimension	T_A	T_B
2D	$\dfrac{O(\lambda_A)}{\log L}$	$\dfrac{O(\lambda_B)}{\log L}$
3D	$\dfrac{O(\lambda_A)}{\log L}$	$O(\lambda_B)$
4D	$O(\lambda_A)$	$O(\lambda_B)$

FIG. 1. Upper panel: phase diagram of toric code models. The critical temperatures T_A and T_B corresponding to the proliferation of two types of excitations depend on the spatial dimension. Lower panel: comparison between topological entanglement negativity $E_{N,\text{topo}}$ and topological von Neumann entropy S_{topo} in toric code models of size L. As $L \to \infty$, $E_{N,\text{topo}} = 0$ for $T > \text{Min}(T_A, T_B)$, consistent with the absence of topological order while S_{topo} remains nonzero in the regime $\text{Min}(T_A, T_B) < T < \text{Max}(T_A, T_B)$. When $\text{Min}(T_A, T_B) \neq 0$ as $L \to \infty$, the behavior of $E_{N,\text{topo}}$ shown close to the critical point (=the shaded region) is just a schematic and we do not probe that region.

correction, only when both types of excitations are suppressed, that is, temperature $T < \text{Min}(T_A, T_B)$. If only one type of excitations is suppressed, i.e., $\text{Min}(T_A, T_B) < T < \text{Max}(T_A, T_B)$, the other type of excitation destroys the topological order, and the model can only realize a self-correcting classical memory [11,17,18,23].

Our main result is summarized in Fig. 1. Through an explicit calculation, we find that topological entanglement negativity is nonzero only when the temperature is simultaneously below both critical temperatures associated with the proliferation of two types of excitations, in line with the aforementioned heuristics. In strong contrast, S_{topo} remains nonzero (drops to half of its ground state value) when temperature is between the lower and upper critical temperatures [17,18].

Disentangling toric codes at finite T.—Before discussing topological entanglement negativity for toric code models in detail, here we provide intuition for finite-T topological order by decomposing a mixed state of interest into a convex sum of pure states: $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$. If there exists a decomposition such that each $|\psi_i\rangle$ is short-range entangled, then ρ is not topologically ordered since it can be prepared from a trivial mixed state (a convex sum of product states) using a finite depth quantum circuit [12] (see also [26] for an explicit construction for toric code). One hint for such a decomposition comes from "minimally entangled typical thermal state" (METTS) ansatz [27]: $\rho = e^{-\beta H}/Z = \sum_m p_m |\phi_m\rangle\langle\phi_m|$ where each $|\phi_m\rangle$ is a METTS

obtained from imaginary time evolution of a product state $|m\rangle$: $|\phi_m\rangle \sim e^{-\beta H/2}|m\rangle$. $p_m = \langle m|e^{-\beta H}|m\rangle/Z$ is the probability corresponding to $|\phi_m\rangle$. Using such decomposition, we now show that the Gibbs state of the toric code in arbitrary spatial dimension is not topologically ordered above $\text{Min}(T_A, T_B)$.

First consider METTS obtained from product states $|m\rangle$ in the Z basis: $|\phi_m(T)\rangle \sim e^{\beta/2 \sum_s A_s} e^{\beta/2 \sum_p B_p}|m\rangle \sim e^{\beta/2 \sum_s A_s}|m\rangle$. All such METTS $|\phi_m(T)\rangle$ at temperature $T > T_A$ are short-range entangled since they can be adiabatically connected to the infinite temperature METTS $|\phi_m(T \to \infty)\rangle$, i.e., a product state, without encountering a phase transition or critical point. Therefore ρ is not topologically ordered for $T > T_A$. Similarly, decomposing ρ with METTS obtained from product states in X basis shows that ρ is not topologically ordered for $T > T_B$. Combining these two observations proves the absence of topological order in toric code for temperature $T > \text{min}(T_A, T_B)$. In fact, this result applies to all CSS code Hamiltonians $H = -\lambda_A \sum_i S_i^{(X)} - \lambda_B \sum_i S_i^{(Z)}$ [28,29], where each local commuting term $S_i^{(X/Z)}$ is a product of Pauli-X/Z operators. Using this result and the observation in Refs. [30–33], one immediately proves the absence of finite-T topological order in the more exotic models such as X-cube model [34] and Haah's code [35]. As an aside, each METTS $|\phi_m(T)\rangle$ is the ground state of a local parent Hamiltonian [26], which can be constructed using an approach analogous to Refs. [36,37]. Finally, we note that this approach can be applied to prove the absence of finite-T topological order in Kitaev's two-dimensional quantum double and its three-dimensional generalization as well [38].

General scheme for calculating negativity.—The above calculation using the METTS ansatz shows when a state is *not* topologically ordered. To understand the fate of topological order for $T < \text{Min}(T_A, T_B)$, we now turn to characterizing the mixed state entanglement of the Gibbs state using entanglement negativity [39–41]. Given a density matrix ρ acting on the Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$: $\rho = \sum_{A,B,A',B'} \rho_{AB,A'B'} |A\rangle|B\rangle\langle A'|\langle B'|$, one defines a partially transposed matrix ρ^{T_B} as $\rho^{T_B} = \sum_{A,B,A',B'} \rho_{AB,A'B'} |A\rangle|B'\rangle\langle A'|\langle B|$. The negativity is defined as $E_N = \log \|\rho^{T_B}\|_1 = \log(\sum_i |\lambda_i|)$ where λ_i are the eigenvalues of ρ^{T_B}. Relatedly, we define nth Renyi negativity: $R_n = b_n \log[\text{tr}(\rho^{T_B})^n/\text{tr}\rho^n]$ where $b_n = 1/(1-n)$ for odd n and $1/(2-n)$ for even n. b_n is chosen so that when ρ is a pure state, R_n reduces to Renyi entanglement entropy: $R_n = S_n, S_{n/2}$ for odd n and even n, respectively. Further, negativity $E_N = \lim_{n \to 1} R_n$, assuming n is even. We now turn to study the negativity of toric code in $d = 2, 3, 4$ dimension.

The negativity in 2D toric code at finite temperature was discussed in Ref. [42], which focuses on how finite

temperature excitations decrease quantum correlations, eventually leading to vanishing of negativity above a "sudden death temperature" $T_d > 0$. Here, we instead focus on the topological part $E_{N,\text{topo}}$, which captures the topological order.

We now present an approach motivated by Ref. [43] to taking the partial transpose of Gibbs states for stabilizer code Hamiltonians $H = -\sum_m S_m$, where S_m is a product of Pauli matrices over sites. Using $e^{\beta S_m} = \cosh\beta + S_m \sinh\beta$, we expand the Gibbs state as $e^{-\beta H} \sim \sum_{\{x_m\}} \prod_m (S_m \tanh\beta)^{x_m}$, where $x_m = 0$ or 1 indicates the absence or presence of S_m. Consider a subregion $\mathcal{R}$ and its complement $\bar{\mathcal{R}}$, taking partial transpose over $\bar{\mathcal{R}}$ in computational bases gives $(\prod_m S_m^{x_m})^{T_{\bar{\mathcal{R}}}} = \psi(\{x_m\}) \prod_m S_m^{x_m}$, where the sign $\psi(\{x_m\}) = 1$ or -1 corresponds to even or odd number of Pauli Ys in $\prod_m S_m^{x_m}$ on region $\bar{\mathcal{R}}$. Since stabilizers S_m supported only in $\mathcal{R}$ or $\bar{\mathcal{R}}$ always give even number of Pauli Ys in $\bar{\mathcal{R}}$ for toric codes, the sign ψ is solely determined by the appearance of the stabilizers across the bipartition boundary. This implies the partial transpose only acts on the bipartition boundary part of the Gibbs state: $\rho^{T_{\bar{\mathcal{R}}}} = (1/Z)(e^{-\beta H_{\mathcal{R}\bar{\mathcal{R}}}})^{T_{\bar{\mathcal{R}}}} e^{-\beta(H_{\mathcal{R}}+H_{\bar{\mathcal{R}}})}$, where $H_{\mathcal{R}}/H_{\bar{\mathcal{R}}}$ denotes the part of H supported on $\mathcal{R}/\bar{\mathcal{R}}$, and $H_{\mathcal{R}\bar{\mathcal{R}}}$ denotes the interaction between $\mathcal{R}$ and $\bar{\mathcal{R}}$. Define $\{A_i\}$, $\{B_j\}$ as the star and plaquette operators across the bipartition boundary, one finds

$$e^{-\beta H_{\mathcal{R}\bar{\mathcal{R}}}} \propto \sum_{\substack{\{n_i=0,1\} \\ \{\sigma_j=0,1\}}} \prod_{i=1}^{N_s^{\partial}} [A_i \tanh(\beta\lambda_A)]^{n_i} \prod_{j=1}^{N_p^{\partial}} [B_j \tanh(\beta\lambda_B)]^{\sigma_j}. \quad (1)$$

As mentioned above, taking partial transpose on a Pauli string introduces a sign determined by the number parity of Pauli Ys in region $\bar{\mathcal{R}}$: $\{[\prod_{i=1}^{N_s^{\partial}} A_i^{n_i}][\prod_{j=1}^{N_p^{\partial}} B_j^{\sigma_j}]\}^{T_{\bar{\mathcal{R}}}} = [\prod_{i=1}^{N_s^{\partial}} A_i^{n_i}][\prod_{j=1}^{N_p^{\partial}} B_j^{\sigma_j}] \psi(\{n_i\}, \{\sigma_j\})$. Since Pauli Ys only occur from products of Pauli Xs and Pauli Zs from neighboring star and plaquette operators, we find $\psi(\{n_i\}, \{\sigma_j\}) = \prod_{j=1}^{N_p^{\partial}} (\prod_{i \in \partial j} \tau_i)^{\sigma_j}$ where we have introduced the Ising variables $\tau_i = 1 - 2n_i \in \{\pm 1\}$. One can now sum over the σ_j variables and express $(e^{-\beta H_{\mathcal{R}\bar{\mathcal{R}}}})^{T_{\bar{\mathcal{R}}}}$ as a partition function over τ_i: $(e^{-\beta H_{\mathcal{R}\bar{\mathcal{R}}}})^{T_{\bar{\mathcal{R}}}} = [\cosh(\beta\lambda_A)]^{N_s^{\partial}} \sum_{\{\tau_i\}} e^{-H'(\{\tau_i\}, \{A_i\}, \{B_j\})}$, where

$$-H' = \sum_{i=1}^{N_s^{\partial}} \frac{1-\tau_i}{2} \log[A_i \tanh(\beta\lambda_A)] + \beta\lambda_B \sum_{j=1}^{N_p^{\partial}} B_j \prod_{i \in \partial j} \tau_i. \quad (2)$$

Replacing the commuting operators A_s, B_p with ± 1 gives the eigenspectrum of $\rho^{T_{\bar{\mathcal{R}}}}$. Interestingly, we find that signs of eigenvalues reflect the parity of braids between star and plaquette operators on the bipartition boundary at zero temperature [26]. This formalism can be used to
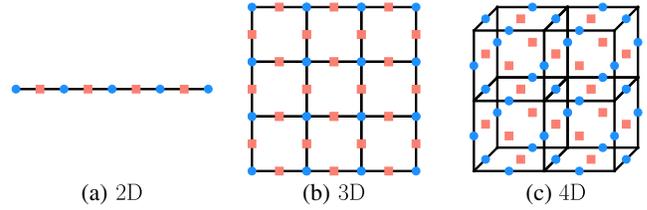
derive the partial transpose of a Gibbs state in other stabilizer models, such as Wen plaquette model [44] or fracton models [34,35]. We provide an alternative derivation of the above result using the matrix product state representation [26].

Two-dimensional toric code.—Recall that the star term $A_s(= \prod_{i \in s} X_i)$ is the product of four Pauli X operators on the star labeled by s, and $B_p(= \prod_{i \in p} Z_i)$ is the product of four Pauli Z operators on the plaquette labeled by p. Defining the model on a 2-torus gives fourfold degenerate ground states, where two qubits can be encoded, and are immune from local perturbation. However, both types of excitations (by flipping signs of A_s and B_p) are pointlike charges, which proliferate at any finite temperature to destroy the topological order and the encoded quantum information in the ground subspace. The topological entanglement negativity is $\log(2)$ at $T = 0$ [45,46], and here we show that the absence of finite-T topological order can be captured by the absence of topological (Renyi) negativity.

In this case, the bipartition boundary is a 1D system of length L with L star and L plaquette operators [see Fig. 2(a)]. Equation (2) implies that H' corresponds to a 1D Ising model in a magnetic field, which yields negativity $E_N = \log\langle|Z(\{A_i\}, \{B_j\})|\rangle$. $Z(\{A_i\}, \{B_j\}) = \{1/[\cosh\beta\lambda_B]^L\} \sum_{\{\tau_i = \pm 1\}} e^{-H'(\{\tau_i\}, \{A_i\}, \{B_j\})}$ with the angled brackets denoting the "disorder average" over the variables $\{A_i = \pm 1\}$ and $\{B_j = \pm 1\}$. This expression was first obtained in Ref. [42] using a replica trick, and we have provided an alternative derivation.

We first consider the limit $\lambda_B \to \infty$, forbidding magnetic charges in the $\mathbb{Z}_2$ gauge theory. Since in this limit the system realizes only a self-correcting classical memory instead of a quantum memory, it is a good starting point to see if the topological entanglement negativity is insensitive to long-distance classical correlations. Considering two connected regions separated by a closed boundary of size L, and defining $x = \tanh(\beta\lambda_A)$, we derive an expression for negativity $E_N = \alpha L - E_{N,\text{topo}}$. The area-law coefficient
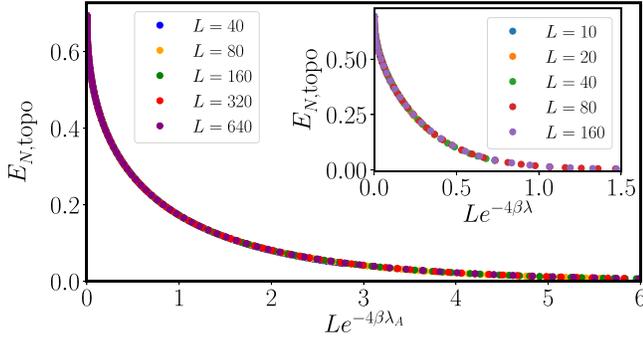

FIG. 2. The boundary operators in toric code for various spatial dimensions. Blue circles and red squares label A_i and B_j operators, respectively. (a) One-dimensional bipartition boundary in 2D toric code, where A_i live on sites, and B_j live on links. (b) Two-dimensional bipartition boundary in 3D toric code, where A_i live on sites, and B_j live on links. (c) Three-dimensional bipartition boundary in 4D toric code, where A_i live on links, and B_j live on faces.

FIG. 3. Scaling collapse of topological negativity in 2D toric code as $\lambda_B \to \infty$ [Eq. (3)]. L is the size of the bipartition boundary, β is the inverse temperature, and λ_A is the coefficient for star operators A_s. Inset: scaling collapse of topological negativity in 2D toric code at $\lambda = \lambda_A = \lambda_B$ using a classical Monte Carlo method combined with a transfer matrix method.

$\alpha = \log(1 + x)$ was first derived in Ref. [42], and here we instead focus on the topological entanglement negativity [26]:

$$E_{N,\text{topo}} = -\log\left\{\frac{1}{2} + \frac{1}{2}(x^{1/2} + x^{-1/2})^{-L}\left(\frac{L}{\frac{L}{2}+1}\right)\right.$$
$$\times\left[\frac{1}{x}{}_2F_1\left(1, -\frac{L}{2}+1; \frac{L}{2}+2; -\frac{1}{x}\right)\right.$$
$$\left.\left. - x{}_2F_1\left(1, -\frac{L}{2}+1; \frac{L}{2}+2; -x\right)\right]\right\}, \quad (3)$$

where ${}_2F_1$ is the hypergeometric function. While $E_{N,\text{topo}} = \log 2$ at zero temperature [45–48], it is exactly zero for any finite temperature as $L \to \infty$. Interestingly, for a finite L at low temperature, $E_{N,\text{topo}}$ only depends on the scaling variable $Le^{-4\beta\lambda_A}$ (Fig. 3), as one may also verify by expanding the hypergeometric function. This scaling variable represents the number of pairs of anyons thermally excited on the boundary. We also obtain analytical expressions for all even and odd Renyi negativities and find they depend respectively on $Le^{-4\beta\lambda_A}$ and $Le^{-2\beta\lambda_A}$ [26]. Next, for general λ_A and λ_B, we combine a classical Monte Carlo method and a transfer matrix method to calculate negativity, and find qualitatively same behavior as in the limit $\lambda_B \to \infty$ (see Fig. 3 inset). We also develop a generalized transfer matrix method to analytically show that the topological (Renyi) negativity is $\log 2$ at zero temperature and vanishes for any finite temperature as $L \to \infty$ [26].

Three-dimensional toric code.—Here the star operator $A_s(= \prod_{i\in s} X_i)$ is the product of six Pauli-X operators on the links emanating from a vertex of the cubic lattice, and the plaquette operator $B_p(= \prod_{i\in p} Z_p)$ is the product of four Pauli-Z operators on the links of a plaquette. Imposing periodic boundary conditions results in eight orthogonal ground states, which can encode three qubits. While flipping the sign of B_p gives looplike excitations, which is suppressed below the critical temperature corresponding to the 3D Z_2 gauge theory confinement transition, flipping A_s gives pointlike excitations, which proliferate at any nonzero temperature to destroy topological order. Now we show that topological negativity can again diagnose finite-T topological order.

Given a bipartition boundary of linear size L, there are L^2 boundary star operators A_i living on the lattice sites of the two dimensional boundary and $2L^2$ plaquette operators B_{ij} living on the links $\langle ij \rangle$ [Fig. 2(b)]. We again utilize the general formalism [Eq. (2)] specialized in this geometry to calculate negativity.

To separately see the effects of pointlike versus looplike excitations, we first consider $\lambda_B \to \infty$ to prohibit looplike excitations. We find negativity is exactly the same as the one in 2D by taking $L \to L^2$ [26], indicating the presence of only pointlike excitations. The topological negativity is exactly given by Eq. (3) by taking $L \to L^2$, and hence it vanishes in the thermodynamic limit at any nonzero temperature.

In contrast, taking $\lambda_A \to \infty$ prohibits pointlike excitations, and thus the finite-T topological order exists up to the critical temperature of 3D classical Z_2 gauge theory, above which the looplike excitations proliferate. Despite Eq. (2) giving the analytical form of $\rho^{T_{\bar{R}}}$, the calculation of negativity is challenging because (1) each eigenvalue of $\rho^{T_{\bar{R}}}$ requires calculating the partition function of 2D Ising model of τ_i spins subject to arbitrary given B_{ij} and A_i (2) plaquette operators B_p cannot be freely chosen for the eigenspectrum due to the local constraint $\prod_{p\in\partial\text{cubic}} B_p = 1$. Therefore, we turn to Renyi negativity R_n, which can be shown as the free energy difference of two statistical mechanics models: $R_n \sim \log \tilde{Z} - \log Z$, where Z is a partition function of 3D Z_2 gauge theory, $\tilde{Z}$ is a partition function of 3D Z_2 gauge theory coupled to n replicas of 2D Ising models. Absence of zero temperature critical point (due to $\lambda_A \to \infty$) allows a low temperature perturbative calculation for R_n: [26]

$$R_n = \alpha L^2 - R_{n,\text{topo}}, \quad (4)$$

where

$$\alpha = \begin{cases} \log 2 - \frac{1}{n-2}\left[\binom{n}{2}e^{-16\beta\lambda_B} + 2\binom{n}{2}e^{-24\beta\lambda_B} + \ldots\right] & \text{for even } n, \\ \log 2 - \frac{n}{n-1}\left[e^{-8\beta\lambda_B} + 2e^{-12\beta\lambda_B} + \frac{9}{2}e^{-16\beta\lambda_B} + \ldots\right] & \text{for odd } n, \end{cases} \quad (5)$$

and $R_{n,\text{topo}} = \log 2$. In fact, using the linked cluster theorem, which demands that only the excitations given by connected spin flips contribute to the logarithm of partition functions, we find those connected "diagrams" only contribute to the area law component of R_n without changing $R_{n,\text{topo}}$. Hence, we expect $R_{n,\text{topo}}$ remains $\log 2$ until the breakdown of the perturbative series, which occurs at the critical point of the 3D Z_2 gauge theory. Since $R_{n,\text{topo}}$ is independent of n, we conclude that $E_{N,\text{topo}} = \log(2)$ as well.

Four-dimensional toric code.—Finally we discuss the toric code in four spatial dimension, which realizes finite-T topological order [9]. Spins reside on each face of the 4D hypercube, and the Hamiltonian reads $H = -\lambda_A \sum_l A_l - \lambda_B \sum_c B_c$, where l, c label links and cubes. A_l is the product of six Pauli-X operators on the faces adjacent to the link l, and B_c is the product of six Pauli-Z operators on the faces around the boundary of the cube c. Flipping the sign of A_l or B_c gives looplike excitations living on the boundary of two-dimensional membranes, whose energy scales with the loop size. Therefore this model has two finite temperature critical points, corresponding to the proliferation of two looplike excitations, and it supports finite-T topological order up to temperature $T_c \propto \text{Min}(\lambda_A, \lambda_B)$.

The boundary of the 4D hypercube is a 3D cubic lattice, where the boundary operators are A_l living on every link and B_f living on every face [Fig. 2(c)]. Using Eq. (2), we find topological negativity at zero temperature is $2 \log 2$, consistent with the topological entanglement entropy [25]. We perform a low temperature perturbative calculation for Renyi negativity of even n, and find $R_n = \alpha L^3 - R_{n,\text{topo}}$ where $\alpha = 2\log 2 - [3n(n-1)/2(n-2)](e^{-16\beta\lambda_A} + e^{-16\beta\lambda_B}) + \cdots$, and the topological part $R_{n,\text{topo}}$ remains the ground state value $2 \log 2$ [25]. Similar to the 3D toric code in $\lambda_A \to \infty$ limit, we expect $R_{n,\text{topo}}$ remains unchanged up to T_c.

Summary.—In this Letter we propose the topological entanglement negativity as a diagnosis for finite-T topological order and correspondingly, a self-correcting quantum memory. We find it successfully detects the absence of finite-T topological order in 2D toric code. We demonstrated the robustness of topological entanglement in 3D toric code when the pointlike excitations are suppressed, and in the 4D toric code, consistent with finite-T topological order. Using METTS ansatz, we also provided an explicit decomposition of the Gibbs state in terms of short-range entangled pure states above $\text{Min}(T_A, T_B)$ where T_A, T_B are defined in Fig. 1.

One application of our proposal is to disentangle quantum correlations from classical ones in realistic models relevant to frustrated magnets. For example, spin-ice systems exhibit emergent photons and monopoles below the degeneracy temperature of classical configurations [49], irrespective of whether the ground state is topologically ordered or not. Negativity provides a precise diagnostic that distinguishes systems where the degenerate states coherently superpose to yield a topologically ordered state [50] ("quantum spin ice"), from systems that exhibit only classical emergent electromagnetism.

An important question remains: what is the critical behavior of topological entanglement negativity across a finite critical point, above which quantum memory is lost? While the thermodynamic criticality of the 4D toric code follows the 4D Ising universality [33], the transition is intrinsically "quantum mechanical" due to the loss of universal, long-distance quantum correlations. Studying the critical behavior of topological entanglement negativity may provide new insights for such a finite temperature "quantum phase transition."

[1] X. G. Wen, Phys. Rev. B **40**, 7387 (1989).
[2] X. G. Wen and Q. Niu, Phys. Rev. B **41**, 9377 (1990).
[3] X. G. Wen, Int. J. Mod. Phys. B **04**, 239 (1990).
[4] Z. Nussinov and G. Ortiz, Phys. Rev. B **77**, 064302 (2008).
[5] Z. Nussinov and G. Ortiz, Proc. Natl. Acad. Sci. U.S.A. **106**, 16944 (2009).
[6] S. Bravyi and B. Terhal, New J. Phys. **11**, 043029 (2009).
[7] O. Landon-Cardinal and D. Poulin, Phys. Rev. Lett. **110**, 090502 (2013).
[8] B. J. Brown, D. Loss, J. K. Pachos, C. N. Self, and J. R. Wootton, Rev. Mod. Phys. **88**, 045005 (2016).
[9] E. Dennis, A. Kitaev, A. Landahl, and J. Preskill, J. Math. Phys. (N.Y.) **43**, 4452 (2002).
[10] R. Alicki, M. Horodecki, P. Horodecki, and R. Horodecki, Open Syst. Inf. Dyn. **17**, 1 (2010).
[11] B. Yoshida, Ann. Phys. (Amsterdam) **326**, 2566 (2011).
[12] M. B. Hastings, Phys. Rev. Lett. **107**, 210501 (2011).
[13] A. Hamma, R. Ionicioiu, and P. Zanardi, Phys. Rev. A **71**, 022315 (2005).
[14] M. Levin and X.-G. Wen, Phys. Rev. Lett. **96**, 110405 (2006).
[15] A. Kitaev and J. Preskill, Phys. Rev. Lett. **96**, 110404 (2006).
[16] C. Castelnovo and C. Chamon, Phys. Rev. B **76**, 174416 (2007).
[17] C. Castelnovo and C. Chamon, Phys. Rev. B **76**, 184442 (2007).

[18] C. Castelnovo and C. Chamon, Phys. Rev. B **78**, 155120 (2008).

[19] S. V. Isakov, M. B. Hastings, and R. G. Melko, Nat. Phys. **7**, 772 (2011).

[20] D. Mazac and A. Hamma, Ann. Phys. (Amsterdam) **327**, 2096 (2012).

[21] B. Swingle and J. McGreevy, Phys. Rev. B **94**, 155125 (2016).

[22] Z. Li and R. S. Mong, arXiv:1910.07545.

[23] D. Poulin, R. G. Melko, and M. B. Hastings, Phys. Rev. B **99**, 094103 (2019).

[24] T.-C. Lu and T. Grover, arXiv:1907.01569.

[25] T. Grover, A. M. Turner, and A. Vishwanath, Phys. Rev. B **84**, 195120 (2011).

[26] See the Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.125.116801 for details.

[27] S. R. White, Phys. Rev. Lett. **102**, 190601 (2009).

[28] A. M. Steane, Phys. Rev. Lett. **77**, 793 (1996).

[29] A. R. Calderbank and P. W. Shor, Phys. Rev. A **54**, 1098 (1996).

[30] J. Haah, arXiv preprint arXiv:1305.6973.

[31] S. Bravyi and J. Haah, Phys. Rev. Lett. **111**, 200501 (2013).

[32] Z. Weinstein, E. Cobanera, G. Ortiz, and Z. Nussinov, Ann. Phys. (Amsterdam) **412**, 168018 (2020).

[33] Z. Weinstein, G. Ortiz, and Z. Nussinov, Phys. Rev. Lett. **123**, 230503 (2019).

[34] S. Vijay, J. Haah, and L. Fu, Phys. Rev. B **94**, 235157 (2016).

[35] J. Haah, Phys. Rev. A **83**, 042330 (2011).

[36] B. Swingle, J. McGreevy, and S. Xu, Phys. Rev. B **93**, 205159 (2016).

[37] S. Ok, K. Choo, C. Mudry, C. Castelnovo, C. Chamon, and T. Neupert, Phys. Rev. Research **1**, 033144 (2019).

[38] This approach only relies on the commutativity of local operators, and thus can be applied to show the absence of finite-T topological order in Kitaev's two-dimensional quantum double and its three-dimensional generalization of any finite groups, where the finite-T transition is absent [51].

[39] J. Eisert and M. B. Plenio, J. Mod. Opt. **46**, 145 (1999).

[40] G. Vidal and R. F. Werner, Phys. Rev. A **65**, 032314 (2002).

[41] M. B. Plenio, Phys. Rev. Lett. **95**, 090503 (2005).

[42] O. Hart and C. Castelnovo, Phys. Rev. B **97**, 144410 (2018).

[43] T.-C. Lu and T. Grover, Phys. Rev. B **99**, 075157 (2019).

[44] X.-G. Wen, Phys. Rev. Lett. **90**, 016803 (2003).

[45] C. Castelnovo, Phys. Rev. A **88**, 042319 (2013).

[46] Y. A. Lee and G. Vidal, Phys. Rev. A **88**, 042318 (2013).

[47] X. Wen, S. Matsuura, and S. Ryu, Phys. Rev. B **93**, 245140 (2016).

[48] X. Wen, P.-Y. Chang, and S. Ryu, J. High Energy Phys. 09 (2016) 012.

[49] C. Castelnovo, R. Moessner, and S. L. Sondhi, Nature (London) **451**, 42 (2008).

[50] M. Hermele, M. P. A. Fisher, and L. Balents, Phys. Rev. B **69**, 064404 (2004).

[51] A. Y. Kitaev, Ann. Phys. (Amsterdam) **303**, 2 (2003).