

Disentangling transitions in topological order induced by boundary decoherence

Tsung-Cheng Lu^{1,2}

¹*Perimeter Institute for Theoretical Physics, Waterloo, Ontario N2L 2Y5, Canada*

²*Department of Physics, Boston University, Boston, MA, 02215, USA*

We study the entanglement structure of topological orders subject to decoherence on the bipartition boundary. Focusing on the toric codes in d space dimensions for $d = 2, 3, 4$, we explore whether the boundary decoherence may be able to induce a disentangling transition, characterized by the destruction of mixed-state long-range entanglement across the bipartition, measured by topological entanglement negativity. A key insight of our approach is the connection between the negativity spectrum of the decohered mixed states and emergent symmetry-protected topological orders under certain symmetry-preserving perturbation localized on the bipartition boundary. This insight allows us to analytically derive the exact results of entanglement negativity without using a replica trick.

Quantum many-body systems with topological orders [1], owing to their long-range entanglement structure, may be used as topological quantum error-correcting codes [2, 3], where the encoded logical information can be protected and processed reliably under sufficiently weak environmental noise. Correspondingly, there exists a decodability transition at a certain critical noise threshold, above which the information is corrupted and no longer decodable. A seminal example is the 2d toric code, where the decodability transition can be understood from the order-disorder transition of the 2d random bond Ising model [3].

While the decodability transition has long been a familiar concept in the field of quantum information, it possesses several intriguing features from the standpoint of quantum phases of matter. For one thing, this transition cannot be detected via any linear observables O of the noisy mixed-state density matrix ρ [4]. Namely, the observable of the form $\text{tr}(\rho O)$ is a smooth function of the noise rate. This observation therefore motivates several outstanding questions. Does the decodability transition separate two distinct mixed-state quantum phases? How should one classify phases of matter in mixed states? Perhaps most intriguingly, does the decodability transition correspond to a certain intrinsically ‘quantum’ phase transition, which only manifests in the entanglement structure, as opposed to any physical observables, of the noisy mixed state? These questions have provided new opportunities to explore mixed-state phases and their transitions [4–30].

In this work, we consider a setup, where the entanglement structure of certain decohered many-body states can be completely understood. We consider a topologically-ordered system divided into two regions \mathcal{A} and \mathcal{B} , and the decoherence *only* acts on the bipartition boundary (see Fig. 1). In this case, one expects the entanglement structure and topological order in the respective bulks of \mathcal{A} and \mathcal{B} to remain intact for arbitrary decoherence strength. Therefore, the essential question is how the boundary decoherence affects the entanglement structure across the interface. In particular, does there exist a ‘disentangling transition’ at a certain critical decoherence strength, across which the long-range entan-

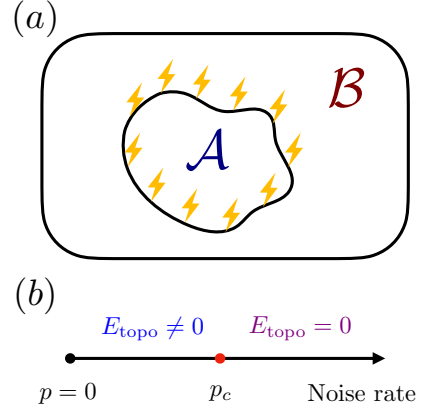


FIG. 1: (a) A topologically-ordered state subject to the decoherence acting *only* on the bipartition boundary between \mathcal{A} and \mathcal{B} . (b) Tuning the noise rate p may induce a disentangling transition at a critical error rate p_c , above (below) which the long-range entanglement between \mathcal{A} and \mathcal{B} , measured by topological entanglement negativity E_{topo} , is zero (non-zero).

glement between \mathcal{A} and \mathcal{B} is destroyed? We will address this question for the toric code models in various space dimensions and explore the entanglement structure via entanglement negativity [31–34], an entanglement measure for mixed states. Specifically, we will employ topological entanglement negativity, a subleading contribution of entanglement negativity, to diagnose long-range entanglement and characterize the disentangling transition.

Our analytical approach is built on Ref.[35], which points out a correspondence between certain bulk topological orders and symmetry-protected topological (SPT) orders [36, 37] emerging from the calculation of entanglement negativity. Recall that entanglement negativity of a bipartite mixed state ρ is defined as $E_N = \log |\rho^\Gamma|_1 = \log (\sum_i |\lambda_i|)$, where ρ^Γ results from taking a partial transpose (i.e. the transpose only acting on the subregion \mathcal{A}) on the density matrix ρ , and $\{\lambda_i\}$ is the eigenspectrum ρ^Γ , also dubbed negativity spectrum. It was found that the negativity spectrum of the ground-subspace density matrix of topological codes in d space dimensions ex-

actly corresponds to an SPT wave function localized on the $d - 1$ dimensional bipartition boundary. Taking the 2d (\mathbb{Z}_2) toric code as an example, the negativity spectrum corresponds to the wave functions in symmetry-charge basis (or equivalently, the strange correlators) of a 1d cluster state that exhibits a $\mathbb{Z}_2 \times \mathbb{Z}_2$ SPT order [38]. Heuristically, this follows from the fact that taking a partial transpose on the toric-code ground-space density matrix proliferates e -particles and m -particles along the bipartition boundary with non-trivial mutual braiding statistics, and this is precisely captured by the braiding phase between two types of the \mathbb{Z}_2 symmetry charges encoded in the SPT.

Built on this result, we find that introducing boundary decoherence amounts to introducing certain symmetry-preserving perturbations on the SPT, and the wave functions of this perturbed SPT completely characterize the negativity spectrum of the decohered mixed state. Importantly, this allows us to analytically compute the entanglement negativity and extract the topological contribution (i.e. topological entanglement negativity) without resorting to a replica trick; for the topological order in d dimensions subject to the boundary decoherence, the entanglement negativity relates to the free energy difference associated with annihilating domain walls in *translationally invariant* statistical mechanics models in $(d - 1)$ -dimensions. While our analytical approaches apply to any qubit-stabilizer models, we will focus on the discussion of the toric codes in various space dimensions, which we summarize below.

For the 2d toric code subject to the boundary decoherence, the entanglement negativity relates to the free energy of the 1d Ising model. Due to the absence of a finite-temperature transition in the 1d Ising model, topological entanglement negativity E_{topo} of the decohered toric code remains the non-zero quantized value $\log 2$ for any non-maximal noise rate. Namely, E_{topo} becomes zero identically only at the maximal decoherence strength $p = \frac{1}{2}$.

For the 3d toric code, when introducing the boundary noise that creates loop-like defects, the emergent statistical mechanics model is a \mathbb{Z}_2 gauge theory in 2d, which does not exhibit any finite-temperature transition. Correspondingly, E_{topo} remains $\log 2$ for any non-maximal noise strength. On the other hand, under the decoherence that creates point-like defects, the entanglement structure is mapped to a 2d Ising model, whose finite-temperature transition indicates the disentangling transition of the decohered toric code at a certain critical decoherence strength p_c , below which $E_{\text{topo}} = \log 2$, and above which, $E_{\text{topo}} = 0$.

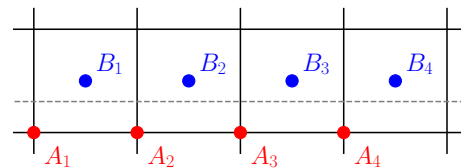
Finally, for the 4d toric code subject to the boundary decoherence, the emergent statistical mechanics model is a \mathbb{Z}_2 gauge theory in 3d, which exhibits a confinement-deconfinement transition at finite temperature. Correspondingly, there exists a disentangling transition at a certain critical decoherence strength p_c , below which $E_{\text{topo}} = 2 \log 2$, and above which, $E_{\text{topo}} = 0$.

Topological entanglement negativity - In this work, we will employ topological entanglement negativity to characterize the bipartite long-range entanglement of topological orders subject to boundary decoherence. Consider a smooth bipartition boundary of size L , one expects the bipartite mixed-state entanglement measured by entanglement negativity follows the form: $E_N = E_{\text{local}} - E_{\text{topo}}$. E_{local} captures the short-range entanglement along the bipartition boundary and exhibits an area-law scaling $E_{\text{local}} \sim \alpha_{d-1} L^{d-1}$ to leading order in L , and E_{topo} is the *topological entanglement negativity*, a universal constant contribution proposed to diagnose the long-range entanglement. Such a scaling form is first proposed in Ref.[39] based on the locality structure of many-body systems, and later demonstrated for finite-temperature Gibbs states [35, 40] and certain topological order under decoherence [4, 16, 23]. When the system of interest is described by a pure state, the negativity coincides with Renyi-1/2 entanglement entropy [34], and the topological entanglement negativity reduces to topological entanglement (Renyi) entropy [41].

2d toric code - Here we discuss the negativity spectrum and the exact result of entanglement negativity in the 2d toric code subject to Pauli-Z noise on the bipartition boundary¹. We will discuss the essential physics and outline the derivation in the main text. A unified, detailed calculation for toric codes of various space dimensions is presented in Appendix.A.

The toric-code Hamiltonian reads $H = -\sum_s A_s - \sum_p B_p$, where the star stabilizer A_s is a product of Pauli-Zs on four edges emanating from a site s , and plaquette stabilizer B_p is a product of Pauli-Xs on four edges around a plaquette p . The density matrix that describes the ground subspace of the toric code is $\rho_0 \propto \prod_s \frac{1+A_s}{2} \prod_p \frac{1+B_p}{2}$.

Consider a 1d closed bipartition boundary of size L ², there are $2L$ alternating star and plaquette stabilizers along the boundary, with the dashed line denoting the bipartition boundary:



As will be convenient, we factorize the mixed state into a bulk part and a boundary part: $\rho_0 \propto \rho_{\text{bulk}} \rho_{\partial,0}$,

¹ We only present the result for Pauli-Z noise as Pauli-X noise will lead to the same mixed-state entanglement structure due to the duality between A_s and B_p stabilizers.

² The boundary can either be contractible (e.g. Fig.1a), or non-contractible, e.g. winding around x direction of the 2d lattice with periodic (open) boundary condition in x (y) direction (the topology of a cylinder). Our calculation applies to both cases.

where the boundary part $\rho_{\partial,0} = \prod_{i=1}^L \frac{1+A_i}{2} \prod_{i=1}^L \frac{1+B_i}{2}$ involves the boundary stabilizers, and $\rho_{\text{bulk}} = \prod_{s \in \text{bulk}} \frac{1+A_s}{2} \prod_{p \in \text{bulk}} \frac{1+B_p}{2}$ involves the bulk stabilizers.

Now we consider the decoherence that only creates the boundary B_i excitations by applying Pauli-Z operators on edges shared by two adjacent B_i s with probability p_z . This leads to a decohered state $\rho = \prod_i N_{z,i}[\rho_0]$, where $N_{z,i}[\cdot]$ is a local noise channel with the action $N_{z,i}[\rho] = (1-p_z)\rho + p_z Z_i \rho Z_i$, and Z_i denotes the Pauli-Z operator on the edge shared between B_i and B_{i+1} . To see the effect of the noise channel, we expand the initial boundary density matrix as a sum over all possible stabilizer strings: $\rho_{\partial,0} \propto \sum_{a,b} \prod_i A_i^{a_i} \prod_i B_i^{b_i}$ with $a \equiv \{a_i\}$, $b \equiv \{b_i\}$, and each $a_i, b_i \in \{0,1\}$ denote the absence/presence of A_i, B_i .

Applying the noise channel amounts to introducing a ‘tension’ so that the weight of a given stabilizer string $\prod_i A_i^{a_i} \prod_i B_i^{b_i}$ is suppressed exponentially in the number of Pauli-Xs acting on the qubits that are subject to decoherence, denoted as $W(X)$. Namely, $\rho_{\partial} \propto \sum_{a,b} \prod_i A_i^{a_i} \prod_i B_i^{b_i} (1-2p_z)^{W(X)}$, and $W(X)$ can be computed as follows. For a given edge, the Pauli-X exists when only one of the two neighboring plaquette stabilizers B_i, B_{i+1} is present. Therefore, $W(X) = \frac{1}{2} \sum_i (1 - \tau_i \tau_{i+1})$, with $\tau_i = 1 - 2b_i \in \{1, -1\}$. By defining $\beta_z = -\frac{1}{2} \log(1-2p_z)$, the boundary decohered state can be written as

$$\rho_{\partial} \propto \sum_{a,b} \prod_i A_i^{a_i} \prod_i B_i^{b_i} e^{\beta_z \sum_{i=1}^L \tau_i \tau_{i+1}}, \quad (1)$$

where the Ising interaction among $\{b_i\}$ (or equivalently, among $\{\tau_i\}$) incorporates the effect of the decoherence.

To proceed, one can expand the decohered mixed state $\rho \propto \rho_{\text{bulk}} \rho_{\partial}$ into a sum over all possible stabilizer strings $\prod_s A_s^{a_s} \prod_p B_p^{b_p}$, with $a_s, b_p = 0,1$. Then taking a partial transpose assigns a sign (1 or -1) depending on whether the number of Pauli-Y operators in the subregion \mathcal{A} is even or odd. This number is only determined by the choice of the boundary stabilizers, and therefore, taking a partial transpose on a subregion acts non-trivially only on the boundary density matrix, i.e. $\rho^{\Gamma} \propto \rho_{\text{bulk}} \rho_{\partial}^{\Gamma}$. Specifically, taking a partial transpose on the boundary stabilizer string gives $(\prod_i A_i^{a_i} \prod_i B_i^{b_i})^{\Gamma} = \psi(a,b) \prod_i A_i^{a_i} \prod_i B_i^{b_i}$ with the sign $\psi(a,b) = \prod_i (-1)^{a_i b_i + b_i a_{i+1}} = 1$ or -1 , because the simultaneous presence of two neighboring star and plaquette boundary stabilizers contributes to a -1 associated with a single Pauli-Y in the subregion \mathcal{A} ³.

As a result, one finds $\rho^{\Gamma} \propto \rho_{\text{bulk}} \rho_{\partial}^{\Gamma}$, with

$$\rho_{\partial}^{\Gamma} \propto \sum_{a,b} \psi(a,b) \prod_i A_i^{a_i} \prod_i B_i^{b_i} e^{\beta_z \sum_{i=1}^L \tau_i \tau_{i+1}} \quad (2)$$

Since the partially-transposed matrix ρ^{Γ} involves only stabilizers, all the eigenvalues (i.e. negativity spectrum of the decohered mixed state) can be obtained by substituting the stabilizers to be 1 or -1 . To have non-vanishing eigenvalues, A_s and B_p in the bulk are fixed to be 1. On the other hand, the boundary stabilizers A_i and B_i can be 1 or -1 , thereby determining the negativity spectrum. In the following, we will abuse the notation by treating ρ^{Γ} as the spectrum (determined by all possible choices of the boundary stabilizer values), rather than a matrix.

It is illuminating to understand the negativity spectrum using a dual description in terms of the variables $\{a_i\}$ and $\{b_i\}$. Introducing a fictitious Hilbert space of $2L$ qubits along the 1d bipartition boundary, with the Z-basis state $|a,b\rangle \equiv |a_1, b_1, a_2, b_2, \dots\rangle$ where $Z_{A,i} |a,b\rangle = (-1)^{a_i} |a,b\rangle$ and $Z_{B,i} |a,b\rangle = (-1)^{b_i} |a,b\rangle$, one can define a state $|\psi\rangle \propto \sum_{a,b} \psi(a,b) |a,b\rangle$, and a trivial state $|+\rangle \propto \sum_{a,b} |a,b\rangle$. Then the negativity spectrum can be expressed as

$$\rho^{\Gamma} \propto \langle + | \prod_i Z_{A,i}^{\frac{1-A_i}{2}} \prod_i Z_{B,i}^{\frac{1-B_i}{2}} e^{\beta_z \sum_i Z_{B,i} Z_{B,i+1}} | \psi \rangle, \quad (3)$$

where distinct eigenvalues in the negativity spectrum are given by distinct choices of $A \equiv \{A_i\}$, $B \equiv \{B_i\}$, which in turn determines the patterns of inserting $Z_{A,i}$ and $Z_{B,i}$, sandwiched in between $|+\rangle$ and $e^{\beta_z \sum_i Z_{B,i} Z_{B,i+1}} | \psi \rangle$ in the fictitious Hilbert space. $|\psi\rangle$ is the 1d cluster state with the parent Hamiltonian $H = -\sum_{i=1}^L Z_{A,i} X_{B,i} Z_{A,i+1} - \sum_{i=1}^L Z_{B,i} X_{A,i+1} Z_{B,i+1}$, which is known to exhibit an SPT order protected by a $\mathbb{Z}_2 \times \mathbb{Z}_2$ symmetry generated by $\prod_{i=1}^L X_{A,i}$ and $\prod_{i=1}^L X_{B,i}$.

As such, the negativity spectrum can be understood as the wave functions of the state $e^{\beta_z \sum_i Z_{B,i} Z_{B,i+1}} | \psi \rangle$ in the X-basis, namely $\prod_i Z_{A,i}^{\frac{1-A_i}{2}} \prod_i Z_{B,i}^{\frac{1-B_i}{2}} |+\rangle$. Alternatively, it can be understood as the strange correlators [42], i.e. the symmetry-charge operators $\prod_i Z_{A,i}^{\frac{1-A_i}{2}} \prod_i Z_{B,i}^{\frac{1-B_i}{2}}$ inserted in between $|+\rangle$ and $e^{\beta_z \sum_i Z_{B,i} Z_{B,i+1}} | \psi \rangle$.

Using the condition $Z_{B,i} X_{A,i+1} Z_{B,i+1} | \psi \rangle = | \psi \rangle$, the negativity spectrum can be simplified as

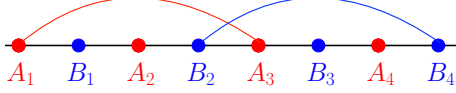
$$\begin{aligned} \rho^{\Gamma} &\propto \langle + | \prod_i Z_{A,i}^{\frac{1-A_i}{2}} \prod_i Z_{B,i}^{\frac{1-B_i}{2}} e^{\beta_z \sum_i X_{A,i}} | \psi \rangle \\ &= \langle + | \prod_i Z_{A,i}^{\frac{1-A_i}{2}} \prod_i Z_{B,i}^{\frac{1-B_i}{2}} | \psi \rangle e^{\beta_z \sum_i A_i}. \end{aligned} \quad (4)$$

Finally, defining $\chi(A,B) = \frac{\langle + | \prod_i Z_{A,i}^{\frac{1-A_i}{2}} \prod_i Z_{B,i}^{\frac{1-B_i}{2}} | \psi \rangle}{\langle + | \psi \rangle}$, the negativity spectrum reads

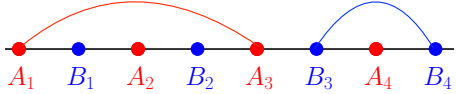
$$\rho^{\Gamma} = \chi(A,B) \frac{e^{\beta_z \sum_i A_i}}{Z}. \quad (5)$$

³ Alternatively, the -1 sign follows from the fact that two neighboring star and plaquette boundary stabilizers anticommute when restricted in the subregion \mathcal{A} .

where Z is the normalization constant to ensure the negativity spectrum sums to 1 (since the trace of a matrix is invariant under partial transpose). $\chi(A, B)$ is non-zero only when the configurations A, B satisfy $\prod_i A_i = \prod_i B_i = 1$ due to the $\mathbb{Z}_2 \times \mathbb{Z}_2$ symmetry of $|\psi\rangle$, and the non-vanishing $\chi(A, B)$ takes the value ± 1 depending on whether there exists non-trivial braiding between A and B . For instance, consider the configuration with only A_1, B_2, A_3, B_4 equal -1 , which corresponds to the following diagram by connecting two points with $A_i = -1$ via a red string and two points with $B_i = -1$ via a blue string:



This pattern gives $\chi(A, B) = -1$, due to the odd number of intersections between two types of strings⁴. On the other hand, $\chi(A, B) = 1$ when there is no braiding between the two types of strings, e.g. the configuration with only A_1, A_3, B_3, B_4 equal -1 as shown below:



Conceptually, $A = \{A_i\}$ and $B = \{B_i\}$ configurations represent the e, m anyons (of the \mathbb{Z}_2 topological order) localized on the bipartition boundary, and the mutual braiding statistics is exactly captured by $\chi(A, B)$, i.e. the wave functions of the 1d cluster state with the $\mathbb{Z}_2 \times \mathbb{Z}_2$ SPT order.

By summing over the absolute values of the negativity spectrum (Eq.5), one can obtain the one-norm of the partially-transposed density matrix $|\rho^\Gamma|_1 = \frac{\sum_{A,B} |\chi(A,B) e^{\beta \sum_i A_i}|}{\sum_{A,B} \chi(A,B) e^{\beta \sum_i A_i}}$, which turns out to be the ratio of two partition functions:

$$|\rho^\Gamma|_1 = \frac{Z_{1d, \text{Ising}}}{\tilde{Z}_{1d, \text{Ising}}}. \quad (6)$$

$Z_{1d, \text{Ising}} = \sum_{\sigma} e^{\beta \sum_i \sigma_i \sigma_{i+1}}$ is the partition function of the 1d Ising model at the inverse temperature $\beta_z = -\frac{1}{2} \log(1 - 2p_z)$, and $\tilde{Z}_{1d, \text{Ising}} = \sum_{\sigma} e^{\beta \sum_i \sigma_i \sigma_{i+1}} \prod_i \delta(\sigma_i \sigma_{i+1} = 1) = 2e^{\beta L}$ is the one without domain walls (i.e. excitations). Hence the entanglement negativity relates to the free energy difference associated with prohibiting domain walls: $E_N = \log |\rho^\Gamma|_1 = \log Z_{1d, \text{Ising}} - \log \tilde{Z}_{1d, \text{Ising}}$, which can be analytically computed:

$$E_N = \alpha L - E_{\text{topo}} \quad (7)$$

The area-law coefficient $\alpha = \log(1 + e^{-2\beta_z})$ captures the short-distance, non-universal part of the entanglement, and topological entanglement negativity $E_{\text{topo}} = \log 2 - \log(1 + (\tanh \beta_z)^L)$ captures the universal, long-range entanglement.

As a sanity check, at zero decoherence ($\beta_z = 0$), one recovers the result that $\alpha = \log 2$ and $E_{\text{topo}} = \log 2$. On the other hand, when increasing decoherence strength (i.e. increasing β_z), α decreases monotonically, consistent with the expectation that decoherence reduces entanglement. The subleading topological contribution E_{topo} is more interesting; for any non-maximal decoherence ($p_z < \frac{1}{2}$, or equivalently, any non-infinite β_z), this quantity remains $\log 2$ in the thermodynamic limit $L \rightarrow \infty$, indicating that the long-range entanglement between the regions \mathcal{A} and \mathcal{B} is robust up to the maximal boundary decoherence.

As the negativity spectrum (Eq.3) is given by the wave function of the state $|\tilde{\psi}\rangle \propto e^{\beta \sum_i Z_{B,i} Z_{B,i+1}} |\psi\rangle$, i.e. the 1d cluster state perturbed by a weak projection $e^{\beta \sum_i Z_{B,i} Z_{B,i+1}}$, one might naturally wonder whether the persistence of the long-range entanglement relates to the robustness of SPT order in $|\tilde{\psi}\rangle$. We find this is indeed the case. Specifically, the SPT order is characterized by two types of string order parameters: $S_A(i, j) = Z_{B,i} X_{A,i+1} \cdots X_{A,j} Z_{B,j}$ and $S_B(i, j) = Z_{A,i} X_{B,i} \cdots X_{B,j-1} Z_{A,j}$ (see e.g. Ref.[43]). Since the perturbation commutes with S_A , one finds $\langle \tilde{\psi} | S_A(i, j) | \tilde{\psi} \rangle = 1$, which is insensitive under the perturbation. On the other hand, $\langle \tilde{\psi} | S_B(i, j) | \tilde{\psi} \rangle = \frac{1}{[\cosh 2\beta_z]^2 [1 + (\tanh 2\beta_z)]^L}$, which vanishes *only* at the limit $\beta_z \rightarrow \infty$. Alternatively, one may consider the wave-function overlap $\langle \tilde{\psi} | \tilde{\psi} \rangle \propto \sum_{\sigma} e^{2\beta \sum_i \sigma_i \sigma_{i+1}}$ (see Appendix.B), which is the partition function of the 1d Ising model, hence non-singular for any $\beta_z < \infty$. This indicates that the SPT order in $|\tilde{\psi}\rangle$ survives for any non-infinite β_z , consistent with the persistence of the long-range entanglement up to maximal decoherence.

3d toric code - Let's now discuss the entanglement structure of the 3d toric code subject to boundary decoherence. The Hamiltonian reads $H = -\sum_v A_v - \sum_f B_f$ with $A_v = ZZZZZZ$ denoting the product of Pauli-Zs on 6 edges around a vertex v , and $B_f = XXXX$ denoting the product of Pauli-Xs on 4 edges on the boundary of a face f . By considering the Pauli-Z noise, which creates loop-like excitations, we will see that the entanglement negativity of the noisy mixed state relates to the 2d classical \mathbb{Z}_2 gauge theory, which does not exhibit a finite-temperature transition, indicating the persistence of long-range entanglement up to the maximal decoherence. On the other hand, considering the decoherence by Pauli-X noise, which creates point-like excitations, the entanglement negativity relates to the 2d classical Ising model, which exhibits a finite-temperature order-disorder

⁴ This can be verified algebraically using $Z_{A,i} X_{B,i} Z_{A,i+1} |\psi\rangle = Z_{B,i} X_{A,i+1} Z_{B,i+1} |\psi\rangle = |\psi\rangle$.

transition. Correspondingly, there exists a disentangling transition in the decohered mixed state at a critical Pauli-X noise strength.

We will consider a 2d plane of size L as the bipartition boundary that separates the two regions \mathcal{A}, \mathcal{B} ⁵. There are L^2 boundary A type operators on the vertices denoted by A_v and $2L^2$ boundary B type operators on edges denoted by B_e .

Under the Pauli-Z decoherence with noise rate p_z that only creates the excitations associated with the boundary B_e operators, using a calculation similar to the 2d toric code, one can derive the negativity spectrum:

$$\rho^\Gamma \propto \langle + | \prod_v Z_{A,v}^{\frac{1-A_v}{2}} \prod_e Z_{B,e}^{\frac{1-B_e}{2}} e^{\beta_z \sum_v \prod_{e \in \partial^* v} Z_{B,e}} | \psi \rangle, \quad (8)$$

where distinct eigenvalues are given by distinct configurations $A \equiv \{A_v = \pm 1\}$, $B \equiv \{B_e = \pm 1\}$, and $\beta_z = -\frac{1}{2} \log(1 - 2p_z)$. We have introduced the fictitious Hilbert space of qubits living on vertices and edges. $|+\rangle$ is the $+1$ Pauli-X product state, and $|\psi\rangle$ is a cluster state in 2d with the parent Hamiltonian $H = -\sum_e X_{B,e} \prod_{v \in \partial e} Z_{A,v} - \sum_v X_{A,v} \prod_{e \in \partial^* v} Z_{B,e}$, where $\prod_{v \in \partial e}$ denotes the product over vertices on the boundary of e , and $\prod_{e \in \partial^* v}$ denotes the product over edges on the coboundary of v (i.e. the product of those edges whose boundaries contain v). $|\psi\rangle$ exhibits an SPT order protected by a \mathbb{Z}_2 0-form $\times \mathbb{Z}_2$ 1-form symmetry generated by $\prod_v X_{A,v}$ and $\prod_{e \in \mathcal{C}} X_{B,e}$ with \mathcal{C} being any closed loops.

With the condition $X_{A,v} \prod_{e \in \partial^* v} Z_{B,e} |\psi\rangle = |\psi\rangle$, the negativity spectrum can be simplified as

$$\rho^\Gamma = \frac{1}{Z} \chi(A, B) e^{\beta_z \sum_v A_v} \quad (9)$$

with $\chi(A, B) = \frac{\langle + | \prod_v Z_{A,v}^{\frac{1-A_v}{2}} \prod_e Z_{B,e}^{\frac{1-B_e}{2}} | \psi \rangle}{\langle + | \psi \rangle}$, and Z being a normalization constant. Using the negativity spectrum, we derive the entanglement negativity

$$E_N = \log Z_{2d, \text{gauge}} - \log \tilde{Z}_{2d, \text{gauge}} \quad (10)$$

where $Z_{2d, \text{gauge}} = \sum_\sigma e^{\beta_z \sum_v \prod_{e \in \partial^* v} \sigma_e}$ is the partition function of the classical 2d \mathbb{Z}_2 gauge theory, and $\tilde{Z}_{2d, \text{gauge}} = \sum_\sigma e^{\beta_z \sum_v \prod_{e \in \partial^* v} \sigma_e} \prod_v \delta(\prod_{e \in \partial^* v} \sigma_e = 1)$ is the one prohibiting domain walls (excitations).

By computing the partition function of the 2d gauge theory, we find $E_N = \alpha L^2 - E_{\text{topo}}$ with the area-law coefficient $\alpha = \log(1 + e^{-2\beta_z})$ and the topological entanglement negativity

$$E_{\text{topo}} = \log 2 - \log \left(1 + (\tanh \beta_z)^{L^2} \right) \quad (11)$$

This indicates that, in the thermodynamic limit $L \rightarrow \infty$, the long-range entanglement survives up to maximal decoherence ($\beta_z \rightarrow \infty$, or equivalently, $p_z = \frac{1}{2}$). Consistent with the persistence of the long-range entanglement, the emergent boundary state $e^{\beta_z \sum_v \prod_{e \in \partial^* v} Z_{B,e}} |\psi\rangle$ that describes the negativity spectrum (Eq.8) exhibits an SPT order for any finite $\beta_z < \infty$, as diagnosed by the wave function overlap (Appendix.B).

Now we consider the Pauli-X decoherence that creates the point-like excitations on vertices with noise strength p_x , and present the negativity spectrum

$$\rho^\Gamma \propto \langle + | \prod_v Z_{A,v}^{\frac{1-A_v}{2}} \prod_e Z_{B,e}^{\frac{1-B_e}{2}} e^{\beta_x \sum_e \prod_{v \in \partial e} Z_{A,v}} | \psi \rangle \quad (12)$$

which can further be simplified as $\rho^\Gamma = \frac{1}{Z} \chi(A, B) e^{\beta_x \sum_e B_e}$ with $\beta_x = -\frac{1}{2} \log(1 - 2p_x)$, and Z being a normalization constant. Using the spectrum, we derive the entanglement negativity

$$E_N = \log Z_{2d, \text{Ising}} - \log \tilde{Z}_{2d, \text{Ising}}. \quad (13)$$

$Z_{2d, \text{Ising}} = \sum_\sigma e^{\beta_x \sum_{\langle ij \rangle} \sigma_i \sigma_j}$ is the partition function of the classical 2d Ising model, and $\tilde{Z}_{2d, \text{Ising}}$ is the one when prohibiting excitations: $\tilde{Z}_{2d, \text{Ising}} = \sum_\sigma e^{\beta_x \sum_{\langle ij \rangle} \sigma_i \sigma_j} \prod_{\langle ij \rangle} \delta(\sigma_i \sigma_j = 1) = 2e^{2\beta_x L^2}$; the prefactor 2 accounts for the two-fold ground-state degeneracy of the Ising model.

In the thermodynamic limit, $\log Z_{2d, \text{Ising}}$ takes the form $L^2 f_x + \log g_x$, where f_x is the free-energy density of the 2d classical Ising model. Below the critical temperature $T_c = 1/\beta_c$, the Ising model spontaneously breaks the global \mathbb{Z}_2 symmetry, hence $g_x = 2$. On the other hand, above the critical temperature, the global symmetry is not broken, giving $g_x = 1$. Therefore, the entanglement negativity reads $E_N = \alpha L^2 - E_{\text{topo}}$ with the area-law coefficient $\alpha = f_x - 2\beta_x$, and topological entanglement negativity

$$E_{\text{topo}} = \begin{cases} \log 2, & \text{for } p_x < p_c \\ 0, & \text{for } p_x > p_c. \end{cases} \quad (14)$$

$p_c = \frac{1 - e^{-2\beta_c}}{2} = 1 - \frac{\sqrt{2}}{2} \approx 0.29$ with $\beta_c = \frac{1}{2} \ln(1 + \sqrt{2})$ being the critical inverse temperature of the 2d Ising model [44]. As such, there exists a disentangling transition diagnosed by the discontinuity of topological entanglement negativity at the critical decoherence strength p_c . Relatedly, $|\tilde{\psi}\rangle \propto e^{\beta_x \sum_e \prod_{v \in \partial e} Z_{A,v}} |\psi\rangle$ that describes the negativity spectrum (Eq.12) exhibit an SPT order up to a critical β_x as diagnosed by the singularity of the wave-function overlap $\langle \tilde{\psi} | \tilde{\psi} \rangle \propto \sum_\sigma e^{2\beta_x \sum_{\langle ij \rangle} \sigma_i \sigma_j}$, which is

⁵ Specifically, we consider a 3d lattice with the periodic boundary condition in x, y directions and open boundary condition in z . The bipartition boundary is a non-contractible 2d plane spanned by x and y directions.

the partition function of the 2d Ising model at the inverse temperature $2\beta_x$. This indicates that the nature of the disentangling transition of the decohered mixed state and the SPT-order transition of the state $|\tilde{\psi}\rangle$ are both described by the 2d Ising universality, albeit at distinct critical temperatures.

4d toric code — Finally, we explore the disentangling transition in the 4d toric code [3], where spins reside on each face (i.e. 2-cell) of a 4-dimensional hypercube. The Hamiltonian is $H = -\sum_e A_e - \sum_c B_c$, where A_e is the product of 6 Pauli-Z operators on the faces adjacent to the edge e , and B_c is the product of 6 Pauli-X operators on the faces around the boundary of the cube c . The bipartition boundary of a 4d hypercube is a 3d lattice, where the boundary stabilizers are A_e living on edges and B_f living on faces⁶.

Under the Pauli-Z noise that creates the excitations associated with the boundary B_f stabilizers with noise rate p_z ⁷, the negativity spectrum of the decohered mixed state reads

$$\rho^\Gamma \propto \langle + | \prod_e Z_{A,e}^{\frac{1-A_e}{2}} \prod_f Z_{B,f}^{\frac{1-B_f}{2}} e^{\beta_z \sum_e \prod_{f \in \partial^* e} Z_{B,f}} | \psi \rangle, \quad (15)$$

where $\beta_z \equiv -\frac{1}{2} \log(1 - 2p_z)$, and we introduce the fictitious Hilbert space of qubits living on edges and faces; $|+\rangle$ is the +1 X-basis product state, and $|\psi\rangle$ is the ground state of the 3d cluster-state Hamiltonian $H = -\sum_f X_{B,f} \prod_{l \in \partial f} Z_{A,l} - \sum_l X_{A,l} \prod_{f \in \partial^* l} Z_{B,f}$ ⁸ [45]. $|\psi\rangle$ exhibits an SPT order protected by the \mathbb{Z}_2 1-form $\times \mathbb{Z}_2$ 1-form symmetry, generated by $\prod_{f \in \partial c} X_f$ and $\prod_{l \in \partial^* v} X_l$.

The negativity spectrum can be further simplified as

$$\rho^\Gamma = \frac{1}{Z} \chi(A, B) e^{\beta_z \sum_e A_e} \quad (16)$$

with $\chi(A, B) = \frac{\langle + | \prod_e Z_{A,e}^{\frac{1-A_e}{2}} \prod_f Z_{B,f}^{\frac{1-B_f}{2}} | \psi \rangle}{\langle + | \psi \rangle}$. Due to the two \mathbb{Z}_2 1-form symmetries, to have non-vanishing $\chi(A, B)$, A and B must form closed loops in the primary lattice and the dual lattice respectively, and $\chi(A, B) = 1, -1$ for even, odd number of times of braiding (linking) between the A loops and B loops. Conceptually, A and B can be thought of as two species of loop excitations of the 4d toric code living on the 3d boundary, and the mutual braiding statistics is encoded in $\chi(A, B)$, i.e. the wave functions of 3d cluster-state SPT in the X basis.

With the negativity spectrum, we derive the entanglement negativity

$$E_N = \log Z_{3d, \text{gauge}} - \log \tilde{Z}_{3d, \text{gauge}}, \quad (17)$$

where $Z_{3d, \text{gauge}} = \sum_\sigma e^{\beta_z \sum_f \prod_{e \in \partial f} \sigma_e}$ is the partition function of the 3d classical \mathbb{Z}_2 gauge theory, and $\tilde{Z}_{3d, \text{gauge}}$ is the same theory but prohibiting excitations: $\tilde{Z}_{3d, \text{gauge}} = \sum_\sigma e^{\beta_z \sum_f \prod_{e \in \partial f} \sigma_e} \prod_f \delta(\prod_{e \in \partial f} \sigma_e = 1) = n e^{3\beta_z L^3}$, with $n = 2^{L^3+2}$ being the number of ground-state configurations of the 3d gauge theory. On the other hand, $\log Z_{3d, \text{gauge}}$ exhibits a confinement-deconfinement transition at a critical temperature $T_c = 1/\beta_c$ [46, 47], and is expected to take the following form in the thermodynamic limit: $L^3 f_z + \log g$. f_z is the free-energy density, $g = 1$ in the confined phase ($\beta_z < \beta_c$), and $g = 4$ in the deconfined phase ($\beta_z > \beta_c$). One way to understand the discontinuity in g is by mapping the 3d classical \mathbb{Z}_2 gauge theory at finite temperature to the 2d quantum \mathbb{Z}_2 gauge theory at zero temperature. The quantity g can then be understood as the ground-state degeneracy in the 2d quantum gauge theory; $g = 4$ in the deconfined phase due to the spontaneous 1-form symmetry breaking, and $g = 1$ in the confined phase due to the restoration of the 1-form symmetry (see Appendix C for detailed discussion). Therefore, there exists a disentangling transition diagnosed by the discontinuity of topological entanglement negativity: $E_{\text{topo}} = 2 \log 2, 0$ for $p_z < p_c$ and $p_z > p_c$ respectively, where $p_c = \frac{1 - e^{-2\beta_c}}{2}$ is critical noise rate, with β_c being the critical inverse temperature of the 3d classical \mathbb{Z}_2 gauge theory.

Discussion - In this work, we present the exact results for the negativity spectrum and the entanglement structure of topological orders subject to decoherence along the bipartition boundary. The negativity spectrum of the decohered mixed state in d space dimensions corresponds to the wave functions of an SPT order in $d - 1$ space dimensions under certain symmetry-preserving perturbations. This understanding allows us to compute entanglement negativity (without resorting to a replica trick) and map it to the free energy of a *translationally-invariant* classical statistical mechanics model in $d - 1$ space dimensions, whose phase transition corresponds to the disentangling transition in the physical systems. While we only discuss the entanglement structure of the toric code, our approach can be straightforwardly applied to any topological qubit-stabilizer models, e.g. X-cube fracton code [48] or Haah's code [49].

As a future question, it would be interesting to explore boundary decoherence from the perspective of separability transitions [12]. Specifically, one can ask: whether a topological order subject to boundary decoherence may exhibit a boundary separability transition at a critical decoherence strength p_c , above (below) which the mixed state ρ can (cannot) be written as $\rho = \sum_i p_i |\psi_i\rangle \langle \psi_i|$, where each $|\psi_i\rangle$ can be made unentangled w.r.t a fixed \mathcal{A}, \mathcal{B} bipartition via a finite-depth local unitary circuit $U_{\text{fd}, i}$. Namely, $|\psi_i\rangle = U_{\text{fd}, i} |\psi_{\mathcal{A}, i}\rangle_{\mathcal{A}} \otimes |\psi_{\mathcal{B}, i}\rangle_{\mathcal{B}}$, with $|\psi_{\mathcal{A}, i}\rangle_{\mathcal{A}}, |\psi_{\mathcal{B}, i}\rangle_{\mathcal{B}}$ being arbitrary states in \mathcal{A}, \mathcal{B} that may exhibit a bulk topological order. If such a boundary separability transition exists, does it coincide with the dis-

⁶ Specifically, we consider a 4d lattice with the periodic boundary condition in x, y, z directions and open boundary condition in w direction. The bipartition boundary is a non-contractible 3d cube spanned by x, y, z directions.

⁷ The Pauli-X noise will lead to the same mixed-state entanglement structure due to the duality between A_e, B_c stabilizers.

⁸ Here we again use ∂ and ∂^* to denote the boundary map and the coboundary map, respectively.

entangling transition witnessed via topological entanglement negativity?

It would also be interesting to explore the connection between the disentangling transition and the channel definition of mixed-state phases of matter [15]. For instance, in the 2d toric code subject to the boundary decoherence, we found that topological entanglement negativity survives up to maximal decoherence. Does it suggest the un-decohered toric code can be reconstructed from the boundary-decohered toric code via a finite-depth quantum channel acting on the bipartition boundary?

More technically, there are two cases where the mixed-state entanglement structure of the decohered toric code remains elusive: (i) the decoherence given by simultaneous Pauli-X and Pauli-Z noise on the bipartition boundary, and (ii) the decoherence given by Pauli-X or Pauli-Z noise channels uniformly throughout the entire system. For the first case, while we present partial progress by deriving the negativity spectrum in Appendix D, we are unable to obtain exact results of entanglement negativity. For the second case, while Ref.[4] derived the moment of the partially-transposed matrix $\text{tr}(\rho^\Gamma)^{2n}$ for integer n , it is unclear how to perform the analytic continuation by taking $n \rightarrow \frac{1}{2}$ to obtain the exact results of entanglement

negativity.

Finally, our approach based on the negativity spectrum and an emergent SPT description may be useful to explore the entanglement structure of intrinsically mixed-state topological orders [16, 29], an intriguing type of topological order that may not be realized in gapped ground states of local Hamiltonians. More generally, it might be fruitful to consider various non-trivial states of matter, e.g. non-abelian topological order, quantum critical states, and chiral topological phases, subject to the decoherence on the bipartition boundary, and explore the entanglement structure of mixed states and their disentangling transitions.

ACKNOWLEDGMENTS

T.-C.L. thanks Zhehao Zhang and Ruihua Fan for helpful comments on the manuscript. T.-C.L. also thanks Claudio Chamon for hosting his visit at Boston University, during which this work is completed. The research at Perimeter Institute is supported in part by the Government of Canada through the Department of Innovation, Science and Economic Development Canada and by the Province of Ontario through the Ministry of Colleges and Universities.

-
- [1] X.-G. Wen, Quantum field theory of many-body systems: from the origin of sound to an origin of light and electrons, *OUP Oxford* (2004).
 - [2] A. Kitaev, Fault-tolerant quantum computation by anyons, *Annals of Physics* **303**, 2 (2003).
 - [3] E. Dennis, A. Kitaev, A. Landahl, and J. Preskill, Topological quantum memory, *Journal of Mathematical Physics* **43**, 4452 (2002).
 - [4] R. Fan, Y. Bao, E. Altman, and A. Vishwanath, Diagnostics of mixed-state topological order and breakdown of quantum memory, arXiv preprint arXiv:2301.05689 (2023).
 - [5] C. de Groot, A. Turzillo, and N. Schuch, Symmetry Protected Topological Order in Open Quantum Systems, *Quantum* **6**, 856 (2022).
 - [6] R. Ma and C. Wang, Average symmetry-protected topological phases, *Phys. Rev. X* **13**, 031016 (2023).
 - [7] J. Y. Lee, Y.-Z. You, and C. Xu, Symmetry protected topological phases under decoherence, arXiv preprint arXiv:2210.16323 (2022).
 - [8] J. Y. Lee, C.-M. Jian, and C. Xu, Quantum criticality under decoherence or weak measurement, *PRX Quantum* **4**, 030317 (2023).
 - [9] Y. Bao, R. Fan, A. Vishwanath, and E. Altman, Mixed-state topological order and the errorfield double formulation of decoherence-induced transitions, arXiv preprint arXiv:2301.05687 (2023).
 - [10] T.-C. Lu, Z. Zhang, S. Vijay, and T. H. Hsieh, Mixed-state long-range order and criticality from measurement and feedback, *PRX Quantum* **4**, 030318 (2023).
 - [11] Y. Zou, S. Sang, and T. H. Hsieh, Channeling quantum criticality, *Phys. Rev. Lett.* **130**, 250403 (2023).
 - [12] Y.-H. Chen and T. Grover, Separability transitions in topological states induced by local decoherence, arXiv preprint arXiv:2309.11879 (2023).
 - [13] Y.-H. Chen and T. Grover, Symmetry-enforced many-body separability transitions, arXiv preprint arXiv:2310.07286 (2023).
 - [14] K. Su, N. Myerson-Jain, C. Wang, C.-M. Jian, and C. Xu, Higher-form symmetries under weak measurement, arXiv preprint arXiv:2304.14433 (2023).
 - [15] S. Sang, Y. Zou, and T. H. Hsieh, Mixed-state quantum phases: Renormalization and quantum error correction, arXiv preprint arXiv:2310.08639 (2023).
 - [16] Z. Wang, Z. Wu, and Z. Wang, Intrinsic mixed-state topological order without quantum memory, arXiv preprint arXiv:2307.13758 (2023).
 - [17] R. Ma, J.-H. Zhang, Z. Bi, M. Cheng, and C. Wang, Topological phases with average symmetries: the decohered, the disordered, and the intrinsic, arXiv preprint arXiv:2305.16399 (2023).
 - [18] P.-S. Hsin, Z.-X. Luo, and H.-Y. Sun, Anomalies of average symmetries: Entanglement and open quantum systems, arXiv preprint arXiv:2312.09074 (2023).
 - [19] K. Su, N. Myerson-Jain, and C. Xu, Conformal field theories generated by chern insulators under decoherence or measurement, *Phys. Rev. B* **109**, 035146 (2024).
 - [20] N. Myerson-Jain, T. L. Hughes, and C. Xu, Decoherence through ancilla anyon reservoirs, arXiv preprint arXiv:2312.04638 (2023).
 - [21] Z. Li and R. S. Mong, Replica topological order in quantum mixed states and quantum error correction, arXiv

- preprint arXiv:2402.09516 (2024).
- [22] K. Su, Z. Yang, and C.-M. Jian, Tapestry of dualities in decohered quantum error correction codes, arXiv preprint arXiv:2401.17359 (2024).
 - [23] A. Lyons, Understanding stabilizer codes under local decoherence through a general statistical mechanics mapping, arXiv preprint arXiv:2403.03955 (2024).
 - [24] J. Y. Lee, Exact calculations of coherent information for toric codes under decoherence: Identifying the fundamental error threshold, arXiv preprint arXiv:2402.16937 (2024).
 - [25] Y.-H. Chen and T. Grover, Unconventional topological mixed-state transition and critical phase induced by self-dual coherent errors, arXiv preprint arXiv:2403.06553 (2024).
 - [26] Y. Guo, J.-H. Zhang, S. Yang, and Z. Bi, Locally purified density operators for symmetry-protected topological phases in mixed states, arXiv preprint arXiv:2403.16978 (2024).
 - [27] R. Ma and A. Turzillo, Symmetry protected topological phases of mixed states in the doubled space, arXiv preprint arXiv:2403.13280 (2024).
 - [28] H. Xue, J. Y. Lee, and Y. Bao, Tensor network formulation of symmetry protected topological phases in mixed states, arXiv preprint arXiv:2403.17069 (2024).
 - [29] R. Sohal and A. Prem, A noisy approach to intrinsically mixed-state topological order, arXiv preprint arXiv:2403.13879 (2024).
 - [30] Z. Wang and L. Li, Anomaly in open quantum systems and its implications on mixed-state quantum phases, arXiv preprint arXiv:2403.14533 (2024).
 - [31] A. Peres, Separability criterion for density matrices, *Phys. Rev. Lett.* **77**, 1413 (1996).
 - [32] M. Horodecki, P. Horodecki, and R. Horodecki, Separability of mixed states: necessary and sufficient conditions, *Physics Letters A* **223**, 1 (1996).
 - [33] J. Eisert and M. B. Plenio, A comparison of entanglement measures, *Journal of Modern Optics* **46**, 145 (1999).
 - [34] G. Vidal and R. F. Werner, Computable measure of entanglement, *Phys. Rev. A* **65**, 032314 (2002).
 - [35] T.-C. Lu and S. Vijay, Characterizing long-range entanglement in a mixed state through an emergent order on the entangling surface, *Phys. Rev. Res.* **5**, 033031 (2023).
 - [36] X. Chen, Z.-C. Gu, and X.-G. Wen, Classification of gapped symmetric phases in one-dimensional spin systems, *Phys. Rev. B* **83**, 035107 (2011).
 - [37] X. Chen, Z.-C. Gu, and X.-G. Wen, Complete classification of one-dimensional gapped quantum phases in interacting spin systems, *Phys. Rev. B* **84**, 235128 (2011).
 - [38] H. J. Briegel and R. Raussendorf, Persistent entanglement in arrays of interacting particles, *Phys. Rev. Lett.* **86**, 910 (2001).
 - [39] T.-C. Lu and T. Grover, Structure of quantum entanglement at a finite temperature critical point, *Phys. Rev. Research* **2**, 043345 (2020).
 - [40] T.-C. Lu, T. H. Hsieh, and T. Grover, Detecting topological order at finite temperature using entanglement negativity, *Phys. Rev. Lett.* **125**, 116801 (2020).
 - [41] A. Kitaev and J. Preskill, Topological entanglement entropy, *Phys. Rev. Lett.* **96**, 110404 (2006).
 - [42] Y.-Z. You, Z. Bi, A. Rasmussen, K. Slagle, and C. Xu, Wave function and strange correlator of short-range entangled states, *Phys. Rev. Lett.* **112**, 247202 (2014).
 - [43] F. Pollmann and A. M. Turner, Detection of symmetry-protected topological phases in one dimension, *Phys. Rev. B* **86**, 125441 (2012).
 - [44] L. Onsager, Crystal statistics. i. a two-dimensional model with an order-disorder transition, *Phys. Rev.* **65**, 117 (1944).
 - [45] R. Raussendorf, S. Bravyi, and J. Harrington, Long-range quantum entanglement in noisy cluster states, *Phys. Rev. A* **71**, 062313 (2005).
 - [46] F. J. Wegner, Duality in Generalized Ising Models and Phase Transitions without Local Order Parameters, *Journal of Mathematical Physics* **12**, 2259 (1971).
 - [47] J. B. Kogut, An introduction to lattice gauge theory and spin systems, *Rev. Mod. Phys.* **51**, 659 (1979).
 - [48] S. Vijay, J. Haah, and L. Fu, Fracton topological order, generalized lattice gauge theory, and duality, *Phys. Rev. B* **94**, 235157 (2016).
 - [49] J. Haah, Local stabilizer codes in three dimensions without string logical operators, *Phys. Rev. A* **83**, 042330 (2011).
 - [50] T. H. Hsieh, From d-dimensional quantum to d+ 1-dimensional classical systems, *Student Review* (2016).

APPENDIX

In Appendix.A, we present a unified derivation of negativity spectrum and negativity for d -dimensional toric codes ($d = 2, 3, 4$) subject to boundary decoherence. Appendix.B discusses the stability of the perturbed SPTs encoded in the negativity spectrum. Appendix.C presents the mapping between a 2d quantum Ising gauge theory at zero temperature and a 3d classical Ising gauge theory at finite temperature. Appendix.D presents the calculation of the negativity spectrum of the 2d toric code subject to both Pauli-X and Pauli-Z decoherence on the boundary.

Appendix A: General formalism of computing negativity spectrum

Here we present the calculation of the negativity spectrum for the toric code in d space dimensions subject to the decoherence acting on the $d - 1$ dimensional bipartition boundary that separates the subregions \mathcal{A} and \mathcal{B} .

The toric code Hamiltonian in d space dimensions takes the form $H = -\sum_i A_i - \sum_j B_j$, where the stabilizer A_i is the product of Pauli-Zs around i , and the stabilizer B_j is the product of Pauli-Xs around j :

2d toric code: qubits are defined on 1-cells, A_i is a product of Pauli-Zs on four 1-cells emanating from a 0-cell, and B_j is a product of Pauli-Xs on four 1-cells on the boundary of a 2-cell.

3d toric code: qubits are defined on 1-cells, A_i is a product of Pauli-Zs on six 1-cells emanating from a 0-cell, and B_j is a product of Pauli-Xs on four 1-cells on the boundary of a 2-cell.

4d toric code: qubits are defined on 2-cells, A_i is a product of Pauli-Zs on the six 2-cells adjacent to a 1-cell, and B_j is the product of Pauli-Xs on the six 2-cells on the boundary of a 3-cell.

The corresponding ground-state subspace density matrix is $\rho_0 \propto \prod_i \frac{1+A_i}{2} \prod_j \frac{1+B_j}{2} \propto \sum_{a,b} \prod_i A_i^{a_i} \prod_j B_j^{b_j}$, with $a \equiv \{a_i\}$, $b \equiv \{b_j\}$, and $a_i, b_j = 0, 1$.

1. Pauli-Z decoherence

We consider the Pauli-Z noise channels applied on the set of qubits (denoted by \mathcal{M}) that only excites the B_j stabilizers on the boundary. At error rate p_z , the decohered mixed state $\rho = \prod_{i \in \mathcal{M}} N_{z,i}[\rho_0]$, with $N_{z,i}[\rho_0] = (1 - p_z)\rho_0 + p_z Z_i \rho_0 Z_i$, may be expressed as

$$\rho \propto \sum_{a,b} \prod_i A_i^{a_i} \prod_j B_j^{b_j} (1 - 2p)^{W(X)}. \quad (\text{A1})$$

The stabilizer string $\prod_j B_j^{b_j}$ may be factorized into a product of the bulk part $\prod_j^{\text{bulk}} B_j^{b_j}$ and the boundary part $\prod_j^{\partial} B_j^{b_j}$, and $W(X)$ is the number of Pauli-Xs belonging to the set \mathcal{M} that appear in the boundary stabilizer string $\prod_j^{\partial} B_j^{b_j}$.

Taking a partial transpose (e.g. a transpose acting on the subregion \mathcal{A}) gives the matrix

$$\rho^\Gamma \propto \sum_{a,b} \prod_i A_i^{a_i} \prod_j B_j^{b_j} (1 - 2p)^{W(X)} \psi(a, b), \quad (\text{A2})$$

with $\psi(a, b) = 1/-1$ when the stabilizer strings $\prod_i A_i^{a_i} \prod_j B_j^{b_j}$ has even/odd number of Pauli-Ys in the subregion \mathcal{A} . Since both $W(X)$ and $\psi(a, b)$ only depends on a_i, b_j of the boundary stabilizers, one can sum over a_i, b_j in the bulk, resulting in

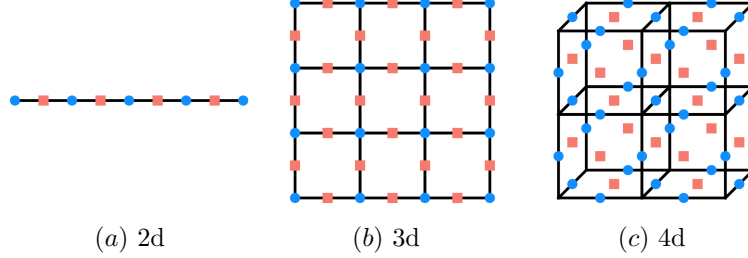
$$\rho^\Gamma \propto \left(\prod_{i \in \text{bulk}} \frac{1+A_i}{2} \prod_{j \in \text{bulk}} \frac{1+B_j}{2} \right) \left(\sum_{a,b} \prod_i A_i^{a_i} \prod_j B_j^{b_j} (1 - 2p)^{W(X)} \psi(a, b) \right) \quad (\text{A3})$$

The negativity spectrum, i.e. the eigenvalues of ρ^Γ , can be obtained by replacing all stabilizers by 1 or -1 . As the stabilizers in the bulk are fixed at 1, the negativity spectrum is solely determined by the values of the boundary stabilizers. Therefore we write the spectrum as

$$\rho^\Gamma \propto \sum_{a,b} \prod_i A_i^{a_i} \prod_j B_j^{b_j} (1-2p)^{W(X)} \psi(a,b), \quad (\text{A4})$$

where we abuse the notation by treating ρ^Γ as the spectrum, rather than a matrix.

To compute $W(X)$ and $\psi(a,b)$, we locate the stabilizers on the $d-1$ dimensional bipartition boundary as follows:



where (a), (b), and (c) depict the boundary of the toric code in 2d, 3d, and 4d, respectively; the set of blue circles (denoted as R_A) specifies the boundary A_i stabilizers, and the set of red squares (denoted as R_B) specifies the boundary B_j stabilizers.

As mentioned above, $W(X)$ counts the number of Pauli-Xs in the set \mathcal{M} appearing in stabilizer string $\prod_j B_j^{b_j}$. This gives $W(X) = \frac{1}{2} \sum_{i \in R_A} (1 - \prod_{j \in \partial^* i} \tau_j)$, where $\tau_j \equiv (-1)^{b_j}$, and $\prod_{j \in \partial^* i}$ denotes the product over $j (\in R_B)$ on the coboundary of $i \in R_A$ (i.e. i belongs to the boundary of j). For example, for 2d boundary of the 3d toric code, $W(X) = \frac{1}{2} \sum_{i \in R_A} (1 - \tau\tau\tau\tau)$, where $\tau\tau\tau\tau$ denotes the product of τ_j s on the four edges emanating from the vertex (0-cell).

On the other hand, $\psi(a,b) = (-1)^{N_Y}$, where N_Y is the number of Pauli-Ys in the subregion \mathcal{A} in the stabilizer string $\prod_{i \in R_A} A_i^{a_i} \prod_{j \in R_B} B_j^{b_j}$. Since the simultaneous presence of neighboring A_i , B_j stabilizers contributes an odd number of Pauli-Ys, one has $\psi(a,b) = \prod_{\langle i \in R_A, j \in R_B \rangle} (-1)^{a_i b_j}$, where $\langle i \in R_A, j \in R_B \rangle$ denotes the pair $i \in R_A, j \in R_B$ neighboring to each other.

With the results of $W(X), \psi(a,b)$, we derive the negativity spectrum:

$$\rho^\Gamma \propto \sum_{a,b} \prod_{i \in R_A} A_i^{a_i} \prod_{j \in R_B} B_j^{b_j} e^{\beta_z \sum_{i \in R_A} \prod_{j \in \partial^* i} \tau_j} \psi(a,b), \quad (\text{A5})$$

with $\beta_z = -\frac{1}{2} \log(1-2p_z)$.

The structure of the negativity spectrum may be understood by introducing the fictitious Hilbert space of qubits living on the sublattices R_A, R_B spanned by the Z basis product states $|a,b\rangle = |\{a_i, b_j\}\rangle$. We can define a state $|\psi\rangle \propto \sum_{a,b} \psi(a,b) |a,b\rangle$ and a trivial state $|+\rangle \propto \sum_{a,b} |a,b\rangle$, then the negativity spectrum can be written as

$$\rho^\Gamma \propto \langle + | \prod_{i \in R_A} Z_{A,i}^{\frac{1-A_i}{2}} \prod_{j \in R_B} Z_{B,j}^{\frac{1-B_j}{2}} e^{\beta_z \sum_{i \in R_A} \prod_{j \in \partial^* i} Z_{B,j}} | \psi \rangle, \quad (\text{A6})$$

where distinct eigenvalues in the negativity spectrum correspond to different patterns of inserting $Z_{i,A}, Z_{j,B}$ acting on the fictitious Hilbert space.

Since $|\psi\rangle = \sum_{a,b} \prod_{\langle i \in R_A, j \in R_B \rangle} (-1)^{a_i b_j} |a,b\rangle$, $|\psi\rangle$ can be obtained by applying $\prod_{\langle i \in R_A, j \in R_B \rangle} \text{CZ}_{i,j}$ on the trivial state $|+\rangle$, with $\text{CZ}_{i,j}$ being the two-qubit controlled-Z gates. It follows that $|\psi\rangle$ is the ground state of the cluster-state Hamiltonian $H_c = -\sum_{i \in R_A} X_{A,i} \prod_{j \in \partial^* i} Z_{B,j} - \sum_{j \in R_B} X_{B,j} \prod_{i \in \partial j} Z_{A,i}$. Therefore, one may employ the condition $\prod_{j \in \partial^* i} Z_{B,j} |\psi\rangle = X_{A,i} |\psi\rangle$ to simplify the negativity spectrum:

$$\begin{aligned} \rho^\Gamma &\propto \langle + | \prod_{i \in R_A} Z_{A,i}^{\frac{1-A_i}{2}} \prod_{j \in R_B} Z_{B,j}^{\frac{1-B_j}{2}} e^{\beta_z \sum_{i \in R_A} X_{A,i}} | \psi \rangle \\ &= \langle + | \prod_{i \in R_A} Z_{A,i}^{\frac{1-A_i}{2}} \prod_{j \in R_B} Z_{B,j}^{\frac{1-B_j}{2}} | \psi \rangle e^{\beta_z \sum_{i \in R_A} A_i}. \end{aligned} \quad (\text{A7})$$

Note that to have non-vanishing eigenvalues, the configurations $A \equiv \{A_i\}, B \equiv \{B_j\}$ need to obey the symmetry constraint resulting from the symmetry of the cluster-state Hamiltonian. Specifically, $|\psi\rangle$ has two types of symmetries: $\prod_{i \in \Gamma_A^*} X_{A,i} = 1$ and $\prod_{j \in \Gamma_B} X_{B,j} = 1$, where Γ_A^* is a cocycle (which has trivial coboundary, i.e. $\partial^* \Gamma_A^* = 0$), and Γ_B is a cycle (which has trivial boundary, i.e. $\partial \Gamma_B = 0$). Hence the configurations A, B must satisfy $\prod_{i \in \Gamma_A^*} A_i = 1$ and $\prod_{j \in \Gamma_B} B_j = 1$, to have non-vanishing eigenvalues in the negativity spectrum.

Finally, the negativity spectrum reads

$$\rho^\Gamma = \frac{1}{Z} \chi(A, B) e^{\beta_z \sum_{i \in R_A} A_i}, \quad (\text{A8})$$

where $Z = \sum_{A, B} \chi(A, B) e^{\beta_z \sum_{i \in R_A} A_i}$ is a normalization constant to ensure that the negativity spectrum sums to one, and $\chi(A, B) = \frac{\langle + | \prod_{i \in R_A} Z_{A,i}^{\frac{1-A_i}{2}} \prod_{j \in R_B} Z_{B,j}^{\frac{1-B_j}{2}} | \psi \rangle}{\langle + | \psi \rangle} = 1, -1$ for $\{A_i, B_j\}$ configurations that satisfy $\prod_{i \in \Gamma_A^*} A_i = 1$ and $\prod_{j \in \Gamma_B} B_j = 1$.

With the negativity spectrum, we proceed to compute entanglement negativity $E_N = \log(|\rho^\Gamma|_1)$, with

$$|\rho^\Gamma|_1 = \frac{1}{Z} \sum_{A, B} \left| \chi(A, B) e^{\beta_z \sum_{i \in R_A} A_i} \right| \quad (\text{A9})$$

First, we compute the denominator Z :

$$\begin{aligned} Z &= \sum_{A, B} \frac{\langle + | \prod_{i \in R_A} Z_{A,i}^{\frac{1-A_i}{2}} \prod_{j \in R_B} Z_{B,j}^{\frac{1-B_j}{2}} e^{\beta_z \sum_{i \in R_A} A_i} \prod_{j \in \partial^* i} Z_{B,j} | \psi \rangle}{\langle + | \psi \rangle} \\ &= \frac{1}{\langle + | \psi \rangle} \langle + | \prod_{i \in R_A} (1 + Z_{A,i}) \prod_{j \in R_B} (1 + Z_{B,j}) e^{\beta_z \sum_{i \in R_A} A_i} \prod_{j \in \partial^* i} Z_{B,j} | \psi \rangle \end{aligned} \quad (\text{A10})$$

Since $|\psi\rangle = \prod_{\langle i \in R_A, j \in R_B \rangle} CZ_{i,j} |+\rangle$, one has $\langle + | \prod_{i \in R_A} (1 + Z_{A,i}) \prod_{j \in R_B} (1 + Z_{B,j}) | \psi \rangle = \langle + | \prod_{i \in R_A} (1 + Z_{A,i}) \prod_{j \in R_B} (1 + Z_{B,j}) |+\rangle = 1$. As a result,

$$Z = \frac{1}{\langle + | \psi \rangle} e^{\beta_z \sum_{i \in R_A} A_i}. \quad (\text{A11})$$

On the other hand, $\sum_{A, B} \left| \chi(A, B) e^{\beta_z \sum_{i \in R_A} A_i} \right| = \sum'_{A, B} e^{\beta_z \sum_{i \in R_A} A_i}$, where $\sum'_{A, B}$ denotes the summation over the A, B configurations with non-zero $\chi(A, B)$. Combining with the calculation of Z , one has

$$|\rho^\Gamma|_1 = \langle + | \psi \rangle \frac{\sum'_A \sum'_B e^{\beta_z \sum_{i \in R_A} A_i}}{e^{\beta_z \sum_{i \in R_A} A_i}} \quad (\text{A12})$$

This can be further simplified by noticing that $\langle + | \psi \rangle \sum'_B = \sum_B \langle + | \prod_{j \in R_B} Z_{B,j}^{\frac{1-B_j}{2}} | \psi \rangle = \langle + | \prod_{j \in R_B} (1 + Z_{B,j}) \prod_{\langle i \in R_A, j \in R_B \rangle} CZ_{i,j} |+\rangle = \langle + | \prod_{j \in R_B} (1 + Z_{B,j}) |+\rangle = 1$. Consequently,

$$|\rho^\Gamma|_1 = \frac{\sum'_A e^{\beta_z \sum_{i \in R_A} A_i}}{\sum'_A e^{\beta_z \sum_{i \in R_A} A_i} \prod_{i \in R_A} \delta(A_i = 1)} \quad (\text{A13})$$

As discussed above, $\{A\}$ needs to obey the constraint $\prod_{i \in \Gamma_A^*} A_i = 1$, and this can be resolved by defining $A_i = \prod_{j \in \partial^* i} \tau_j$, where $\{\tau_j\}$ are independent Ising variables defined on R_B . As a result, $|\rho^\Gamma|_1 = \frac{Z_B}{Z_B}$

$$\begin{aligned} Z_B &= \sum_{\tau} e^{\beta_z \sum_{i \in R_A} A_i} \prod_{j \in \partial^* i} \tau_j \\ \tilde{Z}_B &= \sum_{\tau} e^{\beta_z \sum_{i \in R_A} A_i} \prod_{j \in \partial^* i} \tau_j \prod_{i \in R_A} \delta\left(\prod_{j \in \partial^* i} \tau_j = 1\right) \end{aligned} \quad (\text{A14})$$

Hence the entanglement negativity is the free energy difference associated with forbidding excitations in a certain classical statistical mechanics model:

$$E_N = \log Z_B - \log \tilde{Z}_B \quad (\text{A15})$$

2. Pauli-X decoherence

Here we consider the Pauli-X noise applied on the set of qubits (denoted by \mathcal{K}) that only excites the A_i stabilizers on the boundary. At error rate p_x , the decohered mixed state $\rho = \prod_{i \in \mathcal{K}} N_{x,i}[\rho_0]$, with $N_{x,i}[\rho_0] = (1 - p_x)\rho_0 + p_x X_i \rho_0 X_i$, may be expressed as

$$\rho \propto \sum_{a,b} \prod_i A_i^{a_i} \prod_j B_j^{b_j} (1 - 2p)^{W(Z)}. \quad (\text{A16})$$

$W(Z)$ is the number of Pauli-Zs belonging to the set \mathcal{K} that appear in the boundary stabilizer string $\prod_i^{\partial} A_i^{a_i}$, which can be computed as $W(Z) = \frac{1}{2} \sum_{j \in R_B} (1 - \prod_{i \in \partial j} \sigma_i)$, where $\sigma_i = (-1)^{a_i}$, and $\prod_{i \in \partial j} \sigma_i$ denotes the product of σ_i on the boundary of $j \in R_B$.

With a calculation similar to the Pauli-Z decoherence, we derive the negativity spectrum:

$$\rho^\Gamma \propto \sum_{a,b} \prod_{i \in R_A} A_i^{a_i} \prod_{j \in R_B} B_j^{b_j} e^{\beta_x \sum_{j \in R_B} \prod_{i \in \partial j} \sigma_i} \psi(a, b), \quad (\text{A17})$$

with $\beta_x = -\frac{1}{2} \log(1 - 2p_x)$.

Again, we can introduce the fictitious Hilbert space of qubits living on the lattices R_A, R_B spanned by the Z basis product states $|a, b\rangle = |\{a_i, b_j\}\rangle$. Defining a state $|\psi\rangle \propto \sum_{a,b} \psi(a, b) |a, b\rangle$ and a trivial state $|+\rangle \propto \sum_{a,b} |a, b\rangle$, the negativity spectrum can be written as

$$\rho^\Gamma \propto \langle + | \prod_{i \in R_A} Z_{A,i}^{\frac{1-A_i}{2}} \prod_{j \in R_B} Z_{B,j}^{\frac{1-B_j}{2}} e^{\beta_x \sum_{j \in R_B} \prod_{i \in \partial j} Z_{A,i}} | \psi \rangle, \quad (\text{A18})$$

By employing the condition $\prod_{i \in \partial j} Z_{A,i} |\psi\rangle = X_{B,j} |\psi\rangle$, the negativity spectrum can be simplified as:

$$\begin{aligned} \rho^\Gamma &\propto \langle + | \prod_{i \in R_A} Z_{A,i}^{\frac{1-A_i}{2}} \prod_{j \in R_B} Z_{B,j}^{\frac{1-B_j}{2}} e^{\beta_x \sum_{j \in R_B} X_{B,j}} | \psi \rangle \\ &= \langle + | \prod_{i \in R_A} Z_{A,i}^{\frac{1-A_i}{2}} \prod_{j \in R_B} Z_{B,j}^{\frac{1-B_j}{2}} | \psi \rangle e^{\beta_x \sum_{j \in R_B} B_j}. \end{aligned} \quad (\text{A19})$$

Finally, the negativity spectrum reads

$$\rho^\Gamma = \frac{1}{Z} \chi(A, B) e^{\beta_x \sum_{j \in R_B} B_j}, \quad (\text{A20})$$

where $Z = \sum_{A,B} \chi(A, B) e^{\beta_x \sum_{j \in R_B} B_j}$ is a normalization constant, and $\chi(A, B) = \frac{\langle + | \prod_{i \in R_A} Z_{A,i}^{\frac{1-A_i}{2}} \prod_{j \in R_B} Z_{B,j}^{\frac{1-B_j}{2}} | \psi \rangle}{\langle + | \psi \rangle} = 1, -1$ for $\{A_i, B_j\}$ configurations that satisfy $\prod_{i \in \Gamma_A^*} A_i = 1$ and $\prod_{j \in \Gamma_B} B_j = 1$.

With the negativity spectrum, by following a calculation similar to Appendix A 1, we derive the entanglement negativity

$$|\rho^\Gamma|_1 = \frac{\sum'_B e^{\beta_x \sum_{j \in R_B} B_j}}{\sum'_B e^{\beta_x \sum_{j \in R_B} B_j} \prod_{j \in R_B} \delta(B_j = 1)} \quad (\text{A21})$$

As discussed above, $\{B_j\}$ needs to obey the constraint $\prod_{j \in \Gamma_B} B_j = 1$, and this can be resolved by defining $B_j = \prod_{i \in \partial j} \sigma_i$, where $\{\sigma_i\}$ are independent Ising variables defined on R_A . As a result, $|\rho^\Gamma|_1 = \frac{Z_A}{Z_A}$

$$\begin{aligned}
Z_A &= \sum_{\sigma} e^{\beta_x \sum_{j \in R_B} \prod_{i \in \partial j} \sigma_i} \\
\tilde{Z}_A &= \sum_{\sigma} e^{\beta_x \sum_{j \in R_B} \prod_{i \in \partial j} \sigma_i} \prod_{j \in R_B} \delta\left(\prod_{i \in \partial j} \sigma_i = 1\right).
\end{aligned} \tag{A22}$$

Hence the entanglement negativity is the free energy difference associated with forbidding excitations in a classical statistical mechanics model:

$$E_N = \log Z_A - \log \tilde{Z}_A \tag{A23}$$

Appendix B: Calculation of SPT order parameters

Since the negativity spectrum is given by the wave functions of an SPT-cluster state under certain perturbations, it is natural to explore the connection between the persistence of long-range entanglement and the robustness of the SPTs. Here we show that these two are indeed intimately connected.

1. Pauli-Z decoherence

In this case, the negativity spectrum (Eq.A6) is given by the wave functions in X product-state basis of the perturbed state $|\tilde{\psi}\rangle \propto e^{\beta_z \sum_{i \in R_A} \prod_{j \in \partial^* i} Z_{B,j}} |\psi\rangle$, with $|\psi\rangle = U_{CZ} |+\rangle$ and $U_{CZ} = \prod_{\langle i \in R_A, j \in R_B \rangle} CZ_{i,j}$. The potential singularity of $|\tilde{\psi}\rangle$ may be detected by computing the overlap $\langle \tilde{\psi} | \tilde{\psi} \rangle \propto \langle + | U_{CZ}^\dagger e^{2\beta_z \sum_{i \in R_A} \prod_{j \in \partial^* i} Z_{B,j}} U_{CZ} |+\rangle = \langle + | e^{2\beta_z \sum_{i \in R_A} \prod_{j \in \partial^* i} Z_{B,j}} |+\rangle \propto \sum_{\tau} e^{2\beta_z \sum_{i \in R_A} \prod_{j \in \partial^* i} \tau_j}$, which is the partition function of the statistical mechanics model that characterizes the entanglement negativity of the decohered mixed state (Eq.A15), albeit at different inverse temperatures. This provides a direct connection between the disentangling transition of the decohered state and the stability of the emergent SPTs in the fictitious Hilbert space.

The stability of the emergent SPTs can also be diagnosed by physical observables. For the 1d perturbed boundary state $|\tilde{\psi}\rangle \propto e^{\beta_z \sum_i Z_{B,i} Z_{B,i+1}} |\psi\rangle$ that appears in the negativity spectrum of the 2d toric code, the $\mathbb{Z}_2 \times \mathbb{Z}_2$ SPT order manifests in two types of string order parameters: $S_A(i, j) = Z_{B,i} X_{A,i+1} \cdots X_{A,j} Z_{B,j}$ and $S_B(i, j) = Z_{A,i} X_{B,i} \cdots X_{B,j-1} Z_{A,j}$. For $S_A(i, j)$ operators, since it commutes with the perturbation $e^{\beta_z \sum_i Z_{B,i} Z_{B,i+1}}$, one finds $\langle \tilde{\psi} | S_A(i, j) | \tilde{\psi} \rangle = 1$, which is insensitive to the perturbation. On the other hand, $\langle S_B(i, j) \rangle$ will be perturbed; since $S_B(i, j) e^{\beta_z \sum_n Z_{B,n} Z_{B,n+1}} = e^{\beta_z \sum_n J_n Z_{B,n} Z_{B,n+1}} S_B(i, j)$ where $J_n = -1$ for $n = i-1, j-1$, and $J_n = 1$ otherwise, one finds

$$\langle S_B(i, j) \rangle = \frac{\langle \psi | e^{2\beta_z \sum_{n \neq i-1, j-1} Z_{B,n} Z_{B,n+1}} | \psi \rangle}{\langle \psi | e^{2\beta_z \sum_n Z_{B,n} Z_{B,n+1}} | \psi \rangle} \tag{B1}$$

Using the condition $|\psi\rangle = U_{CZ} |+\rangle$ with U_{CZ} being the depth-1 circuits of neighboring controlled-Z gates that creates the 1d cluster state, one has

$$\begin{aligned}
\langle S_B(i, j) \rangle &= \frac{\langle + | e^{2\beta_z \sum_{n \neq i-1, j-1} Z_{B,n} Z_{B,n+1}} | + \rangle}{\langle + | e^{2\beta_z \sum_n Z_{B,n} Z_{B,n+1}} | + \rangle} \\
&= \frac{\sum_{\tau} e^{2\beta_z \sum_{n \neq i-1, j-1} \tau_n \tau_{n+1}}}{\sum_{\tau} e^{2\beta_z \sum_n \tau_n \tau_{n+1}}},
\end{aligned} \tag{B2}$$

which is the ratio of the two partition functions of the 1d Ising model, with and without the Ising couplings at two endpoints of a string. Due to the absence of the finite-temperature transition of the 1d Ising model, this quantity is non-singular at any finite temperature, indicating the persistence of the SPT order in the state $e^{\beta_z \sum_i Z_{B,i} Z_{B,i+1}} |\psi\rangle$ for any non-infinite β_z . Indeed, it is straightforward to evaluate the partition functions and find $\langle S_B(i, j) \rangle = \frac{1}{[\cosh(2\beta_z)]^2 [1 + \tanh(2\beta_z)]^L} > 0$ for any finite β_z .

2. Pauli-X decoherence

In this case, the negativity spectrum (Eq.A18) is given by the wave functions in X product-state basis of the perturbed state $|\tilde{\psi}\rangle \propto e^{\beta_x \sum_{j \in R_B} \prod_{i \in \partial j} Z_{A,i}} |\psi\rangle$, with $|\psi\rangle = U_{CZ} |+\rangle$ and $U_{CZ} = \prod_{\langle i \in R_A, j \in R_B \rangle} CZ_{i,j}$. Again, the potential singularity of $|\tilde{\psi}\rangle$ may be detected by computing the overlap $\langle \tilde{\psi} | \tilde{\psi} \rangle \propto \langle + | U_{CZ}^\dagger e^{2\beta_x \sum_{j \in R_B} \prod_{i \in \partial j} Z_{A,i}} U_{CZ} | + \rangle = \langle + | e^{2\beta_x \sum_{j \in R_B} \prod_{i \in \partial j} Z_{A,i}} | + \rangle \propto \sum_{\sigma} e^{2\beta_x \sum_{j \in R_B} \prod_{i \in \partial j} \sigma_i}$, which is the partition function of the statistical mechanics model that describes the negativity of the decohered mixed state (Eq.A23), albeit at different inverse temperatures. This provides a direct connection between the disentangling transition of the decohered state and the stability of the emergent SPTs in the fictitious Hilbert space.

Appendix C: Quantum-classical mapping

As discussed in the main text, the topological entanglement negativity of the 4d toric code subject to boundary decoherence relates to the subleading term of the free energy of the 3d classical Ising gauge theory at finite temperature, which is $2 \log 2$ at the low-temperature (deconfined) phase and 0 at the high-temperature (confined) phase. Here, using a standard quantum-classical mapping (see e.g. [47, 50]), we derive this subleading term by mapping the 3d classical Ising gauge theory at finite temperature to the 2d quantum Ising gauge theory at zero temperature.

First, we start from the 2d quantum Ising gauge theory described by the Hamiltonian $H = -\sum_p \prod_{e \in \partial p} Z_e - \Gamma \sum_e X_e$, where the first term denotes the product of Pauli-Zs on the four edges around a plaquette p , and the second term is the onsite transverse field. We compute $\text{tr} [e^{\tilde{\beta} H} G]$, where $G = \prod_v \delta_v (X X X X = 1)$ enforces the Gauss-law constraint that the product of Pauli-Xs on the four edges emanating from any vertex v is fixed at 1. To proceed, one can follow a standard Trotterization by inserting a complete Pauli-Z basis:

$$\begin{aligned} \text{tr} [e^{\tilde{\beta} H} G] &= \lim_{M \rightarrow \infty} \text{tr} \left[G e^{\frac{\tilde{\beta} \Gamma}{M} \sum_e X_e} e^{\frac{\tilde{\beta}}{M} \sum_p \prod_{e \in \partial p} Z_e} \right]^M \\ &= \lim_{M \rightarrow \infty} \sum_{\{s_{e,\tau}^z\}} \prod_{\tau=1}^M \langle \{s_{e,\tau+1}^z\} | G e^{\frac{\tilde{\beta} \Gamma}{M} \sum_e X_e} e^{\frac{\tilde{\beta}}{M} \sum_p \prod_{e \in \partial p} Z_e} | \{s_{e,\tau}^z\} \rangle \end{aligned} \quad (\text{C1})$$

where τ labels the coordinate in the imaginary time direction, with the periodic boundary conditions imposed: $s_{e,M+1}^z \equiv s_{e,1}^z$. For each matrix element, one has

$$\langle \{s_{e,\tau+1}^z\} | G e^{\frac{\tilde{\beta} \Gamma}{M} \sum_e X_e} e^{\frac{\tilde{\beta}}{M} \sum_p \prod_{e \in \partial p} Z_e} | \{s_{e,\tau}^z\} \rangle = \langle \{s_{e,\tau+1}^z\} | G e^{\frac{\tilde{\beta} \Gamma}{M} \sum_e X_e} | \{s_{e,\tau}^z\} \rangle e^{\frac{\tilde{\beta}}{M} \sum_p \prod_{e \in \partial p} s_{e,\tau}^z}. \quad (\text{C2})$$

By inserting a complete Pauli-X basis, the above quantity can be expressed as

$$\begin{aligned} &\sum_{\{s_{e,\tau}^x\}} \langle \{s_{e,\tau+1}^z\} | \{s_{e,\tau}^x\} \rangle \langle \{s_{e,\tau}^x\} | \{s_{e,\tau}^z\} \rangle e^{\frac{\tilde{\beta} \Gamma}{M} \sum_e s_{e,\tau}^x} e^{\frac{\tilde{\beta}}{M} \sum_p \prod_{e \in \partial p} s_{e,\tau}^z} \prod_v \frac{1 + \prod_{e \in \partial^* v} s_{e,\tau}^x}{2} \\ &= 2^{-2L^2} \sum_{\{s_{e,\tau}^x\}} \prod_e (s_{e,\tau}^x)^{\frac{1-s_{e,\tau}^z}{2} + \frac{1-s_{e,\tau+1}^z}{2}} e^{\frac{\tilde{\beta} \Gamma}{M} \sum_e s_{e,\tau}^x} e^{\frac{\tilde{\beta}}{M} \sum_p \prod_{e \in \partial p} s_{e,\tau}^z} 2^{-L^2} \sum_{\{\lambda_{v,\tau+\frac{1}{2}}\}} \prod_v \left(\prod_{e \in \partial^* v} s_{e,\tau}^x \right)^{\frac{1-\lambda_{v,\tau+\frac{1}{2}}}{2}}. \end{aligned} \quad (\text{C3})$$

Note that we introduce $\lambda_{v,\tau+\frac{1}{2}} = \pm 1$ as a Lagrange multiplier to resolve the Gauss law constraint, and these variables will become the spins on the edge along the imaginary time direction, in between the imaginary time τ and $\tau+1$.

Now one can sum over $s_{e,\tau}^x$; for a given $s_{e,\tau}^x$, one computes the summation:

$$\begin{aligned} &\sum_{s_{e,\tau}^x} (s_{e,\tau}^x)^{\frac{1-s_{e,\tau}^z}{2} + \frac{1-s_{e,\tau+1}^z}{2} + \frac{1-\lambda_{v,\tau+\frac{1}{2}}}{2} + \frac{1-\lambda_{v',\tau+\frac{1}{2}}}{2}} e^{\frac{\tilde{\beta} \Gamma}{M} s_{e,\tau}^x} \\ &= e^{\frac{\tilde{\beta} \Gamma}{M}} + s_{e,\tau}^z s_{e,\tau+1}^z \lambda_{v,\tau+\frac{1}{2}} \lambda_{v',\tau+\frac{1}{2}} e^{-\frac{\tilde{\beta} \Gamma}{M}} \\ &= e^{\beta_\tau s_{e,\tau}^z s_{e,\tau+1}^z \lambda_{v,\tau+\frac{1}{2}} \lambda_{v',\tau+\frac{1}{2}}} \end{aligned} \quad (\text{C4})$$

with v, v' being two vertices on the boundary of the edge e , and $\beta_\tau \equiv -\frac{1}{2} \log[\tanh(\frac{\tilde{\beta}\Gamma}{M})]$. This term can be understood as the four-spin coupling around the plaquette spanned by one space direction and one temporal direction. Using the results above, finally, we find

$$\text{tr} [e^{\tilde{\beta}H} G] = \lim_{M \rightarrow \infty} 2^{-3L^2 M} \sum_{\sigma_e} e^{\sum_p K_p \sigma \sigma \sigma \sigma}, \quad (\text{C5})$$

where $K_p = \frac{\tilde{\beta}}{M}$ for plaquettes spanned by two spatial directions, and $K_p = \beta_\tau = -\frac{1}{2} \log[\tanh(\frac{\tilde{\beta}\Gamma}{M})]$ for the plaquettes spanned by one spatial direction and one temporal direction.

Taking the limit $\tilde{\beta} \rightarrow \infty$ with $\tilde{\beta}/M = \beta = O(1)$ fixed, the L.H.S. reduces to the ground-subspace projector, and the R.H.S. becomes the classical 3d Ising Gauge theory at finite temperature β :

$$N_g e^{-\tilde{\beta}E_g} = 2^{-3L^2 \tilde{\beta}\beta} \sum_{\sigma_e} e^{\beta \sum_p \sigma \sigma \sigma \sigma}, \quad (\text{C6})$$

where N_g, E_g are the ground-state degeneracy and energy of the 2d quantum gauge theory, and we have also considered the isotropic limit by setting $\beta = -\frac{1}{2} \log[\tanh(\beta\Gamma)]$. This completes the mapping between the 3d classical Ising gauge theory at finite temperature and the 2d quantum Ising gauge theory at zero temperature. In particular, tuning the inverse temperature β in the 3d classical model amounts to tuning Γ in the 2d quantum model, which exhibits a confinement-deconfinement transition at a critical Γ_c , with $N_g = 4$ ($\Gamma < \Gamma_c$), and $N_g = 1$ ($\Gamma > \Gamma_c$). This implies the free energy of the 3d classical model takes the form $\log Z = L^3 f + \log g$, with $g = 4$ in the low-temperature phase and $g = 1$ in the high-temperature phase.

Appendix D: Toric code subject to both X,Z decoherence

Here we consider the 2d toric code under both Pauli-Z and Pauli-X boundary decoherence with noise rate p_z and p_x , and discuss the negativity spectrum.

In the 2d toric code, along a 1d boundary of size L , there are L star and L plaquette operators, which we label in order: $A_1, B_1, A_2, B_2, \dots$. The Z decoherence occurs on the edge shared between two adjacent B_i, B_{i+1} plaquette stabilizers, and the X decoherence occurs on the vertices shared between two adjacent A_i, A_{i+1} star stabilizers. Correspondingly, the boundary part of the decohered mixed state reads

$$\rho_\partial \propto \sum_{a,b} \prod_i A_i^{a_i} \prod_i B_i^{b_i} e^{\beta_x \sum_{i=1}^L \sigma_i \sigma_{i+1}} e^{\beta_z \sum_{i=1}^L \tau_i \tau_{i+1}} \quad (\text{D1})$$

with $\sigma_i = (-1)^{a_i} \in \{1, -1\}$ and $\tau_i = (-1)^{b_i} \in \{1, -1\}$. Therefore, the negativity spectrum can be expressed as

$$\rho^\Gamma \propto \langle + | \prod_i Z_{A,i}^{\frac{1-A_i}{2}} \prod_i Z_{B,i}^{\frac{1-B_i}{2}} e^{\beta_x \sum_i Z_{A,i} Z_{A,i+1}} e^{\beta_z \sum_i Z_{B,i} Z_{B,i+1}} | \psi \rangle, \quad (\text{D2})$$

where $|\psi\rangle$ is the 1d cluster state defined in the fictitious Hilbert space.

This may be simplified by using the condition $Z_{B,i} Z_{B,i+1} |\psi\rangle = X_{A,i+1} |\psi\rangle$ and $Z_{A,i} Z_{A,i+1} |\psi\rangle = X_{B,i} |\psi\rangle$:

$$\begin{aligned} e^{\beta_x \sum_i Z_{A,i} Z_{A,i+1}} e^{\beta_z \sum_i Z_{B,i} Z_{B,i+1}} |\psi\rangle &= e^{\beta_x \sum_i Z_{A,i} Z_{A,i+1}} e^{\beta_z \sum_i X_{A,i}} |\psi\rangle \\ &= (\cosh \beta_x)^L \sum_\tau \prod_i (\tanh \beta_x Z_{A,i} Z_{A,i+1})^{\frac{1-\tau_i}{2}} e^{\beta_z \sum_i X_{A,i}} |\psi\rangle \\ &= (\cosh \beta_x)^L \sum_\tau e^{\beta_z \sum_i \tau_{i-1} \tau_i X_{A,i}} \prod_i (\tanh \beta_x X_{B,i})^{\frac{1-\tau_i}{2}} |\psi\rangle, \end{aligned} \quad (\text{D3})$$

with $\tau \equiv \{\tau_i = \pm 1\}$. As a result, we obtain the final expression of the negativity spectrum:

$$\rho^\Gamma = \frac{1}{Z} \chi(A, B) \sum_\tau e^{\beta_z \sum_i \tau_{i-1} \tau_i A_i + K_x \sum_i \tau_i} \prod_i \tau_i^{\frac{1-B_i}{2}}, \quad (\text{D4})$$

where $K_x = -\frac{1}{2} \log(\tanh \beta_x)$, Z is a normalization constant to ensure that negativity spectrum sums to one, and $\chi(A, B) = \frac{\langle + | \prod_i Z_{A,i}^{\frac{1-A_i}{2}} \prod_i Z_{B,i}^{\frac{1-B_i}{2}} | \psi \rangle}{\langle + | \psi \rangle}$ is the strange correlator (or equivalently, the ratio of wave functions in X basis) of the 1d cluster state. This expression indicates that the negativity spectrum is given by $\chi(A, B)$ multiplied by the τ spin correlations in a 1d Ising model under a symmetry-breaking field, where A_i, B_i determine the sign of the Ising coupling and spin insertion, respectively. In particular, the sign structure of the negativity spectrum is no longer solely determined by $\chi(A, B)$, which poses a technical difficulty in deriving the entanglement negativity directly from the negativity spectrum.