Strong-to-Weak Spontaneous Symmetry Breaking in Mixed Quantum States

Leonardo A. Lessa,[1,2,†,¶] Ruochen Ma,[3,¶] Jian-Hao Zhang,[4,5,*,¶] Zhen Bi,[4,‡] Meng Cheng,[6,§] and Chong Wang[1,∥]

[1] *Perimeter Institute for Theoretical Physics, Waterloo, Ontario N2L 2Y5, Canada*
[2] *Department of Physics and Astronomy, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada*
[3] *Kadanoff Center for Theoretical Physics, University of Chicago, Chicago, Illinois 60637, USA*
[4] *Department of Physics, The Pennsylvania State University, University Park, Pennsylvania 16802, USA*
[5] *Department of Physics and Center for Theory of Quantum Matter, University of Colorado, Boulder, Colorado 80309, USA*
[6] *Department of Physics, Yale University, New Haven, Connecticut 06511-8499, USA*

(Received 3 July 2024; accepted 9 January 2025; published 6 March 2025)

Symmetry in mixed quantum states can manifest in two distinct forms: *strong symmetry*, where each individual pure state in the quantum ensemble is symmetric with the same charge, and *weak symmetry*, which applies only to the entire ensemble. This paper explores a novel type of spontaneous symmetry breaking (SSB) where a strong symmetry is broken to a weak one. While the SSB of a weak symmetry is measured by the long-ranged two-point correlation function, the strong-to-weak SSB (SWSSB) is measured by the *fidelity correlator*. We prove that SWSSB is a universal property of mixed-state quantum phases, in the sense that the phenomenon of SWSSB is robust against symmetric low-depth local quantum channels. We also show that the symmetry breaking is "spontaneous" in the sense that the effect of a local symmetry-breaking measurement cannot be recovered locally. We argue that a thermal state at a nonzero temperature in the canonical ensemble (with fixed symmetry charge) should have spontaneously broken strong symmetry. Additionally, we study nonthermal scenarios where decoherence induces SWSSB, leading to phase transitions described by classical statistical models with bond randomness. In particular, the SWSSB transition of a decohered Ising model can be viewed as the "ungauged" version of the celebrated toric-code decodability transition. We confirm that, in the decohered Ising model, the SWSSB transition defined by the fidelity correlator is the only physical transition in terms of channel recoverability. We also comment on other (inequivalent) definitions of SWSSB, through correlation functions with higher Rényi indices.

DOI: 10.1103/PRXQuantum.6.010344

I. INTRODUCTION

The notion of *spontaneous symmetry breaking* (SSB) is a cornerstone of modern physics [1,2]. SSB in quantum physics has been well understood for pure states—typically the ground states of some many-body Hamiltonians—and for Gibbs thermal states $\rho = e^{-\beta H}/Z$. Surprisingly, recent advances in understanding phases of matter in mixed quantum states have revealed a novel type of SSB dubbed strong-to-weak SSB (SWSSB) [3,4]. This work aims to establish some key universal aspects of SWSSB and associated phase transitions.

In quantum mechanics, symmetry acts on pure states via multiplication by the corresponding unitary (or antiunitary) U and acts on operators via conjugation. For an open quantum system described by a density operator ρ, however, depending on the coupling to the environment there are two ways that symmetry can act [5–8]. One can have a *strong symmetry* (also called *exact symmetry*), defined as

$$U\rho = e^{i\theta}\rho, \tag{1}$$

which means that, when ρ is viewed as an ensemble, every member state carries the same charge under the symmetry.

*Contact author: Sergio.Zhang@colorado.edu, zhangjianhao11@gmail.com
†Contact author: llessa@pitp.ca
‡Contact author: zjb5184@psu.edu
§Contact author: m.cheng@yale.edu
∥Contact author: cwang4@pitp.ca
¶These authors contributed equally to this work.

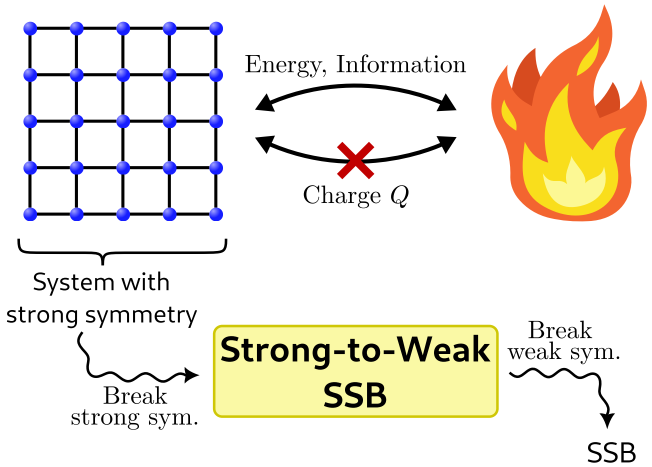

FIG. 1. Main scenario considered in this work: a mixed-state system ρ with strong symmetry $U\rho \propto \rho$, such as one interacting with a (thermal) environment without exchanging charge, that spontaneously breaks the strong symmetry but not the weak symmetry. When this occurs, we call it a strong-to-weak SSB state, detected via a long-range correlator $F_O(x,y) \sim O(1)$. If the weak symmetry is further broken down, then the system exhibits the usual SSB, $\langle O(x)O^\dagger(y)\rangle_\rho \sim O(1)$.

Alternatively, the symmetry can be just *weak* (or *average*), and not strong, defined as

$$U\rho \neq e^{i\theta}\rho, \quad U\rho U^\dagger = \rho, \tag{2}$$

which can happen in an ensemble whose member states have well-defined but different symmetry charges. Physically, strong symmetry arises when the system does not exchange charges with the environment, as illustrated in Fig. 1. These concepts are deeply rooted in equilibrium statistical mechanics: canonical ensembles have "strong" particle-number conservation symmetry, while grand canonical ensembles are only weakly symmetric.

Given these two ways that symmetries can be realized in a mixed state, there is a need to revisit the notion of spontaneous symmetry breaking, which serves as a fundamental organizing principle in quantum many-body physics. In order to clarify the matter, consider a many-body state ρ with global symmetry G, and a charged local operator $O(x)$. Conventional SSB can be defined through the long-range order $\mathrm{Tr}[O(x)O^\dagger(y)\rho] \to \text{const} \neq 0$ when $|x-y| \to \infty$. In this case, regardless of whether the symmetry action g is strong or weak, it is spontaneously broken to nothing. Now given the necessity to distinguish strong and weak symmetries, there is a distinct possibility that a strong symmetry can be spontaneously broken into a weak one. This phenomenon of strong-to-weak SSB (SWSSB) is only possible for mixed states. The relation between the two notions of SSB is illustrated in Fig. 1. In Sec. II we carefully define the notion of SWSSB and study its key properties. Our central result, dubbed the *stability theorem* (Theorems 1 and 2), establishes SWSSB as a universal property of mixed-state quantum phases that is robust against low-depth symmetric local quantum channels.

While it might sound unfamiliar, recall for spin glass order in disordered systems, the ensemble-averaged order parameter vanishes $\overline{\langle O(x)\rangle} = 0$, but each individual state in the disordered ensemble has symmetry-breaking order, reflected in the Edward-Anderson [9] order parameter $\overline{|\langle O(x)\rangle|} \neq 0$. In our language, this means that the global symmetry [under which $O(x)$ is charged] is broken as an exact (strong) symmetry, but is preserved as an average (weak) symmetry. We will explain in Sec. II A that our SWSSB is, in a precise sense, a mixed-state generalization of the Edward-Anderson type of spin glass order.

Another familiar example comes from equilibrium quantum statistical mechanics: the fact that canonical and grand canonical ensembles are completely equivalent in the thermodynamic limit in fact is nothing but SWSSB (we will explain this point in Sec. III A). Therefore, to some extent, SWSSB should be viewed as a fundamental property of thermal states with strong symmetry.

Recently, nonthermal mixed states have received much attention thanks to the rapid progress in engineering and controlling complex many-body states in quantum devices, motivating many investigations into quantum phases in mixed states [3,10–31]. While the full landscape of mixed-state phases is still largely uncharted, novel topological states protected by both strong and weak symmetries have been identified, including examples that have no counterpart in pure states. Given the central role SSB has played in the theory of the ground state (or equilibrium) phases, understanding SWSSB is clearly a foundational issue for mixed-state quantum phases. For example, to define mixed-state symmetry-protected topological phases [4,32–40], it is necessary to impose the absence of any symmetry breaking, including the SWSSB [3,4,37].

The simplest setting of SWSSB in a nonthermal state is to start from a symmetric pure state and introduce local symmetric decoherence. In Secs. III B and III C we study two such examples, one with Ising $\mathbb{Z}_2$ symmetry, the other with $U(1)$. In both cases, we find phase transitions into SWSSB at some finite decoherence strength p_c, with the universality classes described by some classical statistical mechanical models with bond randomness. In particular, the celebrated decodability transition of $\mathbb{Z}_2$ toric code can be viewed as the "gauged" version (through Kramers-Wannier duality) of the Ising SWSSB transition. For the decohered Ising model, we also show (Sec. IV) that the only physical transition is the one involving SWSSB. In particular, using the recently proposed local Petz recovery channel [16], we show that two states in the same phase (both with or without SWSSB) are two-way connected through symmetric low-depth local quantum channels.

The rest of this paper is organized as follows. In Sec. II, we propose formal definitions of SWSSB and establish a few basic universal properties. In particular, we prove that SWSSB, as in Definition 1, is robust under low-depth local channels, and thus can be used as a universal characterization of mixed-state quantum phases. In Sec. III A we discuss SWSSB for thermal states in the canonical ensemble at nonzero temperatures. We then explore a number of examples of SWSSB transitions in decohered and noisy many-body states in Secs. III B and III C. In Sec. IV, we discuss the recoverability of the decohered Ising model. We show that as long as two states belong to the same phase (both with or without the SWSSB order), they are two-way connected through log-depth symmetric local quantum channels. We end with a summary and outlook in Sec. V. Some peripheral details are presented in the appendices.

II. DEFINITION AND PROPERTIES

In this section, we give two inequivalent definitions of SWSSB order parameters: the fidelity correlator and Rényi-2 correlator of local operators carrying strong symmetry charge. Then we discuss some characteristic properties of SWSSB mixed state. We will see that, while the Rényi-2 correlator is in general simpler to calculate, the fidelity correlator offers a more fundamental characterization of the universal properties of SWSSB.

A. Fidelity correlator and stability theorem

Recall that the fidelity [41] between two states ρ and σ is defined as

$$F(\rho,\sigma) = \mathrm{Tr}\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}. \tag{3}$$

When both ρ and σ are pure states, $F(\rho,\sigma)$ becomes the modulus of their overlap. Fidelity is a commonly used measure of distance between mixed states. Motivated by fidelity, we define the *fidelity* average of an operator O in a mixed state ρ as follows:

$$\langle O\rangle_F = F(\rho, O\rho O^\dagger). \tag{4}$$

Again, if $\rho = |\psi\rangle\langle\psi|$, then $\langle O\rangle_F = |\langle\psi|O|\psi\rangle|$, the modulus of the expectation value. But for a generic mixed state, the fidelity average is very different from the usual expectation value defined as $\langle O\rangle = \mathrm{Tr}[\rho O]$. If O is unitary, then the above expression is the actual fidelity between two different density matrices ρ and $\sigma := O\rho O^\dagger$. If O is charged under a strong symmetry of ρ, then $\langle O\rangle_F = 0$. blackNotice that

$$\begin{aligned}\langle O\rangle_F &= \mathrm{Tr}\sqrt{\sqrt{\rho}O\rho O^\dagger\sqrt{\rho}} = \|\sqrt{\rho}O\sqrt{\rho}\|_1\\ &> |\,\mathrm{Tr}\sqrt{\rho}O\sqrt{\rho}| = |\langle O\rangle|,\end{aligned} \tag{5}$$

which implies that if $\langle O\rangle \neq 0$, then we must have $\langle O\rangle_F > 0$.

For a local operator $O(x)$, define the *fidelity correlator* as the fidelity average of $O(x)O^\dagger(y)$, or more explicitly,

$$F_O(x,y) = \langle O(x)O^\dagger(y)\rangle_F\,. \tag{6}$$

Definition 1. A mixed state ρ with a strong symmetry G has SWSSB if the fidelity correlator is long-range ordered:

$$\lim_{|x-y|\to\infty} F_O(x,y) = c > 0, \tag{7}$$

for some charged local operator $O(x)$, and at the same time there is no long-range order in the usual sense, i.e. for any charged local operator O', $\langle O'(x)O'^\dagger(y)\rangle_\rho \to 0$ as $|x-y|\to\infty$.

Hereinafter, we will denote $\lim_{|x-y|\to\infty} F_O(x,y) = c > 0$ by $F_O(x,y)\sim O(1)$, and similarly to other correlators.

The simplest case of a charged local operator $O(x)$ is when it transforms under the symmetry group G as $gO(x)g^{-1} = \lambda_g O(x)$, for some $\lambda_g \in U(1)$. If $O_a(x)$ is part of a higher-dimensional unitary representation of G, such as the spin operators $S^{(\alpha)}$ of SO(3), then we can replace $O(x)O^\dagger(y)$ by the quadratic symmetry-invariant operator $\sum_a O_a(x)O_a^\dagger(y)$.

We now discuss the simplest example in an Ising spin system. The prototypical state with SWSSB is $\rho_0 \propto \mathbb{1} + X$ with a strong $\mathbb{Z}_2$ symmetry generated by $X \equiv \prod_i X_i$. ρ_0 is analogous to the GHZ cat state $|00\cdots\rangle + |11\cdots\rangle$ of the usual $\mathbb{Z}_2$ SSB. In particular, if we write ρ_0 in the Z basis, it is a convex sum of cat states:

$$\rho_0 \propto \sum_s (|s\rangle + X|s\rangle)(\langle s| + \langle s|X), \tag{8}$$

where $s\in\{0,1\}^N$ labels the bit string of spins in the Z basis with N being the total number of spins. It is easy to see that the fidelity correlator $F_Z(x,y)$ for the order parameter $Z(x)$ is exactly equal to 1, independent of x and y.

More generally, let us assume ρ is a mixture of orthogonal pure states: $\rho = \sum_a p_a|\psi_a\rangle\langle\psi_a|$, such that $\langle\psi_a|O(x)O^\dagger(y)|\psi_b\rangle \propto \delta_{ab}$. Then we can write the fidelity correlator as

$$\sum_a p_a\left|\langle\psi_a|O(x)O^\dagger(y)|\psi_a\rangle\right| \equiv \overline{\left|\langle O(x)O^\dagger(y)\rangle\right|}. \tag{9}$$

This is nothing but the Edward-Anderson correlator used for spin glass order if ρ is an ensemble of SSB states. Recall that in a disordered ensemble, the Edward-Anderson order parameter measures exactly the spontaneous breaking of an exact symmetry down to an average symmetry. From this expression we can also see that SWSSB is only applicable to mixed states since the fidelity correlator of pure states is just the (absolute value of) linear correlator, which we assume is not long-range ordered. We

TABLE I. The behavior of SSB diagnostics in a mixed state with a strong symmetry. The operator O is chosen as a generic local operator transformed in a nontrivial representation of the strong symmetry. 0 denotes exponential decay with respect to the distance $|x-y|$.

Symmetry	$F_O(x,y)$	$\mathrm{Tr}(\rho O(x)O^\dagger(y))$
Unbroken	0	0
SWSSB	$O(1)$	0
Completely broken	$O(1)$	$O(1)$

summarize the key patterns of different mixed-state SSB in Table I.

We now study the robustness of SWSSB as we deform our states using low-depth quantum channels. We shall start by defining such channels.

Definition 2. A quantum channel $\mathcal{E}$ is termed a symmetric low-depth channel if it can be purified to a symmetric local unitary operator U acting on $\mathcal{H}\otimes\mathcal{H}^a$, where $\mathcal{H}$ and $\mathcal{H}^a$ are the physical and ancilla Hilbert spaces, respectively. The channel operation $\mathcal{E}(\rho)$ is given by $\mathrm{Tr}_a[U^\dagger(\rho\otimes|0\rangle\langle 0|)U]$, where $|0\rangle$ represents a product state in $\mathcal{H}^a$. Specifically, as follows:

(1) $\mathcal{H}^a=\otimes_i\mathcal{H}_i^a$ is a tensor product of local Hilbert spaces on each site of $\mathcal{H}$. $\dim(\mathcal{H}_i^a)$ is finite for each site.
(2) The unitary circuit U has low depth D. In particular, the channel is called finite depth if $D\sim O(1)$, and logarithmic depth if $D\sim\mathrm{PolyLog}(L)$, where L represents the linear size of the system.
(3) The channel is *strongly symmetric* under the symmetry g if each gate of the purified unitary U commutes with $\tilde{g}\equiv g\otimes I^a$, where I^a is the identity operator on the ancilla space.

The following **stability theorems** establishes SWSSB as a universal property of mixed-state quantum phases.

Theorem 1. If a mixed state ρ has SWSSB [Eq. (7)] and $\mathcal{E}_{\mathrm{SFD}}$ is a strongly symmetric finite-depth (SFD) local quantum channel, then $\mathcal{E}_{\mathrm{SFD}}[\rho]$ will also have SWSSB.

Proof. Recall Uhlmann's theorem [42] that fidelity can be defined in terms of purification as

$$F(\rho,\sigma)=\max_{|\psi_\rho\rangle,|\phi_\sigma\rangle}|\langle\psi_\rho|\phi_\sigma\rangle|, \tag{10}$$

where $|\psi_\rho\rangle,|\phi_\sigma\rangle$ run over all possible purifications of ρ,σ, respectively, and the maximum is taken with respect to all possible purifications. The form of $\sigma=O\rho O^\dagger$ [$O\equiv O(x)O^\dagger(y)$, $O(x)$ and $O(y)$ are charged operators] means that the fidelity can be rewritten as

$$F(\rho,\sigma)=\max_{|\psi_\rho\rangle,|\phi_\rho\rangle}|\langle\psi_\rho|O|\phi_\rho\rangle|. \tag{11}$$

Now we suppose $F(\rho,\sigma)\sim O(1)$ for the mixed state ρ, and consider an SFD channel $\mathcal{E}$, which can be purified to a symmetric finite-depth local unitary operator U acting on the physical system tensor a trivial product state $|0\rangle\langle 0|_a$ in an ancilla Hilbert space $\mathcal{H}^a$. Gathering everything together, we have

$$\begin{aligned}O(1)\sim|\langle\psi_\rho|O|\phi_\rho\rangle| &= |(\langle\psi_\rho|\otimes\langle 0|_a)O(|\phi_\rho\rangle\otimes|0\rangle_a)|\\ &= |(\langle\psi_\rho|\otimes\langle 0|_a)U^\dagger UOU^\dagger U(|\phi_\rho\rangle\otimes|0\rangle_a)|\\ &= |\langle\psi_{\mathcal{E}[\rho]}|O'|\phi_{\mathcal{E}[\rho]}\rangle|,\end{aligned} \tag{12}$$

where $O'=UOU^\dagger=O'(x)O'^\dagger(y)$, and by the SFD nature of U, $O'(x),O'(y)$ are local and carry the same symmetry charge as $O(x),O(y)$. But O' will now act in the ancilla space as well. Decompose $O'=\sum_i O_i\otimes O_i^a$ where O_i acts only in the physical Hilbert space and O_i^a acts only in the ancilla Hilbert space $\mathcal{H}^a$. Crucially, the SFD nature of U means there are only finitely many terms in the decomposition. So there must be at least one i such that

$$\begin{aligned}O(1)&\sim|\langle\psi_{\mathcal{E}[\rho]}|O_i\otimes O_i^a|\phi_{\mathcal{E}[\rho]}\rangle|\\ &\le\|O_i^a\|\cdot|\langle\psi_{\mathcal{E}[\rho]}|O_i\otimes U_a|\phi_{\mathcal{E}[\rho]}\rangle|\\ &\le\|O_i^a\|\cdot\max_{\text{purification}}|\langle\psi'_{\mathcal{E}[\rho]}|O_i|\phi'_{\mathcal{E}[\rho]}\rangle|\\ &=\|O_i^a\|\cdot F\left(\mathcal{E}[\rho],O_i\mathcal{E}[\rho]O_i^\dagger\right),\end{aligned} \tag{13}$$

where U_a is some unitary acting only in the ancilla Hilbert space $\mathcal{H}^a$, and $\|O_i^a\|$ is the operator norm of O_i^a, which is crucially finite. The third line comes from the fact that $U_a|\phi_{\mathcal{E}[\rho]}\rangle$ is just another purification of $\mathcal{E}[\rho]$. So we conclude that the fidelity correlator for O_i on the state $\mathcal{E}[\rho]$ is $O(1)$. By the strong symmetry condition, O_i carries the same charge as O. So $\mathcal{E}[\rho]$ also has SWSSB. ■

Theorem 1 can be extended to more general low-depth channels.

Theorem 2. If a mixed state ρ has SWSSB [Eq. (7)] with $F_O(x,y)\sim O(1)$ for some charged operators $O(x)$, $O(y)$, and $\mathcal{E}_{\mathrm{SFD}}$ is a strongly symmetric depth-D local quantum channel, then $\mathcal{E}_{\mathrm{SFD}}[\rho]$ will have $F_{\tilde{O}}(x,y)\sim\exp(-\mathrm{Poly}(D))$ for some charged local operators $\tilde{O}(x)$, $\tilde{O}(y)$. In particular, for $D\sim\mathrm{PolyLog}(L)$, $F_{\tilde{O}}$ decays subexponentially with system size L.

Proof. We follow the same steps as the proof of Theorem 1 till Eq. (12). Now the evolved operator $O'\equiv UOU^\dagger$ acts nontrivially on two regions X,Y (around the points x, y) with size $|X|, |Y| \sim D^d$, where d is the spatial dimension. We can now decompose the operator O' as

$$O' = \sum_i c_{ik} O_i \otimes O_k^a, \tag{14}$$

where $\{O_i = O_i(x)O_i^\dagger(y)\}$ and $\{O_i^a = O_i^a(x)(O_i^a)^\dagger(y)\}$ are orthogonal families with respect to the Hilbert-Schmidt inner product. We can further choose $\{O_k^a\}$ to be unitary, and $\{O_i\}$ to have the normalization $\mathrm{tr}_{\mathcal{H}}(O_i^\dagger O_j) = \delta_{ij}\, \mathrm{tr}_{\mathcal{H}}(O^\dagger O)$. Recall that $O' = UOU^\dagger$ has the same Hilbert-Schmidt norm as $O \otimes I^a$, so the above normalization requires that $\sum_{ik} |c_{ik}|^2 = 1$. We can now proceed with a refined version of Eq. (13):

$$\begin{aligned} O(1) &\sim \left| \sum_{ik} c_{ik} \langle \psi_{\mathcal{E}[\rho]} | O_i \otimes O_k^a | \phi_{\mathcal{E}[\rho]} \rangle \right| \\ &\leq \sum_{ik} \| c_{ik} O_k^a \| \cdot F\left(\mathcal{E}[\rho], O_i \mathcal{E}[\rho] O_i^\dagger\right), \\ &\leq \left(\sum_{ik} |c_{ik}| \right) \max_i F\left(\mathcal{E}[\rho], O_i \mathcal{E}[\rho] O_i^\dagger\right). \end{aligned} \tag{15}$$

Finally, since $\sum_{ij} |c_{ij}|^2 = 1$, the expression $\sum_{ij} |c_{ij}|$ is upper bounded by the square root of the number of terms in the summation, which is approximately $\exp(\mathrm{Poly}(|X|, |Y|)) \sim \exp(\mathrm{Poly}(D))$. We therefore conclude that for some operator $O_i \equiv \tilde{O}$ acting on the system Hilbert space $\mathcal{H}$, we have

$$F\left(\mathcal{E}[\rho], \tilde{O}\mathcal{E}[\rho]\tilde{O}^\dagger\right) \geq \exp(-\mathrm{Poly}(D)). \tag{16}$$

■

We note that the stability theorems are only concerned with the fidelity correlator, and are not sensitive to the ordinary linear correlator. In fact, the linear correlator in ordinary SSB need not be stable against symmetric finite-depth channels. As a simple example, consider a one-dimensional GHZ state $|\uparrow\uparrow \ldots \uparrow\rangle + |\downarrow\downarrow \ldots \downarrow\rangle$, and perform strong measurements for all the X_i operators. The resulting state is $\mathbb{1} + X$: the strong-to-trivial SSB (measured by ordinary correlator) is eliminated by the strong measurements, but the strong-to-weak SSB survives.

The stability theorem implies that a mixed state ρ with SWSSB cannot be brought to a symmetric pure product state $|00\cdots\rangle$, using a symmetric low-depth channel $\mathcal{E}$, since the latter has $F_O(x, y) = 0$ for any charged operator O. We can make the statement slightly stronger: the state ρ is not only nontrivializable but also *noninvertible*. First let us define the notion of invertible states.

Definition 3. A mixed state on the physical Hilbert space $\rho \in \mathcal{H}$ is *symmetrically invertible* if there exists an ancillary mixed state $\tilde{\rho} \in \tilde{\mathcal{H}}$ (with the symmetry element g acting as $U_g \otimes \tilde{U}_g$), and a pair of symmetric low-depth local channels $\mathcal{E}_T, \mathcal{E}_P : Q(\mathcal{H} \otimes \tilde{\mathcal{H}}) \to Q(\mathcal{H} \otimes \tilde{\mathcal{H}})$ such that

$$\rho \otimes \tilde{\rho} \xrightarrow{\mathcal{E}_T} |00\cdots\rangle\langle 00\cdots| \xrightarrow{\mathcal{E}_P} \rho \otimes \tilde{\rho}, \tag{17}$$

where $|00\cdots\rangle \in \mathcal{H} \otimes \tilde{\mathcal{H}}$ is a symmetric pure product state.

The above definition naturally generalizes the important notion of invertibility in pure states. It has been shown that the notion of invertibility is also important for discussing topological phases in mixed states, in the sense that (1) symmetry-protected (SPT) topological phases in mixed states are well defined only for (bulk) invertible states [4]; and (2) systems with anomalous symmetry (such as those on the boundaries of nontrivial SPT states) cannot be symmetrically invertible [21].

The stability theorems immediately imply that density matrices with SWSSB are not symmetrically invertible:

Corollary 1. A density matrix $\rho \in \mathcal{H}$ with SWSSB is not symmetrically invertible.

Proof. Suppose the (partial) contrary: $\exists \tilde{\rho}$ and a strongly symmetric low-depth quantum channel $\mathcal{E}$, such that

$$\mathcal{E}[\rho \otimes \tilde{\rho}] = |00\cdots\rangle\langle 00\cdots|. \tag{18}$$

Since the pure product state has $F_O(x, y) = 0$ for any local charged operator O, the stability theorem requires $\rho \otimes \tilde{\rho}$ to also have $F_O(x, y) = 0$ for $|x - y|$ much larger than the channel depth D (or exponentially small in $|x - y|$ if the gates in the channel have exponential tails). In particular, for any local charged operator $O_s \equiv O_s(x)O_s^\dagger(y)$ that only acts on the original system $\mathcal{H}$,

$$F(\rho, O_s \rho O_s^\dagger) = F(\rho \otimes \tilde{\rho}, O_s \rho \otimes \tilde{\rho} O_s^\dagger) \to 0, \tag{19}$$

which means the original state ρ cannot have SWSSB. ■

We now comment on some alternative definitions of SWSSB. We note that physically the fidelity defined in Eq. (7) measures how distinguishable ρ and σ are, where $\sigma \equiv O(x)O^\dagger(y)\rho O(y)O^\dagger(x)$. In fact, many other information-theoretic quantities measure the similarity of two mixed states. For example, consider the quantum relative entropy that is defined as

$$S(\rho\|\sigma) = \mathrm{Tr}(\rho \log \rho - \rho \log \sigma). \tag{20}$$

The trace distance is defined as

$$D(\rho, \sigma) = \frac{1}{2}\|\rho - \sigma\|_1. \tag{21}$$

Analogous to Eq. (7), one can define SWSSB based on these measures: a state ρ has SWSSB when the similarity between ρ and σ is nonvanishing and does not depend on $|x-y|$ at large distances. For example, SWSSB in the sense of quantum relative entropy means $S(\rho\|\sigma)$ is upper bounded as $|x-y| \to \infty$.

All these distinguishability measures (fidelity, quantum relative entropy, and trace distance among others) share several properties that make them proper notions of distance between quantum states [43]. Most importantly for our purpose, they satisfy the property known as "data-processing inequality." It says that the distance between two states is monotonic under quantum channels, such that the two states become less distinguishable. Specifically, for two states ρ and σ and a quantum channel $\mathcal{E}$, we have

$$F(\rho, \sigma) \leq F\left(\mathcal{E}[\rho], \mathcal{E}[\sigma]\right). \tag{22}$$

Similar inequalities hold for quantum relative entropy and trace distance. This property directly gives a stability theorem when the order parameter operator O commutes with the quantum channel.

A natural question arises regarding the relationship between SWSSBs defined in terms of the distinguishability measures stated above. Based on the Fuchs-van de Graaf inequality [43]:

$$1 - F(\rho, \sigma) \leq D(\rho, \sigma) \leq \sqrt{1 - F^2(\rho, \sigma)}, \tag{23}$$

the fidelity and trace-distance definitions of SWSSB are entirely equivalent.

Next, we show that if a state ρ has SWSSB defined in terms of quantum relative entropy, it must also have SWSSB defined by fidelity. To this end and for later use, we introduce the sandwiched Rényi relative entropy, defined as

$$\tilde{D}_\alpha(\rho\|\sigma) = \frac{1}{\alpha - 1}\log_2 \mathrm{Tr}[(\sigma^{\frac{1-\alpha}{2\alpha}}\rho\sigma^{\frac{1-\alpha}{2\alpha}})^\alpha]. \tag{24}$$

This assertion that SWSSB defined by quantum relative entropy implies SWSSB defined by fidelity relies on three properties of $\tilde{D}_\alpha$ [44]:

(1) $\tilde{D}_\alpha(\rho\|\sigma)$ converges to the fidelity correlator in Eq. (7) for $\alpha \to \frac{1}{2}$: $\tilde{D}_{\frac{1}{2}}(\rho\|\sigma) = -2\log_2 F(\rho, \sigma)$.
(2) $\tilde{D}_\alpha(\rho\|\sigma)$ converges to the quantum relative entropy $S(\rho\|\sigma)$ for $\alpha \to 1$.
(3) For $0 < \alpha < \beta < 1$, $\tilde{D}_\alpha(\rho\|\sigma) \leq \tilde{D}_\beta(\rho\|\sigma)$.

As a result, we have $F(\rho, \sigma) \geq e^{-\frac{\ln 2}{2}S(\rho\|\sigma)}$. A SWSSB defined in terms of quantum relative entropy thus implies that the fidelity correlator has an $O(1)$ expectation at large $|x-y|$. In fact, as a good distinguishability measure (with data-processing inequality), the sandwiched Rényi relative entropy can also be used to define SWSSB.

We conjecture that all the above definitions are equivalent and will provide supporting evidence in Sec. III.

B. Rényi-2 correlator

The similarity measures of two mixed states are not limited to the above information-theoretic quantities. In this subsection, we introduce an alternative definition of SWSSB that is not equivalent to the definitions in Sec. II A.

A useful way to study the mixed quantum state is utilizing the Choi-Jamiołkowski isomorphism. It maps a density matrix ρ to a pure state $|\rho\rangle\rangle$ in the doubled Hilbert space. More explicitly, the Choi-Jamiołkowski isomorphism of a density matrix $\rho = \sum \lambda_j |\psi_j\rangle\langle\psi_j|$ is the following Choi state in the doubled Hilbert space $\mathcal{H}_d = \mathcal{H}_u \otimes \mathcal{H}_l$:

$$|\rho\rangle\rangle = \frac{1}{\sqrt{\dim(\rho)}}\sum_j p_j |\psi_j\rangle \otimes |\psi_j^*\rangle, \tag{25}$$

where both upper Hilbert space $\mathcal{H}_u$ and lower Hilbert space $\mathcal{H}_l$ are copies of the physical Hilbert space $\mathcal{H}$. The operator $\sigma \equiv O(x)O^\dagger(y)\rho O(y)O^\dagger(x)$ is mapped to the following state in the doubled Hilbert space:

$$|\sigma\rangle\rangle = O_u(x)O_u^\dagger(y)O_l(y)O_l^\dagger(x)|\rho\rangle\rangle, \tag{26}$$

where O_u and O_l are charged operators defined in the upper and lower Hilbert spaces, respectively. The strong symmetry G is mapped to a double symmetry $G_u \times G_l$, where O_u/O_l only carries charge of G_u/G_l.

The natural similarity measure of ρ and σ in the doubled Hilbert space $\mathcal{H}_d$ is simply the overlap between $|\rho\rangle\rangle$ and $|\sigma\rangle\rangle$ (with a normalization), namely

$$\frac{\langle\langle\rho|\sigma\rangle\rangle}{\langle\langle\rho|\rho\rangle\rangle} = \frac{\langle\langle\rho|O_u(x)O_u^\dagger(y)O_l(y)O_l^\dagger(x)|\rho\rangle\rangle}{\langle\langle\rho|\rho\rangle\rangle}. \tag{27}$$

This correlator precisely measures the SSB in doubled Hilbert space that breaks the double symmetry $G_u \times G_l$ to its diagonal subgroup symmetry [3].

Mapping the correlator, Eq. (27), back to the physical Hilbert space, we obtain the following Rényi-2 correlation function of the charged operator $O(x)$, namely

$$R^{(2)}(x,y) := \frac{\mathrm{Tr}\left(O(x)O^\dagger(y)\rho O(y)O^\dagger(x)\rho\right)}{\mathrm{Tr}\rho^2}. \tag{28}$$

In particular, the doubled symmetry $G_u \times G_l$ and its diagonal subgroup symmetry in the doubled Hilbert space are mapped to the strong and weak symmetry in the physical Hilbert space, respectively.

Notice that for the kind of "spin-glass" ensemble discussed in Sec. II A, the Rényi-2 correlator becomes

$$R^{(2)}(x,y) \approx \frac{\sum_a p_a^2 \langle\psi_a|Z_x Z_y|\psi_a\rangle^2}{\sum_a p_a^2}. \tag{29}$$

Gathering everything together, we formulate an alternative definition of SWSSB:

Definition 4. A mixed state ρ with strong symmetry G has **strong-to-weak SSB in the sense of Rényi-2 correlator** if the Rényi-2 correlator of some charged local operator $O(x)$ is nonvanishing:

$$\lim_{|x-y|\to\infty} \frac{\mathrm{Tr}\left(O(x)O^\dagger(y)\rho O(y)O^\dagger(x)\rho\right)}{\mathrm{Tr}\rho^2} \sim O(1), \tag{30}$$

while the conventional correlation function shows no long-range order:

$$\lim_{|x-y|\to\infty} \mathrm{Tr}\left(\rho O(x)O^\dagger(y)\right) = 0. \tag{31}$$

For the simplest example $\rho_0 \propto \mathbb{1} + X$ with a strong $\mathbb{Z}_2$ symmetry, it is easy to see that the Rényi-2 correlator of the order parameter Z is exactly 1, independent of the distance between x and y.

We now demonstrate that two definitions of SWSSB through fidelity correlator (7) and Rényi-2 correlator (30) are *inequivalent* through a simple example. We consider the following state in an Ising spin chain:

$$\rho = \frac{1}{2}|++\cdots\rangle\langle++\cdots| + \frac{1}{2^{L+1}}(\mathbb{1}+X), \tag{32}$$

which is a convex sum of an ideal SWSSB state $\mathbb{1}+X$ and a symmetric pure product state $|++\cdots\rangle$. For the order parameter Z, we have

$$\begin{aligned}\sigma &\equiv Z(x)Z(y)\rho Z(x)Z(y) \\ &= \frac{1}{2}Z(x)Z(y)|++\cdots\rangle\langle++\cdots|Z(x)Z(y) \\ &\quad + \frac{1}{2^{L+1}}(\mathbb{1}+X).\end{aligned} \tag{33}$$

It is then straightforward to check that $F_O(x,y) \to O(1)$ while $R^{(2)}(x,y) \to 0$, namely the state ρ has SWSSB in the fidelity sense, but not in the Rényi-2 sense. The opposite situation, in which $F_O(x,y) \to 0$ but $R^{(2)}(x,y) \to O(1) > 0$, can also happen, as shown in Appendix A. Another example illustrating the inequivalence between the fidelity correlator and the Rényi-2 correlator will be provided in Sec. III, where these two correlators are mapped to observables in distinct statistical models.

The major weakness of the Rényi-2 correlator definition, compared to the one based on fidelity, is the lack of a stability theorem (such as Theorems 1 and 2). In fact, we will discuss an explicit example in Sec. IV where the broken symmetry in the Rényi-2 sense is restored using a low-depth symmetric channel.

Although a stability theorem does not exist for Rényi-2 SWSSB, we can prove that the SWSSB defined through the Rényi-2 correlator Eq. (30) shares the same property of *noninvertibility* with the SWSSB defined through the fidelity Eq. (7).

Theorem 3. A mixed state ρ with SWSSB in the Rényi-2 correlator sense Eq. (30) is not symmetrically invertible.

Proof. Suppose the (partial) contrary: $\exists\tilde{\rho} \in \widetilde{\mathcal{H}}$ and a strongly symmetric local quantum channel $\mathcal{E}$, such that

$$\mathcal{E}[\rho\otimes\tilde{\rho}] = |0\rangle\langle 0| \in \mathcal{H}\otimes\widetilde{\mathcal{H}}. \tag{34}$$

We purify the channel $\mathcal{E}$ with an ancilla Hilbert space $\mathcal{H}^a$ to a low-depth unitary U, namely

$$U\left(\rho\otimes\tilde{\rho}\otimes|0\rangle\langle 0|_a\right)U^\dagger = |0\rangle\langle 0|\otimes\rho_a, \tag{35}$$

where ρ_a is a density matrix in $\mathcal{H}^a$. Then the Rényi-2 correlator in the state $\mathcal{E}(\rho\otimes\tilde{\rho})$ is given by

$$\frac{\mathrm{tr}\left[O'_x O'^\dagger_y(|0\rangle\langle 0|\otimes\rho_a)O'_y O'^\dagger_x(|0\rangle\langle 0|\otimes\rho_a)\right]}{\mathrm{tr}(\rho_a^2)}, \tag{36}$$

where $O'_x = UO(x)U^\dagger$ and $O'_y = UO(y)U^\dagger$. Recall that a unitary does not change the purity of a density matrix. Now because $\mathcal{E}$ is strongly symmetric (and so is the purification U), O' carries the same charge of strong symmetry as O. Expand O'_x as the following form:

$$O'_x = \sum_i o_i(x)\otimes o_i^a(x), \tag{37}$$

where $o_i^a(x)$ acts only on the ancilla $\mathcal{H}^a$ and $o_i(x)$ acts on $\mathcal{H}\otimes\widetilde{\mathcal{H}}$, and the charge of strong symmetry is carried by o_i only. Then the numerator of Eq. (36) is a sum of terms involving the factor $\langle 00\cdots|o_i^\dagger(x)o_j(y)|00\cdots\rangle \sim e^{-|x-y|/\xi}$ (or strictly zero if the gates do not have exponential tails). Therefore, we conclude that the Rényi-2 correlator has to be exponentially small for invertible states. In particular, the original ρ cannot have SWSSB in the Rényi-2 sense. ■

In particular, the above theorem implies that under a symmetric shallow channel, even though the Rényi-2 long-range order may be destroyed, the state cannot be brought to a completely trivial state (a symmetric pure product state)—namely some kind of nontrivial order (for example, SWSSB in the fidelity sense) should persist. Again this will be demonstrated by the explicit example in Secs. III and IV.

To summarize, the overall lessons are the following:

(1) The Rényi-2 correlator, as a characterization of SWSSB, is not as universal as the fidelity correlator. In particular, Stability Theorems 1, 2 do not hold for the Rényi-2 correlator, as will be exemplified in Sec. III B.

(2) Nevertheless, having a nonvanishing Rényi-2 correlator still indicates the nontriviality of a state, due to the noninvertibility Theorem 3. The Rényi-2 correlator remains valuable for this reason, especially given that the Rényi-2 correlator is in general much simpler to calculate than the fidelity correlator.

(3) Experimentally measuring the fidelity correlator is extremely hard in the sense that it generally calls for the full quantum state tomography (QST) of the density matrix, whose complexity is exponential with respect to the system size. Therefore, the other significance of introducing the Rényi-2 correlator is the experimental accessibility: the Rényi-2 correlator can be experimentally measured without performing QST. Recently, the proposed *classical shadow tomography* [45–49] significantly reduces the experimental complexity of measuring nonlinear quantum information-theoretic quantities, especially in open quantum systems. In particular, as indicated in Eq. (30), both the numerator and denominator of the Rényi-2 correlator are polynomials of the density matrix ρ, which emphasizes that both can be efficiently measured by classical shadow tomography.

C. Local indistinguishability and detectability

Recall that pure-state SSB can also be characterized through degenerate states that are locally indistinguishable. For concreteness, consider the following two orthogonal $\mathbb{Z}_2$ SSB states, with different symmetry charges:

$$|\psi_\pm\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow\cdots\uparrow\rangle \pm |\downarrow\cdots\downarrow\rangle\right). \tag{38}$$

For any k-point (k-finite) observable in a simply connected region

$$M_k \equiv M(x_1)M(x_2)\cdots M(x_k), \tag{39}$$

the two states give identical expectation values:

$$\langle\psi_-|M_k|\psi_-\rangle \simeq \langle\psi_+|M_k|\psi_+\rangle. \tag{40}$$

In other words, $|\psi_\pm\rangle$ are *locally indistinguishable*. Furthermore, if we make a superposition, for example, $|\psi\rangle = \frac{1}{\sqrt{2}}(|\psi_+\rangle + |\psi_-\rangle) = |\uparrow\uparrow\cdots\rangle$, then $|\psi\rangle$ explicitly breaks the symmetry, in a way that is locally detectable: $\langle\psi|Z(x)|\psi\rangle = 1$.

Below we demonstrate that a similar notion of local indistinguishability exists for mixed states with SWSSB.

For a density matrix ρ with SWSSB, we construct a new state

$$\tilde{\rho} \equiv \frac{1}{L^d}\sum_x O(x)\rho O^\dagger(x), \tag{41}$$

here we suppose $O(x)$ is a unitary charged operator to ensure that $\tilde{\rho}$ is a legitimate density matrix, for simplicity. The presence of strong symmetry implies that ρ and $\tilde{\rho}$ are orthogonal in the fidelity sense: $F(\rho,\tilde{\rho}) = 0$ because they carry different symmetry charges.

For any k-point observable [Eq. (39)], we have the following simple relations:

$$\begin{aligned}\mathrm{Tr}[M_k(\tilde{\rho}-\rho)] &= \frac{1}{L^d}\mathrm{Tr}\left[M_k\sum_x\left(O(x)\rho O^\dagger(x)-\rho\right)\right]\\ &= \frac{1}{L^d}\sum_x \mathrm{Tr}\left([M_k,O(x)]\rho O^\dagger(x)\right)\\ &= O(k/L^d),\end{aligned} \tag{42}$$

and

$$\begin{aligned}&F\left(\rho, M_k\rho M_k^\dagger\right) - F\left(\tilde{\rho}, M_k\tilde{\rho}M_k^\dagger\right)\\ &\le F\left(\rho, M_k\rho M_k^\dagger\right) - \frac{1}{L^d}\sum_x F\left(O_x\rho O_x^\dagger, M_kO_x\rho O_x^\dagger M_k^\dagger\right)\\ &= F\left(\rho, M_k\rho M_k^\dagger\right) - \frac{1}{L^d}\sum_x F\left(O_x\rho O_x^\dagger, \left([M_k,O_x]+O_xM_k\right)\rho\left(M_k^\dagger O_x^\dagger + \left[O_x^\dagger, M_k^\dagger\right]\right)\right)\\ &= \frac{1}{L^d}\sum_{x\in k} F\left(\rho, M_k\rho M_k^\dagger\right) - F\left(O_x\rho O_x^\dagger, \left([M_k,O_x]+O_xM_k\right)\rho\left(M_k^\dagger O_x^\dagger + \left[O_x^\dagger, M_k^\dagger\right]\right)\right) \sim O\left(\frac{k}{L^d}\right),\end{aligned} \tag{43}$$

where $x \in k$ means that the site x locates inside the support of M_k. The inequalities above came from the joint concavity of fidelity:

$$F\left(\sum_i p_i\rho_i, \sum_i p_i\sigma_i\right) \geq \sum_i p_i F(\rho_i, \sigma_i). \tag{44}$$

We conclude that any finite-point observable cannot distinguish $\tilde{\rho}$ from ρ. To distinguish the two states, the observable M needs to act nontrivially on the entire system—the simplest such observable is the symmetry operator g.

Notice that the conclusion of local indistinguishability does not even need the SWSSB. What does need the SWSSB property is the local detectability of the symmetry breaking. Then we consider the convex sum of ρ and $\tilde{\rho}$,

$$\rho_+ \equiv \frac{\rho + \tilde{\rho}}{2}. \tag{45}$$

The state explicitly breaks the strong symmetry down to a weak symmetry.

The symmetry breaking does become locally detectable using the fidelity average of a local charged operator if the original state ρ has SWSSB:

$$\begin{aligned} &F\left(\rho_+, O^\dagger(y)\rho_+ O(y)\right) \\ &\quad \geq \frac{1}{L^d}\sum_x F\left(\rho, O(x)O^\dagger(y)\rho O(y)O^\dagger(x)\right) \sim O(1), \end{aligned} \tag{46}$$

where the inequality is ensured by the joint concavity of the fidelity (44). It is straightforward to show that the above result on local detectability of explicit strong symmetry breaking for ρ_+ is also valid in the Rényi-2 sense.

III. EXAMPLES

We discuss a few examples in this section. First in Sec. III A we conjecture, with explicit calculations in simple cases and plausibility arguments in general, that a thermal state within a fixed charge sector must have spontaneously broken strong symmetry, for any nonzero temperature. We then discuss two nonthermal examples of SWSSB states in $(2+1)d$ and associated phase transitions: the $(2+1)d$ quantum Ising model under symmetric noise in Sec. III B, and a similar model with rotor degrees of freedom and $U(1)$ symmetry in Sec. III C.

A. SWSSB in thermal states

A standard fact in statistical mechanics is the local equivalence between canonical and grand canonical ensembles. From the open-system perspective, canonical ensembles by definition have strong symmetry, while grand canonical ensembles only have weak symmetry. The local equivalence suggests that the canonical ensembles have SWSSB if the strong symmetry is not spontaneously broken in the conventional sense (i.e., the two-point function of a certain charged operator is long-range ordered).

Conjecture 1. Given a local symmetric Hamiltonian H and $0 < \beta < \infty$, if the Gibbs state $\rho_{\beta,\lambda} \propto P_\lambda e^{-\beta H}$ within a fixed charge sector described by a projector P_λ has no SSB, then it must have SWSSB. In particular, for sufficiently high temperatures, the Gibbs state always has SWSSB.

Here a Gibbs state within a fixed-charge sector is constructed by symmetrizing an ordinary Gibbs state. For example, when the strong symmetry G is a finite group, and $\lambda : G \to U(1)$ is a one-dimensional representation, we express it as

$$\rho = \frac{P_\lambda e^{-\beta H}}{\mathrm{Tr}\left[P_\lambda e^{-\beta H}\right]}, \tag{47}$$

where $P_\lambda = (1/|G|)\sum_{g\in G}\lambda^{-1}(g)g$ is the projector into the charge sector (one-dimensional rep.) $\lambda : G \to U(1)$.

We demonstrate the validity of the above conjecture in a simple commuting projector Hamiltonian of Ising spins in a canonical ensemble. On a lattice, the Hamiltonian takes the following form:

$$H_{\mathrm{CP}} = \sum_i B_i, \tag{48}$$

where B_i satisfies

$$B_i^2 = 1, \quad [B_i, B_j] = 0. \tag{49}$$

We further assume B_i is finitely supported around site i. The strong $\mathbb{Z}_2$ symmetry operator is $U = \prod_j X_j$ and the corresponding charged operator is Z_i, without loss of generality. We assume that

$$[B_i, U] = 0, \quad Z_i B_j = (-1)^{\delta_{ij}} B_j Z_i. \tag{50}$$

For example, we can have $B_i = X_i$, or X_i times any $\mathbb{Z}_2$-invariant functions of Z_i. We also assume that B_i is a complete set of (mutually commuting) observables.

Then one of the corresponding Gibbs canonical ensemble at finite temperature can be formulated as

$$\rho = \frac{1}{Z}\frac{1+U}{2}e^{-\beta H_{\mathrm{CP}}}. \tag{51}$$

If we choose the B_i eigenbasis, the density matrix can be reformulated as

$$\rho = \frac{1}{Z}\sum_{\{b_i\}}{}' e^{-\beta\sum_i b_i}|\{b_i\}\rangle\langle\{b_i\}|, \tag{52}$$

where $b_i = \pm 1$ labels the eigenvalue of B_i, and the partition function $Z = \sum'_{\{b_i\}} e^{-\beta\sum_i b_i}$. Here the prime means that the sum is over configurations satisfying $\prod_i b_i = 1$. Then the fidelity can be calculated easily. For any $i \neq j$ we find

$$F(i,j) = \frac{1}{\cosh^2 \beta}(1 + \tanh^N \beta)^{-1} \overset{N\to\infty}{\approx} \frac{1}{\cosh^2 \beta}. \tag{53}$$

Here we have taken the thermodynamic limit $N \to \infty$. We see that at zero temperature ($\beta \to \infty$), the fidelity $F(i,j)$ vanishes, which is consistent with the ground state of a commuting projector Hamiltonian having no SSB. But for finite temperature, the fidelity $F(i,j)$ will always be positive and independent of the distance between i and j. On the other hand, a $\mathbb{Z}_2$-charged operator, such as Z_i, must anticommute with at least one of B_i. This results in a correlation function $\mathrm{Tr}(\rho Z_i Z_j) = 0$ that vanishes at large $|i-j|$. Therefore, we conclude that a Gibbs canonical ensemble of the Hamiltonian $H = -\sum_i B_i$ at any finite temperature has SWSSB of the $\mathbb{Z}_2$ symmetry.

We now provide a plausibility argument supporting Conjecture 1 for a generic local Hamiltonian H and its Gibbs state ρ in the canonical ensemble. We examine the replicated correlation function

$$\begin{aligned} F^{(n)} &\equiv \frac{\mathrm{Tr}[(\sqrt{\rho}O(x)O^\dagger(y)\rho O(y)O^\dagger(x)\sqrt{\rho})^n]}{\mathrm{Tr}\,\rho^{2n}} = \frac{1}{Z(2n\beta)}\mathrm{Tr}[(O(x)O^\dagger(y)e^{-\beta H}O(y)O^\dagger(x)e^{-\beta H})^n] \\ &= \langle [O(x,\tau=0)O^\dagger(y,\tau=0)][O^\dagger(x,\tau=\beta)O(y,\tau=\beta)][O(x,\tau=2\beta)O^\dagger(y,\tau=2\beta)]\cdots\rangle_{2n\beta}, \end{aligned} \tag{54}$$

where $A(\tau) \equiv e^{-\tau H} A e^{\tau H}$ and $\langle \ldots \rangle_{2n\beta}$ means the thermal (imaginary-time) expectation value at inverse temperature $2n\beta$. Assuming no off-diagonal long-range order at temperature $2n\beta$ (otherwise even the weak symmetry is spontaneously broken), the above correlation function should factorize for large $|x-y|$:

$$\begin{aligned} F^{(n)} &= \langle O(x,\tau=0)O^\dagger(x,\tau=\beta)O(x,\tau=2\beta)\cdots\rangle_{2n\beta} \\ &\quad\cdot \langle O^\dagger(y,\tau=0)O(y,\tau=\beta)O^\dagger(y,\tau=2\beta)\cdots\rangle_{2n\beta}, \end{aligned} \tag{55}$$

which in general should be nonzero ($\sim O(1)$) unless fine tuned. In particular, for low temperature or large β (but still assuming no off-diagonal long-range order), the correlator should decay exponentially in imaginary time, so we expect

$$F^{(n)} \sim e^{-2n\beta\Delta}, \tag{56}$$

where Δ is the energy gap of the excitation created by $O(x)$ above the ground state. As long as $\beta \neq \infty$, the above expression is nonzero.

Analytically continuing to $n = 1/2$, $F^{(n)}$ becomes the fidelity correlator, which demonstrates our original proposal. In particular, Eq. (56) agrees with Eq. (53) with $\Delta = 2$ being the excitation gap.

Another interesting feature of Eq. (56) is that the SWSSB disappears below temperature scale $T^* \sim \Delta/L$. In particular, this means that as far as SWSSB is concerned, the zero-temperature limit $T \to 0^+$ and the thermodynamic limit $L \to \infty$ do not commute. This noncommutativity is absent in ordinary observables, at least when the ground state is gapped $\Delta \sim O(1)$—recall that in the imaginary-time path integral, these two limits are simply infinite size limits in different directions.

B. $\mathbb{Z}_2$ SWSSB in the quantum Ising model

Consider a (2+1)d qubit system on a square lattice, with a strong $\mathbb{Z}_2$ symmetry $U = \prod_i X_i$. The system is initialized in a pure product state $|\psi_0\rangle = \otimes_i |X = +1\rangle_i$. Then we consider a nearest-neighbor dephasing channel that preserves the strong $\mathbb{Z}_2$ symmetry, $\mathcal{E} = \prod_{\langle ij\rangle} \mathcal{E}_{ij}$, with

$$\mathcal{E}_{ij}[\rho] = (1-p)\rho + pZ_iZ_j\rho Z_iZ_j. \tag{57}$$

We note that this model is closely related to the $\mathbb{Z}_2$ toric code under bit-flip noise [11,50]. In fact, under the Kramers-Wannier duality mapping, they are exactly mapped to each other.

To detect the possible SWSSB and identify the universality class of the symmetry-breaking transition, we consider the fidelity correlator of Z, which is charged under the strong $\mathbb{Z}_2$ symmetry.

The starting point is the decohered density matrix in the $\mathbb{Z}_2$ charge basis:

$$\rho = \sum_{\langle X\rangle} Z(X)\,|X\rangle\,\langle X|\,. \tag{58}$$

Here X denotes a configuration of $\mathbb{Z}_2$ charges (i.e., $X = \pm 1$), subject to the constraint that $\prod_i X_i = 1$. In the following, we write $X_i = (-1)^{n_i}$ with $n_i = 0, 1$. To compute the weight $Z(X)$, define an edge $\mathbb{Z}_2$ variable $s_{ij} = 0, 1$, which keeps track of the Kraus operator on the edge ij. When applied to the initial state $|X_i = 1\rangle$, the resulting charge at site i is given by

$$(\nabla \cdot s)_i = \sum_{\mu=x,y} s_{i,i+\mathbf{e}_\mu} + s_{i-\mathbf{e}_\mu,i} \bmod 2. \tag{59}$$

Here $\mathbf{e}_\mu$ is the unit vector along the μ direction. We obtain

$$\begin{aligned} Z(X) &= \sum_{\nabla\cdot s=n} \prod_{\langle ij\rangle} (1-p)^{1-s_{ij}} p^{s_{ij}} \\ &= (1-p)^{N_e} \sum_{\nabla\cdot s=n} \left(\frac{p}{1-p}\right)^{\sum_{\langle ij\rangle} s_{ij}}. \end{aligned} \tag{60}$$

To resolve the constraint $\nabla \cdot s = n$, we denote a special solution of this equation by s^0, and $s_e = s_e^0 + \tilde{s}_p + \tilde{s}_q \bmod 2$ (where e stands for a nearest-neighbor edge, and p, q are the two adjacent square plaquettes). $\tilde{s}_p$ are dual variables living on the plaquettes p. It is more convenient to work with $\mu_e = (-1)^{s_e}$ and $\tau_p = (-1)^{\tilde{s}_p}$, so $\mu_e = \mu_e(X)\tau_p\tau_q$, with $\mu_e(X) = (-1)^{s_e^0}$. Therefore,

$$Z(X) = [p(1-p)]^{N_e/2} \sum_\tau e^{\beta \sum_{\langle pq\rangle} \mu_{pq}(X)\tau_p\tau_q}, \tag{61}$$

where $e^{-2\beta} = p/(1-p)$. In other words, $Z(X)$ is the partition function of a two-dimensional (2D) classical Ising model coupled to background $\mathbb{Z}_2$ gauge fields $\{\mu(X)\}$, with fixed gauge-invariant $\mathbb{Z}_2$ fluxes given by X.

Since ρ is already diagonalized in this basis, we can directly compute the fidelity correlator. We have

$$\begin{aligned} \sigma &= Z_i Z_j \rho Z_i Z_j \\ &= \sum_X Z(X \cdot \delta X) |X\rangle \langle X|, \end{aligned} \tag{62}$$

where $\delta X_k = -1$ if $k = i,j$ and 1 otherwise, and the fidelity correlator is found to be [14]

$$F_{ij} = \mathrm{Tr}\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}} \tag{63}$$

$$\begin{aligned} &= \sum_X \sqrt{Z(X)Z(X\cdot\delta X)} \\ &= \sum_X Z(X) \left(\frac{Z(X\cdot\delta X)}{Z(X)}\right)^{1/2} \\ &= \sum_X Z(X) e^{-\frac{1}{2}\Delta F_{\mathrm{dw}}} \\ &= \langle e^{-\frac{1}{2}\Delta F_{\mathrm{dw}}}\rangle_{\mathrm{RBIM}} \geq e^{-\frac{1}{2}\langle \Delta F_{\mathrm{dw}}\rangle_{\mathrm{RBIM}}}. \end{aligned} \tag{64}$$

Here $\Delta F_{\mathrm{dw}} = -\ln(Z(X\cdot\delta X)/Z(X))$ is the free energy cost of inserting a domain wall of length approximately $|i-j|$. $\langle\cdots\rangle_{\mathrm{RBIM}}$ stands for disorder average in the random-bond Ising model along the Nishimori line [50]. For $p > p_c$, the RBIM is in the paramagnetic phase (i.e., high temperature), so the disorder-averaged free energy cost for a domain wall approaches a constant. Thus we have shown that the fidelity correlator is long-range ordered for $p > p_c$. It also suggests that F_{ij} decays exponentially for $p < p_c$, since the free energy cost diverges linearly with the distance in the ferromagnetic (low-temperature) phase.

We note that the same results follow for the other definition of SSB in terms of sandwiched Rényi entropy, offering supporting evidence for our earlier conjecture in Sec. II A that SWSSB defined using distinct distinguishability measures—namely, quantum relative entropy and the fidelity correlator—ought to be equivalent.

If we instead consider the Rényi-2 correlator (30), following a similar derivation the correlation function can be mapped to the statistical mechanics model originating from evaluating the purity of the density matrix, $\mathrm{Tr}(\rho^2)$, which is exactly the partition function of the $2d$ classical Ising model. Therefore, the transition between $\mathbb{Z}_2$ strongly symmetric state and an SWSSB state induced by a symmetric quantum channel will happen at $p_c^{(2)} \approx 0.178$.

Furthermore, this is another example of the inequivalence between two different definitions of SWSSB through fidelity and Rényi-2 correlators. For this model, the long-range Rényi-2 correlator is a stronger condition than the long-range fidelity correlator, since $p_c < p_c^{(2)}$. We note that it has been proposed in Ref. [3] to use the Rényi-1 quantity, such as the Fisher information or von Neumann relative entropy to define an "intrinsic" transition; based on its correspondence to the decodability transition in toric code [50], this transition was claimed to be the random-bond transition at p_c.

Our discussions are not limited to the Pauli channel like Eq. (57). For example, consider the following nearest-neighbor non-Pauli channel that is $\mathbb{Z}_2$ strongly symmetric,

$$\mathcal{E}_{ij}(\rho) = (1-p)\rho + p e^{i\theta Z_i Z_j} \rho e^{-i\theta Z_i Z_j}, \tag{65}$$

and $\mathcal{E} = \prod_{\langle ij\rangle} \mathcal{E}_{ij}$. Notice that for $\theta = \pi/2$ we recover the previous channel (57). A more general channel similar to Eq. (65) is

$$\mathcal{E}_{ij}(\rho) = \sum_n p_n e^{i\theta_n Z_i Z_j} \rho e^{-i\theta_n Z_i Z_j}, \tag{66}$$

where $p_n \geq 0$ and $\sum_n p_n = 1$.

For these non-Pauli channels, the density matrix is no longer diagonal in the X basis. In the following, we will argue that nevertheless the Ising SWSSB transition is still described by the RBIM universality class, the same as the Pauli case. We will utilize the replica trick to calculate the fidelity, which is not rigorous but still suggestive. In Appendix B we perform the calculations in the Pauli case

$\theta = \pi/2$, and reproduce the same result as the exact calculation above, providing evidence for the validity of the replica trick.

We define the replicated fidelity as

$$F_{m,n}(x,y) = \mathrm{Tr}[(\rho^m \sigma \rho^m)^n], \tag{67}$$

and the standard definition of fidelity (7) is recovered by taking the limit $m, n \to 1/2$. Thus to compute $F_{m,n}(x,y)$, we need to study the replicated partition function $\mathrm{Tr}\,\rho^t$. It is not difficult to see that $t = (2m+1)n$, so the limit $m, n \to 1/2$ corresponds to $t \to 1$.

Let us consider the domain-wall basis (i.e., eigenbasis of Z), in which it is easier to compute the matrix elements of ρ. The initial pure state can be represented in the domain-wall basis as $\rho_0 \propto \sum_{D,D'} |D\rangle\langle D'|$, where the summation is taken over all possible $\mathbb{Z}_2$ domain-wall configurations. The decohered density matrix is

$$\rho \propto \sum_{D,D'} |D\rangle\langle D'| f(D,D'). \tag{68}$$

Here we need to explain $f(D,D')$. The domain-wall configuration can be labeled by $s_e = \pm 1$ on nearest-neighbor lattice edges e ($s_e = 1$ means no domain wall, $s_e = -1$ means a domain wall crossing the bond). Then an explicit calculation gives

$$f(D,D') = \prod_e \alpha^{\frac{s_e - s'_e}{2}} \beta^{\frac{|s_e - s'_e|}{2}}. \tag{69}$$

Here

$$\beta = |1 - p + pe^{2i\theta}| = \sqrt{1 - 4p(1-p)\sin^2\theta}, \qquad \alpha = \frac{1 - p + pe^{2i\theta}}{\beta}. \tag{70}$$

Note that $\beta > 0$ is real, and α is a pure phase factor. For a special point $\theta = \pi/2$, we have $\beta = 1 - 2p, \alpha = 1$. In this case, $\beta^{\frac{|s_e - s'_e|}{2}} = \beta^{\frac{1 - s_e s'_e}{2}}$, and $\prod_e \beta^{\frac{1 - s_e s'_e}{2}} = \beta^{|D - D'|}$, where $|D - D'|$ is the length of the relative domain wall, namely $|D - D'| = \sum_e \frac{1 - s_e s'_e}{2}$.

Naively, the $\alpha^{\frac{s_e - s'_e}{2}}$ part complicates things, and the weight cannot be expressed in terms of relative domain walls alone. However, if we consider the replicated density matrix:

$$\begin{aligned}
\mathrm{Tr}\,\rho^t &\propto \sum_{\{D_s\}} f(D_1, D_2) f(D_2, D_3) \cdots f(D_s, D_1) \\
&= \prod_e \alpha^{\sum_k \frac{1}{2}(s_e^{(k)} - s_e^{(k+1)})} \prod_e \beta^{\sum_k \frac{1}{2}\left|s_e^{(k)} - s_e^{(k+1)}\right|} \\
&= \prod_e \beta^{\sum_k \frac{1}{2}\left|s_e^{(k)} - s_e^{(k+1)}\right|}.
\end{aligned} \tag{71}$$

Crucially, the replicated partition function takes the same form regardless of the value of θ, just with different values of the "tension" β. This is true for each $t \geq 2$, and if we analytically continue to $t \to 1$, it makes sense to conclude that the partition function in the limit $t \to 1$ should have the same universal behavior as the $\theta = \pi/2$ case. Thus if the replica limit exists, the SWSSB transition in the non-Pauli model defined in Eq. (65) is described by the same RIBM universality class.

This can be generalized to the channel in Eq. (66), with

$$\beta = \left|\sum_n p_n e^{2i\theta_n}\right|, \quad \alpha = \frac{\sum_n p_n e^{2i\theta_n}}{\beta}. \tag{72}$$

Then the same result follows. These calculations suggest that the "universality class" of the Ising SWSSB transition is described by the RBIM, independent of microscopic details.

C. $U(1)$ SWSSB in a quantum rotor model

We now consider the analogous transition in a $U(1)$ symmetric model. On each site i, there is a quantum rotor degree of freedom with the 2π-periodic phase variable θ_i and the conjugate number n_i. They satisfy the canonical commutation relation $[\theta_i, n_j] = i\delta_{ij}$. The $U(1)$ symmetry is generated by $Q = \sum_i n_i$. Initially, the system is in a pure product state $|\psi_0\rangle = \otimes_i |n_i = 0\rangle$. Then we consider a nearest-neighbor dephasing channel that preserves the strong $U(1)$ symmetry $\mathcal{E} = \prod_{\langle ij \rangle} \mathcal{E}_{ij}$, with

$$\mathcal{E}_{ij}(\rho) = \sum_{k \in \mathbb{Z}} p_k e^{ik(\theta_i - \theta_j)} \rho e^{-ik(\theta_i - \theta_j)}. \tag{73}$$

Here $0 \leq p_k \leq 1$ should satisfy $\sum_k p_k = 1$. Physically, we expect that p_k should decay rapidly with $|k|$, and it is natural to impose "charge-conjugation" symmetry so that $p_{-k} = p_k$. To make the problem analytically tractable we choose $p_k \propto e^{-\alpha k^2}$, but we expect the results should apply to other p_k as long as it decays sufficiently rapidly as $|k| \to \infty$.

We will study the phase diagram as a function of α. Heuristically, when α is large, the channel is close to the identity, so the system remains close to the pure state. As α decreases, one expects a transition to a $U(1)$ SWSSB phase characterized by the long-range fidelity correlator of the $U(1)$ order parameter $e^{i\theta}$.

It proves useful to work with dual link variables defined as

$$n_i = (\nabla \cdot E)_i, \; \theta_i - \theta_j = A_{ij}. \tag{74}$$

Here $(\nabla \cdot E)_i$ is the lattice divergence: $(\nabla \cdot E)_i = \sum_j E_{ij}$. One can think of A_{ij} as $U(1)$ gauge fields and E_{ij} as the electric fields.

Expanding the quantum channel, we find that the initial state ρ_0 (which has no charge) is acted by paths of $e^{\pm ikA}$ operators, i.e., Wilson lines, which will create some configuration of electric charges. Introduce an integer variable m_e on each edge e (with its canonical orientation), corresponding to acting by the $e^{im_eA_e}$ Kraus operator. We can interpret m as some kind of "electric" field, with the "error" charges given by $(\nabla\cdot m)_i$. Denote the Wilson operator by $W_A(\mathbf{m})=\prod_e e^{im_eA_e}$. Then the decohered density matrix becomes

$$\rho=\sum_{\mathbf{m}}P(\mathbf{m})W_A(\mathbf{m})\rho_0W_A^\dagger(\mathbf{m}),\quad P(\mathbf{m})=\prod_e p_{m_e}. \tag{75}$$

Now we consider tr ρ^t:

$$\operatorname{Tr}\rho^t =\sum_{\mathbf{m}^{(s)}}\prod_{s=1}^{t}P(\mathbf{m}^{(s)})\operatorname{Tr}\left[\prod_{s=1}^{t}W_A(\mathbf{m}^{(s)})|\psi_0\rangle\langle\psi_0|W_A^\dagger(\mathbf{m}^{(s)})\right]. \tag{76}$$

We rewrite the trace as

$$\prod_{s=1}^{t}\langle\psi_0|W_A^\dagger(\mathbf{m}^{(s)})W_A(\mathbf{m}^{(s+1)})|\psi_0\rangle, \tag{77}$$

and see that the trace is nonvanishing only if $W_A(\mathbf{m}^{(s+1)})$ and $W_A(\mathbf{m}^{(s)})$ create the same charges. In other words, $\nabla\cdot(\mathbf{m}^{(s)}-\mathbf{m}^{(s+1)})=0$. Therefore, we may write

$$m_e^{(s+1)}=m_e^{(1)}+v_e^{(s)}, \tag{78}$$

where $v_e^{(s)}$ satisfies $\nabla\cdot v^{(s)}=0$. The "Rényi" partition function becomes

$$\operatorname{Tr}\rho^t=\sum_{\mathbf{m}^{(1)}}P(\mathbf{m}^{(1)})\sum_{\mathbf{v}^{(s)}}\prod_{s=1}^{t-1}P(\mathbf{m}^{(1)}+\mathbf{v}^{(s)}). \tag{79}$$

We can resolve the constraint $\nabla\cdot v=0$ in the following way (suppressing the replica index):

$$\prod_i\delta((\nabla\cdot v)_i)=\int_\phi e^{-i\sum_{i,\hat\mu}\phi_i(v_{i,\hat\mu}-v_{i-\hat\mu,\hat\mu})}. \tag{80}$$

We have defined the short-hand notation

$$\int_\phi\equiv\int_0^{2\pi}\prod_i\frac{d\phi_i}{2\pi}. \tag{81}$$

Thus the partition function becomes

$$\int_\phi\sum_{v^{(s)}}P(m_e+v_e^{(s)})e^{-i\sum_{i,\hat\mu}v_{i,\hat\mu}^{(s)}(\phi_i^{(s)}-\phi_{i+\hat\mu}^{(s)})}. \tag{82}$$

Here we have defined $m_e\equiv m_e^{(1)}$. The summation over the replica index $s=1,2,\ldots,t-1$ is kept implicit. Applying Poisson resummation, we find that the partition function becomes

$$\int_\phi\sum_{\mathbf{m},\boldsymbol{k}^{(s)}}e^{-\alpha\sum_e m_e^2}e^{-\frac{1}{\alpha}H_{\text{eff}}[\mathbf{m}]}, \tag{83}$$

with the effective Hamiltonian

$$H_{\text{eff}}[\mathbf{m}]=\frac{1}{4}\sum_{i,\hat\mu}(\phi_i^{(s)}-\phi_{i+\hat\mu}^{(s)}-2\pi k_{i,\hat\mu}^{(s)}-2i\alpha m_{i,\hat\mu})^2. \tag{84}$$

This can be viewed as the replica partition function of the classical Villain model with imaginary bond disorder given by the $2i\alpha m_{i,\hat\mu}$ term. The effective "temperature" is given by α.

The bond disorder exhibits two unusual features. First it is not a continuous variable, rather valued in $2i\alpha\mathbb{Z}$. Second, it is purely imaginary. We shall however consider two limiting cases. First, for large α, the distribution of m_e is sharply peaked at $m_e=0$, and the effective temperature is also high, so once setting $m_e=0$ the statistical mechanics model belongs to the high-temperature: "disordered" phase of the Villain model, which agrees with our expectation that there is no SWSSB when the channel is very close to the identity. Next, for a small α, it is at least plausible that the physics is qualitatively approximated by treating $a_e=2\alpha m_e$ as a real variable, whose probability distribution is given by $e^{-\frac{1}{4\alpha}a_e^2}$. Since α is small, the disorder again becomes weak, and it is reasonable to expect that the statistical model should be in the "superfluid" phase, i.e. SWSSB, which should be robust to weak disorder at sufficiently low effective temperature. In two dimensions, the low-temperature phase is the Kosterlitz-Thouless (KT) phase, so in this case, the model provides an example of quasi-long-range SWSSB. Therefore under these assumptions, we find a SWSSB phase for small α and a strongly symmetric phase for large α.

To further substantiate the result, we also study the Rényi-2 transition of this model. We find that Tr ρ^2 can be mapped to a variant of the Villain XY model. In fact, for small α, the Boltzmann weight is well approximated by that of the Villain XY model. We thus expect that there is a Rényi-2 $U(1)$ SWSSB phase for sufficiently small α, and a transition at a finite value of α (details can be found in Appendix C).

Let us also compare the results with a rotor model at finite temperature. Following the general argument presented in Sec. III A, in a thermal state the fidelity correlator is long-range ordered, independent of dimensionality. However, the decohered rotor model can exhibit a power-law fidelity correlator in the SWSSB phase in 2D, fundamentally distinct from the thermal state.

IV. RECOVERABILITY

To define concrete phases of matter for mixed quantum states, we need to show that two mixed states belonging to the same phase are two-way connected by low-depth symmetric local quantum channels. In other words, a channel on a state should be "recoverable" if the state remains in the same phase.

Now we explore the recoverability of a mixed state with SWSSB. Specifically, we demonstrate in the Ising example that SWSSB, as defined in terms of the Rényi-2 correlator, could be recovered. This means that there exists a short-depth channel, whose depth is on the order of $\log L$ with L representing the system size, which can transform a state with a long-range Rényi-2 (or more general Rényi-n) order parameter into a state with an exponentially decaying Rényi-2 correlator. However, SWSSB in the fidelity sense is not recoverable due to the stability theorem.

For simplicity, in this section, we focus on a special class of low-depth channels termed circuit channels. A circuit quantum channel comprises layers of disjoint local channels, or "gates," forming a circuit. Such a channel is considered short depth if the depth of the circuit is constant or at most sublinear in L. Moreover, the channel is deemed symmetric if each local gate is strongly symmetric.

Recently, it has been highlighted in Ref. [16] that the effect of a local quantum channel $\mathcal{E}$ supported in a region A as in Fig. 2 can be reversed using a local channel $\mathcal{R}_{\mathcal{E}}$ supported in an enlarged region AB, provided that the discrepancy in the conditional mutual information (CMI) between the initial state ρ_i and the final state $\rho_f = \mathcal{E}(\rho_i)$ decays exponentially with respect to the width l_B of the buffer region B, expressed as

$$I(A;C|B)_{\rho_i} - I(A;C|B)_{\rho_f} \simeq O(e^{-l_B/\xi}), \tag{85}$$

where ξ defines the correlation length of the mixed state. The local recovery channel $\mathcal{R}_{\mathcal{E}}$ is constructed through the Petz recovery channel. Two observations are important for our subsequent discussion: (1) The Petz recovery channel preserves the strong symmetry of the mixed state if the symmetry is on site, which can be verified by the explicit form of the recovery channel (see Appendix D).

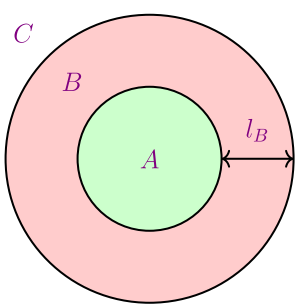

FIG. 2. The regions A, B, and C, with respect to which we calculate the CMI in Eq. (85).

(2) Building on this local recovery, a finite-depth circuit channel acting on the entire system can be recovered region by region, provided that at each intermediate step, the local gates being recovered can be recovered locally, i.e., Eq. (85) holds true for some chosen buffer region. The global recovery channel is also a circuit channel, with depth scaling logarithmically with system size $\log L$. We refer the readers to Appendix D for details regarding the recovery channel.

A. Recoveribility within the SSB phase

We now revisit the Ising example in Sec. III B. We will establish that beginning from a pure symmetric product state, the two resulting mixed states, generated from symmetric dephasing in the form of Eq. (57) with strengths p_1 and p_2, can be mutually connected by a short-depth circuit channel if p_1 and p_2 belong to the same phase of the RBIM. Assuming $p_1 < p_2 < \frac{1}{2}$, and denoting the two states as ρ_{p_1} and ρ_{p_2}, respectively, it is evident that ρ_{p_2} can be prepared from ρ_{p_1} through an additional dephasing channel with strength

$$p = \frac{p_2 - p_1}{1 - 2p_1}. \tag{86}$$

Conversely, to construct a short-depth channel connecting ρ_{p_2} to ρ_{p_1}, it is essential to demonstrate that this dephasing channel with strength p, when applied to ρ_{p_1}, can be recovered by a symmetric short-depth channel. This requirement is equivalent to ensuring that Eq. (85) is met at every intermediate step of the recovery process.

We now establish that within the same phase of the statistical model (i.e., RBIM), Eq. (85) holds. This assertion is supported by the following observation: the CMI, as a combination of von Neumann entropies from different regions, can be mapped to a linear combination of free energies of RBIM. Notably, the von Neumann entropy of a region A (see Fig. 3) can be expressed as

$$S(A)_\rho = \lim_{t\to 1} \frac{1}{1-t} \ln \operatorname{tr} \rho_A^t, \tag{87}$$

where ρ_A represents the reduced density matrix on A. From the expression of the decohered density matrix in Eq. (B2), ρ_A can be computed as

$$\rho_A = (1-p)^{N_A} \sum_{\mathbf{l}_A} (\tanh \tau)^{|\mathbf{l}_A|} |\partial \mathbf{l}_A\rangle\langle \partial \mathbf{l}_A|, \tag{88}$$

where $|\mathbf{l}_A|$ denotes all strings within the region A, potentially including links straddling between region A and its complement. Let us define $Z_t^A = \operatorname{tr}(\rho_A^t)$ as the t-replica

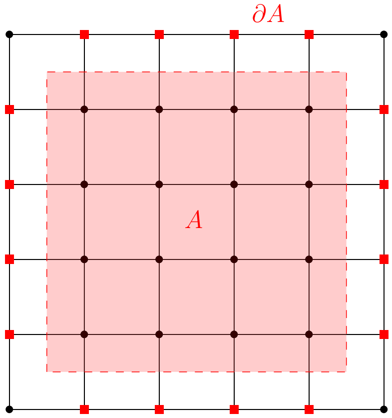

FIG. 3. The region A for which we compute the von Neumann entropy encompasses all sites within the red dashed line. The boundary of A includes all sites not within A but adjacent to sites in A, marked with the red squares above.

partition function, which can be expressed as

$$Z_t^A \propto \sum_{\mathbf{l}_1,\ldots,\mathbf{l}_t} (\tanh\tau)^{|\mathbf{l}_1|+\cdots+|\mathbf{l}_t|}\langle\partial\mathbf{l}_1|\partial\mathbf{l}_2\rangle\cdots\langle\partial\mathbf{l}_t|\partial\mathbf{l}_1\rangle. \quad (89)$$

The nonvanishing contributions to Z_t^A arise from string configurations such that $\mathbf{l}_i$ and $\mathbf{l}_{i+1}$ differ only by closed loops *relative to* the boundary of A. In other words, $\partial(\mathbf{l}_i - \mathbf{l}_{i+1}) \cap A = \emptyset$. Following a derivation analogous to that leading to Eq. (B7), Z_t^A represents the t-replica partition function of an RBIM (on the dual lattice) within region A with a free boundary condition. In the limit $t \to 1$, the von Neumann entropy described in Eq. (87) transforms into the free energy of an RBIM defined on region A with a free boundary condition, denoted as f_A. Consequently, we can express the conditional mutual information as

$$I(A;C|B) = f_{AB} + f_{BC} - f_{ABC} - f_B. \quad (90)$$

When attempting to recover the strength-p dephasing channel that connects ρ_{p_1} to ρ_{p_2} region by region, at each intermediate step, the CMI can be mapped to the linear combination in Eq. (90), computed in an RBIM with a spatially varying strength $p(x)$. Specifically, the regions recovered in previous steps have $p(x) = p_1$, while those that have not been recovered have $p(x) = p_2$. When p_1 and p_2 belong to the same phase of the RBIM, the system, despite its inhomogeneity, belongs to a definite phase (either ordered or disordered) and has a *finite* correlation length. The free energy of a region A thus scales as

$$f_A = \int_A f[p(x)]d^2x + \int_{\partial A} g[p(x)]dl - \gamma_A, \quad (91)$$

where $f(p)$ is the free energy density of the RBIM with effective temperature $\tanh\beta = p/(1-p)$, and $g(p)$ denotes the boundary contribution. γ_A is a constant depending on the phase of the statistical model [51]. Consequently, at each step, Eq. (85) holds, where $\mathcal{E}$ represents the strength-p dephasing within the region A being recovered, and ξ represents the correlation length of RBIM, with all terms in Eq. (91) canceling out in Eq. (90). Physically, the CMI, as a combination of free energies, should exhibit smooth variations within the same phase. Consequently, the left-hand side of Eq. (85) is expected to decay exponentially, given that the system possesses a finite correlation length in both phases of RBIM. The key implication of our argument is that when p_1 and p_2 correspond to the same phase of the RBIM, ρ_{p_1} and ρ_{p_2} can be two-way connected by a symmetric channel within a depth of $\log L$. In particular, the critical decoherence strength for the Rényi-2 correlator, $p_c^{(2)}$, does not obstruct recoverability (Fig. 4).

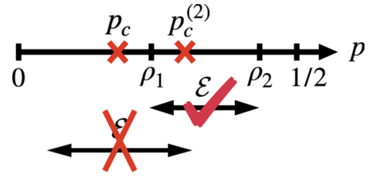

FIG. 4. Phase diagram of the decohered Ising model. Two states ρ_1 and ρ_2 in the same phase (both with $p > p_c$ in the figure) can be two-way connected through some low-depth channels. The two-way connectivity is obstructed if $p_1 > p_c$ and $p_2 < p_c$. The "transition point" for the Rényi-2 correlator, $p_c^{(2)}$, does not obstruct two-way connectivity.

B. Spontaneity and local irrecoverability

In the previous example, the CMI $I(A;C|B)$ in the SWSSB phase is nonzero ($= \ln 2$) at large l_B. We now prove that such nonvanishing CMI is a universal feature of SWSSB states, which is connected to the local nonrecoverability of (symmetry-breaking) local channels.

Theorem 4. If a mixed state ρ has SWSSB in the fidelity correlator sense Eq. (7), then $\exists$ some local channel $\mathcal{E}_i$ (supported around site i) that cannot be recovered locally, namely $\nexists$ local $\mathcal{E}_i'$ such that $\mathcal{E}_i' \circ \mathcal{E}_i[\rho] = \rho$.

Furthermore, for any local $\mathcal{E}_i'$, the trace distance $D(\rho, \mathcal{E}_i' \circ \mathcal{E}_i[\rho]) = O(1) > 0$. As a consequence, the CMI $I(A;C|B)_\rho$ (in the geometry of Fig. 2) must remain nonzero at large distance between A and C, for a region A containing the support of $\mathcal{E}_i$ around site i.

Proof. Let us normalize the order parameter O_i such that the following operation is a valid quantum channel:

$$\mathcal{E}_i[\rho] \equiv \frac{1}{2}\rho + \frac{1}{2}O_i\rho O_i^\dagger + \frac{1}{2}\sum_a K_{i,a}\rho K_{i,a}^\dagger, \quad (92)$$

where the local Kraus operators $\{K_{i,a}\}$ are chosen so that $O_i^\dagger O_i + \sum_a K_{i,a}^\dagger K_{i,a} = I$. For example, if O_i is unitary, then we can choose $K_i = 0$. In the Ising example $O = Z$ and the channel is nothing but a strong Z measurement.

Now consider another arbitrary channel around site i, call it $\mathcal{E}_i'$, and define

$$\tilde{\rho} \equiv \mathcal{E}_i' \circ \mathcal{E}_i[\rho]. \tag{93}$$

To see that $\tilde{\rho} \neq \rho$, it suffices to show that for some site j far away from site i, $F(\tilde{\rho}, O_j^\dagger \tilde{\rho} O_j) = O(1) > 0$, namely $\tilde{\rho}$ does not even have the strong symmetry. This claim follows simply from the joint concavity and the data-processing inequality of fidelity:

$$\begin{aligned} F(\tilde{\rho}, O_j^\dagger \tilde{\rho} O_j) &\geq \frac{1}{2} F(\mathcal{E}_i'[\rho], \mathcal{E}_i'[O_i O_j^\dagger \rho O_j O_i^\dagger]) \\ &\geq \frac{1}{2} F(\rho, O_i O_j^\dagger \rho O_j O_i^\dagger) \\ &= O(1) > 0. \end{aligned} \tag{94}$$

Next we examine more carefully the trace distance $D(\rho, \tilde{\rho})$. Using the triangle inequality of trace distance, we have

$$\begin{aligned} D(\rho, \tilde{\rho}) &\geq D(\rho, O_j^\dagger \rho O_j) - D(\tilde{\rho}, O_j^\dagger \rho O_j) \\ &\geq D(\rho, O_j^\dagger \rho O_j) - D(\tilde{\rho}, O_j^\dagger \tilde{\rho} O_j) \\ &\quad - D(O_j^\dagger \rho O_j, O_j^\dagger \tilde{\rho} O_j). \end{aligned} \tag{95}$$

Now we use several simple observations:

(1) $D(\rho, O_j^\dagger \rho O_j) = 1$ by strong symmetry;
(2) $D(\tilde{\rho}, O_j^\dagger \tilde{\rho} O_j) = 1 - c$ for some $c > 0$ due to Eq. (94);
(3) using the submultiplicative property of trace norm, we have

$$D(O_j^\dagger \rho O_j, O_j^\dagger \tilde{\rho} O_j) \leq \|O_j\|^2 D(\rho, \tilde{\rho}), \tag{96}$$

where $\|O_j\| \sim O(1)$ is the operator norm of O_j. Putting all these relations back into Eq. (95), we obtain

$$D(\rho, \tilde{\rho}) \geq c/(1 + \|O_j\|^2) > 0. \tag{97}$$

The above relation holds for any attempted recovery channel $\mathcal{E}'$. In particular, if we choose $\mathcal{E}_i'$ to be the Petz recovery map $\mathcal{R}_i$, then we have [see Eq. (D7) in Appendix D]

$$D(\rho, \mathcal{R}_i \circ \mathcal{E}_i[\rho]) \leq \sqrt{\ln 2 \cdot I(A; C|B)_\rho}, \tag{98}$$

which immediately implies that the CMI $I(A; C|B)_\rho$ must remain nonzero at large dist(A, C). ■

We dub this theorem *the spontaneity of SWSSB*, for the following reason. Recall in pure-state SSB, the cat (GHZ) state $|\uparrow \cdots \uparrow\rangle + |\downarrow \cdots \downarrow\rangle$ is sometimes called "unphysical" in the literature, which really means that the state is extremely unstable towards local decoherence: if the environment "measures" the Z spin at site i, the state will immediately collapse to either $|\uparrow \cdots \uparrow\rangle$ or $|\downarrow \cdots \downarrow\rangle$. In a real system, even if the measurement strength (the probability for performing the measurement, δP) from the environment is small at each site, the probability for the entire system to remain in the cat state vanishes exponentially with the system size. Even though almost any state will go through some deformation under decoherence, the important feature of cat state is that the postmeasurement state (a product state $|\text{Prod}\rangle$) is very "far away" from the original state (a cat $|\text{GHZ}\rangle$), in the sense that there is no local deformation (local channel) near the measurement location that takes $|\text{Prod}\rangle$ to $|\text{GHZ}\rangle$. We therefore dub such local irrecoverability of local measurements "(weak) spontaneity." We note that weak spontaneity holds even for weak-to-trivial SSB, as the above logic applies equality well to the weakly symmetric state $|\uparrow \cdots \uparrow\rangle\langle\uparrow \cdots \uparrow| + |\downarrow \cdots \downarrow\rangle\langle\downarrow \cdots \downarrow|$.

The above notion of spontaneity is "weak" in the sense that the irrecoverability holds for certain measurement results. If we average over (or simply forget) the measurement results, the postmeasurement state may become locally recoverable. For example, take the weak-to-trivial example $\rho = |\uparrow \cdots \uparrow\rangle\langle\uparrow \cdots \uparrow| + |\downarrow \cdots \downarrow\rangle\langle\downarrow \cdots \downarrow|$, the postmeasurement density matrix is $\mathcal{E}[\rho] = (1/2)\rho + (1/2)Z_i \rho Z_i = \rho$. If the postmeasurement density matrix (without postselecting the measurement outcome) cannot be locally recovered to the original state, we dub the spontaneity "strong"—this is a stronger statement since if the postmeasurement state is locally recoverable for all measurement results, it is simple to construct a recovery channel that recovers the whole density matrix, by taking the convex sum of the corresponding recovery channel for each measurement result.

In this language, what we have shown in Theorem 4 is that the SSB of strong symmetry (either to weak or to trivial) is *strongly spontaneous*. Let us consider a pedagogical example of SWSSB: $\rho_0 \propto \mathbb{1} + X$ as the SWSSB of $\mathbb{Z}_2$ Ising symmetry (Sec. II A). We perform a strong Z measurement on a single qubit and recognize that ρ_0 will immediately collapse to the maximally mixed state $\mathbb{1}$ that no longer has the strong Ising symmetry. Heuristically, the *only* information encoded in the SWSSB density matrix is the global charge sector fixing, i.e., ρ_0 is the maximally mixed state within the even charge sector of the Ising symmetry. Once we perform a strong Z measurement, the information on global charge sector fixing is completely destroyed, and we can never restore this information with any local quantum channel.

Furthermore, the spontaneity theorem also implies a nonvanishing CMI is a universal feature of the SWSSB density matrix. Similar to the stability theorem, this result only requires the fidelity correlator to be nonvanishing, and is insensitive to the ordinary linear correlation function. This means that Theorem 4 also holds for strong-to-trivial SSB—the most familiar example is again the GHZ state, which has CMI $\ln 2$. Also, note that weak-to-trivial SSB does not have a similar property. For example, $\rho = \frac{1}{2}(|\uparrow\cdots\uparrow\rangle\langle\uparrow\cdots\uparrow| + |\downarrow\cdots\downarrow\rangle\langle\downarrow\cdots\downarrow|)$ has vanishing CMI—from earlier discussion, this is because weak-to-trivial SSB is only “weakly spontaneous.”

For pure states, the spontaneity of SSB is quantitatively described by a divergent dynamical susceptibility. For example, the magnetic susceptibility of the GHZ state is divergent with respect to the external Zeeman field along the Z direction. Following the initial version of our manuscript, in Ref. [52] the authors proposed a *fidelity susceptibility*, which diverges in the SWSSB phase and remains finite in the symmetric phase.

V. SUMMARY AND OUTLOOK

In this work, we systematically discussed a new phenomenon in mixed quantum states, namely strong-to-weak spontaneous symmetry breaking. Our key results are as follows:

(1) We defined SWSSB through the long-ranged fidelity correlator Eq. (7) of charged local operators. We proved (Theorems 1 and 2) that SWSSB is stable against strongly symmetric low-depth local quantum channels. These stability theorems established SWSSB as a universal property of mixed-state quantum phases.

(2) We also considered an alternative definition of SWSSB, which appeared in the previous literature [3,4,37], by the long-ranged Rényi-2 correlator Eq. (30) of charged local operators. The two definitions of SWSSB are inequivalent: there are examples with long-range fidelity correlators but short-range Rényi-2 correlators, and vice versa (Appendix A). Moreover, there is no analog of the stability theorems (Theorems 1 and 2) for the Rényi-2 correlator, as we showed through an explicit example in Sec. IV. So the SWSSB defined through the Rényi-2 correlator is not a universal property of mixed-state quantum phases. However, a weaker statement holds: a state with SWSSB in the Rényi-2 sense must still be nontrivial, in the sense that it is not symmetrically invertible (Theorem 3).

(3) We proposed that SWSSB is a generic feature of thermal states at nonzero temperature (Sec. III A). In particular, we conjectured that if the canonical ensemble (i.e., with fixed charge sector) of a local symmetric Hamiltonian without a strong-to-trivial SSB, then it must have SWSSB. We demonstrated the statement in certain commuting-projector models and provided plausibility arguments for general Hamiltonians.

(4) We also studied nonthermal examples of SWSSB, in models where a pure symmetric state is subject to a strongly symmetric finite-depth channel. In particular, in Sec. III B we considered a $(2+1)d$ quantum Ising model under a strongly symmetric nearest-neighbor ZZ channel and found that the Ising SWSSB transition (measured by the fidelity correlator) is described by the $2d$ random-bond Ising model along the Nishimori line. The SWSSB transition is the ungauged version (through Kramers-Wannier duality) of the decodability transition of the $(2+1)d$ toric-code model under the bit-flip noise. In addition, we also considered the $U(1)$ SWSSB in a model of quantum rotors. By mapping to a Villain model with imaginary bond disorder we argued that the model exhibits $U(1)$ SWSSB for a sufficiently strong channel.

(5) Following Ref. [16], we discussed the recoverability of mixed states under symmetric low-depth channels. Specifically, a necessary condition for the recoverability of a channel circuit through a low-depth symmetric channel is the Markov gap condition Eq. (85). For the example of the decohered Ising model in Sec. III B, we showed that the Markov gap condition is violated only across the SWSSB transition (measured by the fidelity correlator). As a consequence, if two states belong to the same side of the SWSSB transition, they belong to the same phase, in the sense that they are two-way connected to each other through Log-depth symmetric channels.

(6) We demonstrated the spontaneity of SWSSB and its relation to the local irrecoverability of local symmetry-breaking channels. Specifically, we proved that for a density matrix with SWSSB, a local quantum channel can induce a global change that cannot be locally recovered (Theorem 4). This further implied that a nonvanishing CMI is a universal feature of SWSSB.

We end this work with some open questions:

(1) Strong-to-weak SSB of higher-form symmetry: A topologically ordered state can be understood as the spontaneous symmetry breaking of some higher-form symmetry [53,54]. Then what about the SWSSB of higher-form symmetry? After the initial arXiv post of our work, the Rényi-2 version of this question has been addressed in Ref. [55].

(2) Ground-state SSB comes with many dynamical features, such as Anderson tower and Goldstone modes (if the symmetry is continuous). It is natural to

ask whether SWSSB has similar dynamical consequences, for example, in Lindbladian dynamics. Several prior works [31,37,56] addressed this issue partially, but a general picture is still lacking at this point. blackIn particular, a follow-up work [57] showed that the Goldstone mode of $U(1)$ SWSSB can be identified as the diffusive hydrodynamic mode, using Schwinger-Keldysh effective field theory.

(3) Conventional SSB features various topological defects, such as domain walls, vortices, and skyrmions. A natural avenue for future work is to consider whether SWSSB can have similar defects and to explore the consequences of such defects.

(4) Can we experimentally measure the fidelity correlator, which has played a central role in our theory, in some practical setup? This may be nontrivial given that fidelity is in general rather challenging to measure. The Rényi correlator may serve as a reasonable (though not perfect) proxy on this regard. Another proxy to the fidelity correlator is the trajectory-resolved Edward-Anderson correlator in monitored dynamics, which has been measured recently in Ref. [58].

Note added.—We would like to draw the reader's attention to a related work [59] on SWSSB to appear in the same arXiv listing.

ACKNOWLEDGMENTS

We thank Yimu Bao, Tarun Grover, Timothy Hsieh, Andy Lucas, Zhu-Xi Luo, Max Metlitski, Shengqi Sang, Brian Swingle, Ruben Verresen, Xiao-Gang Wen, Yichen Xu, Carolyn Zhang, Zhehao Zhang, and Yijian Zou for inspiring discussions. R.M. is especially grateful to Shengqi Sang for many discussions and for explaining Ref. [16]. J.H.Z. and M.C. are grateful to Hong-Hao Tu and Wei Tang for discussions and collaborations on a related project. R.M. was supported in part by Simons Investigator Award No. 990660 and by the Simons Collaboration on Ultra-Quantum Matter, which is a grant from the Simons Foundation (No. 651440). J.H.Z. and Z.B. acknowledge support from NSF under Award No. DMR-2339319. R.M., J.H.Z., Z.B., and C.W. are supported in part by Grant No. NSF PHY-2309135 to the Kavli Institute for Theoretical Physics. L.A.L. acknowledges support from the Natural Sciences and Engineering Research Council of Canada (NSERC) through Discovery Grants. Research at Perimeter Institute (L.A.L. and C.W.) is supported in part by the Government of Canada through the Department of Innovation, Science and Industry Canada and by the Province of Ontario through the Ministry of Colleges and Universities. M.C. acknowledges support from NSF under Award No. DMR-1846109.

APPENDIX A: COUNTEREXAMPLE FOR SWSSB ONLY IN THE RÉNYI-2 SENSE

Here, we are going to show N-qubit states ρ_N that are strongly symmetric under $\mathbb{Z}_2$, exhibit SWSSB in the sense of the Rényi-2 correlator (Definition 4), but not in the fidelity sense (Definition 1). The strategy will be to construct "spin-glass" ensembles $\rho_N \propto \sum_{m=1}^{M_N} p_m \,|\psi_m\rangle\langle\psi_m|$, $\lim_{N\to\infty} M_N = \infty$, such that $\lim_{|x-y|\to\infty}\lim_{N\to\infty} F_{O,N}(x,y) = 0$ for all order parameters O, but with its Rényi-2 correlator satisfying

$$\lim_{|x-y|\to\infty}\lim_{N\to\infty} R^{(2)}_{O',N}(x,y) = \frac{\sum_m p_m^2 |o'_m|^2}{\sum_m p_m^2} > 0, \tag{A1}$$

for the particular order parameter $O' = Z$, with $\lim_{|x-y|\to\infty}\langle\psi_m|Z_x Z_y|\psi_n\rangle = o'_m\delta_{mn}$. Furthermore, we require $\sum_m p_m$ to diverge but $\sum_m p_m|o'_m|$, $\sum_m p_m^2|o'_m|^2$ and $\sum_m p_m^2$ to converge. In that way, at least for o', the fidelity correlator goes to zero, since

$$\lim_{|x-y|\to\infty}\lim_{N\to\infty} F_{O',N}(x,y) = \frac{\sum_m p_m|o'_m|}{\sum_m p_m} = 0, \tag{A2}$$

and we have dropped the normalization condition $\sum_m p_m = 1$ for simplicity of calculation.

Let us consider a qubit chain with $2N$ spins, and define the strong $\mathbb{Z}_2$ symmetry to only act on even spins, meaning the symmetry is generated by $X_e := \otimes_n X_{2n}$. We now consider all length-N bit strings $m_1 m_2\cdots m_N$, and view each bit string as the binary representation of an integer $m\in\{1,2,\ldots,2^N\}$. We can then construct the density matrix $\rho_N = \sum_{m=1}^{2^N} p_m\,|\psi_m\rangle\langle\psi_m|$, where $p_m = m^{-2/3}$ and

$$\begin{aligned}|\psi_m\rangle &= \frac{\mathbb{1}+X_e}{\sqrt{2}}|m_1\tilde{m}_1 m_2\tilde{m}_2\cdots\rangle,\\ |\tilde{m}_i\rangle &:= \sqrt{\frac{1}{2}+\frac{1}{m}}|m_i\rangle + \sqrt{\frac{1}{2}-\frac{1}{m}}X_i|m_i\rangle,\end{aligned} \tag{A3}$$

with $|m_i\rangle$ defined in the Z basis. At large m, $|\tilde{m}_i\rangle \approx |+\rangle + [(-1)^{m_i}/m]|-\rangle$, so we should have $o'_m \sim 1/m^2$. The auxiliary spins on the odd sites are introduced to make sure that different $|\psi_m\rangle$'s are orthogonal with respect to O', allowing the simple forms of the correlators as in Eq. (A1).

Given that $p_m \sim m^{-2/3}$ and $o'_m \sim 1/m^2$, then $\sum_m p_m|o'_m|$, $\sum_m p_m^2|o'_m|^2$ and $\sum_m p_m^2$ all converge to an $O(1)$ value, while $\sum_m p_m \sim m_{\max}^{1/3} \sim 2^{N/3}$. We conclude that while the Rényi-2 remains finite, the fidelity correlator for $O' = Z$ vanishes exponentially in the thermodynamic limit. For other order parameters, the fidelity correlator still goes to zero, as ρ_N converges to the trivial state $\rho_{\mathrm{odd}}\otimes|+\rangle\langle+|_{\mathrm{even}}$, $\rho_{\mathrm{odd}} \propto \sum_m p_m\,|m\rangle\langle m|$, in the fidelity (or trace) distance as

$N \to \infty$:

$$F(\rho_N, \rho_{\mathrm{odd}} \otimes |+\rangle\langle+|_{\mathrm{even}}) \approx \frac{\sum_m p_m(1 - N/2m^2)}{\sum_m p_m} \to 1. \tag{A4}$$

It is clear from this construction, however, that this sequence of states is highly artificial. Thus, it remains a question of whether the same can happen for states arising from more natural contexts, such as thermal or decohered states.

APPENDIX B: ALTERNATIVE DERIVATION OF THE FIDELITY CORRELATOR

We present an alternative derivation of the fidelity correlator in the decohered 2D Ising model Eq. (57) using the replica trick.

Then we utilize the replica trick to calculate the fidelity: the replicated fidelity is defined as

$$F_{m,n}(x,y) = \mathrm{Tr}[(\rho^m \sigma \rho^m)^n], \tag{B1}$$

and the standard definition of fidelity (7) is recovered by taking the limit $m, n \to 1/2$. In the "open string" basis, the density matrix $\mathcal{E}[\rho_0] = \rho$ is given by the following expression:

$$\rho = (1-p)^{2N_v} \sum_{\mathbf{l}} (\tanh\tau)^{|\mathbf{l}|} |\partial\mathbf{l}\rangle\langle\partial\mathbf{l}|, \tag{B2}$$

where $\mathbf{l}$ labels the configurations of "open strings," $\tanh\tau = p/(1-p)$, $|\partial\mathbf{l}\rangle = \prod_{v\in\partial\mathbf{l}} Z_v|++\cdots+\rangle$, and N_v is the total number of vertices. Then our task is the calculation of the (replicated) fidelity Eq. (B1). Define $t = (2m+1)n$. For simplicity, we focus on the special case with $n = 1$ but the derivation can be easily generalized to $n > 1$. We find

$$\begin{aligned}\mathrm{Tr}(\rho^m \sigma \rho^m) &\propto \sum_{\mathbf{l}_1,\ldots,\mathbf{l}_t} (\tanh\tau)^{|\mathbf{l}_1|+\cdots+|\mathbf{l}_t|} \langle\partial\mathbf{l}_1|\partial\mathbf{l}_2\rangle \cdots \langle\partial\mathbf{l}_m|Z_xZ_y|\partial\mathbf{l}_{m+1}\rangle\langle\partial\mathbf{l}_{m+1}|Z_xZ_y|\partial\mathbf{l}_{m+2}\rangle \cdots \langle\partial\mathbf{l}_t|\partial\mathbf{l}_1\rangle \\ &= \sum_{\{\sigma_v^{(1)},\ldots,\sigma_v^{(t)}\}=\pm1} \sigma_x^{(m)}\sigma_y^{(m)} \prod_e \prod_{j=1}^t \left(1 + \tanh\tau \prod_{v\in\partial e} \sigma_v^{(j)}\right) = \langle\sigma_x\sigma_y\rangle_{H_{\mathrm{eff}}},\end{aligned} \tag{B3}$$

here we utilized the following formula:

$$\langle\partial\mathbf{l}_1|\partial\mathbf{l}_2\rangle = \frac{1}{2^{N_v}} \sum_{s_v} \left(\prod_{v_1\in\partial\mathbf{l}_1} s_{v_1}\right)\left(\prod_{v_2\in\partial\mathbf{l}_2} s_{v_2}\right), \tag{B4}$$

and defined $\sigma_{v_j}^{(j)} = s_{v_j}^{(j)} s_{v_j}^{(j+1)}$, which leads to the constraint $\prod_{j=1}^t \sigma_v^{(j)} = 1$ for all sites v. The replicated fidelity has been expressed as the correlation function of the following Hamiltonian at the inverse temperature $\beta = \tau$:

$$H_{\mathrm{eff}} = -\sum_{\langle i,j\rangle}\left(\sum_{k=1}^{t-1} \sigma_i^{(k)}\sigma_j^{(k)} + \prod_{k=1}^{t-1} \sigma_i^{(k)}\sigma_j^{(k)}\right). \tag{B5}$$

For $n > 1$ the derivation is very similar, and we find the replicated fidelity is the following correlation function of the Hamiltonian (B5):

$$F_{m,n}(x,y) = \left\langle \prod_{\alpha=1}^n \sigma_x^\alpha \sigma_y^\alpha \right\rangle_{H_{\mathrm{eff}}}. \tag{B6}$$

The qualitative behavior of $F_{m,n}(\rho,\sigma)$ is essentially determined by the effective Hamiltonian (B5). This can also be understood from the replica partition function $\mathrm{Tr}\,\rho^t$:

$$\mathrm{Tr}(\rho^t) \propto \sum_{\{\sigma_v^{(1)},\ldots,\sigma_v^{(t)}\}} \exp\left\{-\tau\sum_{\langle i,j\rangle}\left(\sum_{k=1}^{t-1}\sigma_i^{(k)}\sigma_j^{(k)} + \prod_{k=1}^{t-1}\sigma_i^{(k)}\sigma_j^{(k)}\right)\right\}, \tag{B7}$$

which is exactly the partition function of the effective Hamiltonian H_{eff}. In particular, if we take the limit $m, n \to 1/2$, the replicated fidelity (B1) will be recovered back to Eq. (7), and effective Hamiltonian (B5) becomes the Hamiltonian of random-bond Ising model (RBIM) along the Nishimori line [11]. Then the transition between a $\mathbb{Z}_2$ strongly symmetric state at $p < p_c$ and an SWSSB state induced by a symmetric quantum channel at $p > p_c$ happens at $p_c \approx 0.109$. We note that the same result for the fidelity correlator can be established directly without using the replica trick, which is explained in Appendix B. Our results are fully consistent with the decodability transition of the toric code [10,11,14,50,60], which can now be interpreted as the "gauged" version of the SWSSB transition through Kramers-Wannier transform—recall that gauge symmetry is naturally strong since the total gauge charge is constrained to be trivial in the Hilbert space.

APPENDIX C: $U(1)$ RÉNYI-2 SWSSB IN THE ROTOR MODEL

In this section, we study the Rényi-2 SWSSB transition in the quantum rotor model.

The initial state ρ_0 is given by

$$\rho_0 \propto \int_{\theta,\theta'} |\theta\rangle\langle\theta'|\,. \tag{C1}$$

It is easy to see that

$$\mathcal{E}(\rho_0) \propto \int_{\theta,\theta'} \prod_{\langle ij\rangle} G_{ij}(\theta,\theta')\,|\theta\rangle\langle\theta'|\,, \tag{C2}$$

where the factor $G_{ij}(\theta,\theta')$ is defined as

$$G_{ij}(\theta,\theta') = \sum_k p_k e^{ik(\theta_i-\theta_j-\theta_i'+\theta_j')}. \tag{C3}$$

Therefore,

$$\mathrm{tr}\,\rho^2 \propto \int_{\theta,\theta'} \prod_{\langle ij\rangle} G_{ij}^2(\theta,\theta'). \tag{C4}$$

Let us write

$$G_{ij}^2(\theta,\theta') = \sum_n f_n e^{in(\theta_i-\theta_j-\theta_i'+\theta_j')}. \tag{C5}$$

Up to an overall constant, it is easy to see that $\mathrm{Tr}\,\rho^2$ is proportional to the following partition function:

$$Z = \int \mathcal{D}\theta \prod_{\langle ij\rangle} \sum_{n_{ij}} f_{n_{ij}} e^{in_{ij}(\theta_i-\theta_j)}. \tag{C6}$$

This can be viewed as a variant of the classical XY model.

For $p_k \propto e^{-\alpha k^2}$, after some algebra we obtain

$$f_n \propto \begin{cases} e^{-\frac{\alpha}{2}n^2}\vartheta_3(e^{-2\alpha}) & n \text{ even} \\ e^{-\frac{\alpha}{2}n^2}\vartheta_2(e^{-2\alpha}) & n \text{ odd} \end{cases}. \tag{C7}$$

Here ϑ_2 and ϑ_3 are Jacobi θ functions, defined as

$$\vartheta_2(q) = \sum_{k\in\mathbb{Z}} q^{(k+\frac{1}{2})^2}, \quad \vartheta_3(q) = \sum_{k\in\mathbb{Z}} q^{k^2}. \tag{C8}$$

A useful fact is that $\vartheta_3(q)$ and $\vartheta_2(q)$ are exponentially close when q is close to 1 (the difference between them is approximately $\sqrt{\frac{\pi}{1-q}}e^{-\frac{\pi^2}{1-q}}$). For small α, we have $e^{-2\alpha} \approx 1$, so $f_n \propto e^{-\frac{\alpha}{2}n^2}$. Using Poisson resummation we find

$$\sum_{n_{ij}\in\mathbb{Z}} e^{-\frac{\alpha}{2}n_{ij}^2} e^{in_{ij}(\theta_i-\theta_j)} = \sqrt{\frac{2\pi}{\alpha}} \sum_{m_{ij}\in\mathbb{Z}} e^{-\frac{1}{2\alpha}(\theta_i-\theta_j+2\pi m_{ij})^2}, \tag{C9}$$

which is the familiar Villain model at temperature α. The approximation becomes asymptotically exact as $\alpha \to 0$. Thus we can conclude that for sufficiently small α there is Rényi-2 SWSSB of the $U(1)$ symmetry.

In addition, in two dimensions the model exhibits Rényi-2 quasi-SWSSB. On a square lattice, numerically it is known that the KT point is $\alpha_c \approx 1.353$ [61]. For this range of α, the approximation $f_n \propto e^{-\alpha n^2/2}$ remains reasonable, so we expect that our model is qualitatively similar to the Villain model, exhibiting a KT transition from the quasi-SWSSB phase to the strongly symmetric phase.

APPENDIX D: THE RECOVERY CHANNEL

In this Appendix, we provide the details of the recovery channel as outlined in Sec. IV and originally proposed in Ref. [16]. In particular, we aim to construct a short-depth recovery channel for a finite-depth quantum channel that preserves the gap of a mixed state. Here the gap of a mixed state is defined as follows.

Definition 5. A mixed state ρ has a Markov gap Δ if the conditional mutual information for the three regions depicted in Fig. 2 decays exponentially, i.e.,

$$\begin{aligned} I(A;C|B)_\rho &\le 2\ln(\min\{d_A,d_C\})\exp(-\Delta\cdot\mathrm{dist}(A,C)), \\ &= \mathrm{const}\cdot\min(|A|,|C|)\exp(-\Delta\cdot\mathrm{dist}(A,C)), \end{aligned} \tag{D1}$$

where $\mathrm{dist}(A,C)$ is the closest distance between the region A and C. $|A|$ and $|C|$ represent the number of sites within regions A and C, respectively.

We observe that any mixed state has a gap $\Delta \ge 0$ due to the local dimension bound.

1. Recovery of a local channel

The first result states that, for a gapped mixed state, the effect of a local channel $\mathcal{E}$ acting on region A can be recovered by another local channel $\mathcal{R}$, supported on a region $A\cup B$ surrounding A, up to exponential decaying errors. The argument goes as follows. For any state ρ, non-negative operator σ and quantum channel $\mathcal{E}$, we have the data-processing inequality:

$$D(\rho\|\sigma) - D(\mathcal{E}(\rho)\|\mathcal{E}(\sigma)) \geq -2\log_2 F(\rho, \mathcal{R}_{\sigma,\mathcal{E}} \circ \mathcal{E}(\rho)), \tag{D2}$$

where F is the fidelity, and

$$\mathcal{R}_{\sigma,\mathcal{E}}(\cdot) := \int_t \beta_0(t)\mathcal{R}^{\frac{t}{2}}_{\sigma,\mathcal{E}}(\cdot) \tag{D3}$$

is the rotated Petz recovery map [62]. Here $\beta_0(t) = \frac{\pi}{2}(\cosh(\pi t) + 1)^{-1}$ and

$$\mathcal{R}^t_{\sigma,\mathcal{E}}(\cdot) := \sigma^{\frac{1}{2}-it}\mathcal{E}^\dagger[\mathcal{E}(\sigma)^{-\frac{1}{2}+it}(\cdot)\mathcal{E}(\sigma)^{-\frac{1}{2}-it}]\sigma^{\frac{1}{2}+it}. \tag{D4}$$

Let us assume: (1) $\mathcal{E}$ acts nontrivially only in region A; (2) $\sigma = \rho_{AB} \otimes \rho_C$. We then have

$$\begin{aligned} &-2\log_2 F(\rho, \mathcal{R}_{\sigma,\mathcal{E}} \circ \mathcal{E}(\rho)) \leq I(AB;C)_\rho - I(AB;C)_{\mathcal{E}(\rho)} \\ &\quad = I(A;C|B)_\rho - I(A;C|B)_{\mathcal{E}(\rho)}. \end{aligned} \tag{D5}$$

Due to the positive-semidefinite nature of the CMI, as well as the data-processing inequality for local channels (acting on the nonconditioning systems), we have

$$0 \leq I(A;C|B)_{\mathcal{E}(\rho)} \leq I(A;C|B)_\rho. \tag{D6}$$

Therefore, the left-hand side of Eq. (D5) is upper bounded by a non-negative number exponentially small in $\mathrm{dist}(A, C) \sim l_B$. By the Fuchs-van de Graaf inequality, at large distances, the constraint on the fidelity can also be expressed as a constraint on the trace distance,

$$\begin{aligned} \|\rho - \mathcal{R}_{\sigma,\mathcal{E}} \circ \mathcal{E}(\rho)\|_1 &\leq \sqrt{4\ln 2 \cdot I(A;C|B)_\rho} \\ &\simeq \mathrm{const} \cdot e^{-\Delta\cdot \mathrm{dist}(A,C)/2}. \end{aligned} \tag{D7}$$

Equation (D7) implies that when the underlying state ρ has a gap $\Delta > 0$, the impact of a local channel can be recovered locally, with an error decaying exponentially in the width of the buffer region B.

2. Symmetries

Now let us show that when the channel $\mathcal{E}$ is strongly symmetric under a unitary on-site symmetry, the associated recovery channel $\mathcal{R}$ is also strongly symmetric. Using the explicit form of the rotated Petz recovery map in Eq. (D4), one has

$$\mathcal{R}^t_{\sigma,\mathcal{E}}(\cdot) = \rho_{AB}^{-it+\frac{1}{2}}\mathcal{E}^\dagger[\mathcal{E}(\rho_{AB})^{it-\frac{1}{2}}(\cdot)\mathcal{E}(\rho_{AB})^{-it-\frac{1}{2}}]\rho_{AB}^{it+\frac{1}{2}}. \tag{D8}$$

If we pick an arbitrary Kraus representation of $\mathcal{E}$, i.e., $\mathcal{E}(\rho) = \sum_i K_i^\dagger \rho K_i$, a corresponding Kraus representation of the recovery map is

$$K^\dagger_{\mathcal{R},i} = \rho_{AB}^{-it+\frac{1}{2}} K_i \mathcal{E}(\rho_{AB})^{it-\frac{1}{2}}. \tag{D9}$$

In order for the recovery channel to be strongly symmetric, we only need ρ_{AB} to be *weakly* symmetric under U_{AB}, which is the symmetry operation restricted to the region AB. This can be explicitly verified as follows:

$$\begin{aligned} U_{AB}\rho_{AB}U^\dagger_{AB} &= U_{AB}\,\mathrm{tr}_{\overline{A\cup B}}(\rho)U^\dagger_{AB} \\ &= \mathrm{tr}_{\overline{A\cup B}}(U\rho U^\dagger) = \rho_{AB}, \end{aligned} \tag{D10}$$

where U is the generator of the global symmetry, and the last line follows from the fact that the state ρ is strongly symmetric. Applying a similar argument, one can also demonstrate that $U_{AB'}\mathcal{E}(\rho_{AB})U^\dagger_{AB'} = \mathcal{E}(\rho_{AB})$, as long as the channel $\mathcal{E}$ (acting on A) is strongly symmetric and we enlarge AB to AB' to accommodate the depth of $\mathcal{E}$. As a summary, each Kraus operator $K^\dagger_{\mathcal{R},i}$ of the recovery channel commutes with U, therefore the recovery map is strongly symmetric.

3. Global recovery

Now, we proceed to demonstrate that a symmetric finite-depth circuit channel $\mathcal{E}$, which acts on the entire system and preserves the gap of the underlying mixed state, can also be recovered by a short-depth channel. We assume that the maximal width of gates in $\mathcal{E}$ is w, and $\mathcal{E}$ has l layers of gates. The depth of $\mathcal{E}$ is then defined to be wl. The global recovery needs two crucial assumptions: (a) After each layer of $\mathcal{E}$, the state, referred to as ρ_k with $0 \leq k \leq l$, has a finite gap lower bounded by 4Δ. (b) Another important requirement is, if we act the state ρ_{k-1} with any subset of gates in kth layer of $\mathcal{E}$, the resulting state remains gapped, also lower bounded by 4Δ. In the context of the RBIM, this assumption implies that in the intermediate states, where a part of the system has been recovered, the entire system, despite its inhomogeneity, remains in a definite phase of the RBIM with a finite correlation length.

The strategy of the global recovery is to recover $\mathcal{E}$ layer by layer. We cut the system into hypercubes of linear size L_0—thus we have $(L/L_0)^d$ hypercubes in total, where d is the spatial dimension of the system. Here L_0 is assumed to be larger than w and the correlation length $1/2\Delta$. We first reorganize each layer of $\mathcal{E}$ into 2^d new layers. In other words, we now perceive the channel $\mathcal{E}$ as a circuit with $h = 2^d l$ layers, but the gates become "sparse." Specifically, in each new layer, within a hypercube of size $2L_0$, only a hypercube of size L_0 contains nontrivial gates. Crucially, for each new layer, these nontrivial hypercubes are separated by a minimum distance of L_0. In the first recovery step, we reverse all gates in the last (hth) layer of the new circuit. We label the hypercubes undergoing recovery as C_x

with $x \in 1, \dots, (L/2L_0)^d$, the gates in the hth layer within C_x as $\mathcal{E}_x^h$, and the corresponding recovery channels as $\mathcal{R}_x^h$. The recovery map $\mathcal{R}_x^h$ is chosen to be the local recovery map mentioned in Sec. D 1, confined to a hypercube C_x' with a linear size of $2L_0$, and sharing the same center as the hypercube x. Note that these recovery channels act on disjoint hypercubes, and can therefore be implemented simultaneously in a single layer. Once more, we denote the state after each new layer of $\mathcal{E}$ as ρ_k, now with $0 \le k \le 2^d l$. We have

$$\begin{aligned}
&\left\|\prod_x \mathcal{R}_x^h(\rho_h) - \rho_{h-1}\right\|_1 \\
&\quad \le \sum_{y=1}^{(L/2L_0)^d} \left\|\prod_{z=1}^{y} \mathcal{R}_z^h \circ \mathcal{E}_z^h(\rho_{h-1}) - \prod_{z=1}^{y-1} \mathcal{R}_z^h \circ \mathcal{E}_z^h(\rho_{h-1})\right\|_1 \\
&\quad \le \sum_{y=1}^{(L/2L_0)^d} \left\|\mathcal{R}_y^h \circ \mathcal{E}_y^h(\rho_{h-1}) - \rho_{h-1}\right\|_1 \\
&\quad \le \mathrm{const}\cdot\left(\frac{L}{2L_0}\right)^d \cdot \mathrm{Poly}(L_0)\cdot e^{-\Delta\cdot L_0},
\end{aligned} \tag{D11}$$

where we have used the triangle inequality and the data-processing inequality of the trace distance (note that recovery channels with distinct labels act in disjoint regions, and therefore commute).

Finally, we proceed to recover the channel $\mathcal{E}$ layer by layer. At each step of the recovery, we reverse gates in $(L/2L_0)^d$ hypercubes with a linear size of L_0. Denote the kth layer of gates in $\mathcal{E}$ as $\mathcal{E}^k$, and the corresponding recovery channel by $\mathcal{R}^k$. We have

$$\begin{aligned}
&\left\|\prod_{k=1}^{h} \mathcal{R}^k \circ \mathcal{E}(\rho) - \rho\right\|_1 \\
&\quad = \left\|\sum_{k=1}^{h}[\prod_{m=1}^{k} \mathcal{R}^m \circ \mathcal{E}^m(\rho) - \prod_{m=1}^{k-1} \mathcal{R}^m \circ \mathcal{E}^m(\rho)]\right\|_1 \\
&\quad \le \sum_{k=1}^{h}\left\|\prod_{m=1}^{k-1} \mathcal{R}^m[\mathcal{R}^k \circ \mathcal{E}^k(\rho_{k-1}) - \rho_{k-1})]\right\|_1 \\
&\quad \le \sum_{k=1}^{h}\left\|\mathcal{R}^k \circ \mathcal{E}^k(\rho_{k-1}) - \rho_{k-1}\right\|_1 \\
&\quad \le \mathrm{const}\cdot\left(\frac{L}{2L_0}\right)^d l\cdot \mathrm{Poly}(L_0)\cdot e^{-\Delta\cdot L_0}.
\end{aligned} \tag{D12}$$

Here, $\rho = \rho_0$ represents the original state, and all layers are time ordered. Specifically, all layers in $\mathcal{E}$ are time ordered, followed by layers in the recovery map with anti-time-ordering. In the derivation we have used (1) the triangle inequality; (2) the data-processing inequality for the trace distance; (3) the assumption that ρ_k has a nonzero gap for all k. This is precisely the assumptions (b) in the first paragraph of this section. In order to recover the error to $O(1/\mathrm{Poly}(L))$, we need $L_0 \simeq \ln(L)/\Delta$. As each gate of the recovery channel has a width of order L_0, the total depth of the recovery map is thus $O(\ln L)$.

4. An illuminating counterexample

We now discuss an illustrative example, in which the global recovery scheme fails due to violation of the condition on CMI Eq. (85) when the channel is acting on parts of the system.

Consider a GHZ state in a quibit chain

$$|\psi_0\rangle = |\uparrow\uparrow\cdots\rangle + |\downarrow\downarrow\cdots\rangle. \tag{D13}$$

Starting from $\rho_0 = |\psi_0\rangle\langle\psi_0|$, we can perform a strong X_i measurement on each site, represented by the (depth-1) channel $\mathcal{E} = \prod_i \mathcal{E}_i$, where

$$\mathcal{E}_i[\rho] = \frac{1}{2}\rho + \frac{1}{2}X_i\rho X_i. \tag{D14}$$

It is easy to verify that the final state is $\rho_f = \mathcal{E}[\rho_0] \propto I + \prod_i X_i$.

The effect of $\mathcal{E}$ on ρ_0 should not be recoverable through a low-depth channel, since the connected correlation function $\mathrm{tr}(\rho Z_x Z_y) - \mathrm{tr}(\rho Z_x)\,\mathrm{tr}(\rho Z_y)$ at large $|x-y|$ is nonzero for ρ_0 but zero for ρ_f. However, naively the CMI calculated in the geometry of Fig. 2 is $\log 2$ for both ρ_0 and ρ_f, which satisfies the requirement from Eq. (85). So what went wrong? It turns out that the CMI bound condition Eq. (85) is violated when only part of the system is being acted by the channel.

Specifically, let us consider the "intermediate" state (below we have two special sites located at two far-separated points $i = 1, x$, the region A is defined as $1 < A < x$, and the region B is defined as $x < B \le L$)

$$\begin{aligned}
\rho' &= \prod_{i\neq 1,x} \mathcal{E}_i[\rho_0] \\
&= \frac{1}{2^{L-1}}\sum_{X_A X_B}[|\uparrow X_A \uparrow X_B\rangle + X_A X_B|\downarrow X_A \downarrow X_B\rangle],
\end{aligned} \tag{D15}$$

where $[\psi]$ stands for $|\psi\rangle\langle\psi|$. Acting on ρ' with $\mathcal{E}_1\cdot\mathcal{E}_x$ will produce the final state $\rho_f \propto I + \prod_i X_i$. But this step from ρ' to ρ_f is not locally recoverable since it eliminates a Bell pair between $i = 1$ and $i = x$. In terms of CMI, it is easy to check that for ρ':

$$I(i = 1; x \cup B|A) = 2\log 2, \tag{D16}$$

which violates the CMI bound Eq. (85).

The lesson is that given the channel and the initial and final states, it is still a nontrivial task to check whether the CMI condition Eq. (85) is satisfied "locally."

[1] L. D. Landau and E. M. Lifshitz, *Statistical Physics, Part 1*, Course of Theoretical Physics, Vol. 5 (Butterworth-Heinemann, Oxford, 1980).
[2] John McGreevy, Generalized symmetries in condensed matter, Annu. Rev. Condens. Matter Phys. **14**, 57 (2023).
[3] Jong Yeon Lee, Chao-Ming Jian, and Cenke Xu, Quantum criticality under decoherence or weak measurement, PRX Quantum **4**, 030317 (2023).
[4] Ruochen Ma, Jian-Hao Zhang, Zhen Bi, Meng Cheng, and Chong Wang, Topological phases with average symmetries: The decohered, the disordered, and the intrinsic, arXiv:2305.16399 [cond-mat.str-el].
[5] Berislav Buča and Tomaž Prosen, A note on symmetry reductions of the Lindblad equation: Transport in constrained open spin chains, New J. Phys. **14**, 073007 (2012).
[6] Victor V. Albert and Liang Jiang, Symmetries and conserved quantities in Lindblad master equations, Phys. Rev. A **89**, 022118 (2014).
[7] Victor V. Albert, Lindbladians with multiple steady states: Theory and applications, arXiv:1802.00010 [quant-ph].
[8] Simon Lieu, Ron Belyansky, Jeremy T. Young, Rex Lundgren, Victor V. Albert, and Alexey V. Gorshkov, Symmetry breaking and error correction in open quantum systems, Phys. Rev. Lett. **125**, 240405 (2020).
[9] S. F. Edwards and P. W. Anderson, Theory of spin glasses, J. Phys. F: Metal Phys. **5**, 965 (1975).
[10] Yimu Bao, Ruihua Fan, Ashvin Vishwanath, and Ehud Altman, Mixed-state topological order and the errorfield double formulation of decoherence-induced transitions, arXiv:2301.05687 [quant-ph].
[11] Ruihua Fan, Yimu Bao, Ehud Altman, and Ashvin Vishwanath, Diagnostics of mixed-state topological order and breakdown of quantum memory, arXiv:2301.05689 [quant-ph].
[12] Yijian Zou, Shengqi Sang, and Timothy H. Hsieh, Channeling quantum criticality, Phys. Rev. Lett. **130**, 250403 (2023).
[13] Yu-Hsueh Chen and Tarun Grover, Symmetry-enforced many-body separability transitions, arXiv:2310.07286 [quant-ph].
[14] Yu-Hsueh Chen and Tarun Grover, Separability transitions in topological states induced by local decoherence, arXiv:2309.11879 [quant-ph].
[15] Shengqi Sang, Yijian Zou, and Timothy H. Hsieh, Mixed-state quantum phases: Renormalization and quantum error correction, arXiv:2310.08639 [quant-ph].
[16] Shengqi Sang and Timothy H. Hsieh, Stability of mixed-state quantum phases via finite Markov length, arXiv:2404.07251 [quant-ph].
[17] Tibor Rakovszky, Sarang Gopalakrishnan, and Curt von Keyserlingk, Defining stable phases of open quantum systems, arXiv:2308.15495 [quant-ph].
[18] Tsung-Cheng Lu, Zhehao Zhang, Sagar Vijay, and Timothy H. Hsieh, Mixed-state long-range order and criticality from measurement and feedback, PRX Quantum **4**, 030318 (2023).
[19] Jong Yeon Lee, Wenjie Ji, Zhen Bi, and Matthew P. A. Fisher, Decoding measurement-prepared quantum phases and transitions: From Ising model to gauge theory, and beyond, arXiv:2208.11699 [cond-mat.str-el].
[20] Guo-Yi Zhu, Nathanan Tantivasadakarn, Ashvin Vishwanath, Simon Trebst, and Ruben Verresen, Nishimori's cat: Stable long-range entanglement from finite-depth unitaries and weak measurements, Phys. Rev. Lett. **131**, 200201 (2023).
[21] Leonardo A. Lessa, Meng Cheng, and Chong Wang, Mixed-state quantum anomaly and multipartite entanglement, arXiv:2401.17357 [cond-mat.str-el].
[22] Yu-Hsueh Chen and Tarun Grover, Unconventional topological mixed-state transition and critical phase induced by self-dual coherent errors, arXiv:2403.06553 [quant-ph].
[23] Zijian Wang, Zhengzhi Wu, and Zhong Wang, Intrinsic mixed-state topological order without quantum memory, arXiv:2307.13758 [quant-ph].
[24] Zijian Wang and Linhao Li, Anomaly in open quantum systems and its implications on mixed-state quantum phases, arXiv:2403.14533 [quant-ph].
[25] Ramanjit Sohal and Abhinav Prem, A noisy approach to intrinsically mixed-state topological order, arXiv:2403.13879 [cond-mat.str-el].
[26] Kohei Kawabata, Ramanjit Sohal, and Shinsei Ryu, Lieb-Schultz-Mattis theorem in open quantum systems, Phys. Rev. Lett. **132**, 070402 (2024).
[27] Yi-Neng Zhou, Xingyu Li, Hui Zhai, Chengshu Li, and Yingfei Gu, Reviving the Lieb-Schultz-Mattis theorem in open quantum systems, arXiv:2310.01475 [cond-mat.str-el].
[28] Po-Shen Hsin, Zhu-Xi Luo, and Hao-Yu Sun, Anomalies of average symmetries: Entanglement and open quantum systems, arXiv:2312.09074 [cond-mat.str-el].
[29] Jinkang Guo, Oliver Hart, Chi-Fang Chen, Aaron J. Friedman, and Andrew Lucas, Designing open quantum systems with known steady states: Davies generators and beyond, arXiv:2404.14538 [quant-ph].
[30] Tsung-Cheng Lu, Disentangling transitions in topological order induced by boundary decoherence, arXiv:2404.06514 [quant-ph].
[31] Sanjay Moudgalya and Olexei I. Motrunich, Symmetries as ground states of local superoperators: Hydrodynamic implications, PRX Quantum **5**, 040330 (2024).
[32] Caroline de Groot, Alex Turzillo, and Norbert Schuch, Symmetry protected topological order in open quantum systems, Quantum **6**, 856 (2022).
[33] Ruochen Ma and Chong Wang, Average symmetry-protected topological phases, arXiv:2209.02723 [cond–mat.str-el].
[34] Jian-Hao Zhang, Yang Qi, and Zhen Bi, Strange correlation function for average symmetry-protected topological phases, arXiv:2210.17485 [cond-mat.str-el].
[35] Jong Yeon Lee, Yi-Zhuang You, and Cenke Xu, Symmetry protected topological phases under decoherence, arXiv:2210.16323 [cond-mat.str-el].
[36] Jian-Hao Zhang, Ke Ding, Shuo Yang, and Zhen Bi, Fractonic higher-order topological phases in open quantum systems, Phys. Rev. B **108**, 155123 (2023).

[37] Ruochen Ma and Alex Turzillo, Symmetry protected topological phases of mixed states in the doubled space, arXiv:2403.13280 [quant-ph].

[38] Yuchen Guo, Jian-Hao Zhang, Shuo Yang, and Zhen Bi, Locally purified density operators for symmetry-protected topological phases in mixed states, arXiv:2403.16978 [cond-mat.str-el].

[39] Hanyu Xue, Jong Yeon Lee, and Yimu Bao, Tensor network formulation of symmetry protected topological phases in mixed states, arXiv:2403.17069 [cond-mat.str-el].

[40] Sanket Chirame, Fiona J. Burnell, Sarang Gopalakrishnan, and Abhinav Prem, Stable symmetry-protected topological phases in systems with heralded noise, arXiv:2404.16962 [quant-ph].

[41] In the literature, the fidelity is often defined as the square of Eq. (3). For our purpose it is more natural to adopt Eq. (3).

[42] A. Uhlmann, The "transition probability" in the state space of a*-algebra, Rep. Math. Phys. **9**, 273 (1976).

[43] Michael A. Nielsen and Isaac L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, 2000).

[44] Martin Müller-Lennert, Frédéric Dupuis, Oleg Szehr, Serge Fehr, and Marco Tomamichel, On quantum Rényi entropies: A new generalization and some properties, J. Math. Phys. **54**, 122203 (2013).

[45] Hsin-Yuan Huang, Richard Kueng, and John Preskill, Predicting many properties of a quantum system from very few measurements, Nat. Phys. **16**, 1050 (2020).

[46] Aniket Rath, Cyril Branciard, Anna Minguzzi, and Benoît Vermersch, Quantum Fisher information from randomized measurements, Phys. Rev. Lett. **127**, 260501 (2021).

[47] Andreas Elben, Steven T. Flammia, Hsin-Yuan Huang, Richard Kueng, John Preskill, Benoît Vermersch, and Peter Zoller, The randomized measurement toolbox, Nat. Rev. Phys. **5**, 9 (2022).

[48] Hong-Ye Hu and Yi-Zhuang You, Hamiltonian-driven shadow tomography of quantum states, Phys. Rev. Res. **4**, 013054 (2022).

[49] Hong-Ye Hu, Soonwon Choi, and Yi-Zhuang You, Classical shadow tomography with locally scrambled quantum dynamics, Phys. Rev. Res. **5**, 023027 (2023).

[50] Eric Dennis, Alexei Kitaev, Andrew Landahl, and John Preskill, Topological quantum memory, J. Math. Phys. **43**, 4452 (2002).

[51] In the ordered and disordered phases, $\gamma = \ln 2$ and 0 when A encompasses the entire system, respectively, and it vanishes when A constitutes only a portion of the system.

[52] Jian-Hao Zhang, Cenke Xu, and Yichen Xu, Fluctuation-dissipation theorem and information geometry in open quantum systems, arXiv:2409.18944 [quant-ph].

[53] Davide Gaiotto, Anton Kapustin, Nathan Seiberg, and Brian Willett, Generalized global symmetries, J. High Energy Phys. **2015**, 172 (2015).

[54] Xiao-Gang Wen, Emergent anomalous higher symmetries from topological order and from dynamical electromagnetic field in condensed matter systems, Phys. Rev. B **99**, 205139 (2019).

[55] Carolyn Zhang, Yichen Xu, Jian-Hao Zhang, Cenke Xu, Zhen Bi, and Zhu-Xi Luo, Strong-to-weak spontaneous breaking of 1-form symmetry and intrinsically mixed topological order, arXiv:2409.17530 [quant-ph].

[56] Olumakinde Ogunnaike, Johannes Feldmeier, and Jong Yeon Lee, Unifying emergent hydrodynamics and Lindbladian low-energy spectra across symmetries, Constraints, and Long-Range Interactions, Phys. Rev. Lett. **131**, 220403 (2023).

[57] Xiaoyang Huang, Marvin Qi, Jian-Hao Zhang, and Andrew Lucas, Hydrodynamics as the effective field theory of strong-to-weak spontaneous symmetry breaking, arXiv:2407.08760 [cond-mat.str-el].

[58] Edward H. Chen, Guo-Yi Zhu, Ruben Verresen, Alireza Seif, Elisa Bäumer, David Layden, Nathanan Tantivasadakarn, Guanyu Zhu, Sarah Sheldon, Ashvin Vishwanath, Simon Trebst, and Abhinav Kandala, Realizing the Nishimori transition across the error threshold for constant-depth quantum circuits, arXiv:2309.02863 [quant-ph].

[59] Pablo Sala, Sarang Gopalakrishnan, Masaki Oshikawa, and Yizhi You, Spontaneous strong symmetry breaking in open systems: Purification perspective, Phys. Rev. B **110**, 155150 (2024).

[60] Chenyang Wang, Jim Harrington, and John Preskill, Confinement-Higgs transition in a disordered gauge theory and the accuracy threshold for quantum memory, Ann. Phys. **303**, 31 (2003).

[61] Wolfhard Janke and Klaus Nather, High-precision Monte Carlo study of the two-dimensional *XY* villain model, Phys. Rev. B **48**, 7419 (1993).

[62] Marius Junge, Renato Renner, David Sutter, Mark M Wilde, and Andreas Winter, in *Annales Henri Poincaré*, Vol. 19 (Springer, 2018), pp. 2955–2978.