

Learning transitions of topological surface codes

Finn Eckstein,^{1,2} Bo Han,^{1,3} Simon Trebst,^{1,*} and Guo-Yi Zhu^{2,†}

¹*Institute for Theoretical Physics, University of Cologne, Zùlpicher StraÙe 77, 50937 Cologne, Germany*

²*The Hong Kong University of Science and Technology (Guangzhou), Nansha, Guangzhou, 511400, Guangdong, China*

³*Department of Condensed Matter Physics, Weizmann Institute of Science, Rehovot 7610001, Israel*

(Dated: December 24, 2025)

For the surface code, topological quantum order allows one to encode logical quantum information in a robust, long-range entangled many-body quantum state. However, if an observer probes this quantum state by performing measurements on the underlying qubits, thereby collecting an ensemble of highly correlated classical snapshots, two closely related questions arise: (i) do measurements decohere the topological order of the quantum state; and (ii) how much of the logical information can one learn from the snapshots? Here we address these questions for measurements in a uniform basis on all qubits. We find that for generic measurement angles, sufficiently far away from the Clifford X , Y , and Z directions (such as the $X + Y + Z$ basis) the logical information is *never lost* in one of the following two ways: (i) for weak measurement, the topological order is absolutely robust; (ii) for projective measurement, the quantum state inevitably collapses, but the logical quantum information is faithfully transferred from the quantum system to the observer in the form of a tomographically complete classical shadow. At these generic measurement angles and in the projective-measurement limit, the measurement ensemble enforced by Born probabilities can be represented by a 2D tensor network that can be fermionized into a disordered, free-fermion network model in symmetry class DIII, which gives rise to a Majorana “metal” phase. When the measurement angle is biased towards the X or Z limits, a critical angle indicates the threshold of a learning transition beyond which the classical shadow no longer reveals full tomographic information (but reduces to a measurement of the logical X or Z state). This learning transition can be described in the language of the network model as a “metal to insulator” transition. For the anisotropic “insulating” angles, there is also a finite learnability threshold in terms of the measurement strength, above which the surface code breaks down. In terms of universality, this latter transition is generically a Nishimori transition. Our results can be generalized from *qubit* measurement to *syndrome* measurement, demonstrating that – in the presence of coherent errors – even syndrome measurements can enable the observer to learn the logical information and collapse the topological order across a learnability threshold.

I. INTRODUCTION

Topological quantum order originates from long-range entanglement [1], which protects logical quantum information against decoherence [2]. The realization of a topological quantum memory in the toric code [3] or its planar variant, the surface code [4], lies at the center of the field of quantum error correction, while also serving as paradigmatic models in quantum statistical physics where their quantum topology is understood as a quantum superposition of a massive amount of loop configurations. These quantum memories are known to be stable against Hamiltonian perturbations [5–7], wave function deformations [8–13], and noise [2, 14–19]. The topological protection mechanism also prevents an observer from learning the encoded information locally [3]. Such learning becomes possible only if the observer is sufficiently entangled with the physical qubits of the system [20], the measurement strength becomes sufficiently strong [21], or the read-out fidelity surpasses a finite threshold [2]. The physics underlying such *learning transitions* has attracted some recent interest [22–32], as its description leads to disordered statistical models which distinguish themselves by their *correlated disorder* – mandated by Born measurement probabilities or Bayesian inference likelihood – and the critical theories

emerging at the phase transitions.

In this manuscript, we study the learning transitions of topological surface codes induced by measuring all physical qubits in a given basis with a tunable measurement strength, i.e. interpolating between weak and projective measurements. After repeated experiments of such an all-qubit measurement, the observer is left with an ensemble of classical snapshots, while the quantum system assumes an ensemble of post-measurement states which, in combination, one can view as a hybrid mixed state [35]. Such a remnant mixed state inherits certain nontrivial correlations from the original topologically ordered, long-range entangled state. For example, when the measurement basis is perfectly aligned along the Z or X directions, and if the measurement is projective, it is known that the quantum memory turns into a classical memory, the quantum loop condensate is dephased into a classical loop gas ensemble [20]. In this case, the learning transition itself, which is driven by tuning the measurement strength, can be mapped to a Nishimori transition [20, 36, 37] and, from a symmetry perspective, can be understood as the strong-to-weak spontaneous symmetry breaking of the underlying 1-form symmetry [15–19]. When, however, the measurement basis is rotated into a generic direction much less is known about the learnability threshold of the surface code. A first indication that there might be an interesting uncharted territory to be explored arose in previous work by the authors [20] where it was shown that the learnability threshold is pushed from finite to *infinite* when the measurement basis is rotated to $X + Z$, half way between the Clifford X and Z angles. At this point

* trebst@thp.uni-koeln.de

† guoyizhu@hkust-gz.edu.cn

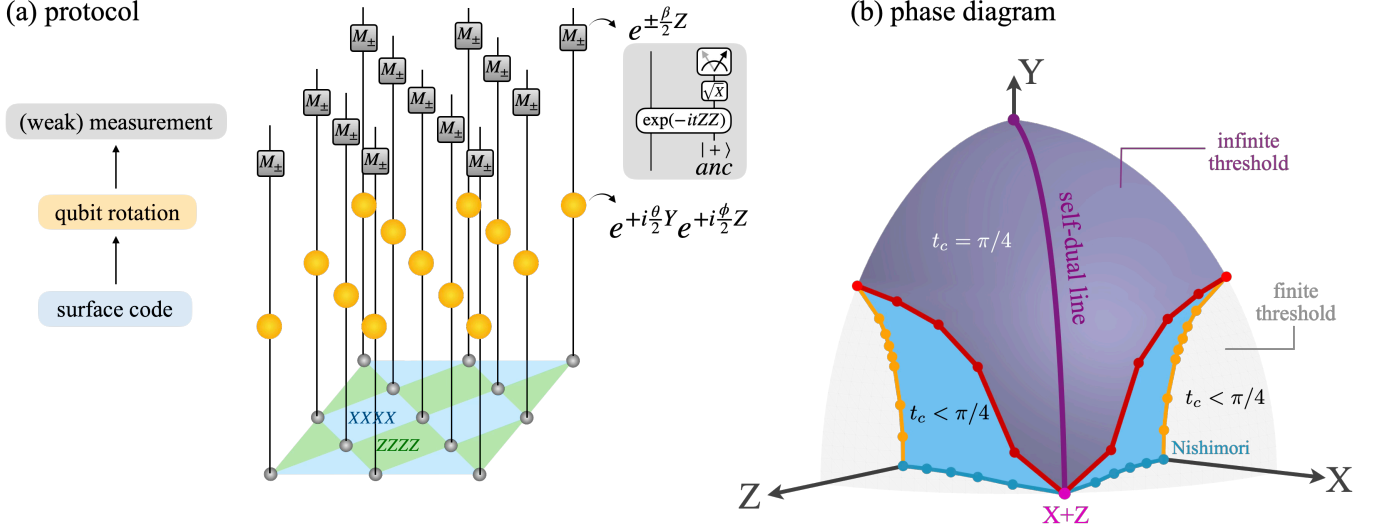


FIG. 1. **(a) Protocol:** A surface code is (weakly) measured in the basis $\hat{\sigma}^{\theta, \phi} \equiv \sin \theta \cos \phi X + \sin \theta \sin \phi Y + \cos \theta Z$ for all qubits. The protocol consists of 3 steps: firstly, prepare a surface code in certain state in the logical space; secondly, apply a layer of coherent rotation to all qubits by $e^{i\theta/2Y} e^{i\phi/2Z}$; thirdly, apply a layer of weak measurement in the computation basis i.e. Pauli Z basis. The weak measurement is implemented by introducing an ancilla prepared in $|+\rangle$ state ($X|+\rangle = |+\rangle$), and entangling it with the physical qubit by Ising interaction evolution up to time $t \in [0, \pi/4]$, and then measuring it out in the Pauli Y basis. As shown in Ref. [21], this ancilla-assisted operation effectively implements a weak measurement gate $e^{\pm\beta Z/2}$ with tunable measurement strength $\beta = \tanh^{-1} \sin(2t) \in [0, \infty)$. In the following, we will also refer to t as the effective measurement strength. In particular, $t = \pi/4$ describes the projective-measurement limit with infinite strength $\beta = \infty$. **(b) Three-dimensional Born average phase diagram of parameter (t, θ, ϕ) :** the radial coordinate corresponds to the compactified measurement strength t , while the direction of the coordinate corresponds to the measurement basis of the physical qubit, akin to the Bloch sphere of a spin-1/2, that can be parametrized by the spherical coordinates (θ, ϕ) . The blue surface denotes the critical boundary of surface code, which sets the learnability threshold. Below the threshold $t < t_c$, the surface code remains topological, and the measurement outcomes do not extract useful information about the logical state in thermodynamic limit. In the conventional measurement basis such as X or Z direction, a finite threshold $t_c = 0.143\pi$ [20, 21, 33, 34] is established by the Nishimori transition (blue dot). The same Nishimori universality describes the whole transition line (blue dots) until the self-dual limit at $X + Z$, which is captured instead by the weak self-dual universality class [35]. Around the self-dual directions (purple line), $t_c = \pi/4$ ($\beta_c = \infty$), the maximal infinite threshold is found to hold for a big phase region (denoted by the purple surface) enveloped by the red dotted line as the boundary.

the self-duality of the surface code (which, in the associated 2D statistical model, can be represented as a non-invertible Kramers-Wannier (KW) duality) was understood to originate this seemingly extraordinary level of robustness [35]. This naturally leads one to ask what happens as one further moves away from these Clifford angles and measures the physical qubits in a basis corresponding to an arbitrary spherical angle on the Bloch sphere. What is the nature of the ensuing mixed state that collects the whole ensemble of the classical snapshots? What can one learn about the logical qubit from such a mixed state? And when the measurement is weak or if measurement readout errors exist, is there (still) a learnability transition? These are precisely the questions that we will address in the following.

Before we provide a short summary of our main results let us quickly mention that one can view the rotation of the measurement angle, or the weakening of the entanglement between the system and the observer, as “coherent errors” [38–44]. Here we prefer a more neutral notation and refer to “coherent rotations” instead as, from the learning perspective that we adopt, such coherent rotations enable varying the learnability threshold by changing the measurement basis.

II. SUMMARY OF OUR RESULTS

Model

Our model consists of a 3-step protocol as illustrated in Fig. 1(a): (i) preparing a topological surface code state $|\psi_\mu\rangle$ where μ labels an arbitrary state of the single logical qubit in the Bloch sphere (e.g. $\mu = \pm$ as eigenstate of logical Pauli X , or $\mu = 0(1)$ as eigenstate of logical Pauli Z) encoded in the surface code; (ii) applying a layer of coherent rotation

$$U(\theta, \phi) = e^{i\frac{\theta}{2} \sum_j Y_j} e^{i\frac{\phi}{2} \sum_j Z_j} \quad (1)$$

uniformly to all qubits; (iii) applying weak measurement of strength $\beta \in [0, \infty]$ in the computation basis, described by the Kraus operator

$$M_s = e^{\frac{\beta}{2} \sum_j s_j Z_j} / \sqrt{e^\beta + e^{-\beta}}, \quad (2)$$

where s labels the measurement outcome in the form of a bit string, e.g. $s = + - + + \dots$. As shown inside the gray box in Fig. 1(a), the weak measurement can be implemented by projectively measuring an ancilla qubit which is first coupled

to the physical qubit by an entangling gate with parameter $t \in [0, \pi/4]$ such that $\tanh \beta = \sin(2t)$ [21, 45]. One can also simply view $t \in [0, \pi/4]$ as a compactified parameter for the measurement strength where $t = \pi/4$ corresponds to the projective-measurement limit $\beta = \infty$.

Putting the three steps together, we arrive at an ensemble of states

$$\{P(\mathbf{s}), |\psi_\mu(\mathbf{s})\rangle\}, \quad (3)$$

where for each classical measurement record \mathbf{s} , the conditional quantum state is given by the following normalized wave function

$$|\psi_\mu(\mathbf{s})\rangle = \frac{M_{\mathbf{s}} U(\theta, \phi) |\psi_\mu\rangle}{\|M_{\mathbf{s}} U(\theta, \phi) |\psi_\mu\rangle\|}, \quad (4)$$

whose probability follows from the Born's rule

$$P(\mathbf{s}) = \langle \psi_\mu(\mathbf{s}) | \psi_\mu(\mathbf{s}) \rangle. \quad (5)$$

The phase diagram we discuss below is for the Born ensemble of states, or equivalently, the mixed state ρ that packs the classical records and the quantum states together into a quantum-classical hybrid mixture [23, 35]

$$\rho = \sum_{\mathbf{s}} |\mathbf{s}\rangle \langle \mathbf{s}| \otimes |\psi_\mu(\mathbf{s})\rangle \langle \psi_\mu(\mathbf{s})|. \quad (6)$$

The task of learning is to infer the logical information μ based on the classical measurement records \mathbf{s} , that are sampled from the probability distribution $P(\mathbf{s})$. In a simple scenario, if the surface code is projectively measured without any coherent rotation in the Z basis, one obtains the classical snapshots in the Z basis, which are known to form a classical loop gas across the lattice, defined by the spin downs $Z = -1$ [46, 47]. Such a measurement also allows one to directly readout the eigenvalue of the logical Z_L operator of the surface code. Similarly, if the coherent rotation rotates Z to the X basis ($\theta = \pi/2, \phi = 0$), the physical qubits are collapsed into the X basis, where the spins with $X = -1$ forms a loop gas across the lattice, which is dual to the Z -loop gas [35]. Now the question is, what happens in between? Specifically, when the qubits are collapsed at the self-dual angle ($\theta = \pi/4, \phi = 0$), i.e. in the $X + Z$ basis, the ensemble of classical snapshots has been shown to be *critical* in Ref. [35], but what does that imply for the associated distribution of logical qubit snapshot? More generally, what happens to these distributions if the measurement angle is rotated to more general directions in the Bloch sphere and if the measurement is weak?

Phase diagram

To address the above questions, we have charted out the spherical phase diagram shown in Fig. 1(b). Akin to a Bloch sphere our phase diagram is characterized by a pair of spherical measurement angles θ, ϕ and a measurement strength t that indicates the radial direction. Its color coding indicates different regimes of finite and infinite *learnability thresholds*

which we have determined by calculating the coherent information [14, 20, 49, 50] as an information-theoretic order parameter of the topological order of the surface code in the presence of measurements. Our underlying numerical computations have charted out the various phase boundaries, most prominently the purple surface around the Y -direction where, akin to the self-dual $X + Z$ point discussed above, the learning transition happens only in the projective-measurement limit. From a mixed-state perspective, surface code remains, for these measurement angles, robust against arbitrarily strong, but non-projective measurements. This should be contrasted with the finite thresholds observed around the corners surrounding the X or Z measurement directions, where already for some finite measurement strength the surface code $\{|\psi(\mathbf{s})\rangle\}$ is destroyed and useful information in $\{P(\mathbf{s})\}$ can be extracted by the observer. Cast in terms of mixed-state terminology, the surface code is dephased for finite measurement strength, with the quantum loop gas loosing its coherence and transitioning into a classical loop gas.

Ensemble of logical states

In the projective-measurement limit, all physical qubits are collapsed into classical states along the measurement direction. In tandem, the logical qubit is also collapsed resulting in a snapshot along a direction, which is determined by the isometry that maps the obtained physical many-qubit snapshot to a logical single-qubit snapshot. Such a logical qubit snapshot can also be materialized by employing a reference qubit R , which in the initial step of preparing the surface code is maximally entangled with the logical qubit, see Fig. 2(a). The collapse of the physical qubits of the surface code then gives rise to a projected ensemble [51] of the reference qubit R as illustrated for various measurement angles in Fig. 2(b). For a conventional measurement basis such as Z or X , the logical snapshots align along the Z or X axis, forming a bimodal distribution. Rotating the measurement angle towards the self-dual angle $X + Z$, we find that the projected ensemble forms a uniform ring covering the “equator” of the Bloch sphere, with the logical state remaining confined to the real plane. As we move out of this plane and towards to the Y -direction, we find that the projected ensemble forms a uniform coverage of the Bloch sphere as representatively illustrated for the $X + Y + Z$ magic angle. This is an important observation as, in terms of the learning problem considered here, it indicates that – even though the measurement basis of the physical qubits is uniform toward a fixed angle – the snapshot of the logical qubit enables *full tomography* for the logical qubit – without the need for randomized measurements [52] or measuring for multiple angles. Such full tomography is possible everywhere on the purple surface, but the Y -direction which is again a Clifford limit where the projected ensemble is distorted back to a bimodal distribution. Notice that such a transition of the projected ensemble from a Haar-measure uniform distribution to a bimodal distribution has recently also been realized in a long-range interacting circuit [53].

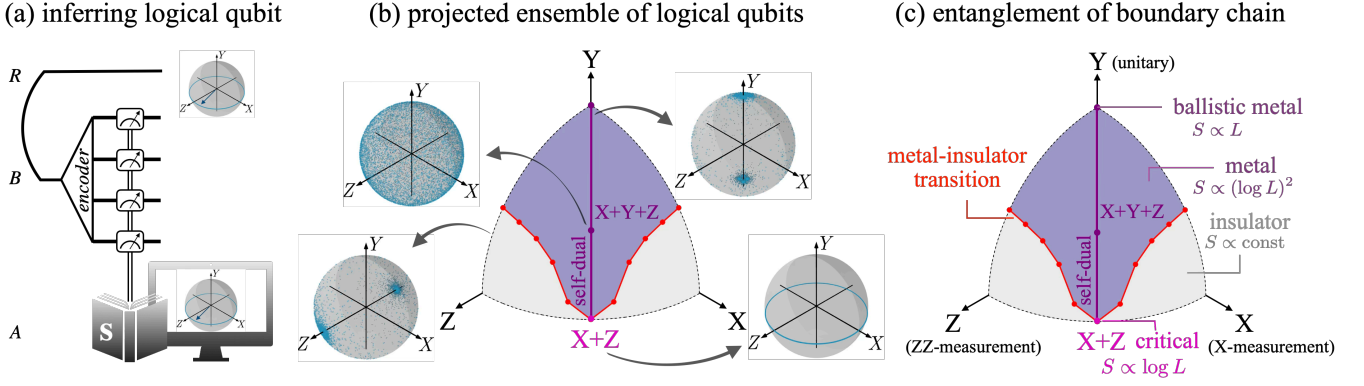


FIG. 2. **projective-measurement limit.** (a) A deterministic isometry matrix encodes the logical qubit to the physical qubits of the surface code. The logical qubit is initially maximally entangled with the reference (R) qubit. The physical qubits are subjected to projective measurement, whose outcome is recorded by an observer (A). For each observed measurement record s , the observer (A) can decode (by inverting the unitary encoder that prepares the surface code state) the corresponding state of reference (R) qubit, which is a snapshot of the logical qubit. These measurement records s form the raw data from which we construct classical shadows of the logical qubit state for tomography [48]. (b) Taking together all the possible measurement records, $\{P(s), |\psi_R(s)\rangle\}$ forms a projected ensemble of R . In the gapped phases (white), the ensemble converges to a bimodal distribution, revealing the logical Z or logical X state. At the self-dual critical point $X + Z$, the ensemble delocalizes to a uniform distribution, but confined to the real XZ plane, forming a ring on the equator of the Bloch sphere. In the metal gapless phase (purple), the ensemble fully delocalizes to a Haar measure uniform coverage of the Bloch sphere. Nevertheless, it crossovers to the bimodal distribution at the Clifford Y limit, which is a singular point. The shown ensembles of states are computed for surface code of code distance $d = 8$. (c) The boundary entanglement entropy has distinct scaling at the phases. In the gapped phases, it exhibits area law. At self-dual critical point, critical logarithmic scaling is found. Into the metal gapless phase, $\log^2 L$ dominates the entanglement entropy scaling. The Y limit corresponds to a ballistic metal, exhibiting volume law scaling. The Y -axis is described by purely unitary evolution. By contrast, the X and Z axes are described by pure measurement evolutions. The circuit and the boundary state is shown in Fig. 3(c).

Boundary entanglement entropy

The measurement-induced mixed state can be characterized by its 2D quantum topology (such as its topological degeneracy or coherent information) or, alternatively, by its (1+1)D boundary entanglement entropy. In more technical terms, the post-measurement state of the surface code can be represented, for every measurement outcome s , by a random 2D PEPS (projected entangled pair state) which can be mapped to a (1+1)D monitored quantum circuit dynamics [35]. The final state of this circuit dynamics captures the complexity of the boundary state. We are particularly interested in characterizing the latter in the projective-measurement limit (at $t = \pi/4$) for arbitrary measurement angles as summarized in the phase diagram of Fig. 2(c). The boundary entanglement entropy exhibits an area law in the surface code phase and the gapped dephasing phases around the X and Z corners, but displays logarithmic scaling $S_{vN} \propto \log L$ at the critical points (pink and red dots in the phase diagram). For generic measurement angles, i.e. for angles corresponding to the purple regime in Fig. 2(c) the boundary entanglement entropy scales $S_{vN} \propto \log^2 L$, which is reminiscent of the entanglement structure of the Goldstone modes in a “metallic” phase [54–58]. The limiting case of measurements along the Y -direction is singular and exhibits volume-law scaling $S_{vN} \propto L$ because the corresponding (1+1)D circuit dynamics consists exclusively of maximally entangling Majorana swap gates, which is a spacetime-dual [59] unitary circuit with ballistic growth of entanglement [60].

Metal-insulator transition

In connecting the monitored (1+1)D fermion circuit in the projective-measurement limit to a free-fermion model, one maps the non-unitary circuit dynamics in (1+1)D dimensions to a unitary, non-interacting fermion model in (2+1)D spatial dimensions with a static Hermitian Hamiltonian [54, 58], coupled to static disorder that follows the Born’s distribution. Microscopically, this is done by relating each non-unitary circuit gate, as a transfer matrix in the time dimension, to a unitary scatterer in two spatial dimensions, as in the standard Chalker-Coddington network model construction [54, 61]. As a consequence, one can identify the entanglement structures stabilized by the (1+1)D monitored fermion circuit, with the ground states (and their respective entanglement structures) of a (2+1)D non-interacting free-fermion model, i.e. a metallic or insulating state. Specifically, this connects the phase with long-range (logarithmic scaling) boundary entanglement (i.e. the purple phase in our phase diagram) to a *metallic* state, which (in its free-fermion formulation) is known to exhibit such logarithmically scaling entanglement [62, 63]. In contrast, the phases near X and Z anisotropic limit exhibit area-law entangled (1+1)D fermion chain, which corresponds to a short-range correlated (2+1)D *insulating* free-fermion system.

In this language, some of the learning transitions discussed in this paper can be recast as a *metal-insulator* transition. For instance, when varying the measurement angle from a generic direction (e.g. around the $X + Y + Z$ point) to one of the Clifford directions near the X and Z corners, one tran-

sitions precisely from a metallic (long-range) entanglement structure to an insulating (short-range) entanglement structure, see the phase diagram in Fig. 2(c). As one tracks this metal-insulator transition (red dots in our phase diagram) as one approaches the self-dual $X + Z$ point (pink dot), the system acquires weak self-dual symmetry resulting in a modified critical point, extensively characterized in Ref. [35]. In comparison, the generic metal-insulator transition in our phase diagram has only been partially studied along the XY line – where the qubit-measurement model considered here can be mapped to a syndrome-measurement model [39] as discussed in Sec. IV B below, and critical exponents and conformal data remain an open question. In this paper we discuss the boundary-entanglement scaling of this generic metal-insulator transition, but leave it to future work for an in-depth investigation of the criticality theory of this point.

Symmetry class

For our model we find that for generic measurement angles the relevant symmetry class of the free-fermion model corresponding to our circuit is symmetry class DIII. Specifically, the transfer matrix of the tensor network can be mapped to that of a Chalker-Coddington network model, which respects time-reversal ($T^2 = -1$), particle-hole ($C^2 = +1$), and chiral symmetries – precisely the symmetries of symmetry class DIII [58]. However, along three special lines in the phase diagram, i.e. the ZX [35], XY , and YZ lines of the phase diagram in Fig. 2(c), the tensor network remains in symmetry class D, where the latter two are related to what has been observed for *syndrome* measurements in Refs. [39, 40]. On these special lines, the tensor network can be decomposed into two decoupled Majorana fermion network models that are time-reversal partners of each other; each copy is in symmetry class D. Such a decomposition, however, does not exist away from the three special lines. We will analyze the XY line further in Sec. IV B and Fig. 11 and, in particular, discuss its equivalence to the error model previously discussed in Refs. [39, 40].

III. LEARNING TRANSITION FOR TUNABLE MEASUREMENT STRENGTH

When the measurement strength is tuned above a certain threshold, the observer obtains certain logical information about the surface code, and meanwhile collapses the surface code into a post-measurement projected ensemble of states – this threshold defines our learnability transition [23]. To diagnose this learning transition in our model, we will rely on two principal observables: (i) the *coherent information* of the 2D surface code; (ii) the *entanglement entropy* of its 1D boundary quantum state. The former literally discriminates two phases: learnable and unlearnable. The latter characterizes the fluctuations and complexity of the remaining quantum system, which can also signal the quantum critical point. But before we introduce these observables in the following, we will briefly review two representations of our surface code model, which

we previously introduced in Ref. [20], that will allow for the efficient calculation of these observables: (i) a 2D (random) tensor network representation; (ii) a (1+1)D quantum circuit which can also be recast in the form of a fermion model.

A. Tensor network and quantum circuit representation

Tensor network representation & Ashkin-Teller model

The 2D surface code can be written as a PEPS [64–66]. In this representation, the weak measurement Kraus operators act as random local non-unitary operators on the physical legs, conditioned upon the classical measurement outcomes. For a visual illustration of the resulting PEPS see Fig. 3. Generic physical observables can be obtained by taking two copies of the PEPS, as the ket and bra, respectively, and contracting out the physical legs with observable operators inserted. The bilayer tensor network, $\langle \psi(s) | \psi(s) \rangle$, then defines the partition function of an associated 2D classical model. This partition function can be readily obtained from the post-measurement tensor network wavefunction $|\psi\rangle$ in Fig. 3(a) as the Born probability

$$P(s) = \langle \psi | M_s^\dagger M_s | \psi \rangle = \sum_{\{Z, \tilde{Z}\}} \exp(-E_{AT}),$$

where [20]

$$E_{AT,ij} = -\frac{Js_{ij}}{2} (Z_i Z_j + \tilde{Z}_i \tilde{Z}_j) + \frac{i\phi}{2} (Z_i Z_j - \tilde{Z}_i \tilde{Z}_j) - \frac{1}{2} \left(2K + \frac{i\pi}{2} (1 - s_{ij}) \right) (Z_i Z_j \tilde{Z}_i \tilde{Z}_j - 1), \quad (7)$$

which is a (2+0)D *classical Ashkin-Teller model* [67] where $Z_i, \tilde{Z}_i = \pm 1$ are the classical Ising variables of the two layers of the tensor network wavefunction $|\psi(s)\rangle$ and $\langle \psi(s)|$, respectively. Here K, J are related to the measurement strength t and coherent rotation angle θ by $\tanh(J) = \sin(2t) \cos(\theta)$ and $e^{-2K} = \sinh(J) \tan(\theta)$. Note that for a finite rotation angle towards the Y direction, $\phi > 0$, this Ashkin-Teller model is augmented by *complex* couplings. In the expressions above, we omit the dependence on the logical qubit and focus on the tensor in the bulk, while the logical qubit can be captured by the boundary virtual leg of the tensor network, see Refs. [20, 35] for more details.

(1+1)D quantum circuit & fermion representation

The 2D classical tensor network can also be interpreted as a (1+1)D monitored quantum circuit, which evolves towards a boundary quantum chain. Its transfer matrix can be written as $T := T_v T_h = \exp(-H_v) \cdot \exp(-H_h)$ where $H_{h,v}$ have been

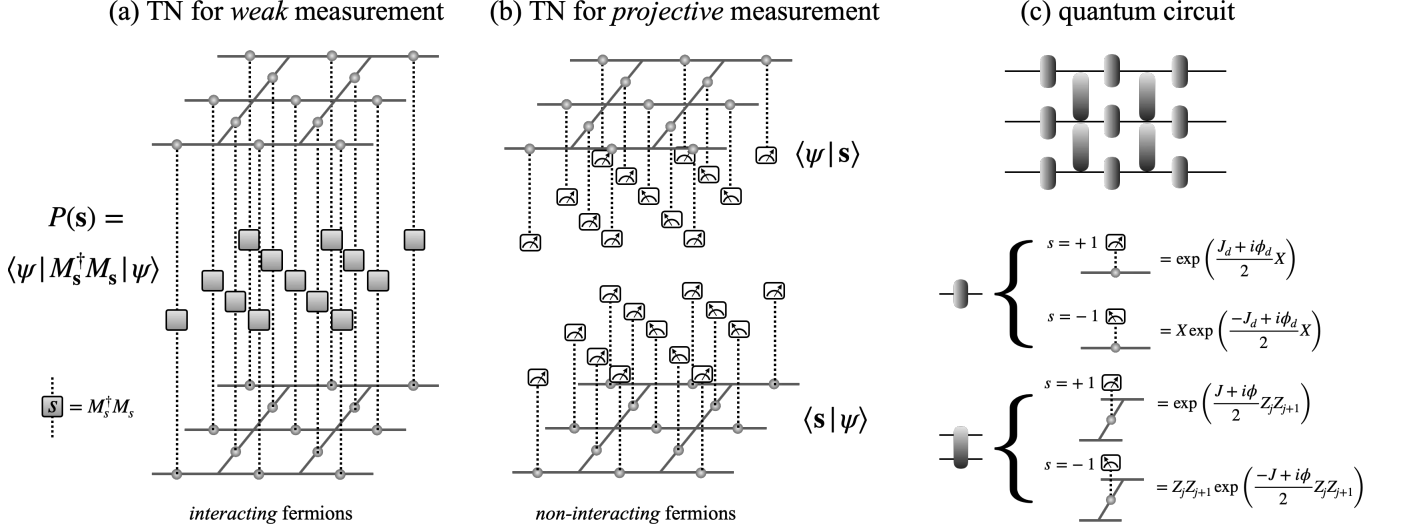


FIG. 3. **Random 2D tensor network and (1+1)D quantum circuit.** The top and the bottom layer of the tensor network captures the bra and ket of the quantum state as a PEPS. (a) For weak measurement, the Kraus operator $M_s^\dagger M_s$ glues the two layers. (b) For projective measurement, $M_s^\dagger M_s$ is a projector, such that $P(s) = |\langle s | \psi \rangle|^2$, which means the two layers can be factorized. As will be shown in Sec. IV B, each layer alone corresponds to a random Ising model with complex coupling, which can be fermionized into a free fermion network model. In this aspect, weakening the measurement strength in (a) induces fermion interactions. Both (a) and (b) as 2D tensor network can be interpreted as (1+1)D (non-unitary) quantum circuit that propagates a 1D quantum chain in the space direction, updated by the transfer matrix. (c) The ket layer of tensor network in (b) can be identified as a monitored (1+1)D quantum circuit. The gate parameters are controlled by the coherent rotation $U(\theta, \phi)$ and the measurement outcome s . Here we show only the ket layer of the state. The gates are derived from Eq. (26) (27).

previously derived [20] as

$$\begin{aligned}
 -H_h = & \sum_j \frac{J s_{2j,2j+1}}{2} (\sigma_{2j}^x \sigma_{2j+1}^x + \sigma_{2j}^y \sigma_{2j+1}^y) \\
 & + i \frac{\phi}{2} (\sigma_{2j}^x \sigma_{2j+1}^x - \sigma_{2j}^y \sigma_{2j+1}^y) \\
 & - \left(K + i \frac{\pi}{2} \frac{1 - s_{2j,2j+1}}{2} \right) (\sigma_{2j}^z \sigma_{2j+1}^z + 1),
 \end{aligned} \quad (8)$$

which in turn can be fermionized as

$$\begin{aligned}
 -H_h = & \sum_j J s_{2j,2j+1} (c_{2j}^\dagger c_{2j+1} + h.c.) + i\phi (c_{2j} c_{2j+1} + h.c.) \\
 & - 2 \left(K + i \frac{\pi}{2} \frac{1 - s_{2j,2j+1}}{2} \right) (2n_{2j} n_{2j+1} - n_{2j} - n_{2j+1} + 1),
 \end{aligned} \quad (9)$$

where $n_j = c_j^\dagger c_j$ is the fermion number operator. H_v takes the same form as H_h , with the replacement $(2j, 2j+1) \rightarrow (2j-1, 2j)$ and the coupling constants transformed into their Kramers-Wannier dualized versions, c.f. Eq. (28) below. Physically, T_h and T_v implement quantum gates on even and odd links, respectively.

In general, the fermionic Hamiltonian is *interacting*, except when the measurement basis is restricted to Z or X axes. However, we find that a *non-interacting* fermion model can capture the projective-measurement limit for arbitrary measurement angles, see Fig. 3(b)(c). We postpone the detailed analysis of this scenario to Sec. IV B below.

B. Coherent information

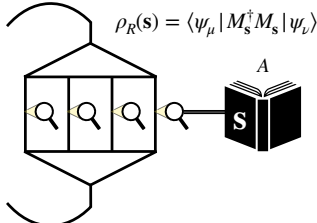
The coherent information, conceptualized in quantum information theory to quantify the size of a coherent logical space through a noisy quantum channel [49], was recently identified as an efficient (scalable) order parameter for measurement-induced phase transitions [68], and for diagnosing topological quantum memory in the presence of incoherent noise [14], under a teleportation channel [20], or for topological mixed states of fermions [69]. Here we employ it as a diagnostic for the topological phase transition of a mixed toric code state under measurement. To keep this manuscript self-contained, let us briefly review this concept, explain it in an intuitive manner and – following its mapping [20] to the tensor network and statistical mechanical models introduced above, which enables us to calculate our phase diagrams to high precision.

The quantum coherent information is a measure for the quantum information in a system, which can be encoded and decoded. For instance, in the surface code one logical qubit can be encoded and used for further manipulations. After the measurement, the amount of quantum information that remains still inside the code is given by the coherent information [49]. To calculate it, we consider three parties of our system as illustrated in Fig. 2(a): (i) a reference qubit R that is initially maximally entangled with the logical space of (ii) a surface code B and (iii) an observer A that extracts some information from the measurement outcomes. Then the coherent information is expressed as the entropy of the reference subtracted by the mutual information between the reference

and the observer

$$I_c = S(R) - I(R : A) = S(RA) - S(A) = S(R|A). \quad (10)$$

As a conditional quantum entropy, it evaluates how much extra information is needed for the observer to learn the reference qubit. Thus $I_c = 0$ means the observer can learn the state of the reference based on the measurement records, while $I_c = \log 2$ means the observer learns nothing about the quantum information based on the measurement outcome. Note that a negative value of I_c cannot appear in our measurement setting (but could in a noisy scenario, where both quantum and classical information are lost [49]). In our case the observer has access to a purely classical state only, which can be cast as a diagonal density matrix $\sum_s P(s) |s\rangle\langle s|$. Further note that the diagnostic depends only on the correlation between R and A (the observer), while the quantum degrees of freedom in B (the surface code) are traced out (discarded):



$$\rho_R(s) = \langle \psi_\mu | M_s^\dagger M_s | \psi_\nu \rangle \quad (11)$$

Therefore the phase diagram depends essentially only on the ensemble of the reference states

$$\{P(s), \rho_R(s)\}, \quad (12)$$

where the state R is obtained by tracing out the measured surface code, whose matrix elements are given by

$$\rho_{R,\mu\nu}(s) = \langle \psi_\mu | M_s^\dagger M_s | \psi_\nu \rangle, \quad (13)$$

where ψ_\pm marks the ground state of the surface code with the respective logical qubit state. Note that this ensemble of states can be viewed as *the ensemble of the topological logical qubit encoded into the surface code*.

We can denote the matrix elements of Eq. (13) as

$$\rho_R(s) = \frac{1}{2P(s)} \begin{pmatrix} P_{++}(s) & P_{+-}(s) \\ P_{-+}(s) & P_{--}(s) \end{pmatrix} \equiv \frac{1 + \vec{\kappa} \cdot \vec{\sigma}}{2}, \quad (14)$$

where the prefactor is necessary for proper normalization. The probability $P(s)$ of a given set of measurement outcomes can be obtained by $P(s) = (P_{++}(s) + P_{--}(s))/2$. On the right-hand side of the equation above we define an effective polarization field in the logical space which results from the observer's measurement

$$\vec{\kappa}(s) = \begin{pmatrix} \frac{P_{++} - P_{--}}{P_{++} + P_{--}}, & \frac{2|\text{Im}(P_{+-})|}{P_{++} + P_{--}}, & \frac{2|\text{Re}(P_{+-})|}{P_{++} + P_{--}} \end{pmatrix}, \quad (15)$$

whose strength can be viewed as the effective learning strength in the long-wavelength limit, after tracing out the whole physical degrees of freedom. Its angle indicates the *effective measurement angle* of the logical qubit, which should

be distinguished from the measurement angle of the physical qubits that is fixed when choosing a measurement basis. Each component of $\vec{\kappa}$ can be related [13] to the anyon condensation and confinement mechanisms at play when decohering the topological quantum order of the toric code: The component κ_X is a sign of m -particle condensation, the component κ_Z a sign of e -particle condensation and the κ_Y component can only arise by a mixture of both mechanisms and therefore has to include measurement of Y operators. (Thus κ_Y vanishes, for instance, in the $X - Z$ plane.)

Expressed in terms of this field, the coherent information becomes the Born average of the conditional reference entropy $I_c = \sum_s P(s) I_s$, where I_s is the von Neuman entropy of the reference or logical qubit conditioned upon the observer's measurement outcome s

$$I_s = -\frac{1 + C_s}{2} \log \frac{1 + C_s}{2} - \frac{1 - C_s}{2} \log \frac{1 - C_s}{2}, \quad (16)$$

with $C_s = |\vec{\kappa}(s)|$ being the strength of the polarization field.

As it is a measure for the information of the logical qubit, the coherent information is $I_c = \log 2$ in the unperturbed surface code phase and $I_c = 0$ in the limit of projective measurement. Where the coherent information drops between these two limiting values, possibly at some intermediate measurement strength, the transition from surface code phase to an ensemble of product states occurs. This information-theoretical observable is found to exhibit minimal finite-size effects [70] and has provided, for instance, the highest-precision determination of the threshold and critical exponents of the Nishimori phase transition [20, 23, 33, 34]. In our interacting fermion case here, this behavior of the coherent information is illustratively shown in the top row of Fig. 4. with panel (a) corresponding to the situation where the transition occurs at *finite* measurement strength, while panel (b) corresponds to a situation where the transition is pushed out all the way to the projective limit.

Recasting the coherent information in the language of a learning transition, the coherent information evaluates the amount of information preserved in the code [49], that is *not* extracted by the observers. In the “thermodynamic limit” of large code distances ($d \rightarrow \infty$), the coherent information converges to a step function that maintains one bit of information ($\log 2$) when the measurement strength remains below the threshold, and zero when it moves above the threshold. The precise threshold of the measurement strength can be determined, e.g. by performing finite-size scaling (in the code distance), and thereby pinpoint the location of the learning phase transition. We map out such thresholds for arbitrary measurement angles and thereby determine the phase diagram in Fig. 1(b) above. For a full spherical illustration see also Fig. 12 below. One can immediately recast this phase diagram also in terms of the post-measurement ensemble of quantum states. Below the measurement threshold, i.e. inside the interior of our “carved sphere”, the quantum states remain topologically ordered, while the states dephased into classical loop gas [35] above the threshold, i.e. on the surface of our sphere and in its “carved-out” regions, see Fig. 12. In the language of spontaneous symmetry breaking (SSB), the topological phase

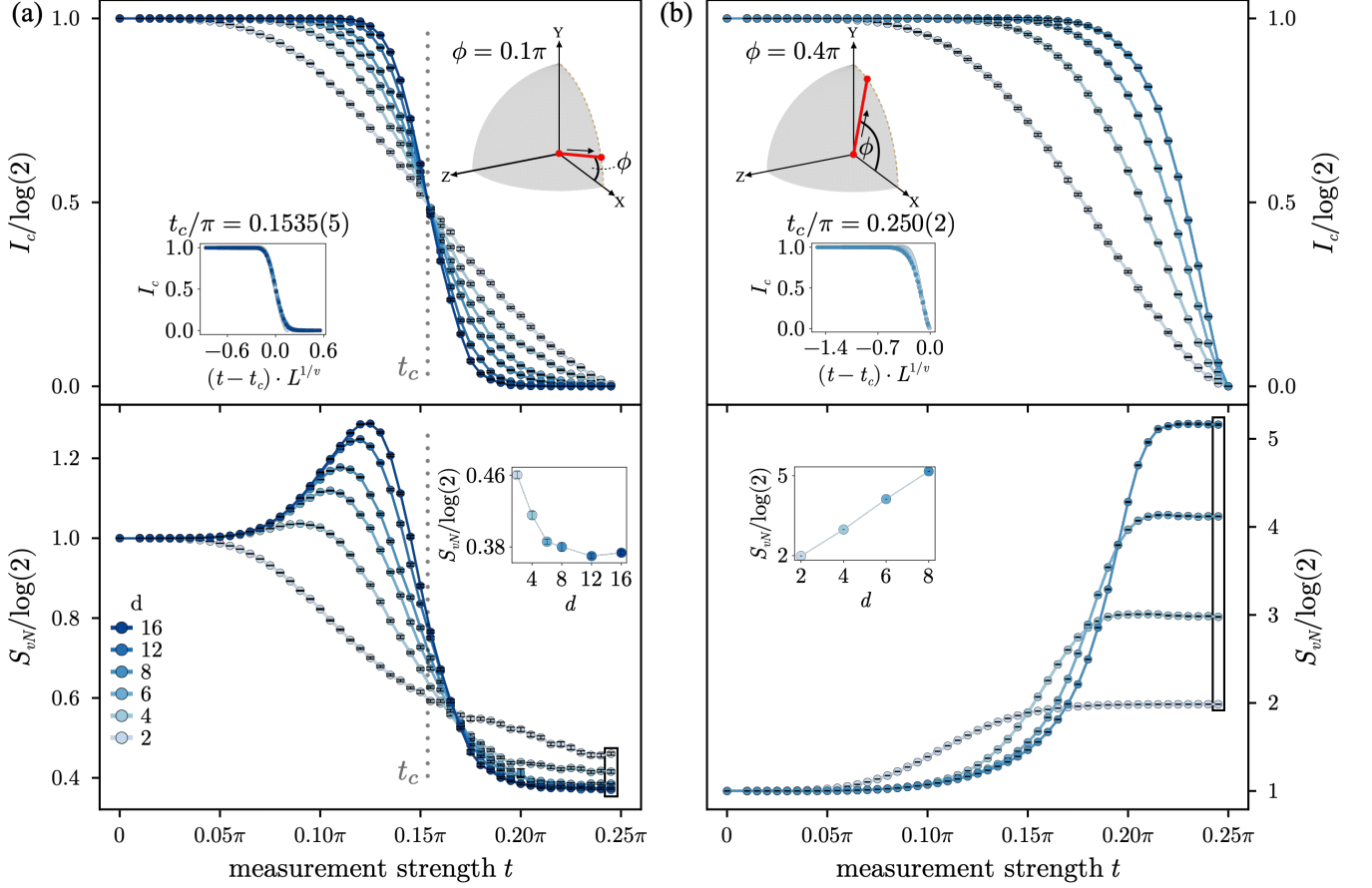


FIG. 4. **Learning transitions by tuning the measurement strength.** Shown are two cuts in the XY plane for *finite* measurement strength t , one far from and one close to the Y -point, shown on the left/right respectively. The top row shows the *coherent information*, while the bottom row shows the *boundary entanglement entropy*. (a) For a cut at an angle $\phi = 0.1\pi$ (close to the X point in the XY plane) the transition occurs at a *finite* measurement strength with the coherent information exhibiting a crossing point at a critical t_c . The boundary entanglement entropy exhibits a peak that approaches the transition point for larger system sizes with $\log(L)$ scaling, characteristic of the scaling at a (1+1)D quantum critical point. In the projective-measurement limit of $t = \pi/4$ the entanglement entropy decreases with system size, indicating an area law phase. (b) For a cut at an angle $\phi = 0.4\pi$ (close to the Y point in the XY plane) the transition occurs only in the *projective* measurement limit ($t = \pi/4$), where the coherent information exhibits a collapse. The boundary entanglement entropy, in contrast, exhibits an entanglement plateau for finite system sizes, whose height scales linearly with code distance, reminiscent of a volume-law scaling.

exhibits SSB of the underlying 1-form symmetries, while the dephased phase exhibits a strong-to-weak SSB of the 1-form symmetries [15–19, 71].

C. Boundary entanglement entropy

Our second diagnostic to detect and characterize the critical theory of the learning transition is the boundary entanglement entropy. The latter is calculated for a one-dimensional chain of boundary qubits that we attach to the surface code and which we leave untouched when measuring out the physical qubits of the surface code. The entanglement structure of the Choi state (obtained by flattening the density matrix to a “wave function” vector [72]) of this boundary can be interpreted as arising from a (1+1)D monitored quantum circuit evolution (induced by the 2D bulk measurements), with the

relevant transfer matrices discussed around Eqs. (8) and (9) above. Alternatively, we can think of the boundary entanglement entropy to characterize the computational complexity of contracting out the 2D tensor network in Fig. 3(a).

Numerical results for the boundary entanglement entropy along the two finite-measurement cuts in the XY plane of the phase diagram are shown in the lower panel of Fig. 4. For the learning transition at finite measurement strength along the first cut (close to the X measurement direction), we observe a non-monotonic behavior of the entanglement entropy with an intermediate peak close to the transition. For increasing code distances, this peak not only moves closer to the transition, but shows a *logarithmic* divergence – indicative of a (1+1)D quantum critical point. In contrast, for the cut close to the Y point where the learning transition is pushed all the way to the projective-measurement limit the boundary entanglement entropy monotonically increases with no indication

of an intermediate transition. Notably, however, an entanglement plateau is visible whose height increases *linearly* with code distance, reminiscent of a volume-law phase. However, since for larger system sizes the width of the plateau also shrinks and possibly vanishes for very large code distances, this plateau might not indicate an intermediate phase, but point to interesting behavior in the projective-measurement limit which we characterize in detail in the following section.

IV. PROJECTIVE-MEASUREMENT LIMIT

One of the most striking observations in mapping out the learning transition in the XZ plane has been that at the self-dual $X + Z$ point the transition is pushed all the way out to the projective limit [20, 35]. This situation where the surface code remains robust or “unlearnable” all the way up to the projective-measurement limit no longer remains a fine-tuned possibility, but instead becomes the generic behavior once we introduce a finite Y -measurement component, see our phase diagram in Fig. 1 or, probably even more convincing, the 3D illustration in Fig. 12 below. But how much do we actually learn about the surface code’s logical state in this projective limit? For this, it will turn out to be quite instructive to look at the reference qubit which, in the projective-measurement limit, captures the purified logical qubit. In particular, we want to understand its distribution when the measurement protocol is repeated for multiple rounds with many copies of the code. We will show that “unlearnable” is a label that only applies to weak measurements, while for projective measurements quite the opposite is true – by performing multiple measurements along the same measurement angle one can perform *full-state tomography* of the logical qubit state encoded in the surface code.

For these generic measurement angles where one is able to learn the logical qubit state via tomography of the reference qubit, one might wonder what the nature of the measured-out surface code is in this projective limit. That is, what can we say about the ensemble of product states or, equivalently, the set of measurement readout snapshots in the projective-measurement limit? Is this ensemble describing a critical state akin to what has been observed for the self-dual $X + Z$ point [35] or is it a general “gapless” state (with long-range entanglement similar to the ground state of a gapless Hamiltonian)? We will show that, generically, one finds the latter situation, i.e. a metal-like gapless Gaussian fermion state.

A. Ensemble of logical snapshots

Let us start our discussion of the projective-measurement limit, by first noting that the coherent information vanishes in this limit, $I_c = 0$, because the surface code state collapses to a classical snapshot $|\mathbf{s}\rangle$. At the same time, the reference qubit R , which we introduced to calculate the coherent information, is purified into a state $|\psi_R(\mathbf{s})\rangle$ where

$$|\psi_R(\mathbf{s})\rangle = \sum_{\mu=\pm} \langle \mathbf{s} | U(\theta, \phi) | \psi_\mu \rangle | \mu \rangle, \quad (17)$$

with $U(\theta, \phi)$ being the coherent rotation (1). The associated density matrix $|\psi_R(\mathbf{s})\rangle\langle\psi_R(\mathbf{s})|$

$$|\psi_R(\mathbf{s})\rangle\langle\psi_R(\mathbf{s})| = \frac{1 + \vec{\kappa}(\mathbf{s}) \cdot \vec{\sigma}}{2} \quad (18)$$

can be visualized as a vector which points at the $\vec{\kappa}(\mathbf{s})$ direction in the Bloch sphere. In particular, for measurement in the Z or X basis directions, the purified reference qubit R points exactly at the $\pm Z$ and $\pm X$ directions, where the sign depends on the measurement outcome. Rotating the measurement axis to the self-dual $X + Z$ direction, the purified qubit R is restricted to lie in the XZ plane (since the imaginary part has to vanish). Plotting our numerical results in Fig. 5(a), we see that one observes a uniform “ring-distribution” in the XZ plane. Moving the measurement direction out of the XZ plane we observe what looks like a uniform spherical distribution for the regimes where the learnability threshold is moved all the way to the projective limit, see Figs. 5(b) and (c).

To quantitatively understand this behavior, let us formally write the distribution of the reference qubit as

$$P(\theta, \phi) = \sum_{\mathbf{s}} P(\mathbf{s}) P(\vec{\kappa}|\mathbf{s}), \quad (19)$$

where $P(\vec{\kappa}|\mathbf{s})$ is a conditional delta-distribution since each measurement outcome \mathbf{s} maps to a deterministic $\mathbf{s} \mapsto \vec{\kappa}(\mathbf{s}) = \cos(\theta)\vec{z} + \sin(\theta)\cos(\phi)\vec{x} + \sin(\theta)\sin(\phi)\vec{y}$. Later we will use θ, ϕ to label $\vec{\kappa}$ in spherical coordinates. This distribution we now want to compare with an ideal reference distribution such as a uniform coverage of the Bloch sphere, i.e. a Haar-measure random distribution of the single logical qubit, which can be parametrized by

$$Q(\theta, \phi) = \frac{1}{4\pi} \sin\theta d\theta d\phi. \quad (20)$$

The “distance” between this reference distribution $Q(\theta, \phi)$ and our observed ensemble $P(\theta, \phi)$ can be quantified by their relative entropy, also known as Kullback-Leibler (KL) divergence

$$D(P||Q) = \int P(\theta, \phi) \log \frac{P(\theta, \phi)}{Q(\theta, \phi)} d\theta d\phi, \quad (21)$$

where $Q(\theta, \phi)$ is the ideal and $P(\theta, \phi)$ is the measured distribution. If two distributions perfectly align, their relative entropy vanishes and the KL divergence is zero.

To evaluate the KL divergence numerically, we perform a Born average over the “free-energy” difference between the measured distribution and the ideal distribution

$$D(P||Q) = \mathbb{E}_P(\log P - \log Q), \quad (22)$$

with more details on our numerical procedure provided in Appendix A. To discuss our numerical data in Fig. 5, let us first mention asymptotic results for some limiting cases. For the Z and X measurement limits which converge towards bimodal distributions, the KL divergence (with regard to the uniform distribution) diverges as $\log n - \log 2$ where n is the number

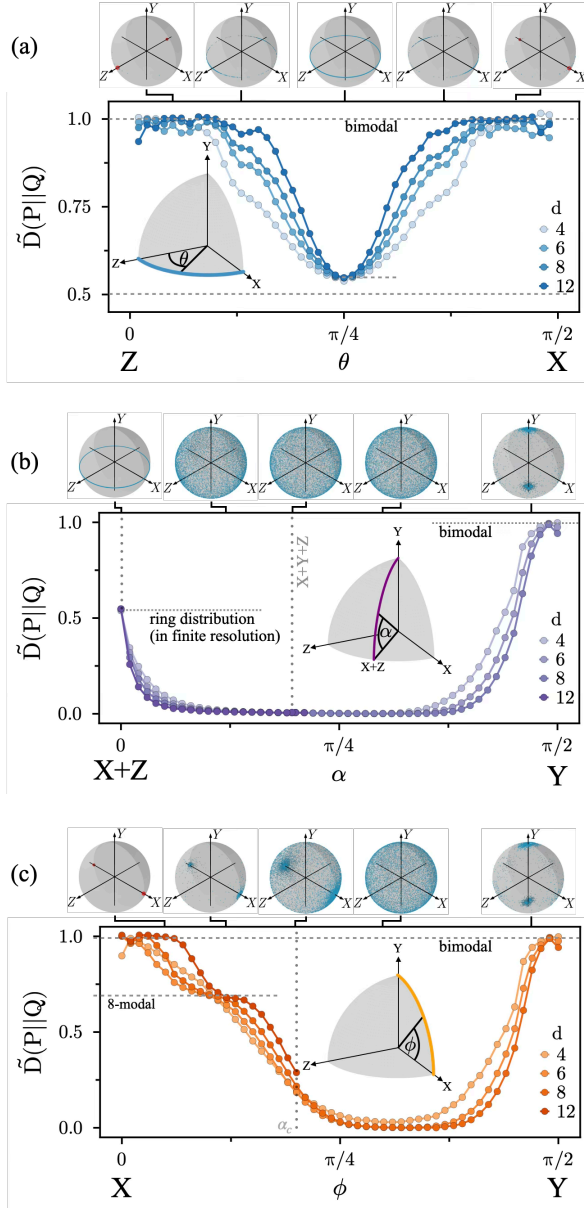


FIG. 5. **Projected ensemble of the reference qubit on (a) the XZ line, (b) the self-dual line, and (c) the XY line.** The spheres above each plot show the projected ensemble distribution for different lines on the Bloch sphere labeled by α in the phase diagram by displaying 10^4 sampled points from the distribution. The X and Z points on the Bloch sphere of highly concentrated densities are marked with small red spheres. In the above plots, the Kullback-Leibler divergence (relative entropy) between the projected ensemble and a uniform (Haar) distribution of the Bloch sphere is normalized by the bimodal distribution value $\tilde{D}(P||Q) = D(P||Q)/(\log n - \log 2)$ where n is the number of discrete patches covering the sphere. For the ring distribution, $\tilde{D}(P||Q)$ becomes $1 - \log(m)/\log(n) \rightarrow 1/2$, where m is the number of patches covering the equator as a 1D line, such that the ratio of dimension is $\log m/\log n \rightarrow 1/2$ for the limit of high resolutions. In the $n \rightarrow \infty$ limit, the dip in (a) at $\theta = \pi/4$ and the first data point at $X + Z$ in (b) reach exactly $\tilde{D}(P||Q) = 1/2$. See Appendix A for more details.

of discretized patches of the sphere (see Appendix A). In contrast, near the self-dual Y limit, it is split into 8 hotspots near the $\pm Y$ poles, which are required by the self-dual symmetry that relates X to Z . Consequently, the KL divergence near Y limit reduces to $\log n - 3 \log 2$, which still diverges but is always smaller by 2 bits than the X and Z limits, due to the symmetry. For the “critical” state at the self-dual $X + Z$ measurement angle, the ideal distribution is a uniformly distributed equator-like ring in the XZ plane

$$Q(\theta, \phi) = \frac{1}{2\pi} [\delta(\phi) + \delta(\phi - \pi)] d\theta d\phi,$$

where $\delta(\phi)$ is the Dirac delta function that integrates to 1. As shown in Fig. 5(a)(b), it approaches a half of the value for the uniform spherical distribution. The basin plateau of the relative entropy in Fig. 5(b)(c), where the KL divergence vanishes, corresponds to the uniform spherical distribution. At these measurement angles, the reference qubit indeed forms a Haar-random post-measurement ensemble.

Let us now explain what such a Haar-random reference qubit distribution allows us to infer about the original logical qubit. In brief, the Haar-randomness of the reference distribution implies that we can effectively measure the logical qubit in *random* directions [52], even though the physical measurement basis is *uniformly fixed* at a certain angle. The protocol is as follows: Instead of starting from a maximally mixed logical state, consider that one prepares a surface code in some pure logical state, which remains unknown to the observer and that he wants to learn about. To this end, the observer performs measurements for all physical qubits along an arbitrary (but uniform) angle, e.g. along the generic $X + Y + Z$ direction, and thereby obtains a measurement snapshot in the form of a bit string $s = ++--\dots$. For each possible s , one can invert the isometry encoder (Fig. 2(a)) to “decode” [73] the corresponding snapshot of the logical qubit $|\psi_L(s)\rangle$, which exactly equals the wave function $|\psi_R(s)\rangle$ in Eq. (17) given the same s . If we denote the isometry encoder as U_{enc} that maps a single qubit $|\mu\rangle$ to a surface code in the logical μ state: $|\psi_\mu\rangle = U_{\text{enc}}|\mu\rangle$, then the decoded snapshot of the logical qubit follows Eq. (17) and Fig. 2(a) to be

$$|\psi_L(s)\rangle = U_{\text{enc}}^\dagger U(\theta, \phi)^\dagger |s\rangle. \quad (23)$$

After this decoding step, one thereby gains a *snapshot of the logical qubit*. Multiple rounds of measurements over many copies of the code then effectively allow measuring the logical qubit in arbitrary direction, forming a classical shadow [48] that is *tomographically complete* [52] in this case. That is, one can perform a full state tomography for an arbitrary prepared logical qubit, based on the measurement snapshots at a single measurement angle (uniformly applied to all physical qubits). Such uniform basis measurement is experimentally quite feasible in a number of quantum platforms, particularly for large-scale neutral Rydberg atom arrays [74].

B. Monitored (1+1)D quantum circuit representation

For each measurement record \mathbf{s} , the reference state (17) or the equivalent logical snapshot (23) can be computed by contracting the tensor network in Fig. 3, with extra legs that capture the reference qubit (see Ref. [20] for details), while the physical qubits are traced out. In the projective-measurement limit of interest here, the bilayer tensor network reduces to two *decoupled* layers, as illustrated in Fig. 3(b). Each layer corresponds to a random classical 2D tensor network state, which can be viewed as a sampled snapshot of the 2D surface code state with coherent error, and the tensor network captures the wave function amplitude of such a snapshot. Remarkably, each layer of the 2D tensor network can further be mapped to a (1+1)D Ising-type quantum circuit, which can then be mapped to a *free-fermion* circuit. Contracting this decoupled tensor network or its free-fermion representation not only gives us direct access to the probability distribution of the measurement outcomes $P(\mathbf{s})$, but also the boundary entanglement entropy of the (1+1)D quantum state. It can also be done in a highly effectively manner [35], allowing us to simulate considerably larger code distances up to $d = 256$ in the projective-measurement limit.

Let us derive the (1+1)D monitored quantum circuit representation, by identifying step-by-step the quantum gates of this circuit from the tensor network in Fig. 3(b). Specifically, an elementary gate in this circuit is generated by measuring a qubit on a given bond (ij) , where the virtual index of the tensor network is labeled by σ_i and σ_j for i, j on the adjacent vertices. Under the duality picture, the Pauli Z of the bond qubit is related to the virtual index by $Z_{ij} = \sigma_i \sigma_j$. Physically, σ_i and Z_{ij} can be interpreted as the order and disorder (domain-wall) parameters, respectively. Therefore, projective measurement of the bond qubit yields the following weight

$$\langle s_{ij} = \pm | e^{i\frac{\theta}{2}Y} e^{i\frac{\phi}{2}Z} | \sigma_i \sigma_j \rangle \propto \exp \left(\frac{J s_{ij} + i(\phi - \pi \frac{1-s_{ij}}{2})}{2} \sigma_i \sigma_j \right), \quad (24)$$

where the effective measurement strength

$$J(\theta) = \tanh^{-1} \cos \theta \quad (25)$$

ranges from ∞ to 0 as θ goes from 0 to $\pi/2$ i.e. upon varying the measurement angle from Z to the XY plane.

Cast in the (1+1)D setting, one can directly write down the 2-body interacting gate (Fig. 3(c))

$$M_Z = \exp \left(\frac{J s_{ij} + i(\phi - \pi \frac{1-s_{ij}}{2})}{2} Z_j Z_{j+1} \right), \quad (26)$$

which is a complex time evolution generated by Ising interaction and can be decomposed into a unitary part controlled by ϕ and a non-unitary (weak measurement) part controlled by θ , followed by a conditional π -pulse unitary as a local feedback. The negative measurement outcome effectively changes the sign of the real coupling constant and adds a π pulse to the imaginary coupling constant, which is consistent with the spatial inversion operation for the physical spin eigenstate pointing towards spherical angle $(\theta, \phi) : \theta \rightarrow \pi - \theta, \phi \rightarrow \phi - \pi$.

The transverse-field-generated evolution gate M_X can be derived by viewing σ_i and σ_j as the states at two adjacent time slices

$$M_X \propto 1 + e^{-(J_d s + i\phi_d)} s X \propto \exp \left(\frac{J_d s + i(\phi_d - \pi \frac{1-s}{2})}{2} X \right), \quad (27)$$

which is essentially a *Kramers-Wannier dual* version of the M_Z gate where

$$\frac{J_d + i\phi_d}{2} = \text{KW} \left(\frac{J + i\phi}{2} \right) = -\frac{1}{2} \log \tanh \frac{J + i\phi}{2} \quad (28)$$

is the Kramers-Wannier (KW) duality relation. More explicitly,

$$J_d = -\log \frac{\sqrt{1 + e^{4J} - 2e^{2J} \cos(2\phi)}}{1 + e^{2J} + 2e^J \cos \phi}, \quad (29)$$

$$\phi_d = -\tan^{-1} \frac{2e^J \sin \phi}{1 - e^{2J}}.$$

Kramers-Wannier duality

Let us make some general comments on the Kramers-Wannier (KW) duality in our circuit and its manifestation in the phase diagram. Firstly, note some useful properties of the KW map for the coupling constants

$$\text{KW}(x^*) = \text{KW}(x)^*, \quad \text{KW}(-x) = \text{KW}(x) - i\pi/2, \quad (30)$$

which means $\text{KW}(\cdot)$ commutes with complex conjugation, but flipping the sign is dualized to a phase shift by $\pi/2$. Therefore, taking the complex conjugate flips the sign of ϕ and results in the sign change of ϕ_d . For the XZ plane, $\phi = 0$, and as such, the self-duality condition $\text{KW}(x) = x$ is *real*, with the unique solution $x = \log \sqrt{1 + \sqrt{2}}$ corresponding to the $X + Z$ point. Once we turn on the Y -component, the electric-magnetic self-duality condition becomes *complex* $\text{KW}(x) = x^*$, which leads to an *extended self-dual line* from the $X + Z$ point to the Y point, along which the measurement angle is symmetric between X and Z

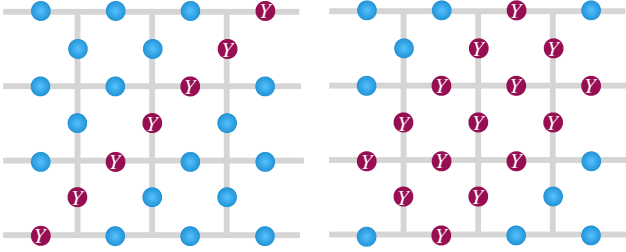
$$\text{KW}(x) = x^* \Leftrightarrow \cos \phi = \cot \theta = \sinh J. \quad (31)$$

As is evident from the quantum gates M_Z and M_X , in the projective-measurement limit, the (1+1)D quantum circuit can be modeled as a monitored Ising chain, which can be equivalently described by a monitored *non-interacting* fermion chain, thereby simplifying the analytical analysis substantially. We will utilize this facilitation to study the non-Hermitian dynamics of the monitored fermion chain under post-selection in Sec. V.

The 2D tensor network state and its corresponding (1+1)D monitored quantum circuit share the same probability distribution function

$$P(\mathbf{s}) = \langle \psi(\mathbf{s}) | \psi(\mathbf{s}) \rangle = \| M_X M_Z \cdots M_X M_Z | \phi \rangle \|^2, \quad (32)$$

(a) different realizations of logical Y operators



(b) Majorana worldlines in ballistic metal

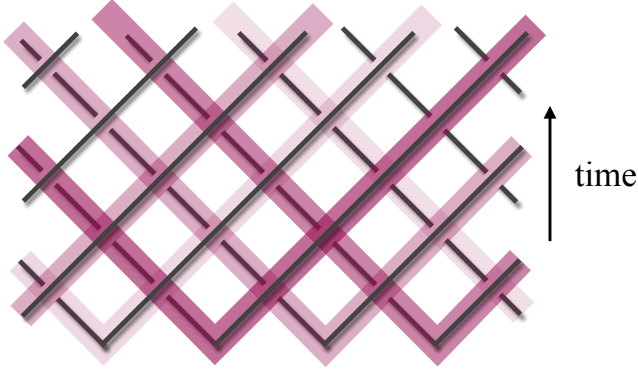


FIG. 6. **Projective Y-measurement limit.** (a) Logical Y operators. By measuring the Pauli Y operators on the purple qubits, one can learn the logical Y information, i.e. $Y_L = \pm 1$. (b) Ballistic metal from measuring toric codes in the Y basis. The light cones in different shades illustrate the effective (1+1)D dynamics of Majorana. The measurement operators in Eq. (34) become purely unitary and act as swap operators between maximally entangled Majorana pairs. As a result, the entangled Majorana pairs labeled by the same color propagate ballistically in time along the light cones.

where $|\psi\rangle$ denote the 2D quantum state (neglecting the logical qubit dependence for simplicity here), while $|\phi\rangle$ denotes the 1D quantum state evolving from the left to the right boundary by the random transfer matrix M_X and M_Z conditioned upon the measurement outcomes s . Such “monitored trajectories” that follow Born’s probability rule can be efficiently sampled [35] and are in one-to-one correspondence with the measured bit strings s of the 2D quantum state.

Before we proceed to a presentation of our numerical results for this, let us look at the structure of the quantum gates M_X and M_Z of the (1+1)D monitored circuit for three special limits: (i) the self-dual line from the self-dual $X + Z$ to the Y point, (ii) the Y point, and (iii) the XY line from the X to the Y point.

Self-dual line

Along the self-dual line from the $X + Z$ to the Y point, parametrized by Eq. (31), the (1+1)D quantum circuit exhibits Kramers-Wannier self-duality. It generalizes the

measurement-only circuit at the $X + Z$ point, extensively studied in Ref. [35], to one which, in addition, includes unitary gates. This can be seen by plugging the self-dual condition into Eq. (29), such that $J_d = J$ and $\phi_d = -\phi$, and the quantum gates of the (1+1)D quantum circuit turn into

$$\begin{aligned} M_Z &= \exp\left(\frac{Js_{ij} + i(\phi - \pi \frac{1-s_{ij}}{2})}{2} Z_j Z_{j+1}\right), \\ M_X &= \exp\left(\frac{Js - i(\phi - \pi \frac{1-s}{2})}{2} X\right), \end{aligned} \quad (33)$$

which are invariant under translation and sublattice complex conjugation. Whereas at the $X + Z$ point, where $\phi = 0$ the circuit reduces to a measurement-only quantum circuit, once we turn on the Y component along the self-dual line, the circuit becomes a hybrid one of measurements and unitary evolutions.

Exactly solvable dual unitary dynamics in the Y -limit

In the Y -only measurement limit, $J = 0 = J_d$ and $\phi = -\phi_d = \pi/2$, the gates reduce to purely unitary gates of $\pi/2$ pulses (maximally entangling gates) with random signs

$$M_Z \rightarrow e^{\pm i \frac{\pi}{4} Z_j Z_{j+1}}, \quad M_X \rightarrow e^{\pm i \frac{\pi}{4} X}. \quad (34)$$

This circuit dynamics is shown in Majorana representation in Fig. 6(b) where the “light cone” dynamics illustrates the formation of a *ballistic metal*. It is a special case of the dual-unitary circuit, which is unitary along both the “time” direction and the “space” direction. Consequently, the entanglement entropy grows as

$$S_{vN} = t \log 2, \quad (35)$$

for $d > t$.

In this limit, all the physical qubits are collapsed onto the Y basis. Without extra noise, one can read the logical Y information from the measurement outcomes supported on the logical Y operators. See Fig. 6(a).

XY line

Along the XY line, $\theta = \pi/2$, which leads to $J = 0$ and $J_d = -\log \tan \frac{\phi}{2}$ and $\phi_d = -\pi/2$. Consequently, the quantum gates reduce to

$$\begin{aligned} M_Z &= \exp\left(i \frac{\phi}{2} Z_j Z_{j+1}\right), \\ M_X &= \exp\left(\frac{s J_\phi}{2} X\right), \end{aligned} \quad (36)$$

where $J_\phi \equiv -\log(i \tan \frac{\phi}{2}) = \tanh^{-1} \cos \phi - i \frac{\pi}{2}$. Here all the random Z_j operators inserted into the circuit are propagated

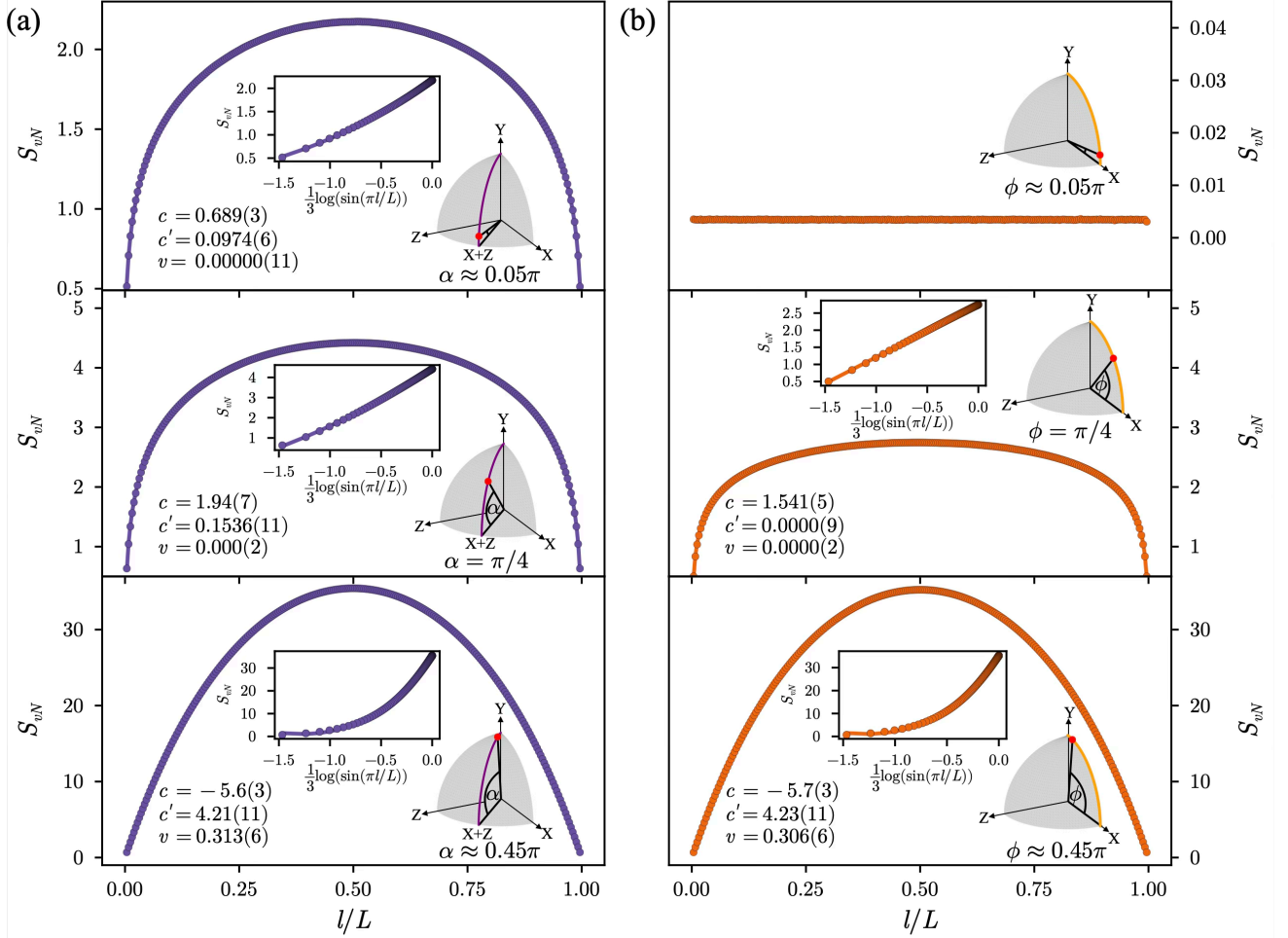


FIG. 7. **Entanglement arcs.** Shown are numerical data for the boundary entanglement entropy $S(l)$ as a function of subsystem size l along the boundary of total length $L = 256$ at different points in the projective-measurement limit. The solid lines are the fitting curves using the entanglement arc formula of Eq. (39). The prefactors obtained by these fits are given and the insets show the data rescaled by $1/3 \log(\sin(\pi l/L))$. The entanglement arcs are shown for measurement angles on (a) the self-dual line and (b) the XY line, where their measurement angle values are indicated by red dots in the inset.

to the final time slice, by flipping certain signs of the X evolution. The corresponding 2D statistical model can then be simply written as

$$\mathcal{Z}(s) \propto \sum_{\sigma} \exp \left(\frac{i}{2} \sum_{ij} \left(\phi - \pi \frac{1 - s_{ij}}{2} \right) \sigma_i \sigma_j \right), \quad (37)$$

which can also be easily derived from the tensor network [c.f. Fig. 3]. For example, the Boltzmann weight for a single bond can be derived as follows

$$\begin{aligned} \begin{array}{c} \text{X} \\ \diagup \quad \diagdown \\ \text{---} \oplus \text{---} \end{array} &= \begin{array}{c} \left(\begin{array}{cc} 1 & e^{-i\phi} \\ e^{i\phi} & 1 \end{array} \right) \\ \text{---} \text{---} \end{array}, \quad s = +1 \\ \\ \begin{array}{c} \text{X} \\ \diagup \quad \diagdown \\ \text{---} \oplus \text{---} \end{array} &= \begin{array}{c} \text{---} \text{---} \end{array}, \quad s = -1. \end{aligned} \quad (38)$$

Let us close this detailed description of our circuit model, with the observation that the exact same tensor network and statistical model also describes the phase transition in a model where the coherent rotation is followed by a subsequent *syndrome* measurement [38–40]. We will further discuss such connections with other models in our general discussion section at the end of the manuscript.

C. Boundary entanglement entropy

Let us now investigate the boundary entanglement entropy of the (1+1)D quantum circuit and map out the phase diagram shown in Fig. 2(c). Numerical results for the boundary entanglement entropy in the projective-measurement limit along (a) the self-dual line and (b) the XY line are shown in Figs. 7 and 8, respectively. To identify different contributions to their entanglement structure we fit the entanglement entropy to a formula distinguishing volume-law contributions for $v \neq 0$,

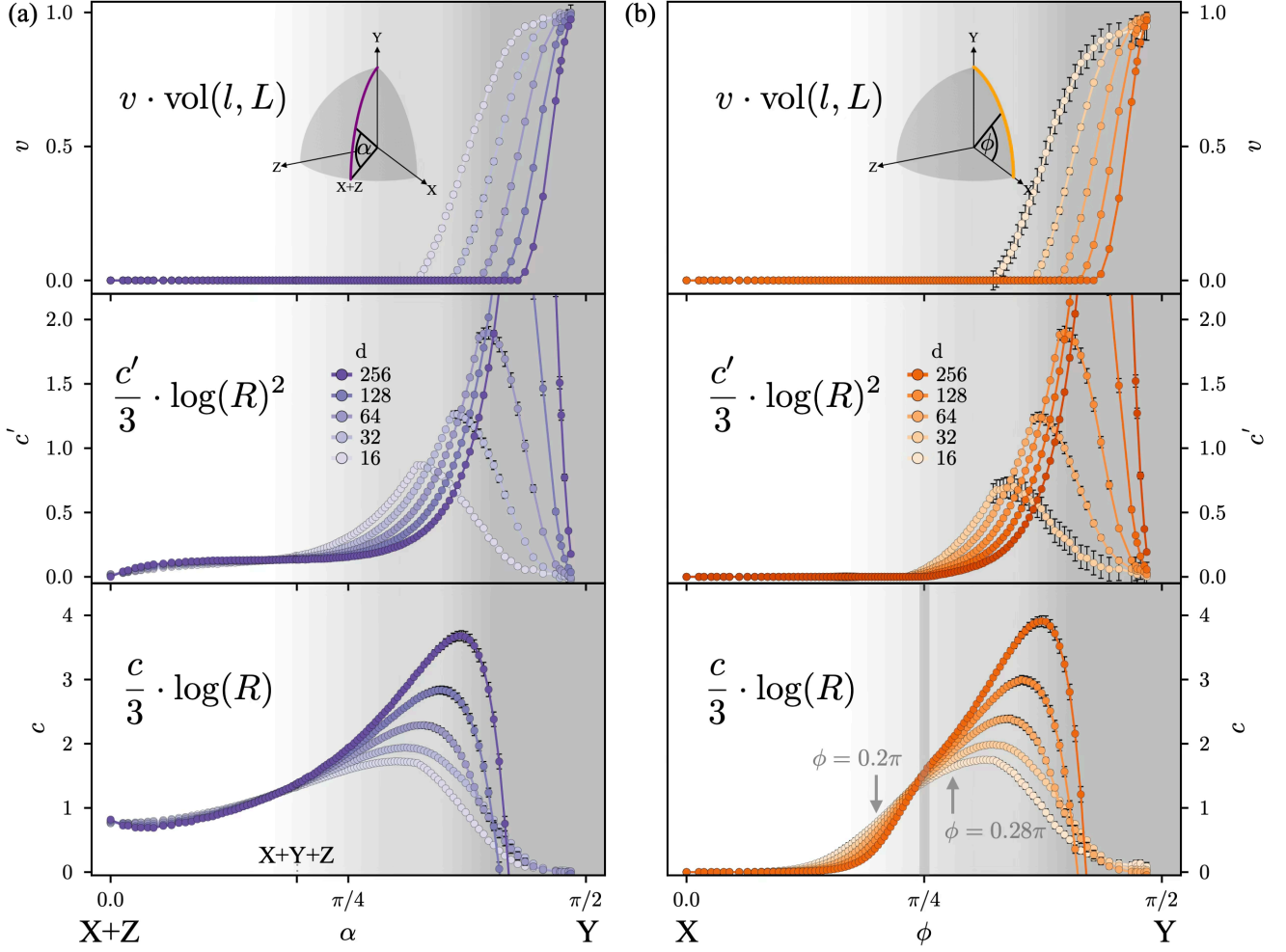


FIG. 8. **Boundary entanglement entropy for the projective-measurement limit on (a) the self-dual line and (b) the XY line.** The prefactors are the result of fitting the numerical data for Eq. (39). (a) On the self-dual line, the system always shows the critical behavior $c \cdot \log(R)$ in the entanglement entropy before reaching the volume-law entanglement regime in a finite system size. In addition, we also find a non-zero contribution from $c' \cdot \log(R)^2$ once we turn on Y-component for $\alpha > 0$. Both of these contributions are stable for small α and diverge before vanishing in the volume-law entanglement phase. There is a clear volume-law entanglement phase around $\alpha = \pi/2$. However, this regime shrinks with increasing system size. In the thermodynamic limit, the volume-law behavior only appears at the Y-point. The specialty of the Y-point is discussed in detail in Sec. IV B. For the XY line, the crossing point of c is taken as an indication of a metal-insulator transition around $\phi_c \approx 0.25\pi$. For comparison, we also indicate the critical value $\phi = 0.28\pi$ obtained by conductivity calculations in Ref. [39], and the phase boundary of the infinite-threshold phase $\phi_c \approx 0.2\pi$ from the small-size interacting-fermion coherent information calculations in Fig. 1(b).

from metallic $c' \neq 0$ and critical $c \neq 0$ contributions

$$S_{vN}(l, L) = v \cdot \text{vol}(l, L) + \frac{c'}{6} \cdot (\log R)^2 + \frac{c}{6} \cdot \log R + a, \quad (39)$$

where $R = \frac{L}{\pi} \sin(\pi \cdot \frac{l}{L})$ is the chord length of a length- l segment on a circle of circumference L , and $\text{vol}(l, L)$ is the Page curve of Gaussian fermions [75]. As discussed in Sec. II, the $\text{vol}(l, L)$ contribution stems from the ballistic spreading of entanglement under unitary evolutions. The $(\log R)^2$ term is known to arise in the Clifford-dynamics of monitored Majorana, which is a setting that can be mapped to the physics of completely packed loop models with crossings (CPLC) [57, 76–78]. This contribution is also found to be indispensable in our *non-Clifford* scenario of Gaussian fermions

(which cannot be captured by a statistical loop model). The remaining $\log R$ term in Eq. (39) is the standard universal logarithmic behavior associated with an underlying conformal criticality of the quantum circuit [79, 80].

As Fig. 7 and 8 show, the onset of Y component generally increases the entanglement entropy. For the self-dual line, small deviations from $X+Z$ critical point (which exhibits logarithmic entanglement scaling only) quickly turn on squared logarithmic scaling as the leading term. This readily indicates that the whole self-dual line is immersed in the gapless phase (Fig. 2(c) purple region). For the XY line, on the other hand, it takes a finite deviation $\phi > 0$ to leave the gapped phase transitioning into the gapless phase – where we use “gapped” and

“gapless” to refer to short-range entangled and long-range entangled phases, respectively. Locating the precise angle where this entanglement transition occurs, can be done in two complementary ways: From the entanglement-entropy scaling in the projective-measurement limit, we obtain a (rough) estimate of the transition point $\phi_c \approx 0.25\pi$, where it displays an almost perfect logarithmic scaling (as shown in the middle panel of Fig. 7(b)). This point also coincides with a crossing point in the finite-size scaling fit of c (for code distances up to $d = 256$), see the bottom panel in Fig. 8(b) above. In comparison, our earlier estimate of the boundary of the phase with infinite (measurement-strength) threshold was $\phi_c \approx 0.2\pi$ (see Fig. 4(b)), where we looked at codes up to distance $d = 16$. The same transition point has been estimated from conductivity calculations in Ref. [39] to be around $\phi_c \approx 0.28\pi$. The discrepancy here is likely due to the strong finite-size effects of the underlying metal-insulator-transition, which requires further investigation. Put together, the existing estimates indicate that the transition point should lie somewhere within the window $0.20\pi \lesssim \phi_c \lesssim 0.28\pi$ as we have indicated by the two arrows in the lower panel of Fig. 8(b). Coming close to the Y point, there appears to be an intermediate volume-law scaling regime, although the volume-law prefactor vanishes in the thermodynamic limit everywhere except the exact Y point. This is attributed to the formation of a ballistic 2D metal in the Y -limit as discussed above.

Stepping back and looking at our overall phase diagram, let us make the following general observations: The most striking feature of our phase diagram is an extended phase with long-range entanglement. Its origin can be best understood as a metallic phase that arises in the following way: From the original tensor network representation of our learning problem, we have derived a (1+1)D monitored free-fermion circuit (introduced in Sec. IV B above). Its non-unitary dynamics in (1+1)D can be mapped to the Hermitian ground state physics of a Born-disordered fermion system in (2+1)D spacetime [35, 58, 81]. In this mapping, the majority of our phase diagram, the phase with long-range entanglement charted out above, corresponds to a *Majorana metal* phase (generally in symmetry class DIII in the Altland-Zirnbauer classification) [54–56, 58], described by the $SO(N \rightarrow 1)$ nonlinear σ model. In contrast, the gapped phase with short-range entanglement corresponds to an insulator, which undergoes a topological transition [82] across the self-dual $X + Z$ point, which separates the insulating phases in the X and Z limits, respectively [35]. Note that the XZ line [35], XY line [39] and ZY line are special limits that fall into symmetry class D with broken time reversal symmetry (for each layer).

One should also draw a distinction of our phase diagram from the Clifford phase diagrams [77, 83] that is mapped to a classical statistical loop model [76]. In particular, the Clifford phase diagram describes a distinct problem to ours, where the bulk qubits of the surface code are subjected to random Clifford projective measurements in the X, Y, Z directions [83]. As a result, their resulting phase diagram and the universality class of the phase transitions are distinctive from ours.

V. NON-HERMITIAN DYNAMICS FROM POST-SELECTION

Post-selection can provide a powerful complement to the Born-averaged description of quantum dynamics with measurement that we have discussed so far. While Born averaging yields an unconditional, trace-preserving evolution that captures what happens in reality, it washes out structures that exist at the level of individual trajectories. In contrast, conditioning on measurement outcomes (i.e. performing post-selection), especially in the projective-measurement limit, turns the same measurement process into a selective probe that isolates specific sectors of the trajectory ensemble, making otherwise invisible dynamical patterns transparent. As a result, post-selection yields a deterministic, “clean” trajectory with greater analytic tractability than the Born-averaged evolution. In this sense, post-selection can act as a microscope to resolve the microstructure of monitored dynamics beyond typical (Born-averaged) runs.

A. Non-unitary Floquet operator

Building on the projective-measurement limit discussed in Sec. IV, we investigate the post-selection limit, where the dynamics are conditioned on a fixed sequence of measurement outcomes rather than Born-averaging over all possible measurement outcomes $\{s\}$ in our qubit measurements of the surface code. To be specific, we focus on the post-selected configuration where $s = +1$ for all measurement outcomes in the following discussion. In this case, the effective (1+1)D quantum circuit in Fig. 3(c) is simplified to a Floquet operator

$$U_F = \exp\left(\frac{J + i\phi}{2} \sum_j Z_j Z_{j+1}\right) \exp\left(\frac{J_d + i\phi_d}{2} \sum_j X_j\right), \quad (40)$$

which realizes a non-Hermitian Floquet Majorana / Ising model [84, 85], c.f. Eq. (26) and (27). In the Majorana representation, the Floquet operator becomes

$$U_F = \exp\left(\frac{J + i\phi}{2} \sum_j i\gamma_{2j}\gamma_{2j+1}\right) \exp\left(\frac{J_d + i\phi_d}{2} \sum_j i\gamma_{2j-1}\gamma_{2j}\right). \quad (41)$$

The phase diagram of the non-Hermitian Floquet Ising chain characterized by the boundary entanglement entropy of the steady state is shown in Fig. 9(a). Notice that when $|J + i\phi| = |J_d + i\phi_d|$, the model is (complex) Kramers-Wannier self-dual in the sense of Eq. (30) and can be fermionized as a Majorana chain with complex hopping constants and with no dimerization. In the post-selection limit, volume-law scaling only appears on the self-dual line except at the $X + Z$ point, at which the system is critical with log behavior in the entanglement entropy. Away from the self-dual line, the dark and light gray regimes represent long-range and short-range entangled area-law phases, respectively. By contrast, in the Born-averaged case, we have seen that finite code distances lead to a small, but finite regime of volume-law behavior in the entanglement

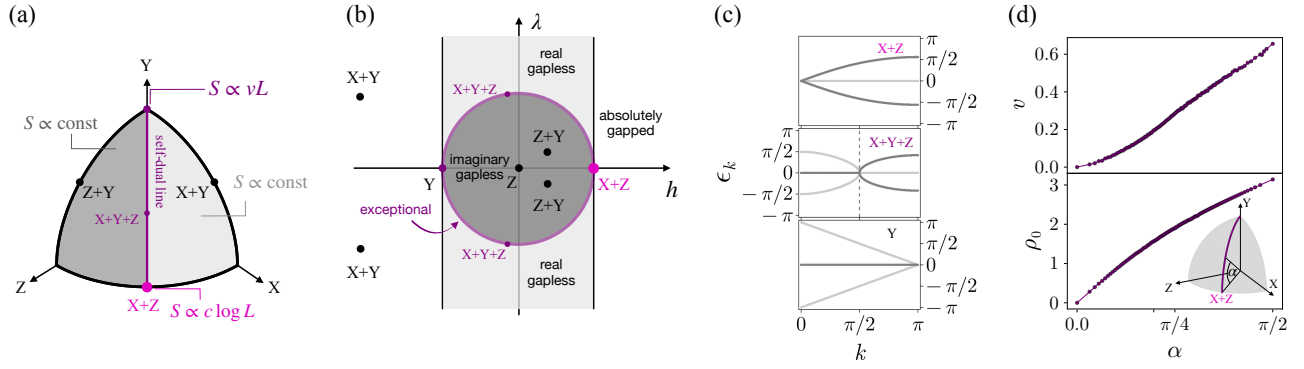


FIG. 9. **Non-Hermitian physics from post-selection.** (a) Phase diagram characterized by the boundary entanglement entropy for post-selection limit [c.f. Fig. 2 (b) in comparison with Born-average limit]. The purple color, for finite Y -component, indicates a *volume-law* behavior in boundary entanglement entropy which persists even in the thermodynamic limit, which is absent in the Born scenario. For the post-selection phase diagram, dark and light gray regimes are long-range and short-range entangled area-law phases in the qubit representation, respectively. In the Majorana fermion representation, they correspond to topological and trivial Majorana chains, respectively. (b) Phase diagram of the effective non-Hermitian Majorana chain Hamiltonian Eq. (42). In particular, the self-dual line in (a) maps to the unit circle in the non-Hermitian Majorana phase diagram, respectively. (c) Spectrum of the Floquet operator $U_F = \prod_k e^{-iH_k}$ along the self-dual line, where the Majorana fermion is described by the non-Hermitian Hamiltonian H_k , with k denoting the momentum. The dark and light gray lines represent the real and imaginary parts of the spectrum ϵ_k of H_k . (d) The volume-law prefactor v from boundary entanglement entropy [Eq. (39)] and the density of steady states ρ_0 from the spectrum of Floquet operator [Eq. (40)] along the self-dual line are correlated. The boundary entanglement entropy was calculated with system size $L = 128$ and the spectrum of the (1+1)D free fermion Floquet operator was calculated with system size $L = 10,000$.

entropy (which disappears only in the “thermodynamic limit” of large code distances). Notably, the post-selection and Born-average pictures are naturally connected in *replica field theory*. Specifically, the post-selection limit and Born-average correspond to the $n \rightarrow \infty$ and $n \rightarrow 1$ limits, respectively, where n is the replica number. As n increases, the volume-law regime shrinks and eventually turns into a single (self-dual) line as $n \rightarrow \infty$.

B. Effective non-Hermitian Hamiltonian and its phase diagram

Besides the boundary entanglement entropy, the spectrum of the Floquet operator U_F also reveals rich structures. In particular, the *effective Hamiltonian* of the Floquet operator U_F in Eq. (41) in the limit $|J + i\phi|, |J_d + i\phi_d| \ll 1$ can be obtained via the Baker-Hausdorff formula as

$$H_{\text{eff}} = - \sum_j i\gamma_{2j}\gamma_{2j+1} - (h + i\lambda) \sum_j i\gamma_{2j-1}\gamma_{2j}, \quad (42)$$

where $h + i\lambda = \frac{J_d + i\phi_d}{J + i\phi}$. This limit can be realized, e.g. by choosing distinct measurement angles for the qubits on the horizontal and vertical bonds of the surface code, by recalling $J(\theta) = \tanh^{-1} \cos \theta$ from Eq. (25) and Eq. (28). The phase diagram of H_{eff} is shown in Fig. 9(b). Specifically, the dark gray region inside the unit circle is gapless in the imaginary part but gapped in the real part of the spectrum, while the light gray region outside the unit circle is gapless in the real part but gapped in the imaginary part. By contrast, for large magnetic field, $h > 1$, the spectrum is absolutely gapped, i.e. gapped

in both real and imaginary parts. In the effective Hamiltonian H_{eff} , $|h + i\lambda| = 1$ is the *complex* self-duality condition, while $h = 1$ is the *real* self-duality condition. [See also the discussion below Eq. (30)]. Particularly, the self-dual line maps to the unit circle of *exceptional* critical points of H_{eff} ; the long-range and short-range area-law phases in Fig. 9(a) map to the inside and outside the unit circle, respectively. Notice that the lower half of (h, λ) -plane is redundant, as $\lambda \rightarrow -\lambda$ is equivalent to $(J, \phi) \rightarrow -(J, \phi)$, which does not change the phase. In particular, at the self-dual $X + Z$ point, $J = J_d, \phi = \phi_d = 0$, corresponding to the $(1, 0)$ point in the (h, λ) plane; at the Y point, $J = J_d = 0, \phi = -\phi_d = \pi/2$, corresponding to the $(0, 1)$ point; at the $X + Y + Z$ point, $J = J_d = \tanh^{-1}(1/\sqrt{3})$ and $\phi = -\phi_d = \pi/4$, which is also labeled on the unit circle in Fig. 9(b). Other special points off the self-dual line in Fig. 9(a) can be readily determined as well. Specifically, the X point is mapped to infinity on the (h, λ) plane; the Z point is mapped to the origin; the $X + Y$ point is mapped to the absolutely gapped region and the $Z + Y$ point is mapped inside the self-dual unit circle.

C. Quasi-energy spectrum and its connection to boundary entanglement entropy

The Floquet operator for the non-interacting Majorana fermions can be readily diagonalized in momentum space as $U_F = \prod_k e^{-iH_k}$, where H_k is the Floquet Hamiltonian with quasi-energies ϵ_k . (A detailed analysis of the Floquet operator is relegated to Appendix B.) Notably, zero modes of $\text{Im}(\epsilon_k)$ correspond to the steady states that contribute to entanglement

spreading. At the self-dual $X + Z$ point, the (1+1)D Floquet Majorana state exhibits a logarithmic entanglement entropy scaling with the code distance d , which is related to the fact that the Floquet spectrum has a steady state only at $k = 0$ as shown in Fig. 9(c); other imaginary modes correspond to the forced-measurement-only circuit that evolves the quantum state to the low-energy area-law corner of the Hilbert space. By contrast, the onset of a Y component ($\phi, \phi_d \neq 0$) turns the imaginary time into a complex-time evolution. As such, there exist states whose quasi-energies are purely real, [c.f. Fig. 9(d) at the $X + Y + Z$ and Y points]. As discussed before, these are the steady states that contribute to the spreading of entanglement entropy. Consequently, the late-time-evolved state has volume-law entanglement entropy behavior. At the Y point, $\text{Im}(\epsilon_k)$ always vanish. As such, the system has the largest density of steady states contributing the entanglement spreading. Notice that the real energy gap closes at the $k = \pi$ momentum point. At this momentum point, we have Majorana zero modes ($\epsilon_{k=\pi} = 0$); U_F is reduced to the exchange operators of these Majorana fermions, as discussed below Eq. (34).

To understand further the connection between the steady states from the spectrum and the volume-law behavior in the boundary entanglement entropy, we show the volume law scaling prefactor v and the density of steady states defined as $\rho_0 = \frac{1}{\pi} \sum_k \delta(\text{Im}(\epsilon_k))$ along the self-dual line in Fig. 9(d). Notably, the steady state density is nonzero *only* along the self-dual line, which agrees with the nonzero volume-law factor v obtained from the boundary entanglement entropy [c.f. Fig. 9(a)]. Notice that the increase of ρ_0 also qualitatively agrees with that of v . Intuitively, the steady states of the non-Hermitian dynamics do not decay under the dissipative dynamics. Consequently, their ballistic motion contributes to the entanglement spreading described by quasiparticle picture [86]. The quantitative deviation comes from the fact that the steady states do not contribute equal amount of entanglement entropy [87].

D. Non-Hermitian interacting fermion chain

For the finite-strength weak measurement, the post-selected effective Hamiltonian is a non-Hermitian but interacting fermion system. The partition function after post-selecting $s_{ij} = 1$ can be readily obtained from Eq. (7) as

$$P(\mathbf{s} = 1) = \langle \psi | M_{\mathbf{s}=1}^\dagger M_{\mathbf{s}=1} | \psi \rangle = \sum_{Z, \tilde{Z}} \exp(-E_{AT}(\mathbf{s} = 1)),$$

where

$$E_{AT,ij}(s_{ij} = 1) = -\frac{J}{2}(Z_i Z_j + \tilde{Z}_i \tilde{Z}_j) + \frac{i\phi}{2}(Z_i Z_j - \tilde{Z}_i \tilde{Z}_j) - K(Z_i Z_j \tilde{Z}_i \tilde{Z}_j - 1). \quad (43)$$

Under the quantum-classical statistical mechanics mapping, the (2+0)D classical AT model becomes a (1+1)D quantum

XXZ chain. Specifically,

$$P(\mathbf{s} = 1) \propto \exp\left(-\sum_j H_{2j,2j+1}\right) \cdot \exp\left(-\sum_j H_{2j-1,2j}\right), \quad (44)$$

where [20]

$$\begin{aligned} H_{2j,2j+1} &= -\frac{J}{2}(\sigma_{2j}^x \sigma_{2j+1}^x + \sigma_{2j}^y \sigma_{2j+1}^y) \\ &\quad + i\frac{\phi}{2}(\sigma_{2j}^x \sigma_{2j+1}^x - \sigma_{2j}^y \sigma_{2j+1}^y) - K(\sigma_{2j}^z \sigma_{2j+1}^z + 1), \\ H_{2j-1,2j} &= -\frac{J_d}{2}(\sigma_{2j-1}^x \sigma_{2j}^x + \sigma_{2j-1}^y \sigma_{2j}^y) \\ &\quad + i\frac{\phi_d}{2}(\sigma_{2j-1}^x \sigma_{2j}^x - \sigma_{2j-1}^y \sigma_{2j}^y) - K_d(\sigma_{2j-1}^z \sigma_{2j}^z + 1), \end{aligned} \quad (45)$$

where σ are the spin-1/2 operators of the effective quantum XXZ chain. Here K, J are related to the measurement strength t and coherent rotation angle θ by $\tanh(J) = \sin(2t) \cos(\theta)$ and $e^{-2K} = \sinh(J) \tan(\theta)$, and related to K_d, J_d via Eq. (29). The system has a $U(1)$ symmetry when either J or ϕ is zero; on the other hand, when both J and ϕ are nonzero, the $U(1)$ symmetry of the XXZ chain is broken to a Z_2 symmetry. This $U(1)$ to Z_2 symmetry breaking can be more transparently seen from the fermion representation of the XXZ chain. Under a Jordan-Wigner transformation, the XXZ chain can be mapped to a complex fermion chain c_j with non-Hermitian pairing terms and density-density interactions. Specifically, Eq. (45) becomes

$$\begin{aligned} H_{2j,2j+1} &= J(c_{2j}^\dagger c_{2j+1} + \text{h.c.}) + i\phi(c_{2j} c_{2j+1} + \text{h.c.}) \\ &\quad - 2K(2n_{2j} n_{2j+1} - n_{2j} - n_{2j+1} + 1), \\ H_{2j-1,2j} &= J_d(c_{2j-1}^\dagger c_{2j} + \text{h.c.}) + i\phi_d(c_{2j-1} c_{2j} + \text{h.c.}) \\ &\quad - 2K_d(2n_{2j-1} n_{2j} - n_{2j-1} - n_{2j} + 1), \end{aligned} \quad (46)$$

where $n_j = c_j^\dagger c_j$ is the fermion number operator. As is evident from the spin-1/2 representation in Eq. (45), the $U(1)$ symmetry (corresponding to the fermion number conservation) is broken to Z_2 symmetry (corresponding to the fermion number *parity* conservation) when both J and ϕ are nonzero. Notice that when $J = 0$ and $\phi \neq 0$, the system still has a $U(1)$ symmetry after a relabeling of particles and holes on a sublattice, e.g. $c_{2j} \rightarrow c_{2j}^\dagger$.

Let us conclude this section with some comments on the post-selection limit. Post-selected trajectories can qualitatively capture the Born-averaged behavior when the chosen trajectory is typical, i.e., it carries a probability weight that does not decay exponentially with system size or time. However, its entanglement dynamics can drastically differ from that of the Born-averaged case if the post-selection trajectory is atypical. Nevertheless, as we have shown here, the post-selection limit provides a useful lens to understanding the entanglement dynamics of monitored quantum circuits via the non-Hermitian Hamiltonian framework.

VI. DISCUSSION

Let us close our manuscript, by putting our results into a broader context. We start by discussing the connection of our learning phase diagram to studies of noisy channels, the ensuing decoding problems, and their computational complexity. We also draw a connection to other models discussed in the literature. This includes models studying the effect of syndrome measurements in the presence of coherent rotations, which we recast to learning problems and establish a one-to-one correspondence to our model for a limited set of parameters. We also comment on connections of our model to explorations of non-Hermitian, interacting fermion chains. We then briefly touch on how our learning phase diagram can be put to work in engineering thresholds for teleportation protocols that have to deal with imperfectly created Bell pairs. Finally, we compare the generic measurement-induced deformation of the toric code and its phase diagram studied in this manuscript to previously studied deformations, particularly Hamiltonian and wavefunction deformations, and describes similarities and distinctions in their respective phase diagram. We end with a hypothesis of how our results might generalize for arbitrary CSS codes.

Monitored dynamics, noisy channel, and decoding

If one discards the measurement record, e.g. by summing over the measurement outcomes, then our model turns into a non-Clifford channel $\otimes_j \mathcal{N}_j$ where each qubit is subjected to a dephasing in the $\sigma_j^{\theta,\phi}$ basis

$$\mathcal{N}_j(\cdot) = \cos^2(t)(\cdot) + \sin^2(t)\sigma_j^{\theta,\phi}(\cdot)\sigma_j^{\theta,\phi}. \quad (47)$$

In the projective-measurement limit $t = \pi/4$, the maximally dephased mixed state is purely “classic” and becomes equivalent to the ensemble of post-measurement states – the core object of this paper. It implies that the toric code subject to maximal dephasing in a generic basis (far away from X and Z axes) turns into a 2D metal state. The phase diagram for finite probability of dephasing, however, is not exactly solved yet and it remains an open question to be discussed in (our) future work. Relatedly, a 2-replica phase diagram of a noisy toric code where each qubit suffers from an i.i.d. single-qubit unitary rotation was studied in Ref. [88]. It is important to note that such a 2-replica problem is quite distinct from the Born measurement scenario with its correlated disorder structure and distinctive critical phenomena described by (non-unitary) conformal field theories [35].

For the noisy channel, one can recast the learning problem as a decoding challenge: If the quantum information is still coherently encoded in the dephasing mixed state, can one decode the state and correct to transform it back to a pristine surface code? A straight-forward approach, which however is not guaranteed to produce an optimal result, is to adopt a minimum-weight-perfect-matching (MWPM) decoder [2]. Based on a sequence of syndrome measurements it identifies

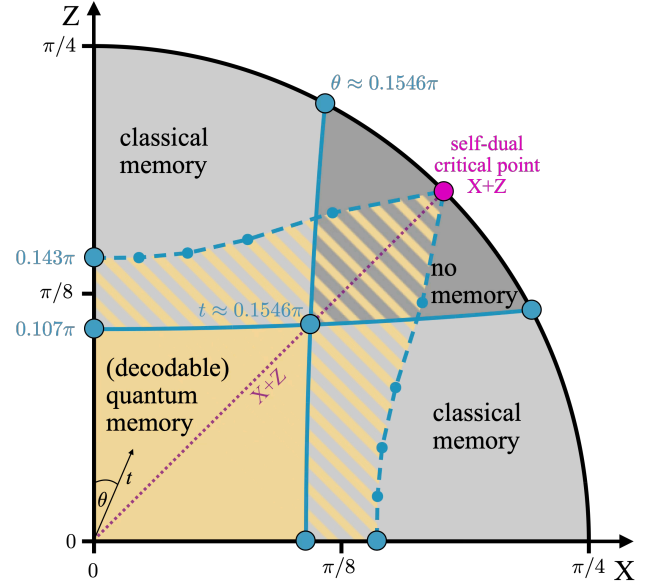


FIG. 10. **Phase diagram for decodability of the logical information of the surface code.** The dashed blue lines show the transitions for the optimal threshold, and the solid blue lines show the minimal weight perfect matching (MWPM) decoder: In the yellow region MWPM can retrieve the full logical information. In the light grey region, the surface code cannot be used as quantum memory anymore, but still serves as a classical memory using the respective encodings for the X and the Z direction. In the dark grey region, the surface code does not serve as either quantum or classical memory. The regime with yellow stripes between the solid and the dash blue lines mark the regime where the logical information generally is still decodable, but a MWPM decoder fails. All blue lines and points are characterized by Nishimori transitions. The self-dual critical point ($X + Z$) is captured by an unstable RG fixed point in a distinct universality class, which flows away from this unstable fixed point along the Nishimori points under RG flows.

Pauli string operators to pair up the e and m particles separately, so as to correct the surface code after the noise (47) back to the pristine surface code. Since one pairs up e and m separately, the phase diagram for such a decoder generally consists of four phase regions, see Fig. 10 for the XZ plane for example [89]. The solid blue lines correspond to the Nishimori decoding transition for either e or m string proliferation – which, however, sits *below* the coherent information threshold (dashed blue line) indicating a suboptimal decoding. This suboptimality is particularly striking when turning to the self-dual $X + Z$ direction (dotted purple line) where the threshold is slightly enhanced (compared to the decoding for X or Z noise only), but significantly smaller than the maximal threshold obtainable for an optimal decoder (as indicated by the coherent information threshold at $t = \pi/4$ along the self-dual line). Instead, for sufficiently large t , the logical quantum memory is completely lost (not even allowing for a classical memory phase). The reason is that by discarding the entire measurement record (of single-qubit measurements) and treating the e and m particles (syndromes) separately, one cannot capture any of their correlations which become partic-

ularly pronounced in this region of the phase diagram. This points to an interesting open question to develop a more efficient decoder that takes into account both syndrome and qubit measurements and capable enough to detect their underlying correlations.

Decoding complexity

An alternative approach to the decoding problem would be to employ a tensor network decoder. Provided the measurement outcome fed, e.g. from a quantum experiment, one can contract out the random tensor network in Fig. 3 to decode its bulk correlations and thereby the observed snapshot of the logical qubit. With the analytic insight that the tensor network for projective measurements (i.e. in the $t = \pi/4$ limit) can be free-fermionized, one can always contract it with *polynomial* complexity. However, for more general scenarios (where one does not have such insight), or if $t < \pi/4$ where the fermions become interacting, the contraction complexity generically depends on the bond dimension of the tensor network representation and thus upon the entanglement entropy of the boundary state [21, 90]. In particular, in the self-dual plane spanned by the $X + Z$ and Y bases, there are different categories of contraction complexity as summarized in Tab. I below. Overall, staying closer to the $X + Z$ direction implies lower complexity of tensor network contraction, due to the proximity with the “gapped insulating” phases in the X and Z neighborhoods. In contrast, while the Y limit is integrable, its entanglement entropy grows and a naive tensor network contraction is exponentially difficult.

angle	$X + Z$	$X + Z + \epsilon Y$	Y
S_{vN} scaling	$\log L$	$(\log L)^2$	L
MPS bond dimension D	$L^{\epsilon/6}$	$L^{\log L}$	e^L
contraction complexity	polynomial	superpoly	exponential

TABLE I. **Tensor network computational complexity for sampling the toric code in the self-dual directions.** The top row specifies the measurement direction. In the first row, S_{vN} indicates the boundary entanglement entropy, and D refers to the virtual bond dimension of the boundary MPS when contracting out the tensor network. While the whole self-dual plane yields infinite threshold, the decoding is not necessarily equally hard. This decoding complexity is indicated in the final row, in terms of the tensor network simulation of the toric code state measured in the given basis, which ranges from polynomially easy in the exact $X + Z$ direction, to superpolynomial when the direction deviates from $X + Z$ with finite Y component.

Connection to other models

Learning from syndrome measurements

While we have concentrated on learning the toric code via measurements of its physical qubits, we find that there is a closely related setting where, instead, one performs standard

syndrome measurements. At first sight, these situations might look entirely unrelated, since syndrome measurements are typically performed with the idea of learning the noise instead of the logical qubit in order to perform quantum error correction, i.e. to restore the logical state of the toric code without learning it (so as to preserve it). However, if one performs a coherent rotation on the physical qubits and then performs syndrome measurements, this situation can drastically alter – with syndrome measurements ultimately destroying the toric code state and instead allowing an observer to learn the logical state via the set of collected syndromes.

Specifically, we have been able to establish a one-to-one correspondence (of the bulk ensemble) between qubit and syndrome measurement for the case of projectively measuring the toric code in the $\cos \phi X + \sin \phi Y$ single-qubit basis (i.e. along the XY line in our phase diagram), which leads to exactly the same underlying statistical mechanics model with that of a toric code under coherent error $e^{i\phi Z/2}$ followed by a *syndrome* measurement, as illustrated in Fig. 11 above. This syndrome measurement setting has been studied in Refs. [38–40], while similar setups with general coherent errors are also studied in Refs. [41, 44]. This not only means they share the same phase transition and criticality, but they also lead to exactly corresponding Born samples, which allow one to learn about the logical information. In the previous works [39, 40], the phase transition induced by cranking up the rotation angle was interpreted as an error-induced breakdown of the logical qubit. However, distinct from the incoherent-noise induced transition [2] where information is completely lost into the environment, in the coherent error scenario the logical information is not all lost, but part of it can be reconstructed by the syndrome observer in the form of a classical shadow. This might sound contradictory to the usual notion that the syndrome space is designed to be orthogonal with the logical space, such that it only reveals the noise information without harming the logical space. However, the coherent error rotates the basis and overlaps the syndrome with the logical qubit. One can therefore expect that beyond a rotation angle threshold, the Born samples from the *rotated* syndrome measurement allow one to construct a tomographically complete classical shadow for the logical qubit – a scenario which we will address, in more detail, in future work.

Post-selected non-hermitian interacting fermion chain

Away from the projective-measurement limit, the post-selected (1+1)D quantum circuit is effectively described by a non-Hermitian *interacting* fermion chain, see Eq. (46) above. In this regime, the standard quasiparticle picture of entanglement spreading is no longer reliable, since interactions invalidate a simple free-quasiparticle description. Nevertheless, we believe that a robust volume-law phase persists: instead of being confined to the single self-dual line shown in Fig. 9(a), the region exhibiting volume-law scaling extends over a finite portion of the post-selected phase diagram. Figure 4(b) illustrates this behavior on the XY plane, where the Born-averaged entanglement entropy displays clear volume-

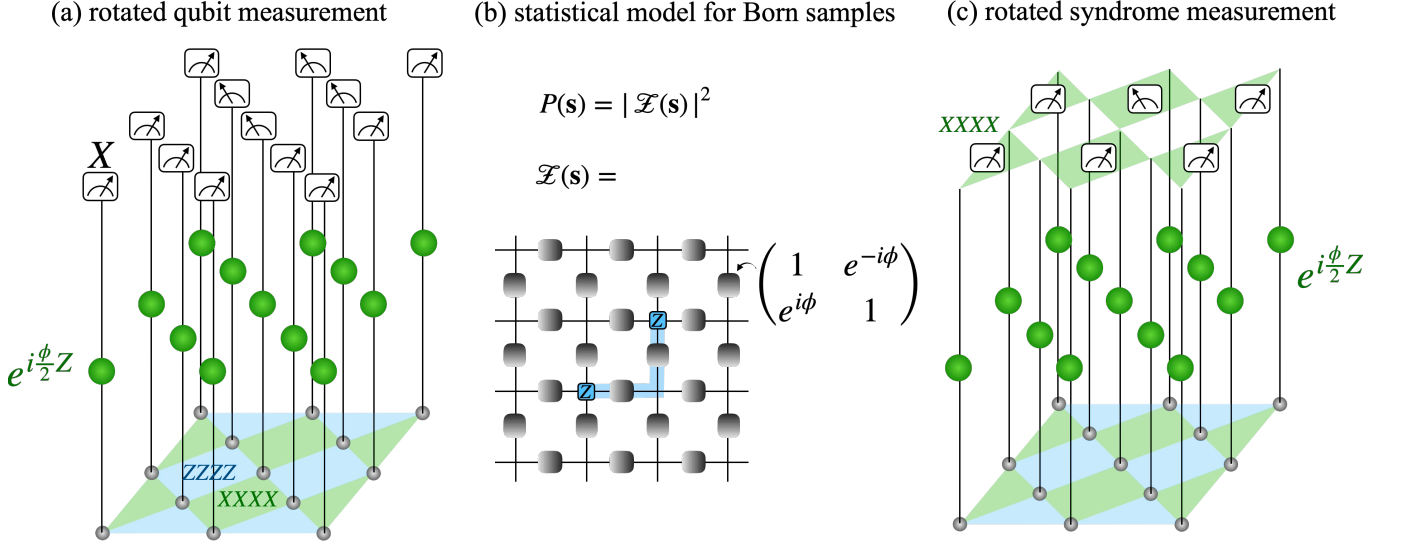


FIG. 11. **Qubit versus syndrome measurements.** The same classical statistical model (b) describes the Born sample of measuring either (a) the qubit or (c) the syndrome, of the surface code with a coherent Pauli Z rotation (for all qubits). The tensor network of the statistical model [c.f. Fig. 3(a)] is shown in (b), where the random disorder corresponds to injecting Pauli Z operator, or equivalently, a classical Ising variable, to the site where one observes an ϵ particle as $XXXX = -1$ syndrome in (c), or where the string of $X = -1$ measurement outcomes terminates in (a).

law scaling away from the projective-measurement limit in finite system sizes. For properly chosen post-selected measurement trajectories – for example, by conditioning on the most probable measurement configuration – we therefore expect this volume-law behavior to be captured quantitatively within the effective non-Hermitian interacting fermion chain description.

Teleportation transition threshold engineering

Our global phase diagram in Fig. 1 not only describes the toric code under weak measurement, but also captures the teleportation transition of a toric code with non-maximally entangled resources [20, 89]. In that setting, the parameter t tunes a coherent error in the entangling gate of Bell state preparation for a teleportation protocol that employs one entangled Bell pair to teleport the physical qubits of the toric code and thereby the full logical qubit state in a robust manner [20]. Recast in this teleportation protocol, our finding of a substantial infinite-threshold phase region for generic measurement angles, dictates that an optimal way to teleport the topological toric (or surface) code is to rotate the relative basis of any imperfectly entangled Bell pairs. Then for generic angle far from the X and Z limits it would suffice to teleport with even *arbitrarily weakly entangled pairs*.

Comparison to other toric code deformations

Let us close by comparing our measurement-induced deformation of the toric code with the deformations previously

studied in the literature, i.e. the Hamiltonian and wavefunction deformation of the toric code [11], and compare the respective phase diagrams. The most general Hamiltonian phase diagram [5–7], induced by applying a 3-component magnetic field, is a three-dimensional generalization of the long-known Fradkin-Shenker phase diagram of the Z_2 gauge theory [91], giving rise to a box-like phase region [7] in which the toric code remains robust. The phase transitions at these finite thresholds are either continuous (in the XZ plane) or first-order (in the presence of a Y -component). By contrast, the wave function deformation phase diagram – which corresponds to a post-selection scenario from the perspective of measurements – exhibits not only finite thresholds but also a critical line extending all the way out to infinity (i.e. the projective-measurement limit) along the self-dual direction within the XZ plane [13]. Adding a Y -component to the measurement basis, this critical point crosses over into a line (in the post-selected, projective-measurement limit) with volume-law entanglement entropy scaling in its 1D boundary state.

The Born measurement phase diagram, which we have explored in this manuscript, is drastically distinct from these previous cases. Instead of breaking down into trivial paramagnet or critical states at high symmetry points, the toric code and its topological quantum order are found to remain maximally stable against non-projective measurements along generic angles. This maximal stability of the topological quantum state has a nice interpretation in the context of the learning setup described in this manuscript, as it results from the maximal input that a dual classical observer requires in order to learn such a macroscopically entangled quantum state.

Finally, let us mention that there are two natural Born measurement variants of our model and phase diagram: (i) Given that qubit loss or leakage error are relatively common in to-

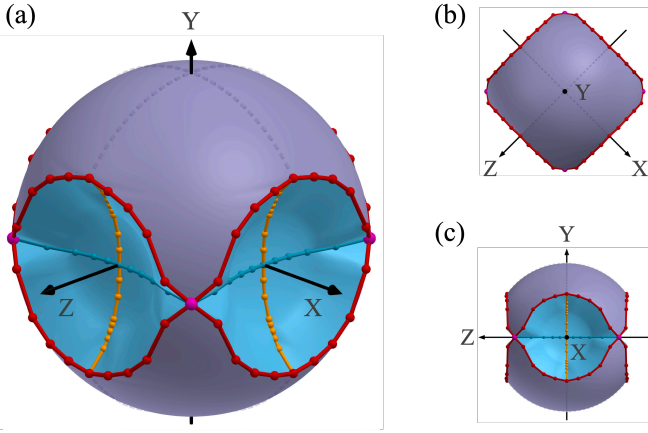


FIG. 12. **Three-dimensional phase diagram** of our model, viewed from different angles. The four blue regions carved out of the sphere correspond to the learnable phases, where the observer successfully learns the logical Z or X information (while collapsing the quantum loop condensate wave function to the corresponding classical loop gas ensemble at some finite measurement strength). For generic directions, the learnability threshold is pushed to the projective-measurement limit (indicated by the purple sphere), with the observer obtaining, in the absence of any readout errors, full tomographic information of the logical qubit from repeated measurements, without the need for randomized measurements. From the free-fermion perspective the purple sphere corresponds to a metallic phase in symmetry class DIII – with the exclusion of the two dashed “meridian” lines crossing in the Y point indicating symmetry class D, which, in addition, also applies to the “equator”, i.e. the projective-measurement limit in the XZ plane.

day’s NISQ platforms, one might wonder what happens to our model when replacing the measurement *strength* by a measurement *probability* [33, 92]. In such a modified protocol, the Nishimori learnability transition along X or Z measurement axes is replaced by a percolation transition at 50% probability [33, 92]. But for a coherent rotation to a generic measurement direction, when the problem is no longer describable by a Clifford circuit, the ensuing criticality and potential renormalization group flow remain an open question. (ii) In further modifying (and simplifying) our model, one could study a *Clifford version* of our spherical phase diagram, where also the coherent rotation is replaced by simply unbalancing the projective-measurement probabilities of X , Y and Z only. While the “surface” of such a phase diagram, i.e. $p(X) + p(Y) + p(Z) = 1$, was studied in Ref. [78], a global phase diagram similar to our three-dimensional one has not been explored to the best of our knowledge. We leave the exploration of these alternative models to future study.

Generalization to CSS code

Taking a step back and considering our learning phase diagram on the three-dimensional Bloch sphere as illustrated in Fig. 12, it becomes strikingly apparent that the learning threshold is *generically* infinite. That is, the quantum code

is typically stable in an absolute manner, because the learning observer (or an adversary eavesdropper) needs full and faithful (projective) measurements in order to read out the complex information encoded in the long-range entanglement of the toric code. The exceptions, however, are the corners surrounding the X and Z axes, where the finite thresholds carved out of the sphere, indicate that an incomplete (weak) measurement already allows for the readout of partial information in the form of the logical X or Z state, while losing the complementary information in the collapse of the quantum code. This qualitative feature might be generalized to generic CSS codes [93, 94], built from X and Z stabilizers. Namely, they could also be stable against generic non-Clifford measurements away from the X and Z corners. We leave this as an interesting subject for future study.

ACKNOWLEDGMENTS

ST and GYZ would like to thank Andreas Ludwig for illuminating discussions of the metal phase and symmetry class and for collaboration on related projects. Two of us (ST and GYZ) are grateful to the Kavli Institute for Theoretical Physics (KITP), which is supported by the National Science Foundation by grant NSF PHY-2309135, and the KITP Programs “Noise-robust Phases of Quantum Matter” and “Learning the Fine Structure of Quantum Dynamics in Programmable Quantum Matter,” where part of this work was carried out. The Cologne research group is supported, in part, by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany’s Excellence Strategy—Cluster of Excellence Matter and Light for Quantum Computing (ML4Q) EXC 2004/1 – 390534769 as well as within the CRC network TR 183 (Project Grant No. 277101999) as part of subproject B01. GYZ acknowledges the support of NSFC-Young Scientists Fund (No. 12504181) and Start-up Fund of HKUST(GZ) (No. G0101000221). Our numerical simulations were performed on the JUWELS cluster at the Forschungszentrum Juelich, the Noctua2 cluster at PC2 in Paderborn, and the RAMSES cluster at RRZK Cologne.

DATA AVAILABILITY

The numerical data shown in the figures and for the ensembles of the reference qubit are available on Zenodo [95]. Additionally, visualizations of the data, including an interactive version of the three-dimensional phase diagram of Fig. 12, are provided there in the programming language Julia.

Appendix A: KL divergence and numerical discretization

As given in the main manuscript the Kullback-Leibler divergence is defined as

$$D(P||Q) = \int P(\theta, \phi) \log \frac{P(\theta, \phi)}{Q(\theta, \phi)} d\theta d\phi, \quad (A1)$$

where $P(\theta, \phi)$ is the probability distribution of the measurement outcomes on the Bloch sphere, and $Q(\theta, \phi)$ is the uniform distribution on the Bloch sphere. For numerical calculation of the probability distribution $P(\theta, \phi)$ from the points given by the calculated samples, we discretize the unit sphere into N patches with (approximately) equal area $\Delta\Omega = 4\pi/N$. In this discretization the equal distribution over the sphere has the density $Q(\theta, \phi) = 1/N$ in each patch. The probability distribution $P(\theta, \phi)$ is estimated by counting the number of samples in each patch divided by the total number of samples and the area of the patch. With this discretization the Kullback-Leibler divergence can be approximated as

$$D(P||Q) \approx \sum_{i=1}^N P_i \log \frac{P_i}{Q_i} = \sum_{i=1}^N P_i \log(P_i N) \quad (A2)$$

where P_i is the estimated probability in patch i .

If the probability distribution $P(\theta, \phi)$ is uniform over the sphere, the estimated probability in each patch is $P_i = 1/N$ and the Kullback-Leibler divergence is zero. If the probability distribution $P(\theta, \phi)$ is a delta function at one point on the sphere, all samples fall into one patch and the estimated probability in that patch is $P_i = 1$ and zero in all other patches. In this case the Kullback-Leibler divergence sums to $\log(N)$. With a bimodal distribution with two delta functions at opposite points on the sphere, the Kullback-Leibler divergence sums correspondingly to $\log(N/2)$ and if the distribution is over eight points in different patches to $\log(N/8)$.

The plots shown in this manuscript are calculated with $N = 192$. Thus the maximal Kullback-Leibler divergence is $\log(192) \approx 5.26$. The bimodal distribution has a Kullback-Leibler divergence of $\log(96) \approx 4.56$ and the distribution over eight points has a Kullback-Leibler divergence of $\log(24) \approx 3.18$. The discretization into $N = 192$ patches is a compromise between the resolution of the probability distribution and the number of samples needed to estimate the probability distribution in each patch, since for finite samples even the perfect uniform distribution will show some deviation from a Kullback-Leibler divergence of zero. This deviation can be estimated by calculating the Kullback-Leibler divergence for a perfect uniform distribution with finite samples. The accuracy of the calculation is only dependent on the number of samples over the number of bins: $\frac{\text{samples}}{N}$. This is shown in Fig. 13. For the calculations shown in this manuscript we used 10^4 samples which gives a deviation of the Kullback-Leibler divergence from zero of about 0.01 (currently not yet all points converged to the full number of samples, so slightly higher deviations of same order of magnitude possible).

When the Kullback-Leibler divergence is normalized to the value of the bimodal distribution $\log(N/2)$, the value one

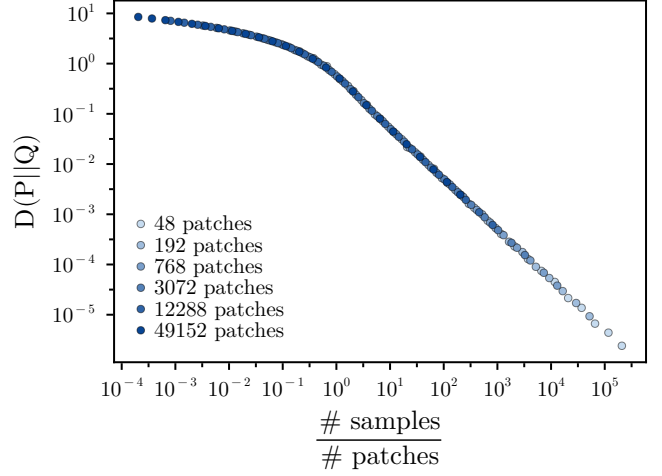


FIG. 13. **Kullback-Leibler divergence** of the uniform distribution with finite samples sampled according to the uniform distribution in a finite resolution. Plotted is the Kullback-Leibler divergence of the perfect uniform distribution with different resolutions of the HEALPix scheme (Hierarchical Equal Area isoLatitude Pixelation of a sphere) with $N = 12 \cdot 4^n$ patches of equal size, where n is the order of the discretization and N the number of patches in the respective discretization. Due to the finite number of samples even the Kullback-Leibler divergence of points sampled according to the uniform distribution takes small finite values. Since all discretizations follow the same curve, the Kullback-Leibler divergence of the perfect uniform distribution only depends on the ratio of number of samples over number of patches. In the limit of infinite samples the Kullback-Leibler divergence always goes to zero as the points are sampled by the distribution with which they are compared.

corresponds to a bimodal distribution and the value zero to a uniform distribution. This normalized Kullback-Leibler divergence is used in Fig. 3 of the main manuscript. In such a normalization the limit of the ring distribution in the X-Z plane corresponds to exactly $\frac{1}{2}$. Both bimodal and ring distribution diverge to infinity in that limit, since these distributions are clearly distinct and thus far away from the uniform distribution over the whole sphere. That the ratio between these distributions becomes exactly $\frac{1}{2}$ can be understood how we discretize into patches and how this number of patches scales with each step. In the resolution used for the calculations here, there are 192 patches of (in very good approximation) equal area. The bimodal distribution concentrates on exactly two of these patches, while the ring distribution covers the 18 patches around the X-Z plane. Thus the expected value here would be $\frac{\log(192/18)}{\log(192/2)} \approx 0.51$, which is very close to the numerically observed value of 0.5. Increasing the number of patches N will increase the number of patches by a factor of four in every step, while the number of patches covered by the ring distribution increases by a factor of two. If n is the number of additional steps in discretization, we obtain

$$K(\text{bimodal}) = \log\left(\frac{192}{2} \cdot 4^n\right), \quad (A3)$$

for the bimodal distribution and

$$K(\text{ring}) = \log\left(\frac{192 \cdot 4^n}{18 \cdot 2^n}\right) = \log\left(\frac{192}{18} \cdot 2^n\right), \quad (\text{A4})$$

for the ring distribution. That means for the ratio in the limit following:

$$\begin{aligned} \tilde{K} &= \lim_{n \rightarrow \infty} \frac{\log\left(\frac{192}{18} \cdot 2^n\right)}{\log\left(\frac{192}{2} \cdot 4^n\right)} \\ &= \lim_{n \rightarrow \infty} \frac{\log(2^n)}{\log(4^n)} = \frac{\log(2)}{\log(4)} = \frac{1}{2}. \end{aligned} \quad (\text{A5})$$

Appendix B: Detailed Analysis of non-Hermitian dynamics in the post-selection limit

In this appendix, we provide more details on the spectrum of the post-selected Floquet operator in Eq. (40) following Ref. [87]. The Floquet operator in Eq. (40) can be written in a more symmetric form as

$$U_F = \exp\left(\frac{J_d + i\phi_d}{4} \sum_j X_j e^{\frac{J_d + i\phi}{2} \sum_j Z_j Z_{j+1}} e^{\frac{J_d + i\phi_d}{4} \sum_j X_j}\right). \quad (\text{B1})$$

Define $H_1 = \sum_j Z_j Z_{j+1}$ and $H_0 = \sum_j X_j$. Under the Jordan-Wigner transformation

$$X_j = 1 - 2c_j^\dagger c_j, \quad c_j^+ = \left(\prod_{\ell=1}^{j-1} X_\ell\right) \sigma_j^+, \quad c_j^- = \left(\prod_{\ell=1}^{j-1} X_\ell\right) \sigma_j^-, \quad (\text{B2})$$

where $\sigma_j^\pm = \frac{1}{2}(Z_j \pm iY_j)$, we have

$$Z_j = \sigma_j^+ + \sigma_j^- = \left[\prod_{\ell=1}^{j-1} (1 - 2c_\ell^\dagger c_\ell)\right] (c_j + c_j^\dagger). \quad (\text{B3})$$

As a result,

$$H_1 = \sum_{j=1}^L (c_j^\dagger - c_j)(c_{j+1}^\dagger + c_{j+1}), \quad (\text{B4})$$

where we impose the anti-periodic boundary condition $c_{L+1} = -c_1$ for the even fermion parity sector $e^{i\pi N} = e^{i\pi \sum_j c_j^\dagger c_j} = \prod_{j=1}^L (1 - 2c_j^\dagger c_j) = 1$ and the periodic boundary condition $c_{L+1} = c_1$ for the odd fermion parity sector $e^{i\pi N} = -1$. Subsequently, Fourier transform of the complex fermion operator $c_j = \frac{1}{\sqrt{L}} \sum_k e^{ikj} c_k$ leads to

$$H_1 = 2 \sum_{k>0} \begin{pmatrix} c_k^\dagger & c_{-k} \end{pmatrix} \begin{pmatrix} \cos k & i \sin k \\ -i \sin k & -\cos k \end{pmatrix} \begin{pmatrix} c_k \\ c_{-k}^\dagger \end{pmatrix}. \quad (\text{B5})$$

Similarly, we have

$$\begin{aligned} H_0 &= \sum_j X_j = \sum_j (1 - 2c_j^\dagger c_j) \\ &\sim -2 \sum_{k>0} (c_k^\dagger c_k + c_{-k}^\dagger c_{-k}) \\ &= -2 \sum_{k>0} \begin{pmatrix} c_k^\dagger & c_{-k} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} c_k \\ c_{-k}^\dagger \end{pmatrix}, \end{aligned} \quad (\text{B6})$$

up to a complex constant. As a result, the Floquet operator can be written in the momentum space as

$$U_F \propto \prod_{k>0} e^{-\frac{J_d + i\phi_d}{2} \sigma_z} \cdot e^{(J_d + i\phi) \sigma_k} \cdot e^{-\frac{J_d + i\phi_d}{2} \sigma_z}, \quad (\text{B7})$$

where $\sigma_k = \sigma_z \cos k - \sigma_y \sin k$. In particular, at Y point of the phase diagram, we have $J = J_d = 0, \phi = -\phi_d = \pi/2$. Consequently, the Floquet operator reduces to

$$\begin{aligned} U_F^Y &= \exp\left(\frac{i\pi}{4} \sum_j Z_j Z_{j+1} e^{-\frac{i\pi}{4} \sum_j X_j}\right) \\ &= \prod_{k>0} e^{i\frac{\pi}{2} \sigma_k} e^{i\frac{\pi}{2} \sigma_z} = \prod_{k>0} (-e^{-ik\sigma_x}), \end{aligned} \quad (\text{B8})$$

which is a *unitary* operator.

-
- [1] X. Chen, Z.-C. Gu, and X.-G. Wen, Local unitary transformation, long-range quantum entanglement, wave function renormalization, and topological order, *Phys. Rev. B* **82**, 155138 (2010).
 - [2] E. Dennis, A. Kitaev, A. Landahl, and J. Preskill, Topological quantum memory, *Journal of Mathematical Physics* **43**, 4452 (2002).
 - [3] A. Y. Kitaev, Fault-tolerant quantum computation by anyons, *Annals of Physics* **303**, 2 (2003).
 - [4] S. B. Bravyi and A. Y. Kitaev, Quantum codes on a lattice with boundary, arXiv e-prints [quant-ph/9811052](https://arxiv.org/abs/quant-ph/9811052).
 - [5] S. Trebst, P. Werner, M. Troyer, K. Shtengel, and C. Nayak, Breakdown of a Topological Phase: Quantum Phase Transition in a Loop Gas Model with Tension, *Phys. Rev. Lett.* **98**, 070602 (2007).
 - [6] I. S. Tupitsyn, A. Kitaev, N. V. Prokof'ev, and P. C. E. Stamp, Topological multicritical point in the phase diagram of the toric code model and three-dimensional lattice gauge Higgs model, *Phys. Rev. B* **82**, 085114 (2010).
 - [7] S. Dusuel, M. Kamfor, R. Orús, K. P. Schmidt, and J. Vidal, Robustness of a Perturbed Topological Phase, *Phys. Rev. Lett.* **106**, 107203 (2011).
 - [8] E. Ardonne, P. Fendley, and E. Fradkin, Topological order and conformal quantum critical points, *Annals of Physics* **310**, 493 (2004).
 - [9] C. L. Henley, From classical to quantum dynamics at Rokhsar-Kivelson points, *Journal of Physics: Condensed Matter* **16**, S891 (2004).
 - [10] C. Castelnovo and C. Chamon, Quantum topological phase transition at the microscopic level, *Phys. Rev. B* **77**, 054433 (2008).

- (2008).
- [11] C. Castelnovo, S. Trebst, and M. Troyer, Topological Order and Quantum Criticality, in [Understanding Quantum Phase Transitions](#) (Taylor & Francis, 2010) pp. 169–192, [arXiv:0912.3272](#).
 - [12] J. Haegeman, V. Zauner, N. Schuch, and F. Verstraete, Shadows of anyons and the entanglement structure of topological phases, [Nature Communications](#) **6**, 8284 (2015).
 - [13] G.-Y. Zhu and G.-M. Zhang, Gapless Coulomb State Emerging from a Self-Dual Topological Tensor-Network State, [Phys. Rev. Lett.](#) **122**, 176401 (2019).
 - [14] R. Fan, Y. Bao, E. Altman, and A. Vishwanath, Diagnostics of Mixed-State Topological Order and Breakdown of Quantum Memory, [PRX Quantum](#) **5**, 020343 (2024).
 - [15] J. Y. Lee, C.-M. Jian, and C. Xu, Quantum Criticality Under Decoherence or Weak Measurement, [PRX Quantum](#) **4**, 030317 (2023).
 - [16] R. Ma, J.-H. Zhang, Z. Bi, M. Cheng, and C. Wang, Topological Phases with Average Symmetries: The Decohered, the Disordered, and the Intrinsic, [Phys. Rev. X](#) **15**, 021062 (2025).
 - [17] P. Sala, S. Gopalakrishnan, M. Oshikawa, and Y. You, Spontaneous strong symmetry breaking in open systems: Purification perspective, [Phys. Rev. B](#) **110**, 155150 (2024).
 - [18] C. Zhang, Y. Xu, J.-H. Zhang, C. Xu, Z. Bi, and Z.-X. Luo, Strong-to-weak spontaneous breaking of 1-form symmetry and intrinsically mixed topological order, [Phys. Rev. B](#) **111**, 115137 (2025).
 - [19] L. A. Lessa, R. Ma, J.-H. Zhang, Z. Bi, M. Cheng, and C. Wang, Strong-to-Weak Spontaneous Symmetry Breaking in Mixed Quantum States, [PRX Quantum](#) **6**, 010344 (2025).
 - [20] F. Eckstein, B. Han, S. Trebst, and G.-Y. Zhu, Robust Teleportation of a Surface Code and Cascade of Topological Quantum Phase Transitions, [PRX Quantum](#) **5**, 040313 (2024).
 - [21] G.-Y. Zhu, N. Tantivasadakarn, A. Vishwanath, S. Trebst, and R. Verresen, Nishimori’s Cat: Stable Long-Range Entanglement from Finite-Depth Unitaries and Weak Measurements, [Phys. Rev. Lett.](#) **131**, 200201 (2023).
 - [22] L. Zdeborová and F. Krzakala, Statistical physics of inference: thresholds and algorithms, [Adv. Phys.](#) **65**, 453 (2016).
 - [23] M. Pütz, S. J. Garratt, H. Nishimori, S. Trebst, and G.-Y. Zhu, Learning transitions in classical Ising models and deformed toric codes, [arXiv e-prints](#) (), [2504.12385](#).
 - [24] F. Barratt, U. Agrawal, A. C. Potter, S. Gopalakrishnan, and R. Vasseur, Transitions in the Learnability of Global Charges from Local Measurements, [Phys. Rev. Lett.](#) **129**, 200602 (2022).
 - [25] Y. Bao, S. Choi, and E. Altman, Theory of the phase transition in random unitary circuits with measurements, [Phys. Rev. B](#) **101**, 104301 (2020).
 - [26] U. Agrawal, J. Lopez-Piqueres, R. Vasseur, S. Gopalakrishnan, and A. C. Potter, Observing Quantum Measurement Collapse as a Learnability Phase Transition, [Phys. Rev. X](#) **14**, 041012 (2024).
 - [27] M. Ippoliti and V. Khemani, Learnability Transitions in Monitored Quantum Dynamics via Eavesdropper’s Classical Shadows, [PRX Quantum](#) **5**, 020304 (2024).
 - [28] A. A. Akhtar, H.-Y. Hu, and Y.-Z. You, Measurement-induced criticality is tomographically optimal, [Phys. Rev. B](#) **109**, 094209 (2024).
 - [29] H. Singh, R. Vasseur, A. C. Potter, and S. Gopalakrishnan, Mixed-state learnability transitions in monitored noisy quantum dynamics, preprint (2025), [arXiv:2503.10308](#).
 - [30] A. Nahum and J. L. Jacobsen, Bayesian critical points in classical lattice models, [Phys. Rev. B](#) **112**, 235113 (2025).
 - [31] S. Gopalakrishnan, E. McCulloch, and R. Vasseur, Monitored Fluctuating Hydrodynamics, [arXiv e-prints](#) [2504.02734](#).
 - [32] S. W. P. Kim, C. von Keyserlingk, and A. Lamacraft, Measurement-induced phase transitions in quantum inference problems and quantum hidden Markov models, [arXiv e-prints](#) [2504.08888](#).
 - [33] M. Pütz, R. Vasseur, A. W. W. Ludwig, S. Trebst, and G.-Y. Zhu, Flow to Nishimori universality in weakly monitored quantum circuits with qubit loss, [arXiv e-prints](#) (), [2505.22720](#).
 - [34] Z.-Q. Wan, X.-D. Dai, and G.-Y. Zhu, Revisiting Nishimori multicriticality through the lens of information measures, [arXiv e-prints](#) [2511.02907](#).
 - [35] Q. Wang, R. Vasseur, S. Trebst, A. W. W. Ludwig, and G.-Y. Zhu, Decoherence-induced self-dual criticality in topological states of matter, preprint (2025), [arXiv:2502.14034](#).
 - [36] H. Nishimori, Internal Energy, Specific Heat and Correlation Function of the Bond-Random Ising Model, [Progress of Theoretical Physics](#) **66**, 1169 (1981).
 - [37] H. Nishimori, [Statistical Physics of Spin Glasses and Information Processing](#) (Oxford University Press, 2001).
 - [38] S. Bravyi, M. Englbrecht, R. König, and N. Peard, Correcting coherent errors with surface codes, [npj Quantum Information](#) **4**, 55 (2018).
 - [39] F. Venn, J. Behrends, and B. Béri, Coherent-Error Threshold for Surface Codes from Majorana Delocalization, [Phys. Rev. Lett.](#) **131**, 060603 (2023).
 - [40] J. Behrends, F. Venn, and B. Béri, Surface codes, quantum circuits, and entanglement phases, [Phys. Rev. Res.](#) **6**, 013137 (2024).
 - [41] J. K. Iverson and J. Preskill, Coherence in logical quantum channels, [New Journal of Physics](#) **22**, 073066 (2020).
 - [42] J. Behrends and B. Béri, The Surface Code beyond Pauli Channels: Logical Noise Coherence, Information-Theoretic Measures, and Errorfield-Double Phenomenology, [PRX Quantum](#) **6**, 040350 (2025).
 - [43] J. Behrends and B. Béri, Statistical Mechanical Mapping and Maximum-Likelihood Thresholds for the Surface Code under Generic Single-Qubit Coherent Errors, [PRX Quantum](#) **6**, 040305 (2025).
 - [44] H. Liu and X. Chen, Coherent error induced phase transition, [arXiv e-prints](#) [2506.00650](#).
 - [45] E. H. Chen, G.-Y. Zhu, R. Verresen, A. Seif, E. Bäumer, D. Layden, N. Tantivasadakarn, G. Zhu, S. Sheldon, A. Vishwanath, S. Trebst, and A. Kandala, Nishimori transition across the error threshold for constant-depth quantum circuits, [Nature Physics](#) **21**, 161 (2025).
 - [46] M. A. Levin and X.-G. Wen, String-net condensation: A physical mechanism for topological phases, [Phys. Rev. B](#) **71**, 045110 (2005).
 - [47] X.-G. Wen, Choreographed entanglement dances: Topological states of quantum matter, [Science](#) **363**, eaal3099 (2019).
 - [48] H.-Y. Huang, R. Kueng, and J. Preskill, Predicting many properties of a quantum system from very few measurements, [Nature Physics](#) **16**, 1050 (2020).
 - [49] B. Schumacher and M. A. Nielsen, Quantum data processing and error correction, [Phys. Rev. A](#) **54**, 2629 (1996).
 - [50] S. Lloyd, Capacity of the noisy quantum channel, [Phys. Rev. A](#) **55**, 1613 (1997).
 - [51] M. Ippoliti and W. W. Ho, Dynamical Purification and the Emergence of Quantum State Designs from the Projected Ensemble, [PRX Quantum](#) **4**, 030322 (2023).
 - [52] A. Elben, S. T. Flammia, H.-Y. Huang, R. Kueng, J. Preskill, B. Vermersch, and P. Zoller, The randomized measurement toolbox, [Nature Reviews Physics](#) **5**, 9 (2023).

- [53] C. Liu, M. Ippoliti, and W. W. Ho, Coherence-induced deep thermalization transition in random permutation quantum dynamics, arXiv e-prints [2510.18369](#).
- [54] C.-M. Jian, B. Bauer, A. Keselman, and A. W. W. Ludwig, Criticality and entanglement in nonunitary quantum circuits and tensor networks of noninteracting fermions, *Phys. Rev. B* **106**, 134206 (2022).
- [55] M. Fava, L. Piroli, T. Swann, D. Bernard, and A. Nahum, Nonlinear Sigma Models for Monitored Dynamics of Free Fermions, *Phys. Rev. X* **13**, 041045 (2023).
- [56] J. Merritt and L. Fidkowski, Entanglement transitions with free fermions, *Phys. Rev. B* **107**, 064303 (2023).
- [57] K. Klocke and M. Buchhold, Majorana Loop Models for Measurement-Only Quantum Circuits, *Phys. Rev. X* **13**, 041028 (2023).
- [58] C.-M. Jian, H. Shapourian, B. Bauer, and A. W. W. Ludwig, Measurement-induced entanglement transitions in quantum circuits of non-interacting fermions: Born-rule versus forced measurements, arXiv e-prints [2302.09094](#).
- [59] The duality here refers to exchanging space and time dimension, which should be distinguished from Kramers Wannier duality [35] for the whole line between Y and $X + Z$ discussed in this paper.
- [60] L. Piroli, B. Bertini, J. I. Cirac, and T. c. v. Prosen, Exact dynamics in dual-unitary quantum circuits, *Phys. Rev. B* **101**, 094304 (2020).
- [61] F. Merz and J. T. Chalker, Two-dimensional random-bond Ising model, free fermions, and the network model, *Phys. Rev. B* **65**, 054425 (2002).
- [62] D. Gioev and I. Klich, Entanglement Entropy of Fermions in Any Dimension and the Widom Conjecture, *Phys. Rev. Lett.* **96**, 100503 (2006).
- [63] M. M. Wolf, Violation of the Entropic Area Law for Fermions, *Phys. Rev. Lett.* **96**, 010404 (2006).
- [64] Z.-C. Gu, M. Levin, B. Swingle, and X.-G. Wen, Tensor-product representations for string-net condensed states, *Phys. Rev. B* **79**, 085118 (2009).
- [65] O. Buerschaper, M. Aguado, and G. Vidal, Explicit tensor network representation for the ground states of string-net models, *Phys. Rev. B* **79**, 085119 (2009).
- [66] N. Bultinck, M. Mariën, D. Williamson, M. Şahinoğlu, J. Haegeman, and F. Verstraete, Anyons and matrix product operator algebras, *Annals of Physics* **378**, 183 (2017).
- [67] J. Ashkin and E. Teller, Statistics of Two-Dimensional Lattices with Four Components, *Phys. Rev.* **64**, 178 (1943).
- [68] M. J. Gullans and D. A. Huse, Scalable Probes of Measurement-Induced Criticality, *Phys. Rev. Lett.* **125**, 070606 (2020).
- [69] Z.-M. Huang, L. Colmenarez, M. Müller, and S. Diehl, Coherent information as a mixed-state topological order parameter of fermions, *Phys. Rev. Res.* **7**, 043009 (2025).
- [70] L. Colmenarez, Z.-M. Huang, S. Diehl, and M. Müller, Accurate optimal quantum error correction thresholds from coherent information, *Phys. Rev. Res.* **6**, L042014 (2024).
- [71] Y.-J. Liu, W.-T. Xu, F. Pollmann, and M. Knap, Detecting emergent 1-form symmetries with quantum error correction, arXiv e-prints [arXiv:2502.17572](#) (2025).
- [72] M.-D. Choi, Completely positive linear maps on complex matrices, *Linear Algebra Appl.* **10**, 285 (1975).
- [73] This decoder is the inverse of the encoder that maps a logical qubit to a surface code state, so as to map a physical bitstring to a logical snapshot. It should not be confused as the usual decoder in quantum error correction that learns the logical error from syndrome.
- [74] D. Bluvstein, S. J. Evered, A. A. Geim, S. H. Li, H. Zhou, T. Manovitz, S. Ebadi, M. Cain, M. Kalinowski, D. Hangleiter, J. P. Bonilla Ataides, N. Maskara, I. Cong, X. Gao, P. Sales Rodriguez, T. Karolyshyn, G. Semeghini, M. J. Gullans, M. Greiner, V. Vuletić, and M. D. Lukin, Logical quantum processor based on reconfigurable atom arrays, *Nature* **626**, 58 (2024).
- [75] E. Bianchi, L. Hackl, and M. Kieburg, Page curve for fermionic Gaussian states, *Phys. Rev. B* **103**, L241118 (2021).
- [76] A. Nahum, P. Serna, A. M. Somoza, and M. Ortuño, Loop models with crossings, *Phys. Rev. B* **87**, 184204 (2013).
- [77] S. Sang, Y. Li, T. Zhou, X. Chen, T. H. Hsieh, and M. P. Fisher, Entanglement Negativity at Measurement-Induced Criticality, *PRX Quantum* **2**, 030313 (2021).
- [78] A.-R. Negari, S. Sahu, and T. H. Hsieh, Measurement-induced phase transitions in the toric code, *Phys. Rev. B* **109**, 125148 (2024).
- [79] P. Calabrese and J. Cardy, Entanglement entropy and quantum field theory, *Journal of Statistical Mechanics: Theory and Experiment* **2004**, 06002 (2004).
- [80] P. Calabrese and J. Cardy, Entanglement entropy and conformal field theory, *Journal of Physics A Mathematical General* **42**, 504005 (2009).
- [81] M. S. Foster, H. Guo, C.-M. Jian, and A. W. W. Ludwig, Free-Fermion Measurement-Induced Volume- to Area-Law Entanglement Transition in the Presence of Fermion Interactions, arXiv e-prints [2510.23706](#).
- [82] N. Read and D. Green, Paired states of fermions in two dimensions with breaking of parity and time-reversal symmetries and the fractional quantum Hall effect, *Phys. Rev. B* **61**, 10267 (2000).
- [83] S. Sang and T. H. Hsieh, Measurement-protected quantum phases, *Phys. Rev. Research* **3**, 023200 (2021).
- [84] V. Ravindranath and X. Chen, Robust Oscillations and Edge Modes in Nonunitary Floquet Systems, *Phys. Rev. Lett.* **130**, 230402 (2023).
- [85] L. Su, A. Clerk, and I. Martin, Dynamics and phases of nonunitary Floquet transverse-field Ising model, *Phys. Rev. Res.* **6**, 013131 (2024).
- [86] P. Calabrese and J. Cardy, Evolution of entanglement entropy in one-dimensional systems, *Journal of Statistical Mechanics: Theory and Experiment* **2005**, 04010 (2005).
- [87] T.-C. Lu and T. Grover, Spacetime duality between localization transitions and measurement-induced transitions, *PRX Quantum* **2**, 040319 (2021).
- [88] S. Lee and E.-G. Moon, Mixed-State Topological Order under Coherent Noise, *PRX Quantum* **6**, 030355 (2025).
- [89] Y. Zou, H. Xia, A. Zhang, and *et al.*, to appear (2026).
- [90] J. C. Napp, R. L. La Placa, A. M. Dalzell, F. G. S. L. Brandão, and A. W. Harrow, Efficient Classical Simulation of Random Shallow 2D Quantum Circuits, *Phys. Rev. X* **12**, 021021 (2022).
- [91] E. Fradkin and S. H. Shenker, Phase diagrams of lattice gauge theories with Higgs fields, *Phys. Rev. D* **19**, 3682 (1979).
- [92] D. Lee and B. Yoshida, Randomly Monitored Quantum Codes, arXiv e-prints [2402.00145](#).
- [93] A. R. Calderbank and P. W. Shor, Good quantum error-correcting codes exist, *Phys. Rev. A* **54**, 1098 (1996).
- [94] A. M. Steane, Error correcting codes in quantum theory, *Phys. Rev. Lett.* **77**, 793 (1996).
- [95] F. Eckstein, B. Han, S. Trebst, and G.-Y. Zhu, Data for “Learning transitions of topological surface codes” (2025).