Decoding the projective transverse field Ising model

Felix Roser ⬤,[*] Hans Peter Büchler ⬤, and Nicolai Lang ⬤

Institute for Theoretical Physics III and Center for Integrated Quantum Science and Technology, University of Stuttgart,
70550 Stuttgart, Germany

The competition between noncommuting projective measurements in discrete quantum circuits can give rise to entanglement transitions. It separates a regime where initially stored quantum information survives the time evolution from a regime where the measurements destroy the quantum information. Here we study one such system—the projective transverse field Ising model—with a focus on its capabilities as a quantum error correction code. The idea is to interpret one type of measurement as an error and the other type as a syndrome measurement. We demonstrate that there is a finite threshold below which quantum information encoded in an initially entangled state can be retrieved reliably. In particular, we implement the maximum likelihood decoder to demonstrate that the error correction threshold is distinct from the entanglement transition. This implies that there is a finite regime where quantum information is protected by the projective dynamics, but cannot be retrieved by using syndrome measurements.

I. INTRODUCTION

Entanglement lies at the heart of quantum mechanics and plays a key role in various fields, two important examples being *quantum error correction* [1] and *entanglement transitions* [2]. Because of decoherence, quantum error correction has been identified as an indispensable step towards scalable, universal quantum computation [3,4], and the existence of quantum codes demonstrated by Shor [5] is the bedrock on which the promises of quantum computation rest [5]. As it turns out, entanglement between the physical qubits of a quantum code is a necessary ingredient for quantum error correction. Independent of these considerations, the notion of *monitored quantum circuits* has gained traction in recent years [2,6–14]. The idea is to study quantum circuits built from random unitary gates, which typically lead to volume-law entangled states, and monitor them with local, projective measurements. The latter counter the entanglement growth and their competition can lead to transitions in the entanglement structure of the system. Strikingly, there are regimes in which a finite rate of projective measurements can be tolerated while still preserving long-range entanglement [2,6,7]. It then seems natural to study the error correction capabilities of such random systems. In this paper, we study a specific model that has been shown to feature an entanglement transition, and view it as a quantum code to illuminate the relation between quantum error correction and entanglement.

The rationale of quantum error correction is the following [15]: First, the quantum information to be protected (the *logical qubits*) is mapped to a subspace of the full system in such a way that local operations on few physical qubits can neither leak information about the amplitudes nor change them. This subspace is known as *code space* and necessarily comprises states that share entanglement between the physical qubits. If such a system is projectively monitored by the environment, two things can happen: Either the environment gains access to the amplitudes and they are lost irretrievably or it does not and the amplitudes are still hidden in the state of the system. However, as the interaction with the environment injects entropy into the system, these hidden amplitudes cannot be accessed right away. The next step of quantum error correction is therefore to measure observables that do not destroy the encoded amplitudes but retrieve information about the errors induced by the environment, thereby lowering the entropy of the system. The extracted information is called *syndrome* and can be used to reconstruct the unknown errors to access the hidden amplitudes; this last step is referred to as *decoding* and the method used to convert the syndrome into a tentative error pattern is the *decoding algorithm*. One can use different decoding algorithms for the same quantum code. Whether decoding for a particular instance succeeds depends on the error pattern, the knowledge of the decoder (the syndrome), and the decoder itself. The maximum error rate up to which decoding with a given decoder succeeds on average is known as *error threshold*, a quantity that in many cases can only be approximated numerically. Prominent examples for quantum codes are Shor's nine-qubit code [16], the seven-qubit Steane code [17], and scalable codes derived from topologically ordered systems like the toric/surface codes [18–21]. For all of these, efficient decoding algorithms and numerical results for their error thresholds are known [22–28]; the particularly high thresholds of surface codes make them promising candidates for real-world implementations [29,30].

The notion of monitored quantum circuits that give rise to entanglement transitions was introduced in Refs. [2,6,9] and originated from the question whether extensive, coherent subsystems can be stabilized in the presence of environmental

---

*felix.roser@itp3.uni-stuttgart.de

noise. Typically, one starts with a product state of an extensive number of qubits equipped with a geometry and applies a random sequence of 2-local unitaries (to model coherent, local interactions) interspersed with single-qubit projective measurements; the relative rate of unitaries and projective measurements is the parameter of the system. The unitaries are ideally drawn from the Haar measure, but often restricted to the Clifford group instead to make use of the stabilizer formalism for efficient simulation [15,31–35]. The findings of Refs. [2,6,7,9] show that there is a finite, critical rate for the projective measurements at which a continuous transition in the entanglement structure takes place: below the critical rate, the unitary evolution dominates and stabilizes the system in a volume-law phase, whereas, above the critical rate, the projective measurements prevail so no long-range entanglement can build up (indicated by area-law states with only local entanglement). Later, it was shown that similar transitions can be found in systems where the unitaries are replaced by projective measurements that do not commute with the projective measurements of the environment [36–40]. Depending on the measurements [38], the entanglement transitions in such purely projective models often separate phases of different area laws (instead of volume-law and area-law phases). One such model, derived from the transverse field Ising model and dubbed *projective transverse field Ising model (PTIM)* was studied in Ref. [37]. It features a critical monitoring rate dictated by bond percolation. Below this rate, the system exhibits long-range entangled states with an area law; above, it transitions into product states with short-range entanglement. Here we use this model to study the connection between quantum error correction and entanglement transitions.

That these two fields are related is not hard to see: If one initializes a monitored quantum circuit not in a product state but an entangled state that encodes the amplitudes of a logical qubit nonlocally, one can ask how long these amplitudes survive the evolution of the system and how the average lifetime scales with the system size. It turns out that this is an alternative characterization of the entanglement transition [10,37,41–47]: below the critical monitoring rate (in the entangling phase), the lifetime grows exponentially with the system size so the amplitudes survive indefinitely in macroscopic systems; by contrast, the growth is subexponential (typically logarithmic) in the disentangling phase. In this sense, monitored quantum circuits in the entangling phase can be seen as random quantum error correction codes that protect amplitudes by scrambling them into distributed degrees of freedom faster than the environment can extract information. This relates to the first (encoding) stage of quantum error correction explained above. What is not so clear is the second (decoding) stage: Under which conditions and how is it possible to *retrieve* the scrambled amplitudes? As for quantum codes, one may expect an error threshold for the projective measurements of the environment below which retrieval is possible by some decoding algorithm. However, it is unclear how the error threshold relates to the entanglement threshold except that the latter poses an upper bound on the former (in the disentangling phase, the amplitudes are lost and cannot be retrieved).

In this paper, we answer these questions specifically for the one-dimensional PTIM with a detailed study of several

decoding algorithms. We analyze the decoders numerically making use of a mapping between trajectories with projective measurements and "classical" trajectories with unitary operations. We introduce quantities to evaluate the performance of decoders quantitatively. Our analysis includes a naïve decoder based on *majority voting* [the maximum likelihood decoder (MLD) for classical repetition codes], which is an instructive approach that, unfortunately, fails. We then discuss and evaluate a decoder based on minimum weight perfect matching (MWPM), an algorithm that is often used for the decoding of topological surface codes. The algorithm has been applied successfully to the decoding of the PTIM by Li and Fisher [48]; we verify their results qualitatively and find a finite error threshold below which the decoder reliably retrieves the encoded quantum information. We then extend these results and implement the MLD of this system, i.e., the provably optimal decoder. We show that its decoding threshold provides only a marginal improvement over the MWPM decoder. Furthermore, we show that the decoding threshold of the MLD is distinct from the entanglement transition of the PTIM. We conclude that there is a finite range of parameters where the encoded amplitudes *survive* the monitoring by the environment but cannot be *retrieved* without having access to the full system dynamics.

The remainder of this paper is structured as follows. In Sec. II, we start with a description of the model and focus on its interpretation as a quantum code. We continue by formalizing the task of decoding and introduce quantities to gauge the performance of decoders. In Sec. III, we present a brief discussion of a naïve decoder based on majority voting and demonstrate why it fails. Taking guidance from this failure, we construct a decoder based on MWPM in Sec. IV and study its performance numerically; we find a finite error threshold. In Sec. V, we implement the MLD, which we find to perform slightly better than the MWPM decoder. We conclude in Sec. VI with a summary and outlook. In Appendixes A–D we provide technical details and proofs for some claims in the main text.

II. SETTING

A. The model

We start by introducing the PTIM and its interpretation as a quantum code. Consider an open chain of L qubits prepared in the initial state $|\Psi(t=0)\rangle = |\Psi_0\rangle$. To propagate the system from time step t to $t+1$, we first loop through all sites and measure with probability $p \in [0, 1]$ the operator

$$E_i = \sigma_i^x. \tag{1}$$

Here, σ_i^α denotes the Pauli matrix $\alpha = x, y, z$ acting on qubit $i = 1, \ldots, L$. We will refer to these measurements as *error measurements*; they describe the monitoring of the system by the environment. Next, we loop through all edges $e = (i, i+1)$ between adjacent sites and measure

$$S_e = \sigma_i^z \sigma_{i+1}^z \tag{2}$$

with probability $1 - q \in [0, 1]$. We will refer to these as *syndrome measurements* as they are the stabilizers of the quantum repetition code (see below). Note that $q \in [0, 1]$ refers to the probability that a stabilizer is *not* measured. Repeating
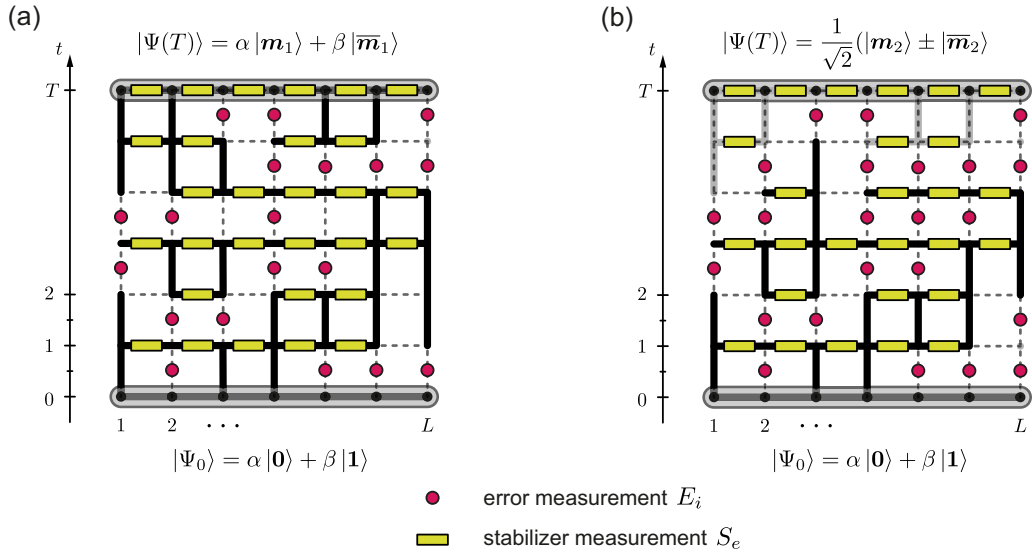
FIG. 1. The model. The projective transverse Ising model (PTIM) is a stochastic process where in each time step $t \mapsto t + 1$ first error measurements $E_i = \sigma_i^x$ occur with probability p per site, and then syndrome measurements occur with probability $1 - q$ per edge $e = (i, i + 1)$ along a one-dimensional open chain of L qubits. In this paper, we study sample averages over many trajectories $\mathcal{T} = \{\Psi(t)\}$ for initial states $|\Psi(t = 0)\rangle = |\Psi_0\rangle = \alpha|0 \ldots 0\rangle + \beta|1 \ldots 1\rangle$ that belong to the code space of the quantum repetition code. We are interested in whether the logical amplitudes $\alpha, \beta \in \mathbb{C}$ survive the PTIM evolution for $t = T \sim L$ time steps. To test this, we finalize each trajectory with a complete syndrome measurement. The final state $|\Psi(T)\rangle$ can then have only two forms: (a) Trajectory where the amplitudes survive, which corresponds to the bond percolation of the initial cluster highlighted in black. While horizontal edges are active if there is a stabilizer measurement, vertical edges are active if there is *no* error measurement. The cluster connected to the initial state at $t = 0$ carries the encoded amplitudes. (b) Trajectory where the initial cluster does not percolate. In this case, the encoded amplitudes are lost. The final qubit patterns $\boldsymbol{m}_i$ depend on the trajectory $\mathcal{T}$.

this two-step process generates a trajectory $\mathcal{T} = \{\Psi(t)\}$ of the PTIM, see Fig. 1(a) for an example. For a fixed initial state, the trajectory is uniquely determined by the space-time patterns E^p and S^p of error- and syndrome measurements and their respective measurement outcomes E^r and S^r. With a slight abuse of notation, we refer to a trajectory as $\mathcal{T} = (E, S)$ where $E = (E^p, E^r)$ and $S = (S^p, S^r)$. The parameters of this model are the probabilities p for an error and q for a missing stabilizer measurement, and we are interested in properties of this stochastic process when sampled over many trajectories for large system sizes $L \to \infty$. Note that there are two types of randomness in the model: the *classical* randomness of measurement choices encoded in the patterns E^p and S^p, and the *quantum* randomness from the measurement outcomes E^r and S^r.

For an initial product state of the form $|\Psi_0\rangle = |+\rangle \otimes \cdots \otimes |+\rangle \equiv |+ \cdots +\rangle$ (where $|+\rangle$ denotes the state with $\sigma^x|+\rangle = |+\rangle$), it was shown in Ref. [37] that this process features an entanglement transition at $p + q = 1$ that is dictated by (anisotropic) bond percolation on the square lattice: For $p + q > 1$ (error and/or failed stabilizer measurements dominate) the state $\Psi(t \to \infty)$ remains short-range entangled, whereas for $p + q < 1$ (low error rate and/or syndrome measurements dominate), stable long-range entanglement emerges.

By contrast, here we are interested in the error correction capabilities of this model. To this end, we consider (typically entangled) initial states of the form

$$|\Psi_0\rangle = \alpha|0 \ldots 0\rangle + \beta|1 \ldots 1\rangle, \tag{3}$$

with $\sigma^z|0\rangle = |0\rangle$ and $\sigma^z|1\rangle = -|1\rangle$; $\alpha, \beta \in \mathbb{C}$ with $|\alpha|^2 + |\beta|^2 = 1$ are the amplitudes of the logical qubit to be protected. The two-dimensional subspace $\mathcal{C} \ni |\Psi_0\rangle$ spanned by $|\boldsymbol{0}\rangle \equiv |0 \ldots 0\rangle$ and $|\boldsymbol{1}\rangle \equiv |1 \ldots 1\rangle$ is the code space of the quantum repetition code with stabilizer $\mathcal{S} = \langle\{S_e\}\rangle$, i.e., $\hat{S}|\Psi\rangle = |\Psi\rangle$ for all $|\Psi\rangle \in \mathcal{C}$ and $\hat{S} \in \mathcal{S}$. This code can only correct bit flip errors $E_i = \sigma_i^x$ but not phase errors σ_i^z (the latter correspond to parity-violating terms in its fermionic representation; this is not of relevance in the following).

To comply with the concept of quantum error correction, we modify the evolution of the PTIM, initialized in Eq. (3), by demanding that in the final time step $t = T$ (where T is fixed beforehand and typically $T \sim L$), *all* stabilizers $\{S_e\}$ are measured, irrespective of the failure probability q. This forces the final state $|\Psi(t = T)\rangle$ into the two-dimensional subspace $\mathcal{C}_{\boldsymbol{m}}$ spanned by $|\boldsymbol{m}\rangle$ and $|\overline{\boldsymbol{m}}\rangle = \prod_i \sigma_i^x|\boldsymbol{m}\rangle$ where $\boldsymbol{m} = (m_1, \ldots, m_L)$ with $m_i \in \{0, 1\}$ is a qubit configuration that corresponds to the syndrome measurements at $t = T$ and $\overline{\boldsymbol{m}}$ is its globally flipped configuration (which necessarily also matches the syndrome). It was shown in Ref. [37] that the final state has only two possible forms,

$$|\Psi(T)\rangle = \begin{cases} \frac{1}{\sqrt{2}}|\boldsymbol{m}\rangle \pm \frac{1}{\sqrt{2}}|\overline{\boldsymbol{m}}\rangle \\ \alpha|\boldsymbol{m}\rangle + \beta|\overline{\boldsymbol{m}}\rangle, \end{cases} \tag{4}$$

where $\boldsymbol{m} = \boldsymbol{m}(\mathcal{T})$ is some qubit configuration that depends on the trajectory $\mathcal{T}$. In the first case, $|\Psi(T)\rangle = \frac{1}{\sqrt{2}}|\boldsymbol{m}\rangle \pm \frac{1}{\sqrt{2}}|\overline{\boldsymbol{m}}\rangle$, the environment gained access to the logical qubit through its measurements $E = (E^p, E^r)$ and no recovery of the encoded amplitudes is possible. The probability for this outcome
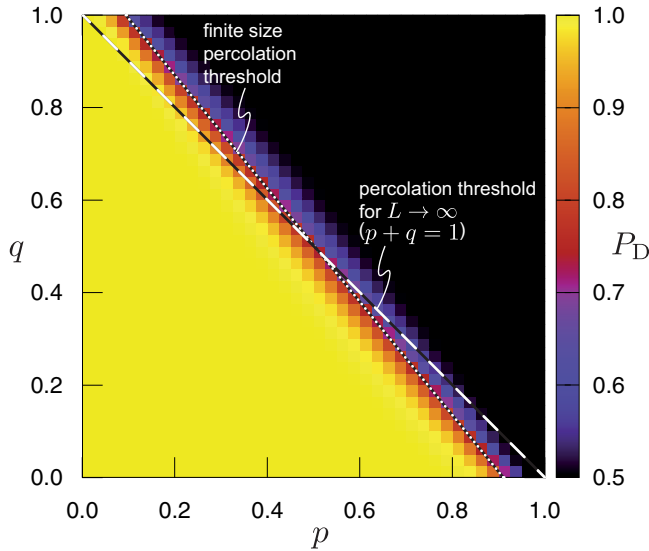
FIG. 2. Decoding probability with full knowledge. Probability P_D to successfully decode the PTIM with full knowledge of the error and stabilizer measurements $\mathcal{T} = (E, S)$ as a function of the error probability p and the stabilizer failure probability q for $L = T = 51$. For $p + q > 1$, decoding fails almost surely because in the disentangling phase the amplitudes are lost due to the monitoring of the environment. Because of finite-size effects, the true percolation threshold is slightly tilted and smeared out. For each datapoint, we sampled 20 000 trajectories.

approaches unity for $L, T \to \infty$ if $p + q > 1$, i.e., in the disentangling phase of the PTIM. This is so because the cluster that carries the amplitudes does not percolate, as shown in Fig. 1(b). In the second case, $|\Psi(T)\rangle = \alpha|\boldsymbol{m}\rangle + \beta|\overline{\boldsymbol{m}}\rangle$, the cluster *does* percolate [Fig. 1(a)] and the amplitudes survive the monitoring by the environment; however, they are now encoded in a rotated basis $|\boldsymbol{m}\rangle = \hat{C}^\dagger|\mathbf{0}\rangle$, where $\hat{C} = \prod_i (\sigma_i^x)^{C_i}$, $C = (C_1, \dots, C_L) \in \{0, 1\}^L$, is a qubit flip pattern that describes the effect of the measurements by stabilizers and environment. The probability for this to happen approaches unity for $L, T \to \infty$ in the entangling phase of the PTIM, i.e., for $p + q < 1$.

The goal of this paper is to recover the encoded amplitudes from $|\Psi(T)\rangle$, which is tantamount to finding a correction string $C = C(\mathcal{T})$ such that

$$\hat{C}|\Psi(T)\rangle = \alpha|\mathbf{0}\rangle + \beta|\mathbf{1}\rangle = |\Psi_0\rangle. \tag{5}$$

We refer to this as *decoding of the noisy quantum repetition code*, which we discuss in the next subsection.

B. Decoding

The entanglement transition at $p + q = 1$ acts as an upper bound in the (p, q) parameter space up to which decoding can possibly succeed, and we will focus on $p + q < 1$ henceforth. If we have access to the full trajectory $\mathcal{T} = (S, E)$, it is straightforward to construct the decoding string C deterministically because we can simulate the evolution $|\Psi(t)\rangle$ efficiently using the stabilizer formalism (there are even more efficient ways to achieve this, but this shall not be our focus here). This is illustrated in Fig. 2 where we plot the

decoding probability P_D as a function of the error probability p and the stabilizer failure probability q (we will introduce the quantity P_D formally below). Note that the threshold of the entanglement transition is slightly tilted and smeared out due to finite-size effects; we checked that for $L, T \to \infty$ the transition gets sharper and approaches the off-diagonal $p + q = 1$ predicted by percolation theory.

However, given our description of the measurements above, it is reasonable to assume that we neither have access to *when and where* error measurements occurred (E^p) nor to their *results* (E^r), as these are performed by the environment; hence, we have only access to the syndrome data $S = (S^p, S^r)$. In Fig. 3, we illustrate the trajectory from Fig. 1(a) and construct an abstract, reduced representation that illustrates the knowledge of the decoder.

This lack of knowledge makes the construction of C a nontrivial problem. However, the following observation helps us in this regard: Remember that the syndrome data S includes the outcomes of the full syndrome measurement in the last step. As a consequence, we do know the rotated code space $\mathcal{C}_{\boldsymbol{m}}$ spanned by $|\boldsymbol{m}\rangle$ and $|\overline{\boldsymbol{m}}\rangle$. Given the amplitudes survived, the final state can have only two forms:

$$|\Psi(T)\rangle = \begin{cases} \alpha|\boldsymbol{m}\rangle + \beta|\overline{\boldsymbol{m}}\rangle \\ \alpha|\overline{\boldsymbol{m}}\rangle + \beta|\boldsymbol{m}\rangle. \end{cases} \tag{6}$$

In the first case, the right choice for the correction string that defines $\hat{C}$ is $C \equiv \boldsymbol{m}$, whereas, in the second case, it is $\overline{C} \equiv \overline{\boldsymbol{m}}$. Decoding therefore amounts to choosing a decoding string from a set $\{C, \overline{C}\}$ of two possible strings; selecting the wrong one leads to a bit flip error on the logical qubit.

So, although we do not have access to an extensive amount of information, namely, $E = (E^p, E^r)$, we are actually missing only a single bit, namely, whether to choose C or $\overline{C}$ to decode $|\Psi(T)\rangle$. A *decoder* or *decoding algorithm* picks one of these two as a function of the syndrome data S:

$$D : S = (S^p, S^r) \mapsto D(S) \in \{C, \overline{C}\}. \tag{7}$$

Note that in the following it is sufficient to start with an initial state $|\Psi_0\rangle = |\mathbf{0}\rangle$ to quantify the performance of the decoder, as the system dynamics commutes with the logical X-operator $X = \prod_i \sigma_i^x$. Thus, any initial state which is polarized in the X direction will never change its polarization, and it is best to choose an initial state with vanishing X polarization.

For a given trajectory $\mathcal{T} = (E, S)$ with final state $|\Psi_f(\mathcal{T})\rangle \equiv |\Psi(T)\rangle$, we can define the overlap

$$f^{qm}(\mathcal{T}, C) := |\langle\Psi_0|\hat{C}|\Psi_f(\mathcal{T})\rangle|^2 \tag{8}$$

to identify the correct string C to decode the quantum information. This expression returns 1 (0) for the correct (incorrect) string when the amplitudes survive; in cases where the amplitudes are lost, one finds $f^{qm} = \frac{1}{2}$, irrespective of the chosen correction string. This allows us to determine the performance of a decoder D. For a single trajectory $\mathcal{T} = (E, S)$, we define

$$f_D^{qm}(\mathcal{T}) := f^{qm}(\mathcal{T}, C = D(S)); \tag{9}$$

the performance of the decoder is then quantified by the probability to correctly decode the quantum information,

$$P_D(p, q; L, T) := \langle\!\langle f_D^{qm} \rangle\!\rangle = \sum_{\mathcal{T}} P^{qm}(\mathcal{T}) f_D^{qm}(\mathcal{T}), \tag{10}$$
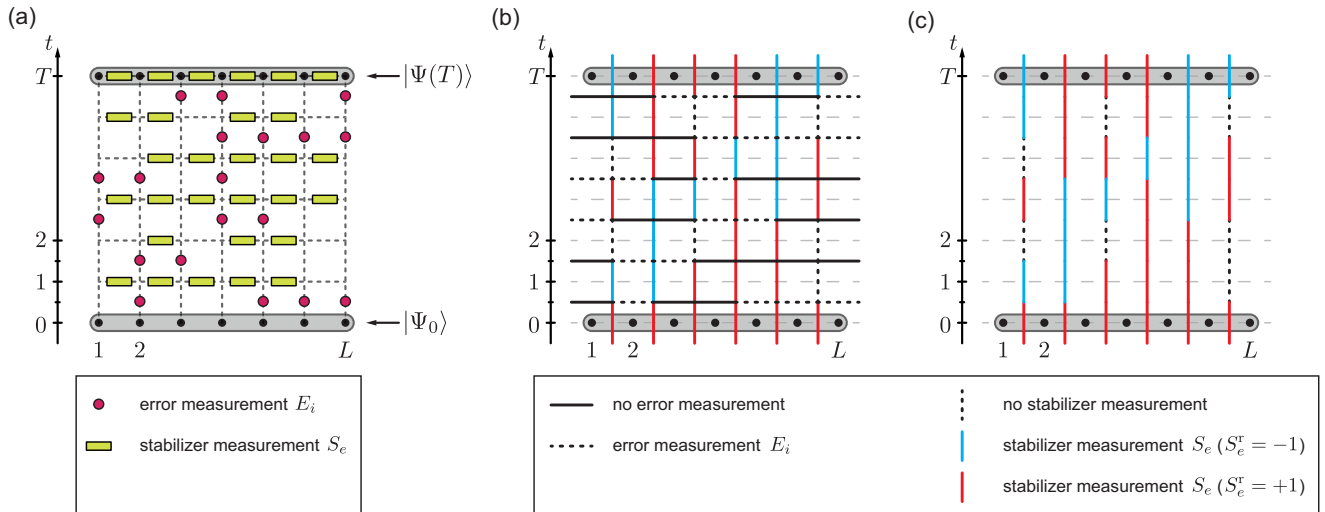
FIG. 3. Trajectory representations of the PTIM. (a) Full trajectory of the PTIM reproduced from Fig. 1(a) where the amplitudes survive the monitoring by the environment. (b) Switching to the dual of the dashed space-time lattice in (a), we can replace stabilizer measurements S^p by solid *vertical* edges, the colors of which encode the measurement outcomes S^r. *Absent* error measurements are replaced by solid *horizontal* edges. The emergent cluster structure is dual to the percolation cluster in Fig. 1(a). This representation on the dual lattice will be useful later for constructing decoders. (c) Decoders only have access to the syndrome data $S = (S^p, S^r)$, i.e., the vertical edges of the representation in (b). *Decoding* amounts to reconstructing the full pattern in (b) from the reduced one in (c) (or an equivalent one in a suitably defined sense).

where $\langle\langle\bullet\rangle\rangle$ denotes the sample average over trajectories and $P^{\mathrm{qm}}(\mathcal{T})$ is the probability for the trajectory $\mathcal{T} = (E, S)$. Note that $P^{\mathrm{qm}}(\mathcal{T})$ is a highly nontrivial quantity: While the classical probability distribution that governs E^p and S^p is straightforward to describe, the distributions of measurement *outcomes* E^r and S^r are not obvious as they depend on the history of the trajectory and are not independent.

The quantity P_D is our main figure of merit to evaluate the performance of a decoder. It parametrically depends on p and q, and we are interested in its behavior for large systems $L \to \infty$. A value of $P_D = 1$ indicates that the decoder D correctly decodes every trajectory and therefore allows us to restore the encoded amplitudes reliably. For high error rates and/or few stabilizer measurements, P_D is expected to drop to $\frac{1}{2}$ because either the decoder fails to decode the system (viz. it tosses a coin to decide on a correction string) or the system is in the disentangling phase and the amplitudes are lost to the environment.

In addition to the decoding probability P_D, we will use the *mean time to first failure (MTFF)* T_D to compare different decoders. To define T_D, consider a (ideally, infinitely long) trajectory $\mathcal{T}$ and for $t = 0, 1, \ldots$ define $\mathcal{T}|_t$ as the first t time steps of $\mathcal{T}$ terminated by a full stabilizer measurement. We can then apply a decoder D to $\mathcal{T}|_t$ and compute $f_D^{\mathrm{qm}}(\mathcal{T}|_t)$. Let $t_D^{\mathrm{qm}}(\mathcal{T})$ denote the first time step t where $f_D^{\mathrm{qm}}(\mathcal{T}|_t) < 1$ and define the MTFF as the sample average of this quantity:

$$T_D(p, q; L) := \langle\langle t_D^{\mathrm{qm}}\rangle\rangle. \tag{11}$$

We will refer to decoders as *succeeding* for a given set of parameters p and q if the MTFF grows superpolynomially in the limit $L \to \infty$; by contrast, the MTFF of a *failing* decoder grows sublinearly (typically logarithmically) in this limit.

C. Efficient numerical simulations

To evaluate P_D numerically by sampling, for each sample two steps are required: First, a trajectory $\mathcal{T}$ of the PTIM must be generated and, second, based on the syndromes S, the decoder must be evaluated to compute $f^{\mathrm{qm}}(\mathcal{T}, C = D(S))$. The second step depends on the decoding algorithm used and will be commented on in the respective sections where we study the performance of different decoders. The sampling of PTIM trajectories is described in the following.

Since both E_i and S_e belong to the Pauli group on L qubits, and $|\Psi_0\rangle = |\mathbf{0}\rangle$ is the unique state stabilized by $\{\sigma_i^z\}$, the PTIM evolution can be simulated exactly within the stabilizer formalism with only polynomial overhead [31,33–35], as stated by the Gottesmann-Knill theorem [15].

However, there is a more efficient method by mapping the trajectory to a *classical* stochastic process derived from the PTIM that operates on a chain of L classical *bits* initialized in the configuration $\boldsymbol{m}_0 = \mathbf{0}$ (we drop the "classical" in the following). Instead of performing error measurements E_i with probability p, we flip every bit in every time step with probability $p/2$. This results in a bit flip pattern $\tilde{E}^p$ on the space-time lattice which determines the final bit configuration $\boldsymbol{m}_T$ after T time steps. We then perform stabilizer measurements on the classical system just as we would on the PTIM (only there is no collapse due to the measurement so we can do this en bloc after generating the bit flip pattern); this yields a pattern S^p and the corresponding measurement results S^r. Given the data $S = (S^p, S^r)$ sampled in this fashion, we can apply a given decoder D to compute the correction string $C = D(S) \in \{0, 1\}^L$. With the (trivial) evaluation function

$$f^{\mathrm{bi}}(\tilde{E}^p, C) := \begin{cases} 1, & C = \boldsymbol{m}_T \\ 0, & \text{otherwise,} \end{cases} \tag{12}$$

we can evaluate every decoder also on the classical system. [The superscript ᵇⁱ stands for "bits" to distinguish this evaluation function from its quantum counterpart defined in Eq. (8)].

The crucial point, which we prove in Appendix A, is the relation

$$\langle\langle f_D^{\mathrm{qm}}\rangle\rangle_{\mathrm{qm}} = \langle\langle f_D^{\mathrm{bi}}\rangle\rangle_{\mathrm{bi}}, \tag{13}$$

which holds for all decoders D (the subscripts to the sample averages indicate sampling with respect to the full quantum model and the classical analog described above, respectively). This allows us to sample the classical process described above to evaluate the performance of decoders instead of performing the full quantum mechanical simulation of the PTIM. This approach results in a considerable speedup compared to the stabilizer formalism, which enables us to study larger systems and/or larger samples to reduce statistical fluctuations. We cross-checked the validity of Eq. (13) numerically on smaller systems using a full-fledged stabilizer simulation.

III. DECODING WITH MAJORITY VOTING

A. Algorithm

We will now introduce and discuss our first decoder D. This decoder will have the threshold $p = 0$ except for the special point $q = 0$, and therefore will fail to decode the stored quantum information for the PTIM. It serves as preparation and motivation for the more sophisticated decoders in Secs. IV and V.

Let us first focus on the special case with $q = 0$, where at every time step *all* stabilizers are measured and projective errors occur with probability $p \in (0, 1)$. This is the usual situation of quantum error correction. Note, however, that our *projective* error measurements with probability p effectively result in qubit flips with probability $p/2$. This is in contrast to conventional treatments where errors are modeled by *unitary* operators, in which case the probability for an error to occur and for a qubit to flip are identical.

If we exclude the exponentially unlikely situation where in a single time step an error occurs on every qubit, it is straightforward to check that, starting from $|\Psi_0\rangle = |\mathbf{0}\rangle$, every round of stabilizer measurements projects the system into a product state of the form $|\Psi(t)\rangle = |\mathbf{m}_t\rangle$. In this situation, the PTIM evolution becomes basically classical and guessing the qubit flips from the stabilizer measurements S^r is equivalent to decoding the classical repetition code: To compute the final correction string C that recovers the initial state, $\hat{C}|\Psi(T)\rangle = |C \oplus \mathbf{m}_T\rangle = |\mathbf{0}\rangle$, we split $C = \oplus_{t=1}^{T} C_t$ into correction strings C_t that correct the qubit flips that occurred in time step t. To compute C_t from the accumulated syndrome data $S^r = (S_t^r)_{t=0,\dots,T}$, we compare the syndrome measurements S_{t-1}^r at time $t-1$ with the ones S_t^r in the subsequent time step. Since the syndromes are complete (all stabilizers were measured, $q = 0$), this allows for only two consistent flip patterns $C_t \in \{C_{t,1}, C_{t,2}\}$ with $C_{t,2} = \overline{C_{t,1}}$ (the bar denotes the complementary bit string). The premise of the *majority voting decoder (MVD)* is to choose the one with *fewer* flips, as this is the more likely one for $p/2 < 1/2$. This choice is unique for chains of odd length L; for consistency, we will stick to odd L throughout the paper. It is well-known that this

decoder is perfect, i.e., it succeeds for all $p/2 < 1/2$ almost surely in the limit $L \to \infty$, which is therefore also true for the PTIM evolution at $q = 0$. Note that in contrast to conventional error models, where p denotes the probability for (unitary) qubit flips, the decoding transition in the PTIM appears at $p = 1$, i.e., when *all* qubits undergo an error measurement within each time step. Our goal is now to generalize the MVD decoder to the case where some stabilizers fail to be measured with probability $q > 0$.

The basic procedure remains unchanged, i.e., we decompose $C = \oplus_{t=1}^{T} C_t$ into corrections per time step and try to derive C_t from the syndrome measurements. Because of $q > 0$, typically there will be gaps in the syndromes S_t^r where no measurement was performed. This lack of knowledge enlarges the set of consistent flip patterns $C_t \in \{C_{t,1}, C_{t,2}, C_{t,3}, \dots\}$. To construct this list efficiently, we define a "tentative" qubit configuration $\tilde{\mathbf{m}}_{t-1} \equiv \oplus_{\tau=1}^{t-1} C_\tau \oplus \mathbf{0}$. By construction, $\tilde{\mathbf{m}}_{t-1}$ is consistent with the (partial) syndrome S_{t-1}^r. To construct C_t, we list all consistent flip patterns $C_{t,k}$ such that $\tilde{\mathbf{m}}_{t,k} \equiv C_{t,k} \oplus \tilde{\mathbf{m}}_{t-1}$ is consistent with the partial syndromes S_t^r. Following the rationale of majority voting, we then choose for C_t the $C_{t,k}$ with the fewest flips. This defines our decoding algorithm $D : S \mapsto C = D(S)$, which we will analyze in the next subsection.

B. Results

To assess the performance of majority voting, we computed the MTFF T_{MV} as a function of the system size L for fixed parameters p and q. The results are shown in Fig. 4(a) for the representative parameters $p = 0.2 = q$ and L up to 501. These results demonstrate that the MTFF does *not* scale superpolynomially with the system size L, i.e., the decoding algorithm does not decode the system efficiently. We checked that varying the parameters p and q does not alter the result qualitatively, i.e., there seems to be no parameter regime with $q > 0$ where T_{MV} grows superpolynomially with L.

This behavior can be made plausible with a single trajectory that highlights a crucial weakness of the MVD; see Fig. 4(b). The sketched trajectory has only a single error measurement E_9 between $t = 0$ and $t = 1$; furthermore, there is a single missing stabilizer measurement at $t = 1$ that creates a two-qubit segment connected by a syndrome measurement $S_e = -1$. Consequently, the decoder can only toss a coin to decide on the location of the error. In 50% of the cases, its choice is incorrect, so the internal model $\tilde{\mathbf{m}}_1$ deviates from the true state $|\Psi(t = 1)\rangle = |\mathbf{m}_1\rangle$ on this segment. In the next steps, no errors occur but there is a missing stabilizer in each time step. Because the majority voting is restricted to segments of contiguous stabilizer measurements, the decoder is forced to enlarge the discrepancy between $\tilde{\mathbf{m}}_t$ and $\mathbf{m}_t$ until at $t = 4$ the internal model is wrong everywhere. In total, the decoder has to assume eight flips for its internal model $\{\tilde{\mathbf{m}}_t\}$ whereas in reality a single error occurred. Because of its time-local mode of operation to determine C_t, the MVD has no chance to find the true error pattern deterministically.

There is also a more abstract perspective on this. Remember that the decoder *does work* for $q = 0$, i.e., for the conventional repetition code. For this, it is crucial that majority voting is applied to an extensive set of qubits (namely, L).
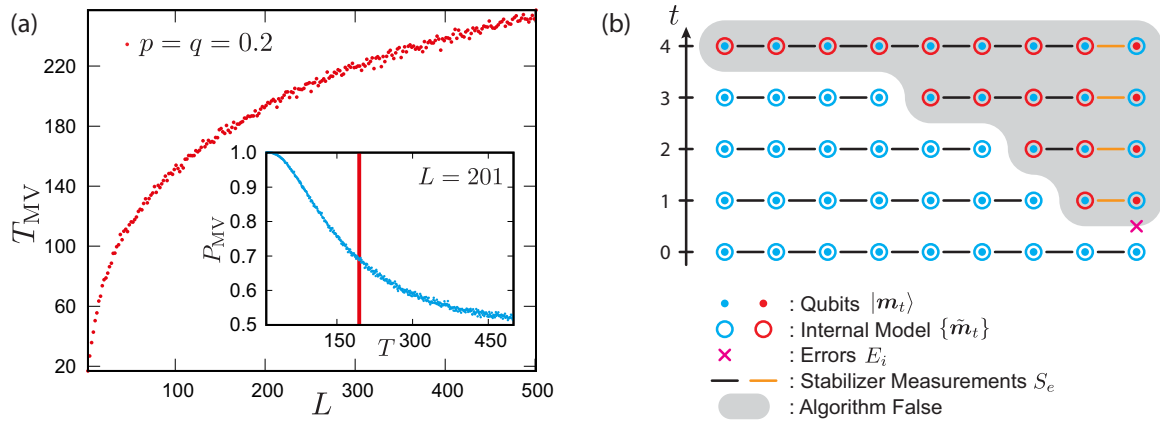
FIG. 4. Majority voting decoder (MVD). (a) Mean time to first failure (MTFF) T_{MV} for the MVD for $p = q = 0.2$ averaged over 5000 trajectories as a function of the chain length L (odd L). Inset: Decoding probability $P_{MV}(p, q; L, T)$ for the same p and q and $L = 201$ averaged over 10 000 trajectories as a function of the decoding time T. The vertical red line indicates the MTFF $T_{MV}(p, q; L)$ for these parameters. These results demonstrate that the MTFF does *not* grow exponentially with the system size. We checked that these results are representative and do not change for other parameters p and q. (b) Example trajectory that illustrates a crucial weakness of the MVD. The tentative bit pattern $\{\tilde{m}_t\}$ constructed by the decoder is indicated by circles; the true state of the system $\{m_t\}$ by bullets. Note that this is an example where the state $|\Psi(t)\rangle$ has product form $|m_t\rangle$ for all times t. The decoding result is $\hat{C}|\Psi(T = 4)\rangle = |\mathbf{1}\rangle \neq |\mathbf{0}\rangle = |\Psi_0\rangle$, so $f^{qm}(\mathcal{T}, C = D(S)) = 0$.

For $q > 0$, the probability to find a segment of l contiguous stabilizer measurements in a single time step is $q^2(1 - q)^l$; the average length of such a cluster is therefore $\bar{l} = (1 - q)/q$. The MVD, as defined above, performs majority voting on each of these segments separately—it cannot keep track of correlations between the segments. Because $\bar{l}$ is independent of L and finite for $q > 0$, these $\bar{l}$-local decisions do not improve for $L \to \infty$.

Combining these findings, the crucial flaw of the MVD seems to be that it composes the final correction string C out of stepwise corrections C_t that are the result of a time-local minimization procedure. Especially, the decoder does not take into account the full syndrome data S *globally* but slices it into independent pieces; it operates, in a sense, one-dimensional and ignores the two-dimensional space-time geometry that comes with S. Thus, it cannot exploit correlations between disjoint segments of contiguous syndrome measurements at any given time step. We will overcome this issue with the much more sophisticated MWPM decoder in the next section.

IV. DECODING WITH MINIMUM WEIGHT PERFECT MATCHING

A. Algorithm

The MWPM decoder takes into account the full two-dimensional space-time geometry and derives from the syndromes S a possible error pattern E^p with a *globally* minimal number of errors. This error pattern is then used to decide on one of the two decoding strings $\{C, \overline{C}\}$. Our approach is motivated by the use of MWPM for the decoding of two-dimensional surface codes [49,50], which is conceptually similar to the decoding of a noisy, one-dimensional quantum repetition code [21] in that nodes on a two-dimensional lattice must be matched pairwise.

The decoder D is defined by three steps: First, the error syndrome is used to construct an abstract graph on which then, in the second step, MWPM is performed. Finally, the

found MWPM is used to select one of the two decoding strings $\{C, \overline{C}\}$. We now describe these three steps in detail, following the illustrations in Fig. 5; a motivation for this algorithm is given below.

(1) In the first step, we construct from the space-time pattern of syndromes S in Fig. 5(a) the reduced dual space-time lattice (graph) in Fig. 5(b), augmented by dummy nodes and dummy edges. To this end, we keep *all* horizontal edges of the lattice [dashed gray in (a)], but *only* the vertical edges without syndrome measurement [dashed black in (a)]. We then highlight all nodes of the graph in (b) where a blue vertical line with $S_i = -1$ in (a) terminates. To allow for matchings to the boundaries, we add gray dummy nodes on the endpoints of horizontal edges. If the total number of highlighted nodes (blue and gray) is odd, we add an additional dummy node to the graph. We then connect all dummy nodes pairwise by dummy edges [for the sake of clarity, we do not show these in Fig. 5(b)]. Finally, we assign integer weights to all edges of the graph which count the number of horizontal segments (= qubits) traversed by these edges. In particular, all vertical and all dummy edges have weight zero.

(2) In the second step, we compute a MWPM of the blue and gray nodes on this graph. A *perfect* matching is given by pairwise connections of *all* highlighted nodes along edges of the graph. A perfect matching has *minimum weight* if the total sum of weights of all used edges is minimal. Minimum weight perfect matchings can be efficiently computed using the *Blossom algorithm* [51]; here we use an optimized implementation by Kolmogorov, known as `Blossom V` [52]. The result is a MWPM, illustrated by orange paths in Fig. 5(c), where every horizontal edge traversed contributes 1 to the total weight. The horizontal edges of the MWPM are the positions where the decoder assumes that a qubit was flipped.

(3) In the third and final step, we use the MWPM to decide on the correction string $C = (c_1, \ldots, c_L) = D(S)$. We set $c_i = 0$ if the qubit at position i crosses an *even* number of horizontal edges in the MWPM, and $c_i = 1$ if it crosses an *odd*
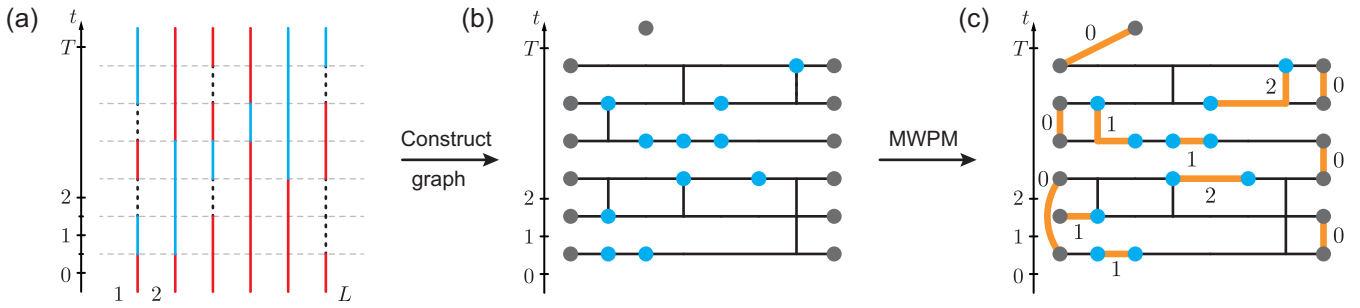
FIG. 5. Minimum weight perfect matching (MWPM) decoder—algorithm. (a) Syndrome data of the trajectory in Fig. 3(b) on the dual space-time lattice; this is the input of the decoder. [The panel is mostly identical to Fig. 3(c); we reproduce it here for convenience.] (b) Construction of the reduced, dual space-time lattice (graph) as the input for the MWPM algorithm. It consists of *all* horizontal edges but only the *dashed* vertical edges of the dual space-time lattice in (a). The endpoints of blue line segments in (a) are labeled by blue nodes. To allow matchings to the edges, we add dummy nodes (gray) on the boundary and an additional dummy node (top) to make the total number of nodes even. The dummy nodes are fully connected by dummy edges (not shown). All edges are weighted by their number of horizontal segments of the dual space-time lattice (= qubits). This weighted graph allows for a perfect matching of all blue and gray nodes by construction and is the input to the MWPM algorithm. (c) Example of a perfect matching (orange lines). The weight of each matching path is given by the number of traversed horizontal edges (which correspond to assumed error measurements). The shown perfect matching is a MWPM as it minimizes the total weight (= number of assumed errors).

number. In other words: the decoder assigns $c_i = 0$ ($c_i = 1$) to qubit i if it assumes an even (odd) number of flips on this qubit.

The motivation for this algorithm is that a perfect matching on the constructed graph describes a *possible* error pattern E^p (if one ignores the matchings on dummy edges between dummy nodes). This is so because the paths of the perfect matching in combination with the blue vertical edges in Fig. 5(a) (where syndrome measurements $S_i = -1$ signal flipped adjacent qubits) form the closed boundaries of space-time regions where the qubits might have been flipped. A given perfect matching therefore presupposes projective errors on all its horizontal segments of the space-time lattice. In particular, the total weight of the matching (= sum of all horizontal segments) corresponds to the number of required errors. Computing an error pattern E^p that (1) explains the observed syndrome S and (2) minimizes the total number of errors is therefore equivalent to finding a MWPM on the reduced space-time lattice. The matchings to dummy nodes on boundaries are necessary because the endpoint of error strings that terminate on boundaries is not signaled by a syndrome measurement.

Note that this approach is very similar to the decoding of topological quantum memories like the surface code [18,19,21], where decoding boils down to matching (= fusing) pairs of anyonic excitations with a minimal amount of (unitary) errors. It is well-known that MWPM is an efficient method to achieve this [49,50]. Indeed, the situation here can be interpreted as the decoding problem of an anisotropic version of the surface code with peculiar boundary conditions [21].

B. Results

We start our analysis by computing the decoding probability P_{MWPM} for $L = T = 51$ as a function of p and q, Fig. 6(a). Note that it is reasonable to scale $T \sim L$ when studying the MWPM decoder; for simplicity, we set $T = L$

in the following. The results demonstrate that there is a finite region in parameter space where decoding succeeds. As expected, this region is fully contained in the regime where the PTIM is in the entangling phase (cf. Fig. 2); it is considerably smaller though. There is a clear transition between the region where decoding succeeds ($P_{\text{MWPM}} \approx 1$) and the region where it fails ($P_{\text{MWPM}} \approx 0.5$). We checked that this transition becomes sharper for larger systems (it also shifts slightly; we discuss this in more detail below). Let us comment on a few peculiarities. First, on the q axis ($p = 0$) the decoder always succeeds because there are no errors to be corrected. This intuition is confirmed by our results. On the p axis ($q = 0$), one would expect the decoder to be successful as well because MWPM is equivalent to global majority voting for $q = 0$, i.e., when every stabilizer is measured (and we know that the MVD succeeds in this special case for all $p/2 < 1/2$). However, our results suggest that the decoder *fails* for $p \gtrsim 0.7$. We confirmed that this is a (strong) finite-size effect which can even be described analytically (see Appendix D). For $L = T \rightarrow \infty$, the decoding region will indeed grow on the p axis until it reaches $p = 1.0$.

To assess the finite-size scaling on the diagonal, we computed P_{MWPM} as a function of $p = q$ for increasing system sizes $L = T$, see Fig. 6(b). We find a clear crossing with almost negligible finite-size shift (in contrast to $q = 0$) that becomes sharper in the limit $L = T \rightarrow \infty$. We conclude that the decoding phase is indeed a proper subset of the entangling phase and pinpoint the decoding threshold at $p_{\text{thr}}^{\text{MWPM}} = q_{\text{thr}}^{\text{MWPM}} \approx 0.324$. Note that on the diagonal the entanglement transition takes place at $p_c = q_c = 0.5$. In the intermediary regime $p_{\text{thr}} < p = q < p_c$ we face the peculiar situation that the encoded amplitudes of the logical qubit survive the monitoring by the environment but are inaccessible using only the syndrome data S and MWPM decoding.

Now that we know the phase diagram of the MWPM decoder, we can also evaluate the MTFF T_{MWPM}. The decoding transition is also visible in this quantity. To demonstrate this, we pick three values for p on the diagonal $p = q$: $p_1 = 0.280$
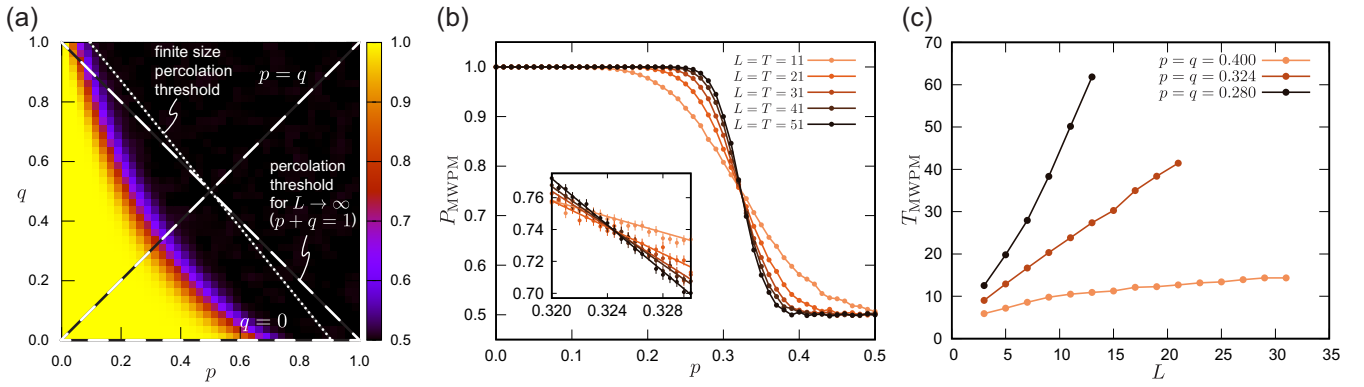
FIG. 6. Minimum weight perfect matching (MWPM) decoder—results. (a) Decoding probability P_{MWPM} of the MWPM decoder as a function of $p, q \in [0, 1]$ for $L = T = 51$, computed from a sample of 20 000 trajectories for every datapoint. For comparison, we reproduce the percolation thresholds from Fig. 2 (dotted: numerical for finite-size system; dashed: exact for $L \to \infty$); they indicate the entanglement transition of the PTIM. The decoding phase is a proper subset of the entangling phase of the PTIM. (b) Decoding probability P_{MWPM} along the diagonal $p = q$ in (a) for increasing system sizes $T = L = 11, \dots, 51$ and sampled over 50 000 trajectories for every datapoint. We find a crossing with negligible finite-size shift that gets sharper for $L = T \to \infty$ (inset), indicating an error threshold $p_{\text{thr}}^{\text{MWPM}} = q_{\text{thr}}^{\text{MWPM}} \approx 0.324$ away from the entanglement transition at $p_c = q_c \approx 0.5$. The error bars in the inset represent the standard deviation of the samples. (c) Mean time to first failure (MTFF) T_{MWPM} as a function of system size L in the decoding phase ($p_1 = 0.280$), at the error threshold ($p_2 = 0.324$), and outside the decoding phase but in the entangling phase ($p_3 = 0.400$), sampled over 5 000 trajectories for every datapoint. The MTFF grows exponentially with the system size in the decoding phase [cf. Fig. 4(a)].

in the decoding phase, $p_2 = 0.324$ on the phase boundary, and $p_3 = 0.400$ outside the decoding phase (but in the entangling phase). In Fig. 6(c), we plot the MTFF T_{MWPM} as a function of system size $L = T$. It shows the expected exponential behavior in the decoding phase, while it seems to grow linearly ($T_{\text{MWPM}} \sim L$) at criticality and subalgebraically [$T_{\text{MWPM}} \sim \log(L)$] away from the decoding phase.

V. MAXIMUM LIKELIHOOD DECODING

So far, we have demonstrated the existence of a nontrivial decoding threshold for our model, i.e., a transition at a critical error rate up to which it is possible to retrieve the stored quantum information knowing only the syndrome S. While the entanglement transition provides an upper bound on this critical error rate, the MWPM decoder presented in the previous section provides a lower bound on this threshold. Decoders that achieve this threshold are MLDs. These decoders are defined as the ones that, given the syndrome data S, choose the correction string that is most likely correct. This is done by calculating probabilities for classes of trajectories that yield the same syndrome data S and final state $\hat{C}|\Psi_0\rangle$. We define the probability of such a class as

$$P_{f^{\text{qm}}}^{\text{qm}}(C|S) := \sum_{\mathcal{T}|_S} P^{\text{qm}}(\mathcal{T}) f^{\text{qm}}(\mathcal{T}, C), \tag{14}$$

where $\mathcal{T}|_S$ restricts the summation to trajectories with the syndrome S. The output of the MLD is the correction string C that maximizes Eq. (14) for a given syndrome S. Clearly, the MLD provides the highest possible decoding probability

$$P_{\text{MLD}}(p, q; L, T) = \sup_D P_D(p, q; L, T), \tag{15}$$

and therefore is the best possible decoder. While in many cases it is impossible (or unknown how) to implement a maximum likelihood decoder efficiently for a given error

correction code, below we demonstrate the efficient implementation of such a decoder for the PTIM.

Performing the sum in Eq. (14) is highly nontrivial because it is constrained by the syndrome and typically contains an exponentially large number of terms. Recently, an efficient implementation of an MLD for the (perfect) surface code has been demonstrated by Bravyi *et al.* [26] using a clever resummation technique that is based on an equivalent formulation in terms of a quadratic fermion theory. Here we implement the maximum likelihood decoding for our system in two steps:

(1) We introduce a classical stochastic process with random bit flips and stabilizer measurements, and prove that defining the MLD based on the decoding probability $P_{f^{\text{bi}}}^{\text{bi}}(C|S)$ for this classical model is equivalent to implementing the MLD of its quantum mechanical pedant.

(2) The decoding probability $P_{f^{\text{bi}}}^{\text{bi}}(C|S)$ for this classical model is then efficiently evaluated using the resummation techniques developed by Bravyi *et al.* [26].

The classical stochastic process derived from the PTIM has already been introduced in Sec. II C, and is reviewed here again. It describes a chain of L classical *bits* with all bits initialized in state 0. The dynamics is governed by random bit flips with probability $p/2$, resulting in a bit flip pattern $\tilde{E}^p$ on the space-time lattice. In every time step, we flip the bits according to $\tilde{E}^p$ and then perform stabilizer measurements with probability $1 - q$, which yields the stabilizer pattern S^p. The input of the decoder is the combination of the stabilizer pattern and the corresponding results $S = (S^p, S^r)$, just as for the quantum system.

The crucial point is that there is a straightforward way to adapt Vargo's algorithm to the classical variant of our system. We then use this algorithm to compute the total probability $P_{f^{\text{bi}}}^{\text{bi}}(C|S)$ of a fixed correction string C by summing over all consistent bit patterns. This results in an algorithm with a runtime that scales polynomially [like $O(L^4)$] with system size $L = T$, and depends parametrically on the error probability p
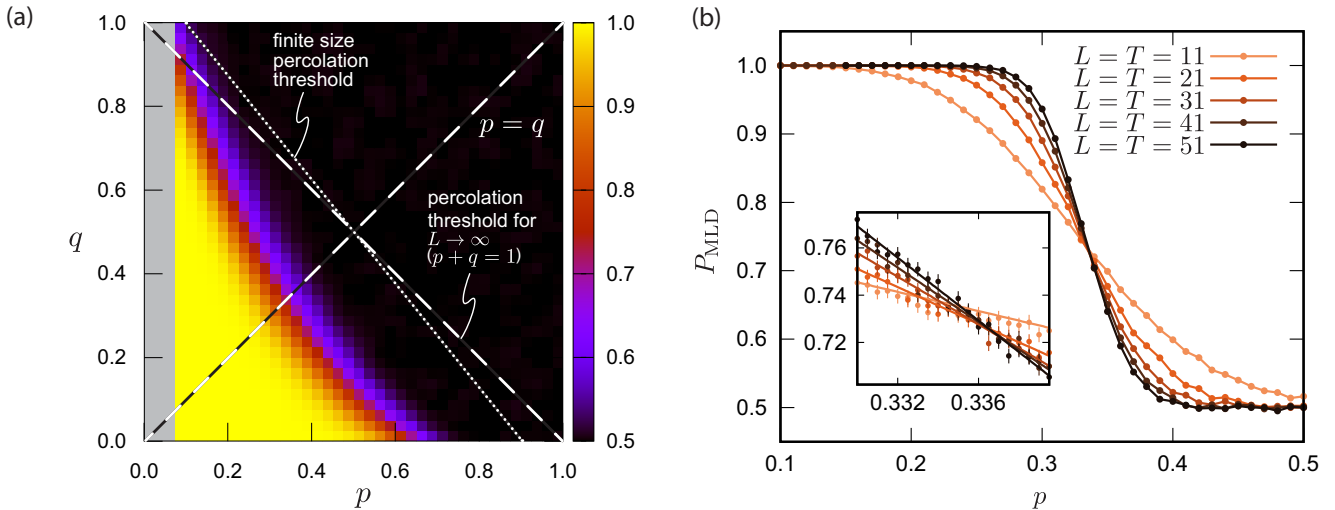
FIG. 7. Maximum likelihood decoder (MLD). (a) Decoding probability P_{ML} of the ML decoder as a function of $p, q \in [0, 1]$ for $L = T = 41$, computed from a sample of 20 000 trajectories for every datapoint. For comparison, we reproduce the percolation thresholds from Fig. 2 (dotted: numerical for finite-size system; dashed: exact for $L \rightarrow \infty$); they indicate the entanglement transition of the PTIM. As for the MWPM decoder, the decoding phase is a proper subset of the entangling phase of the PTIM. We omit data for $p \leqslant 0.05$ because in this regime the MLD algorithm becomes numerically unstable and the results are not trustworthy. (b) Decoding probability P_{MWPM} along the diagonal $p = q$ in (a) for increasing system sizes $T = L = 11, \ldots, 51$ and sampled over 50 000 trajectories for every datapoint. Again we find a crossing with negligible finite-size shift that gets sharper for $L = T \rightarrow \infty$ (inset), indicating an error threshold $p_{thr}^{ML} = q_{thr}^{ML} \approx 0.336$ away from the entanglement transition at $p_c = q_c \approx 0.5$. The error bars in the inset represent the standard deviation of the samples.

and syndrome failure rate q. The decoder $D_{ML}^{bi}(S)$ then returns the correction string that maximizes $P_{f^{bi}}^{bi}(C|S)$.

At this point, we have a working MLD for the classical system. Due to the equivalence between the classical and the quantum trajectories [recall Eq. (13)], one can prove rigorously that this algorithm satisfies the conditions for a MLD also for the projective error model, i.e., $\langle\langle f_{D_{ML}^{qm}}^{qm}\rangle\rangle_{qm} = \langle\langle f_{D_{ML}^{bi}}^{qm}\rangle\rangle_{qm}$. This means that the MLD of the classical system D_{ML}^{bi} performs just as well as the MLD of the quantum model D_{ML}^{qm} when applied on quantum trajectories; hence we will just refer to the classical decoder D_{ML}^{qm} as MLD in the following. As the proof is rather lengthy and quite technical, we defer it to Appendix B.

A. Results

As for the MWPM decoder, we compute the decoding probability P_{ML} for the ML decoder as a function of $p, q \in [0, 1]$ for a square space-time lattice $L = T = 41$, see Fig. 7(a). Note that the ML algorithm is computationally more expensive than the MWPM algorithm, which is why we stick to smaller systems for reliable statistics. The results are very similar to MWPM decoding. Most importantly, the decoding phase still seems to be a proper subset of the PTIM entangling phase. Contrary to the MWPM decoder, we encounter numerical instabilities for very low error rates $p \lesssim 0.05$ where the results become erratic and seemingly random. This instability is a consequence of the very small probabilities of specific error patterns and has already been noted in Ref. [26]. We evade this technical issue by omitting the unreliable data in Fig. 7(a), as we are mainly interested in the phase boundary anyway.

To check for finite-size effects, we plot P_{ML} on the diagonal $p = q$ as a function of the system size $L = T = 11, \ldots, 51$, Fig. 7(b). As for the MWPM decoder, there is a clear crossing with only a small finite-size shift (inset); the transition gets sharper for larger systems and indicates an error threshold $p_{thr}^{ML} = q_{thr}^{ML} \approx 0.336$, slightly larger than the MWPM threshold $p_{thr}^{MWPM} \approx 0.324$. This difference is small but nonetheless shows that the ML decoder performs slightly better than the MWPM decoder, as expected. However, these findings also show that the MWPM decoder is already a near-optimal decoder and cannot be improved significantly. The rather small improvement of ML over MWPM decoding is in agreement with previous results for surface codes [26,50].

VI. SUMMARY

In this paper, we studied the error correction capabilities of the PTIM, a stochastic model of two competing classes of projective measurements that is characterized by an entanglement transition. We interpret the competing measurements as stabilizer measurements of the quantum repetition code and error measurements by the environment, respectively. In the entangling phase of the PTIM, the system acts as a quantum memory that preserves the amplitudes of a single logical qubits from being accessed by the environment. Our goal was to study methods for retrieving these amplitudes without having access to the measurements by the environment.

In a first attempt, we generalized the MVD—which is known to work for (quantum) repetition codes—to our setting. Numerical results suggested that this approach fails for generic parameters of the PTIM. We provided an intuition for this failure and used it as a starting point to construct

our second decoding algorithm based on MWPM. This decoder makes use of the full syndrome data and numerics revealed that it successfully retrieves the encoded amplitudes for a nontrivial range of parameters which, however, does not exhaust the complete entangling phase of the PTIM. This result suggested the existence of a parameter regime where quantum amplitudes cannot be accessed by the environment but, at the same time, remain inaccessible without having full access to the state of the system. To assess this hypothesis rigorously, we introduced a third decoder—the MLD—and showed that it improves only slightly on MWPM. Because the MLD is the optimal decoder for our system, we concluded that there is indeed an intermediary regime where the encoded amplitudes are neither accessible to the environment (through error measurement) nor to the observer (through syndrome measurements). This result also shows that it is impossible to pinpoint the entanglement transition by measuring the syndrome data alone. However, from a practical point of view, the MWPM decoder seems to be the better choice because it is already a near-optimal and computationally less expensive than the MLD.

APPENDIX A: CLASSICAL SIMULATION OF THE QUANTUM SYSTEM

In the main text, we used the quantity $P_D(p, q; L, T)$ to evaluate the performance of decoders. Here we intend to simulate realistic quantum systems to test our decoders and evaluate the decoding probability $P_D^{\mathrm{qm}}(p, q; L, T) = \langle\langle f_D^{\mathrm{qm}}\rangle\rangle$, which is a functional of the decoder D. In the following, we will prove that $\langle\langle f_D^{\mathrm{qm}}\rangle\rangle_{\mathrm{qm}} = \langle\langle f_D^{\mathrm{bi}}\rangle\rangle_{\mathrm{bi}}$, where we also distinguish between the sampling of a quantum and a classical system. This allows us to evaluate the performance of a decoder D on the quantum system by simulating classical systems with bits.

Before we start the proof, we remind our readers that any quantum system is defined by the pattern of error measurements E^p, the pattern of stabilizer measurements S^p, and the corresponding measurement results E^r and S^r. However, in our use case the results of the error measurements are not important, so we will drop them in our calculations. On the other hand, a classical trajectory of bits is fully determined by the pattern of bit flips $\tilde{E}^p$ but there still exists a pattern of stabilizer measurements S^p and the corresponding results S^r. To facilitate our calculations, we want to introduce the number of measurements in a pattern which we denote as $|S^p|$, $|E^p|$, and $|\tilde{E}^p|$. Furthermore, we introduce the concept of reduced measurement patterns E_{red}^p and S_{red}^p, which are subsets of the original patterns but only include the measurements without a predetermined measurement result. For example, measuring a stabilizer twice in a row will not change the result, so the second measurement would not be part of the reduced pattern.

For any system, the sample average is

$$\langle\langle f_D^\alpha\rangle\rangle_\alpha = \sum_{\mathcal{T}^\alpha \in \{\mathcal{T}^\alpha\}} P^\alpha(\mathcal{T}^\alpha) f^\alpha(\mathcal{T}^\alpha; D(S)), \quad (A1)$$

where $\mathcal{T}^\alpha$ indicates a trajectory in the system α and $\{\mathcal{T}^\alpha\}$ is the set of all possible trajectories. $P^\alpha(\mathcal{T}^\alpha)$ is the probability of $\mathcal{T}^\alpha$ to occur in the system α and f^α denotes the evaluation function for D in α, where we make clear that the decoder $D(S)$ is a function which depends only on the stabilizers S which are part of the trajectory.

Now focus on the quantum system. The probability of a single trajectory $\mathcal{T}^{\mathrm{qm}}$ occurring is

$$P^{\mathrm{qm}}(\mathcal{T}^{\mathrm{qm}}) := \underbrace{p^{|E^p|}(1-p)^{LT-|E^p|}}_{P^{\mathrm{qm}}(E^p)} \underbrace{q^{(L-1)T-|S^p|}(1-q)^{|S^p|}}_{P^{\mathrm{qm}}(S^p)}$$

$$\times \underbrace{\left(\frac{1}{2}\right)^{|S_{\mathrm{red}}^p|}}_{P^{\mathrm{qm}}(S^r|E^p, S^p)}, \quad (A2)$$

with the number of stabilizer and error measurements $|S^p|$ and $|E^p|$. We can insert this for the sample average:

$$\langle\langle f_D^{\mathrm{qm}}\rangle\rangle = \sum_{\mathcal{T}^{\mathrm{qm}} \in \{\mathcal{T}^{\mathrm{qm}}\}} P^{\mathrm{qm}}(\mathcal{T}^{\mathrm{qm}}) f^{\mathrm{qm}}(\mathcal{T}^{\mathrm{qm}}; D(S)) \quad (A3)$$

$$= \sum_{E^p} P^{\mathrm{qm}}(E^p) \sum_{S^p} P^{\mathrm{qm}}(S^p) \sum_{S_{|E^p, S^p}^r} P^{\mathrm{qm}}(S^r|E^p, S^p)$$

$$\times f^{\mathrm{qm}}(\mathcal{T}^{\mathrm{qm}}; D(S)). \quad (A4)$$

Here we sample the trajectories by first sampling over all error patterns, then sampling over all stabilizer patterns, and lastly sampling over the different possible results of the stabilizer measurements. (The results of the error measurements have no effect on any of our observations and can thus be ignored.) We can do the same for a classical system of bits. Here the probability of a classical trajectory $\mathcal{T}^{\mathrm{bi}}$ is

$$P^{\mathrm{bi}}(\mathcal{T}^{\mathrm{bi}}) = \underbrace{\left(\frac{p}{2}\right)^{|\tilde{E}^p|}\left(1-\frac{p}{2}\right)^{LT-|\tilde{E}^p|}}_{P^{\mathrm{bi}}(\tilde{E}^p)}$$

$$\times \underbrace{q^{(L-1)T-|S^p|}(1-q)^{|S^p|}}_{P^{\mathrm{bi}}(S^p)}. \quad (A5)$$

While it is possible to sample bit flip patterns $\tilde{E}^p$ with the probability $p/2$ on every site in every time step, it will prove useful to use a different approach. The idea is that bit flips can also be sampled by first sampling a pattern E^p of *potential* bit flips using the probability p, and successively sampling $\tilde{E}^p$ from E^p using the probability $1/2$. To do so, we add the pattern E^p to the trajectory $\mathcal{T}^{\mathrm{bi}}$ and define the following adjusted probability:

$$P^{\mathrm{bi}}(\mathcal{T}^{\mathrm{bi}}; E^p) := \underbrace{p^{|E^p|}(1-p)^{LT-|E^p|}}_{P^{\mathrm{bi}}(E^p)} \underbrace{q^{(L-1)T-|S^p|}(1-q)^{|S^p|}}_{P^{\mathrm{bi}}(S^p)}$$

$$\times \underbrace{\left(\frac{1}{2}\right)^{|E^p|}}_{P^{\mathrm{bi}}(\tilde{E}^p|E^p)}. \quad (A6)$$

This probability is related with Eq. (A5) via

$$P^{\mathrm{bi}}(\mathcal{T}^{\mathrm{bi}}) = \sum_{E^p_{|\mathcal{T}^{\mathrm{bi}}}} P^{\mathrm{bi}}(\mathcal{T}^{\mathrm{bi}}; E^p). \tag{A7}$$

Here we take the sum over all patterns E^p which are consistent with the classical trajectory $\mathcal{T}^{\mathrm{bi}}$. This relation is proven in Appendix C.

The new quantity defined in Eq. (A6) allows us to write the classical sampling average:

$$\langle\!\langle f_D^{\mathrm{bi}} \rangle\!\rangle_{\mathrm{bi}} = \sum_{\mathcal{T}^{\mathrm{bi}}} P^{\mathrm{bi}}(\mathcal{T}^{\mathrm{bi}}) f^{\mathrm{bi}}(\mathcal{T}^{\mathrm{bi}}; D(S)) \tag{A8}$$

$$= \sum_{S^p} \sum_{\tilde{E}^p} \sum_{E^p_{|\mathcal{T}^{\mathrm{bi}}}} P^{\mathrm{bi}}(\mathcal{T}^{\mathrm{bi}}; E^p) f^{\mathrm{bi}}(\mathcal{T}^{\mathrm{bi}}; D(S)) \tag{A9}$$

$$= \sum_{E^p} P^{\mathrm{bi}}(E^p) \sum_{S^p} P^{\mathrm{bi}}(S^p) \tag{A10}$$

$$\times \sum_{\tilde{E}^p \in E^p} P^{\mathrm{bi}}(\tilde{E}^p | E^p) f^{\mathrm{bi}}(\mathcal{T}^{\mathrm{bi}}; D(S)). \tag{A11}$$

Here we used the fact that $\sum_{\tilde{E}^p} \sum_{E^p_{|\tilde{E}^p}} \cdots = \sum_{E^p} \sum_{\tilde{E}^p_{|E^p}} \cdots$ and wrote $\sum_{E^p_{|\mathcal{T}^{\mathrm{bi}}}} \cdots = \sum_{E^p_{|\tilde{E}^p}} \cdots$ and $\sum_{\tilde{E}^p_{|E^p}} \cdots = \sum_{\tilde{E}^p \in E^p} \cdots$.

Comparing the quantum system in Eq. (A4) and the classical system in Eq. (A11), we see that the samplings only differ in the last sum, where in the quantum case the sampling goes over the possible stabilizer results and in the classical case the sampling goes over the possible bit flips. In the following, we will show that the following equation:

$$\sum_{S^r_{|E^p, S^p}} P^{\mathrm{qm}}(S^r | E^p, S^p) f^{\mathrm{qm}}(\mathcal{T}^{\mathrm{qm}}; D(S))$$

$$= \sum_{\tilde{E}^p \in E^p} P^{\mathrm{bi}}(\tilde{E}^p | E^p) f^{\mathrm{bi}}(\mathcal{T}^{\mathrm{bi}}; D(S)) \tag{A12}$$

holds. As a first step, we insert the probabilities defined above:

$$\sum_{S^r_{|E^p, S^p}} \left(\frac{1}{2}\right)^{|S^p_{\mathrm{red}}|} f^{\mathrm{qm}}(\mathcal{T}^{\mathrm{qm}}; D(S))$$

$$= \sum_{\tilde{E}^p \in E^p} \left(\frac{1}{2}\right)^{|E^p|} f^{\mathrm{bi}}(\mathcal{T}^{\mathrm{bi}}; D(S)). \tag{A13}$$

First, we focus on the classical right-hand side and make use of the fact that flipping a bit twice with a $50\,\%$ probability yields the same sampling as flipping it once with the same probability. This means that we can sample $\tilde{E}^p$ just on the reduced error pattern E^p_{red} without changing the result. On the quantum mechanical left-hand side of the equation, we can see that the sampling over all possible stabilizer measurement results is in reality just a sampling over the reduced stabilizer results. All other results are determined by that. Thus,

we find

$$\sum_{S^r_{\mathrm{red}|E^p, S^p}} \left(\frac{1}{2}\right)^{|S^p_{\mathrm{red}}|} f^{\mathrm{qm}}(\mathcal{T}^{\mathrm{qm}}; D(S))$$

$$= \sum_{\tilde{E}^p \in E^p_{\mathrm{red}}} \left(\frac{1}{2}\right)^{|E^p_{\mathrm{red}}|} f^{\mathrm{bi}}(\mathcal{T}^{\mathrm{bi}}; D(S)). \tag{A14}$$

We know that the patterns E^p and S^p determine whether or not the original cluster survives in the quantum system. Therefore, we can consider two cases now.

(1) *Cluster survives.* If the sampling of E^p and S^p is such that the original Bell cluster in the quantum mechanical system survives, it is $|S^p_{\mathrm{red}}| = |E^p_{\mathrm{red}}|$. Both sums have an equal number of terms and we can identify them with each other in pairs, as we will explain now. First note that if the original cluster survives, the evaluation function f^{qm} either returns value 1 or 0. Consider one term on the left side of Eq. (A14), which corresponds to a single quantum mechanical trajectory [see Fig. 8(a)]. On the plaquettes of $\mathcal{L}$ on which the original cluster lives, we can easily determine the spin orientations as they are polarized in z direction. This is the area marked gray in Fig. 8(b). As we find the final state $|\psi_{\mathrm{f}}\rangle$ of the trajectory, this process already determines the value of the evaluation function $f^{\mathrm{qm}}(\mathcal{T}^{\mathrm{qm}}; D(S))$. Note that the trajectory on the gray plaquettes in Fig. 8(b) has classical characteristics. To associate the quantum trajectory with a classical trajectory, the spin orientations on gray plaquettes can be artificially backtracked in time [without crossing error measurements, see Fig. 8(c)]. Any missing plaquettes can be eliminated by deleting all trivial error measurements $E^p \setminus E^p_{\mathrm{red}}$ from $\mathcal{L}$. This allows us to backtrack those plaquettes too and assign every position in $\mathcal{L}$ with a spin orientation in the z direction (which is *not* the real quantum trajectory). However, it is clear that this spin configuration can be considered as a valid trajectory on classical bits and be found as one term on the right side of Eq. (A14), which obviously shares the same value of the evaluation function $f^{\mathrm{qm}}(\mathcal{T}^{\mathrm{qm}}; D(S)) = f^{\mathrm{bi}}(\mathcal{T}^{\mathrm{bi}}; D(S))$. This procedure allows us to determine a corresponding classical trajectory for every quantum trajectory and thus pair up terms on both sides of Eq. (A14) which share the same value. Thus, Eq. (A14) holds.

(1) *Cluster does not survive.* If the sampling of E^p and S^p is such that the original Bell cluster in the quantum mechanical system does not survive, we already know that [for any reasonable decoder $D(S)$] $f^{\mathrm{qm}}(\mathcal{T}^{\mathrm{qm}}, D(S)) = 1/2$. On the other hand, it is $|S^p_{\mathrm{red}}| = |E^p_{\mathrm{red}}| - 1$. We insert those relations into Eq. (A14) and find

$$\sum_{S^r_{\mathrm{red}|E^p, S^p}} \left(\frac{1}{2}\right)^{|S^p_{\mathrm{red}}|} \frac{1}{2} = \sum_{\tilde{E}^p \in E^p_{\mathrm{red}}} \left(\frac{1}{2}\right)^{|S^p_{\mathrm{red}}+1|} f^{\mathrm{bi}}(\mathcal{T}^{\mathrm{bi}}; D(S)). \tag{A15}$$

However, there are twice as many classical trajectories than there are quantum trajectories and half of them are evaluated as $f^{\mathrm{bi}}(\mathcal{T}^{\mathrm{bi}}; D(S)) = 1$ (for the other half of the trajectories, the evaluation function vanishes). Therefore, also in this case Eq. (A14) holds.
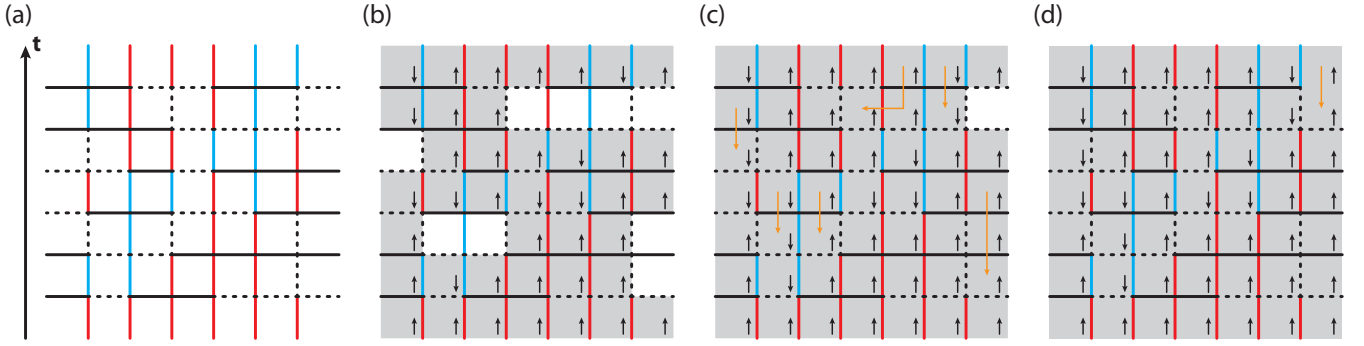
FIG. 8. Visual guide to identify terms in Eq. (A14) if the original Bell cluster survives in the trajectory. (a) Here we show a grid $\mathcal{L}$ with the measurements of a quantum mechanical trajectory. The colors are defined as in Fig. 3. (b) Wherever the original product state lives (gray plaquettes), squares can be associated with a spin orientation (we indicate them with arrows). (c) By backtracking the spin orientations in time, more plaquettes of $\mathcal{L}$ can be associated with a spin orientation (which does not correspond to the quantum trajectory). (d) All remaining plaquettes can be backtracked by removing all $E^p \backslash E_{red}^p$ from $\mathcal{L}$. The associated configuration of spins on $\mathcal{L}$ can be identified with a classical trajectory in Eq. (A14).

This proves, that $\langle\langle f_D^{qm} \rangle\rangle_{qm} = \langle\langle f_D^{bi} \rangle\rangle_{bi}$ independent on the decoder D. As a consequence the performance of any decoder D on the quantum system, can be evaluated on classical systems.

APPENDIX B: PROOF OF THE MAXIMUM LIKELIHOOD DECODER

We will prove that the MLD introduced in Sec. V, for which we assumed the system to be entirely classical with bit flips instead of error measurements, deserves its name also in the quantum mechanical system. Thus we will prove that the decoder yields the most likely quantum mechanical spin configuration and therefore gives us the best probability of decoding the system that can be achieved.

1. Some thoughts in advance

Consider two discrete variables A and B, a function $P(B) \geqslant 0 \forall B$ and a function $f(A; B)$. Then we can find an upper bound to the following supremum:

$$\sup_A \left[\sum_B P(B) f(A; B) \right] \leqslant \sum_B P(B) \sup_A [f(A; B)]. \quad (B1)$$

Here we consider the sum over all B and find the supremum out of all values of A.

However, if we now consider $A(B)$ to be a function out of the set of all functions that take the values B as input, we can actually reach the upper bound:

$$\sup_A \left[\sum_B P(B) f(A(B); B) \right] = \sum_B P(B) \sup_A [f(A(B); B)]. \quad (B2)$$

This is due to the fact that we can construct a function $A(B)$ that for every value of B maximizes the function f (no matter the form of f). We will use this later.

2. The proof

In general, we evaluate a decoder $D(S)$ in a system α via the function

$$P_D^\alpha(p, q; L, T) = \langle\langle f_D^\alpha \rangle\rangle_\alpha, \quad (B3)$$

which is the average value of the evaluation function f^α for many trajectories $\mathcal{T}^\alpha$. We can express this value by summing over all possible trajectories $\{\mathcal{T}^\alpha\}$ with the probability $P^\alpha(\mathcal{T}^\alpha)$ of a single trajectory occurring:

$$\langle\langle f_D^\alpha \rangle\rangle_\alpha = \sum_{\mathcal{T}^\alpha \in \{\mathcal{T}^\alpha\}} P^\alpha(\mathcal{T}^\alpha) f^\alpha(\mathcal{T}^\alpha; D(S)). \quad (B4)$$

In the following, we will show that a MLD as defined in using Eq. (14) maximizes the sample average $\langle\langle f_D^\alpha \rangle\rangle$:

$$\sup_D \left[\langle\langle f_D^\alpha \rangle\rangle_\alpha \right] = \sup_D \left[\sum_{\mathcal{T}^\alpha \in \{\mathcal{T}^\alpha\}} P^\alpha(\mathcal{T}^\alpha) f^\alpha(\mathcal{T}^\alpha; D(S)) \right]. \quad (B5)$$

It is important to note that here we consider all possible decoders of the form $D(S)$ which take stabilizer measurements S as an input and return a correction string. Now we realize that the trajectories $\mathcal{T}^\alpha$ can be grouped into equivalence classes $[\mathcal{T}] = S(\mathcal{T}^\alpha)$ via their respective stabilizers S. This allows us to rewrite the sum by first considering all possible stabilizers S and for each one of them summing over all trajectories:

$$= \sup_D \left[\sum_S \sum_{\mathcal{T}^\alpha \in S} P^\alpha(\mathcal{T}^\alpha) f^\alpha(\mathcal{T}^\alpha; D(S)) \right]. \quad (B6)$$

In the following, we use Eq. (B2) to take the supremum into the outer sum. We can do this because we just require the decoder D to be any function that takes stabilizers S as input and returns correction strings. Thus, we find

$$= \sum_S \sup_D \left[\sum_{\mathcal{T}^\alpha \in S} P^\alpha(\mathcal{T}^\alpha) f^\alpha(\mathcal{T}^\alpha; D(S)) \right].$$

As the decoder $D(S)$ is just a function that returns a correction string C depending on the input S and we just fixed S, we can

now just write the supremum as

$$= \sum_S \sup_C \left[\sum_{\mathcal{T}^\alpha \in S} P^\alpha(\mathcal{T}^\alpha) f^\alpha(\mathcal{T}^\alpha; C) \right].$$

We can now consider our definition of a MLD in Eq. (14) and rewrite as follows: *For a given S, choose the correction string C_i such that*

$$P^\alpha_{f^\alpha}(C_i | S) P(S^p) = \sum_{\mathcal{T}^\alpha \in S} P^\alpha(\mathcal{T}^\alpha) f^\alpha(\mathcal{T}^\alpha; C_i) \quad \text{(B7)}$$

is maximized. Here the sum does only consider trajectories $\mathcal{T}^\alpha \in S$. We can thus use the MLD $D^\alpha_{\mathrm{ML}}(S)$ to finalize our calculation:

$$\sup_D \left[\langle\langle f^\alpha_D \rangle\rangle_\alpha \right] = \sum_S \sum_{\mathcal{T}^\alpha \in S} P^\alpha(\mathcal{T}^\alpha) f^\alpha\left(\mathcal{T}^\alpha; \mathcal{D}^\alpha_{\mathrm{ML}}(\mathcal{S}) \right)$$

$$= \langle\langle f^\alpha_{D^\alpha_{\mathrm{ML}}} \rangle\rangle_\alpha. \quad \text{(B8)}$$

From Appendix A, we know that $\langle\langle f^{\mathrm{bi}}_D \rangle\rangle_{\mathrm{bi}} = \langle\langle f^{\mathrm{qm}}_D \rangle\rangle_{\mathrm{qm}}$. Thus, we find the following connection between the classical and the quantum MLD. It is

$$\sup_D \left[\langle\langle f^{\mathrm{qm}}_D \rangle\rangle_{\mathrm{qm}} \right] = \langle\langle f^{\mathrm{qm}}_{D^{\mathrm{qm}}_{\mathrm{ML}}} \rangle\rangle_{\mathrm{qm}} \quad \text{(B9)}$$

and

$$\sup_D \left[\langle\langle f^{\mathrm{bi}}_D \rangle\rangle_{\mathrm{bi}} \right] = \langle\langle f^{\mathrm{bi}}_{D^{\mathrm{bi}}_{\mathrm{ML}}} \rangle\rangle_{\mathrm{bi}}. \quad \text{(B10)}$$

Using Appendix A, we find

$$\langle\langle f^{\mathrm{qm}}_{D^{\mathrm{qm}}_{\mathrm{ML}}} \rangle\rangle_{\mathrm{qm}} = \sup_D \left[\langle\langle f^{\mathrm{qm}}_D \rangle\rangle_{\mathrm{qm}} \right] \quad \text{(B11)}$$

$$= \sup_D \left[\langle\langle f^{\mathrm{bi}}_D \rangle\rangle_{\mathrm{bi}} \right] \quad \text{(B12)}$$

$$= \langle\langle f^{\mathrm{bi}}_{D^{\mathrm{bi}}_{\mathrm{ML}}} \rangle\rangle_{\mathrm{bi}} \quad \text{(B13)}$$

$$= \langle\langle f^{\mathrm{qm}}_{D^{\mathrm{bi}}_{\mathrm{ML}}} \rangle\rangle_{\mathrm{qm}}. \quad \text{(B14)}$$

This proves that the classical MLD yields the exact same results as the quantum mechanical MLD. ∎

APPENDIX C: PROOF OF THE SAMPLING IN EQ. (A6)

Here we show a short proof for the sampling discussed in Eq. (A6). We do so by inserting Eq. (A6) in Eq. (A7) to obtain Eq. (A5).

Consider the term

$$\frac{1}{P^{\mathrm{bi}}(S^p)} \sum_{E^p_{|\mathcal{T}^{\mathrm{bi}}}} P^{\mathrm{bi}}(\mathcal{T}^{\mathrm{bi}}; E^p), \quad \text{(C1)}$$

where we take the sum over all error patterns E^p that contain the bit flip pattern $\tilde{E}^p$ which is part of $\mathcal{T}^{\mathrm{bi}}$. Inserting Eq. (A6) yields

$$= \sum_{E^p_{|\mathcal{T}^{\mathrm{bi}}}} p^{|E^p|} (1-p)^{LT-|E^p|} \left(\frac{1}{2} \right)^{|E^p|} \quad \text{(C2)}$$

$$= \sum_{E^p_{|\mathcal{T}^{\mathrm{bi}}}} \left(\frac{p}{2} \right)^{|E^p|} (1-p)^{LT-|E^p|} \quad \text{(C3)}$$

$$= \left(\frac{p}{2} \right)^{|\tilde{E}^p|} (1-p)^{LT-|\tilde{E}^p|} \sum_{E^p_{|\mathcal{T}^{\mathrm{bi}}}} \left(\frac{p}{2(1-p)} \right)^{|E^p|-|\tilde{E}^p|}. \quad \text{(C4)}$$

Now consider the summation over all possible patterns E^p consistent with the trajectory $\mathcal{T}^{\mathrm{bi}}$. This is equivalent to a summation over all virtual patterns $E^p_{\mathrm{v}} := E^p \setminus \tilde{E}^p$. This pattern is constructed such that $0 \leqslant |E^p_{\mathrm{v}}| = |E^p| - |\tilde{E}^p| \leqslant LT - |\tilde{E}^p|$ and allows us to rewrite the sampling as

$$= \left(\frac{p}{2} \right)^{|\tilde{E}^p|} (1-p)^{LT-|\tilde{E}^p|} \sum_{E^p_{\mathrm{v}}} \left(\frac{p}{2(1-p)} \right)^{|E^p_{\mathrm{v}}|}. \quad \text{(C5)}$$

It becomes clear now that sampling over all possible virtual patterns E^p_{v} is equivalent to choosing all possible subsets of a set with $LT - |\tilde{E}^p|$ elements. This can be formulated as

$$= \left(\frac{p}{2} \right)^{|\tilde{E}^p|} (1-p)^{LT-|\tilde{E}^p|}$$

$$\times \sum_{k=0}^{LT-|\tilde{E}^p|} \binom{LT-|\tilde{E}^p|}{k} \left(\frac{p}{2(1-p)} \right)^k \quad \text{(C6)}$$

$$= \left(\frac{p}{2} \right)^{LT} \sum_{k=0}^{LT-|\tilde{E}^p|} \binom{LT-|\tilde{E}^p|}{k} 1^k \left(\frac{2(1-p)}{p} \right)^{(LT-|\tilde{E}^p|)-k}. \quad \text{(C7)}$$

Using the binomial theorem, we find

$$= \left(\frac{p}{2} \right)^{LT} \left(1 + \frac{2(1-p)}{p} \right)^{LT-|\tilde{E}^p|} \quad \text{(C8)}$$

$$= \left(\frac{p}{2} \right)^{|\tilde{E}^p|} \left(1 - \frac{p}{2} \right)^{LT-|\tilde{E}^p|} \quad \text{(C9)}$$

$$= \frac{P^{\mathrm{bi}}(\mathcal{T}^{\mathrm{bi}})}{P^{\mathrm{bi}}(S^p)}. \quad \text{(C10)}$$

This proves the relation in Eq. (A7) and thus proves that the sampling in Eq. (A6) is correct. ∎

APPENDIX D: THERMODYNAMIC LIMIT OF THE MWPM DECODER FOR $q = 0$

Here we want to shortly discuss the performance of the MWPM decoder for $q = 0$ in the thermodynamic limit. For this special case, analytical calculations can be made.

For $q = 0$, all stabilizers are measured and the lattice produced by the MWPM decoder as shown in Fig. 5(b) only features unconnected horizontal lines. Essentially, the MWPM decoder is now a MVD as discussed in Sec. III. Now if in a classical system of bits the number of flips in a single step randomly surpasses $L/2$ in a time step (which has a nonzero likelihood in finite systems), the decoder will match the nodes falsely and fail in this time step. If this happens in an even number of time steps, the decoder yields the correct result after the last step, otherwise it fails. Therefore, we can
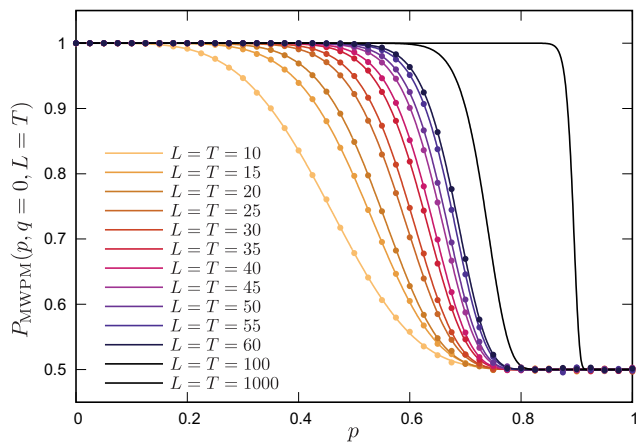
FIG. 9. Analytical investigation of the MWPM decoder for the case $q = 0$. The lines in this plot show the analytical results of the MWPM decoder in the case $q = 0$. The data points are gained from simulations with 100 000 samples.

calculate the decoding probability analytically via

$$P_{D_{\mathrm{MWPM}}}(p, q = 0; L, T)$$

$$= \sum_{t=0, t \text{ even}}^{T} \binom{T}{t} (P(L, p))^t (1 - P(L, p))^{T-t}, \qquad \text{(D1a)}$$

$$P(L, p)$$

$$= \sum_{b=\lceil L/2 \rceil}^{L} \binom{L}{b} \left(\frac{p}{2}\right)^b \left(1 - \frac{p}{2}\right)^{L-b} \begin{cases} \frac{1}{2} & \text{if } b = \frac{L}{2} \\ 1 & \text{else} \end{cases}. \qquad \text{(D1b)}$$

The function $P(L, p)$ calculates the probability of more than $L/2$ bits being flipped in a single time step. If exactly $L/2$ bits are flipped, the MWPM decoder randomly guesses correctly or falsely. Therefore, we multiply this probability by $1/2$.

The analytical function was plotted for different system sizes in Fig. 9. The dots in the plot represent simulation data. They clearly fit the theory. As can be seen, the decoding probability exhibits a strong finite-size effect but in the thermodynamic limit $L = T \to \infty$, the decoding probability slowly approaches 1 for all $p < 1$.

[1] B. M. Terhal, Quantum error correction for quantum memories, Rev. Mod. Phys. **87**, 307 (2015).

[2] B. Skinner, J. Ruhman, and A. Nahum, Measurement-Induced Phase Transitions in the Dynamics of Entanglement, Phys. Rev. X **9**, 031009 (2019).

[3] E. Knill, R. Laflamme, and W. Zurek, Threshold accuracy for quantum computation, arXiv:quant-ph/9610011.

[4] D. Aharonov and M. Ben-Or, Fault-tolerant quantum computation with constant error, in *Proceedings of the Twenty-Ninth Annual ACM Symposium on Theory of Computing, STOC'97*, (Association for Computing Machinery, New York, 1997), p. 176–188.

[5] P. W. Shor, Algorithms for quantum computation: Discrete logarithms and factoring, in *Proceedings of the 35th Annual Symposium on Foundations of Computer Science, Santa Fe, NM, USA* (IEEE, 1994), pp. 124–134.

[6] Y. Li, X. Chen, and M. P. A. Fisher, Quantum Zeno effect and the many-body entanglement transition, Phys. Rev. B **98**, 205136 (2018).

[7] A. Chan, R. M. Nandkishore, M. Pretko, and G. Smith, Unitary-projective entanglement dynamics, Phys. Rev. B **99**, 224307 (2019).

[8] M. Szyniszewski, A. Romito, and H. Schomerus, Entanglement transition from variable-strength weak measurements, Phys. Rev. B **100**, 064204 (2019).

[9] Y. Li, X. Chen, and M. P. A. Fisher, Measurement-driven entanglement transition in hybrid quantum circuits, Phys. Rev. B **100**, 134306 (2019).

[10] M. J. Gullans and D. A. Huse, Dynamical Purification Phase Transition Induced by Quantum Measurements, Phys. Rev. X **10**, 041020 (2020).

[11] Q. Tang and W. Zhu, Measurement-induced phase transition: A case study in the nonintegrable model by density-matrix renormalization group calculations, Phys. Rev. Res. **2**, 013022 (2020).

[12] C.-M. Jian, Y.-Z. You, R. Vasseur, and A. W. W. Ludwig, Measurement-induced criticality in random quantum circuits, Phys. Rev. B **101**, 104302 (2020).

[13] X. Turkeshi, R. Fazio, and M. Dalmonte, Measurement-induced criticality in $(2 + 1)$-dimensional hybrid quantum circuits, Phys. Rev. B **102**, 014315 (2020).

[14] Y. Bao, S. Choi, and E. Altman, Theory of the phase transition in random unitary circuits with measurements, Phys. Rev. B **101**, 104301 (2020).

[15] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University Press, Cambridge, UK, 2010).

[16] P. W. Shor, Scheme for reducing decoherence in quantum computer memory, Phys. Rev. A **52**, R2493 (1995).

[17] A. Steane, Multiple-particle interference and quantum error correction, Proc. R. Soc. London, Ser. A **452**, 2551 (1996).

[18] S. B. Bravyi and A. Y. Kitaev, Quantum codes on a lattice with boundary, arXiv:quant-ph/9811052.

[19] A. Kitaev, Fault-tolerant quantum computation by anyons, Ann. Phys. **303**, 2 (2003).

[20] M. H. Freedman and D. A. Meyer, Projective plane and planar quantum codes, Found. Comput. Math. **1**, 325 (2001).

[21] E. Dennis, A. Kitaev, A. Landahl, and J. Preskill, Topological quantum memory, J. Math. Phys. **43**, 4452 (2002).

[22] G. A. Paz-Silva, G. K. Brennen, and J. Twamley, Fault Tolerance with Noisy and Slow Measurements and Preparation, Phys. Rev. Lett. **105**, 100501 (2010).

[23] D. Crow, R. Joynt, and M. Saffman, Improved Error Thresholds for Measurement-Free Error Correction, Phys. Rev. Lett. **117**, 130503 (2016).

[24] A. G. Fowler, A. M. Stephens, and P. Groszkowski, High-threshold universal quantum computation on the surface code, Phys. Rev. A **80**, 052312 (2009).

[25] J. R. Wootton and D. Loss, High Threshold Error Correction for the Surface Code, Phys. Rev. Lett. **109**, 160503 (2012).

[26] S. Bravyi, M. Suchara, and A. Vargo, Efficient algorithms for maximum likelihood decoding in the surface code, Phys. Rev. A **90**, 032326 (2014).

[27] A. G. Fowler, Proof of Finite Surface Code Threshold for Matching, Phys. Rev. Lett. **109**, 180502 (2012).

[28] D. S. Wang, A. G. Fowler, A. M. Stephens, and L. C. L. Hollenberg, Threshold error rates for the toric and planar codes, Quantum Inf. Comput. **10**, 456 (2010).

[29] S. Krinner, N. Lacroix, A. Remm, A. Di Paolo, E. Genois, C. Leroux, C. Hellings, S. Lazar, F. Swiadek, J. Herrmann, G. J. Norris, C. K. Andersen *et al.*, Realizing repeated quantum error correction in a distance-three surface code, Nature **605**, 669 (2022).

[30] Y. Zhao, Y. Ye, H.-L. Huang, Y. Zhang, D. Wu, H. Guan, Q. Zhu, Z. Wei, T. He, S. Cao, F. Chen, T.-H. Chung *et al.*, Realization of an Error-Correcting Surface Code with Superconducting Qubits, Phys. Rev. Lett. **129**, 030501 (2022).

[31] D. Gottesman, Class of quantum error-correcting codes saturating the quantum hamming bound, Phys. Rev. A **54**, 1862 (1996).

[32] D. Gottesman, Stabilizer codes and quantum error correction, arXiv:quant-ph/9705052.

[33] D. Gottesman, Theory of fault-tolerant quantum computation, Phys. Rev. A **57**, 127 (1998).

[34] D. Gottesman, The Heisenberg representation of quantum computers, arXiv:quant-ph/9807006.

[35] S. Aaronson and D. Gottesman, Improved simulation of stabilizer circuits, Phys. Rev. A **70**, 052328 (2004).

[36] A. Nahum and B. Skinner, Entanglement and dynamics of diffusion-annihilation processes with Majorana defects, Phys. Rev. Res. **2**, 023288 (2020).

[37] N. Lang and H. P. Büchler, Entanglement transition in the projective transverse field Ising model, Phys. Rev. B **102**, 094204 (2020).

[38] M. Ippoliti, M. J. Gullans, S. Gopalakrishnan, D. A. Huse and V. Khemani, Entanglement Phase Transitions in Measurement-Only Dynamics, Phys. Rev. X **11**, 011030 (2021).

[39] A. Lavasani, Y. Alavirad, and M. Barkeshli, Measurement-induced topological entanglement transitions in symmetric random quantum circuits, Nat. Phys. **17**, 342 (2021).

[40] S. Sang and T. H. Hsieh, Measurement-protected quantum phases, Phys. Rev. Res. **3**, 023200 (2021).

[41] D. Aharonov, Quantum to classical phase transition in noisy quantum computers, Phys. Rev. A **62**, 062311 (2000).

[42] S. Choi, Y. Bao, X.-L. Qi, and E. Altman, Quantum Error Correction in Scrambling Dynamics and Measurement-Induced Phase Transition, Phys. Rev. Lett. **125**, 030505 (2020).

[43] Y. Li and M. P. A. Fisher, Statistical mechanics of quantum error correcting codes, Phys. Rev. B **103**, 104306 (2021).

[44] R. Fan, S. Vijay, A. Vishwanath, and Y.-Z. You, Self-organized error correction in random unitary circuits with measurement, Phys. Rev. B **103**, 174309 (2021).

[45] Y. Nakata, E. Wakakuwa, and H. Yamasaki, One-shot quantum error correction of classical and quantum information, Phys. Rev. A **104**, 012408 (2021).

[46] M. J. Gullans, S. Krastanov, D. A. Huse, L. Jiang, and S. T. Flammia, Quantum Coding with Low-Depth Random Circuits, Phys. Rev. X **11**, 031066 (2021).

[47] L. Fidkowski, J. Haah, and M. B. Hastings, How dynamical quantum memories forget, Quantum **5**, 382 (2021).

[48] Y. Li and M. P. A. Fisher, Robust decoding in monitored dynamics of open quantum systems with Z_2 symmetry, arXiv:2108.04274 (2021).

[49] G. Duclos-Cianci and D. Poulin, Fast Decoders for Topological Quantum Codes, Phys. Rev. Lett. **104**, 050504 (2010).

[50] A. G. Fowler, A. C. Whiteside, and L. C. L. Hollenberg, Towards Practical Classical Processing for the Surface Code, Phys. Rev. Lett. **108**, 180501 (2012).

[51] J. Edmonds, Paths, trees, and flowers, Can. J. Math. **17**, 449 (1965).

[52] V. Kolmogorov, Blossom V: A new implementation of a minimum cost perfect matching algorithm, Math. Program. Comput. **1**, 43 (2009).