

Topological order and memory time in marginally-self-correcting quantum memoryKarthik Siva^{1,*} and Beni Yoshida^{2,†}¹*Institute for Quantum Information and Matter, California Institute of Technology, Pasadena, California 91125, USA*²*Perimeter Institute for Theoretical Physics, Waterloo, Ontario, Canada N2L 2Y5*

(Received 22 April 2016; published 22 March 2017)

We examine two proposals for marginally-self-correcting quantum memory: the cubic code by Haah and the welded code by Michnicki. In particular, we prove explicitly that they are absent of topological order above zero temperature, as their Gibbs ensembles can be prepared via a short-depth quantum circuit from classical ensembles. Our proof technique naturally gives rise to the notion of free energy associated with excitations. Further, we develop a framework for an ergodic decomposition of Davies generators in CSS codes which enables formal reduction to simpler classical memory problems. We then show that memory time in the welded code is doubly exponential in inverse temperature via the Peierls argument. These results introduce further connections between thermal topological order and self-correction from the viewpoint of free energy and quantum circuit depth.

DOI: [10.1103/PhysRevA.95.032324](https://doi.org/10.1103/PhysRevA.95.032324)**I. INTRODUCTION**

To create a usable memory system, classical or quantum, one must be able to write, store, and read out information reliably. The main difficulty in this endeavor is the protection of the information written against thermal noise during storage. Therefore, one natural question is the feasibility of self-correcting quantum memory, a device which would correct itself by dissipating energy to the thermal bath [1]. Instead of requiring active error correction, the device would only require initialization of the system in a ground state and a final round of error correction upon readout. Such a self-correcting quantum memory device could serve as the equivalent hard drive quantum computing technology. Here we define a self-correcting quantum memory as a system with local interactions, nondivergent terms, perturbative stability, and diverging quantum memory time at nonzero temperature [2]. See Refs. [3,4] for recent reviews.

The mechanism by which self-correction could be implemented is inspired by memory systems in classical physics. In classical memory, such as the two-dimensional Ising ferromagnet below a critical temperature, the energy cost associated with the erroneous flip dominates its entropic advantage so that it is likely to flip back and avoid proliferating. Thus a necessary ingredient for self-correcting quantum memory would be that the topologically ordered ground states of a quantum memory system are separated by large energy barriers. These ideas have inspired several proposals for self-correcting quantum memory [2,5–10]. However, at this time, these are not verified or are not fully self-correcting in light of the aforementioned definition. We refer to these as marginally self-correcting. Additionally, several no-go results are known [11–14]. Namely, it has been shown that systems with both translation and scale invariance cannot be self-correcting in three dimensions [15].

Implementation of codewords as topologically ordered ground states of quantum many-body systems is especially attractive as codewords store logical qubits nonlocally with

stability against local perturbations. However, the presence of topological order in the ground states is not sufficient for self-correcting behaviors. For example, the ground states of the two-dimensional toric code are topologically ordered as ground states are locally indistinguishable and cannot be prepared from product states via a short-depth circuit. Yet, due to pointlike mobile excitations associated with stringlike logical operators, it is not self-correcting [16,17]. On the other hand, the four-dimensional toric code, which is provably a self-correcting quantum memory [18], is also topologically ordered above zero temperature in a sense that its Gibbs ensemble cannot be prepared via a short-depth circuit [19]. This raises another important question concerning whether topological order can survive above zero temperature in three spatial dimensions, or in other words, whether Gibbs ensembles of three-dimensional physical systems can be prepared immediately, in light of the definition of thermal topological order proposed by Hastings [19]. It is currently not known whether the aforementioned proposals of self-correcting quantum memories are topologically ordered at nonzero temperature or not.

The question of topological order at nonzero temperature is expected to be equivalent to that of the feasibility of self-correcting quantum memory. However, the connection has been established only for simple models such as the two-dimensional and four-dimensional toric code. In this paper, we study static and dynamical properties of marginally-self-correcting quantum memory models and further establish the connection between these two notions. We shall mainly focus on two specific examples, the cubic code and the welded code, but our techniques may extend to a larger class of systems. Our work also serves as a nontrivial test of the definition of thermal topological order proposed by Hastings.

Main results. The cubic code is a three-dimensional stabilizer code with fractal-shaped logical operator and $O(\log(n))$ energy barrier. We begin by proving that the cubic code, which breaks scale invariance, is not topologically ordered above zero temperature. More precisely, we show that the thermal Gibbs ensemble of the cubic code can be approximated from some classical ensemble via an $O(\log(n))$ -range quantum circuit. (Note that nonconstant range is necessary since the approximation error is assessed by trace norm.) Further,

*ksiva@caltech.edu

†byoshida@perimeterinstitute.ca

this result extends to a larger class of quantum codes with fractal-shaped logical operators.

Although the result itself is perhaps not surprising, as evidences of thermal instability of the cubic code have been known, what is interesting is the proof technique and also the physical intuition therein. First, we introduce the concept of an imperfect Hamiltonian $H(p)$ where interaction terms are removed from a stabilizer Hamiltonian H with some probability p which is set by the temperature: $p = \frac{2}{e^{\beta} + 1}$. We then show that, if there (almost surely) exists a short-depth quantum circuit to disentangle $H(p)$, then the system is not topologically ordered at a given temperature. Second, we observe that the missing terms in an imperfect Hamiltonian $H(p)$ act as sinks for excitations so that by moving excitations to locations of missing terms, they can be eliminated from the system. This observation allows one to identify the Gibbs ensemble approximation problem as evaluations of free energies, enabling us to employ the Peierls argument to heuristically assess the absence and the presence of topological order at nonzero temperature. Finally, we prove that isolated excitations in the cubic code can be eliminated by moving them to nearby sinks. The proof utilizes polynomial representations of translation invariant stabilizer Hamiltonians in a crucial manner.

Next, we turn our attention to the welded code which is free from stringlike logical operators, has $O(n^{2/9})$ energy barrier, and breaks translation invariance. We show that quantum memory time of the welded code is upper bounded by $O(\exp[\exp(c\beta)])$ for a fixed temperature. Namely, we show that quantum memory time of the welded code is upper bounded by classical memory time of the Ising model supported on a sparse lattice (Fig. 9). We then argue that memory time can be computed from the Peierls argument. Our derivation of memory time is not mathematically rigorous, and we back up our analytical result by numerical simulations of memory time of the sparse Ising model and confirm doubly exponential behavior.

Again, although the result itself is perhaps not surprising, as relations between the one-dimensional Ising model and the two-dimensional toric code have been noticed, what is interesting is the proof technique and also the physical intuition therein. The key idea is to choose a minimal set of Davies generators that ensure ergodicity in the weak-coupling limit since the gap of the Lindblad operator (in other words, the inverse of memory time) is lower bounded by that of a minimal choice. Thermal dynamics under minimal generators can be effectively described by a simpler classical system which enables a formal reduction to a classical problem. We will provide an algorithm to find such minimal generators which applies to arbitrary CSS stabilizer codes. This generalizes the work by Alicki *et al.* on Davies generators for the two-dimensional toric code to arbitrary CSS stabilizer codes [16]. Applications of this algorithm to the welded code enable reduction to problems concerning the Ising model on a sparse graph. We also reveal that the procedure of the welding has an interesting interpretation as attaching several copies of topological codes along the same gapped boundary.

Plan of the paper. The paper is organized as follows. In Sec. II, we develop key theoretical tools to study topological order at nonzero temperature. In Sec. III, we prove that the cubic code is not topologically ordered at nonzero temperature.

In Sec. IV, we introduce an ergodic decomposition of Davies generators for CSS codes. In Sec. V, we examine memory time of the welded code and provide numerical results.

In Appendix A, we present an ergodic decomposition for the cubic code and show that its memory time is upper bounded by classical memory time of two-dimensional classical fractal spin models. In Appendix B, we present an interpretation of the welding procedure from the viewpoint of anyon condensations on gapped boundaries. In Appendix D, we discuss generic properties of an imperfect Hamiltonian in three dimensions and show that a three-dimensional imperfect Hamiltonian can be coarse-grained into the form of lattice gauge theories.

II. THERMAL TOPOLOGICAL ORDER

Let us recall the definition of topological order at nonzero temperature due to Hastings [19]. Consider a frustration-free commuting Hamiltonian on a D -dimensional lattice of size $n = L^D$:

$$H = - \sum_j Q_j, \quad (1)$$

where Q_j are projectors, $[Q_i, Q_j] = 0$, and ground states satisfy $Q_j|\psi\rangle = |\psi\rangle$ for all j . The Hamiltonian possesses topological order at zero temperature if it requires a unitary circuit with large circuit depth to create ground states. Namely, a ground state $|\psi\rangle$ is said to be (R, ϵ) trivial if there exists a unitary quantum circuit U with range R such that $|\psi - U\psi_{\text{prod}}| \leq \epsilon$ for some product state ψ_{prod} where $|\dots|$ represents the trace norm. By a range- R unitary, we mean a unitary quantum circuit U where the circuit depth multiplied by the maximum range of each unitary in the circuit is bounded by some range R [20].

Topological order at nonzero temperature is a generalization of this notion to a thermal ensemble. Let ρ_β be the Gibbs ensemble of the Hamiltonian H at $\beta = T^{-1}$ with $k_B = 1$; $\rho_\beta = \frac{e^{-\beta H}}{\text{Tr} e^{-\beta H}}$. The system is said to be topologically ordered at temperature T if it requires a large depth quantum circuit to prepare the Gibbs ensemble from a classical ensemble.

Definition 1 (Hastings [19]). Let each site j have an additional degree of freedom \mathcal{K}_j , defining an enlarged Hilbert space $\mathcal{H}_j \otimes \mathcal{K}_j$ on each site. A density matrix ρ is said to be (R, ϵ) trivial if there exists a unitary quantum circuit U with range R such that

$$|\rho_\beta - \text{Tr}_{(\mathcal{K}_j)}(U\rho_{\text{cl}}U^\dagger)| \leq \epsilon, \quad (2)$$

where ρ_{cl} is a classical state of range R , and where the trace is over the added degrees of freedom \mathcal{K}_j . A classical state of range R is defined to be a thermal ensemble such that $\rho_{\text{cl}} = \mathcal{Z}^{-1} \exp(-H_{\text{cl}})$ where H_{cl} is a Hamiltonian with range R which is diagonal in a product basis.

We shall often focus on a special class of quantum memory models known as CSS stabilizer codes (Hamiltonians). A stabilizer code encodes k logical qubits in the ground-state subspace \mathcal{L} of n physical qubits supporting a stabilizer group \mathcal{S} . States spanning \mathcal{L} are defined by their invariance under the action of elements of \mathcal{S} :

$$\mathcal{L} = \{|\psi\rangle \in (\mathbb{C}^2)^{\otimes n} : s|\psi\rangle = |\psi\rangle \forall s \in \mathcal{S}\}. \quad (3)$$

\mathcal{S} is an Abelian subgroup of the Pauli group acting on n qubits \mathcal{P}_n defined as

$$\mathcal{P}_n = \langle iI, X_1, Z_1, X_2, Z_2, \dots, X_n, Z_n \rangle. \quad (4)$$

The invariance of $|\psi\rangle$ under the action of \mathcal{S} implies $-I \notin \mathcal{S}$. Logical operators of the code are given by Pauli operators P such that for the centralizer $\mathcal{C}(\mathcal{S})$, $P = \mathcal{C}(\mathcal{S}) \setminus \mathcal{S}$. Additionally, one can always choose generators S_i of \mathcal{S} so that $\mathcal{S} = \langle S_1, S_2, \dots, S_N \rangle$, and \mathcal{L} is the space of ground states of a Hamiltonian $H = -\sum_{i=1}^N S_i$ with $N \geq n - k$. A CSS code is a stabilizer code for which every generator of \mathcal{S} can be written as a product of X -type or Z -type stabilizer elements. The corresponding Hamiltonian can take the following special form:

$$H = -\sum_{i=1}^{N_X} S_i^{(X)} - \sum_{i=1}^{N_Z} S_i^{(Z)}, \quad (5)$$

where $S_i^{(X)}$ is a product of solely Pauli X operators and $S_i^{(Z)}$ is a product of solely Pauli Z operators.

A. Imperfect Hamiltonian

To study the presence (or the absence) of topological order at nonzero temperature, it is convenient to consider an *imperfect Hamiltonian*. Consider a stabilizer Hamiltonian $H = -\sum_{j=1}^N \frac{1+S_j}{2}$ where S_j are Pauli stabilizer generators.

Definition 2. Let $k_j = 0, 1$ be binary integers for $j = 1, \dots, N$ and define the imperfect Hamiltonian

$$H(\vec{k}) = -\sum_{j=1}^N k_j S_j. \quad (6)$$

The maximally mixed ensemble in the ground state space of $H(\vec{k})$ is denoted by $\rho(\vec{k})$. Define the free ensemble $\rho_{\text{free}}(\beta)$ by

$$\rho_{\text{free}}(\beta) = \sum_{\vec{k}} \text{Prob}(\vec{k}) \cdot \rho(\vec{k}), \quad (7)$$

where

$$\text{Prob}(\vec{k}) = \prod_j (1-p)^{k_j} p^{(1-k_j)} \quad p = \frac{2}{e^\beta + 1}. \quad (8)$$

In other words, we remove interaction terms from the stabilizer Hamiltonian with some probability p . In the next subsection, we will discuss the physical intuitions of the removed terms.

The thermal ensemble ρ_β can be approximated by the free ensemble $\rho_{\text{free}}(\beta)$ under a certain condition. In general, stabilizer generators S_j are not independent since an identity operator can be formed by multiplying S_j . Let us introduce N -component binary vectors, denoted by $\vec{m}^{(i)}$ for $i = 1, \dots, 2^M$, such that

$$\prod_{j=1}^N (S_j)^{m_j^{(i)}} = I. \quad (9)$$

Here we take $\vec{m}^{(1)}$ to be $m_j^{(1)} = 0$ for all j . Note that $k = n - (N - M)$.

Definition 3. Let $w(\vec{m}^{(i)})$ be the Hamming weight (the number of nonzero entries) of a binary vector $\vec{m}^{(i)}$. A stabilizer

Hamiltonian is said to be (α, γ) -free if there exist positive constants $\alpha > \gamma > 0$ such that

$$w(\vec{m}^{(i)}) \geq O(L^\alpha), \quad i \neq 1, \quad (10)$$

and $M \leq O(L^\gamma)$.

In computing $\rho_{\text{free}}(\beta)$, each stabilizer generator S_j was removed randomly with probability $p(\beta)$ as if S_j were independent operators. However, Eq. (9) implies that S_j are correlated with each other. The (α, γ) -free condition assures that correlations are weak and S_j can be treated as independent operators. Note that the cubic code (and quantum fractal codes) are $(2, 1)$ -free and the two-dimensional toric code is $(2, 0)$ -free.

Lemma 1. Let H be an (α, γ) -free stabilizer Hamiltonian. Then

$$|\rho(\beta) - \rho_{\text{free}}(\beta)| \leq c \cdot \exp(-c' \cdot L^\alpha) \quad (11)$$

where $c, c' > 0$ are some constants independent of β, L .

The above lemma can be generalized to stabilizer Hamiltonians with qudits (spins with d internal states). Namely, for a Hamiltonian $H = -\sum_j Q_j$ with $Q_j = \frac{1}{d}(I + S_j + S_j^2 + \dots + S_j^{d-1})$, we set the probability function by $p = \frac{d}{e^\beta - 1 + d}$.

Proof. The Gibbs ensemble is proportional to

$$\begin{aligned} \frac{e^{-\beta H}}{(e^\beta + 1)^N} &= \prod_{j=1}^N \left([1 - p(\beta)] \frac{S_j + I}{2} + p(\beta) \frac{I}{2} \right) \\ &= \sum_{\vec{k}} \text{Prob}(\vec{k}) \cdot \mathcal{M}(\vec{k}) \end{aligned} \quad (12)$$

with $\mathcal{M}(\vec{k}) = \prod_{j=1}^N \left(\frac{S_j + I}{2} \right)^{k_j} \left(\frac{I}{2} \right)^{1-k_j}$. Here $\mathcal{M}(\vec{k})$ is proportional to $\rho(\vec{k})$ and can be written as $\mathcal{M}(\vec{k}) = F(\vec{k})\rho(\vec{k})$ where $F(\vec{k}) = \text{Tr} \mathcal{M}(\vec{k})$. Consider an N -component binary vector \vec{k}' . We say that \vec{k}' is included in \vec{k} if $k'_j \leq k_j$ for all j , and write $\vec{k}' \subseteq \vec{k}$. Then $\mathcal{M}(\vec{k})$ can be written as

$$\mathcal{M}(\vec{k}) = \frac{1}{2^N} \sum_{\vec{k}' \subseteq \vec{k}} \prod_j (S_j)^{k'_j}. \quad (13)$$

Nonzero contributions to $\text{Tr} \mathcal{M}(\vec{k})$ come from \vec{k}' such that $\prod_j (S_j)^{k'_j} = I$ since the trace of nonidentity Pauli operators is zero. So, $F(\vec{k})$ is equal to the number of $\vec{k}' \subseteq \vec{k}$ such that $\prod_j (S_j)^{k'_j} = I$. Then we have

$$\rho_\beta \propto \sum_{i=1}^{2^M} \sum_{\vec{k} \supseteq \vec{m}^{(i)}} \text{Prob}(\vec{k}) \cdot \rho(\vec{k}). \quad (14)$$

We split the above sum into two parts:

$$\sum_{\vec{k}} \text{Prob}(\vec{k}) \cdot \rho(\vec{k}) + \sum_{i=2}^{2^M} \sum_{\vec{k} \supseteq \vec{m}^{(i)}} \text{Prob}(\vec{k}) \cdot \rho(\vec{k}). \quad (15)$$

Since the Hamming weight of $\vec{m}^{(i)}$ for $i > 1$ is at least L^α , we have

$$\sum_{i=2}^{2^M} \sum_{\vec{k} \supseteq \vec{m}^{(i)}} \text{Prob}(\vec{k}) \leq 2^M \cdot [1 - p(\beta)]^{L^\alpha} \quad (16)$$

which approaches zero exponentially since $\alpha > \gamma$. The first term in Eq. (15) is $\rho_{\text{free}}(\beta)$, so $\rho_\beta \approx \rho_{\text{free}}(\beta)$. This completes the proof. ■

B. Cleaning excitations

Consider a CSS stabilizer Hamiltonian: $H = -\sum_j S_j^{(X)} - \sum_j S_j^{(Z)}$. We hope to transform this Hamiltonian into a classical Hamiltonian with Pauli-Z terms only. A X -type stabilizer generator $S_j^{(X)}$ is said to be R -removable if there exists a Pauli Z operator V_j which is exclusively supported inside a box of linear length R surrounding $S_j^{(X)}$ and satisfies

$$[S_j^{(X)}, V_i] = 0 \quad (i \neq j), \quad \{S_j^{(X)}, V_j\} = 0. \quad (17)$$

The operator V_j can eliminate an excitation associated with $S_j^{(X)}$ since V_j can flip the sign of $S_j^{(X)}$ without affecting other stabilizer generators.

Lemma 2. If all the X terms $S_j^{(X)}$ in a CSS-type Hamiltonian H are R -removable, then there exists a range R unitary U such that UHU^\dagger is a classical Hamiltonian.

Proof. Let V_j be the removers for $S_j^{(X)}$. Consider the following range- R unitary transformation:

$$U = \prod_j \exp\left(\frac{\pi}{4} S_j^{(X)} V_j\right). \quad (18)$$

Since $S_j^{(Z)} \rightarrow S_j^{(Z)}$ and $S_j^{(X)} \rightarrow V_j$, one has

$$UHU^\dagger = -\sum_j S_j^{(Z)} - \sum_j V_j \quad (19)$$

which is a classical Hamiltonian. ■

Intuitively, the lemma says that if one can eliminate $S_j^{(X)}$ excitations by unitary operations acting inside a box of size R , then the Hamiltonian can be shown to be R trivial.

Let us show that two-dimensional toric code is not topologically ordered at nonzero temperature by using this lemma. This statement has been already proven by Hastings, but we provide a slightly different proof. Assume that a vertex operator A_v at a vertex v is missing in $H(p)$. Then the missing vertex v behaves as a sink of charge excitations. Namely, for a single isolated charge at arbitrary location, let us move it to the location of missing vertex by applying some appropriate stringlike Pauli Z operators. Once it reaches the missing vertex v , the charge excitation can disappear. Therefore, vertices with missing vertex terms can absorb charge excitations.

Imagine that we split the entire lattice into a grid of $\sqrt{c \log(L)} \times \sqrt{c \log(L)}$ qubits. Then, for sufficiently large but finite c , there almost surely exists at least one missing vertex term per grid as in Fig. 1. Namely, the probability of having no missing vertex in a grid is $(1 - p)^{c \log(L)} = L^{c \log(1-p)}$ which is polynomially vanishing since $\log(1 - p) < 0$. Thus, one can bring charge excitations in each grid to the sink within the same grid. Let g denote a grid and $n(g)$ denote the number of sinks in g . Since there are in total $\frac{L^2}{c \log(L)}$ grids, the probability of having at least one missing vertex in each grid g is given as follows:

$$\text{Prob}[n(g) > 0 \forall g] = (1 - L^{c \log(1-p)})^{L^2/c \log(L)}. \quad (20)$$

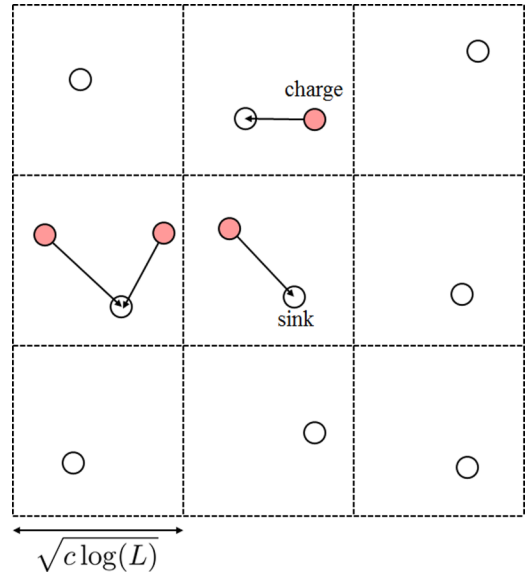


FIG. 1. White dots represent missing vertex terms which act as sinks of excitations. Red dots represents charge excitations that can be absorbed into a sink. The entire system is split into a grid of $\sqrt{c \log(L)} \times \sqrt{c \log(L)}$ qubits such that there is at least one sink per patch.

The right-hand side of Eq. (20) is lower bounded by $1 - \frac{L^2}{c \log(L)} L^{c \log(1-p)}$, so if $c > -\frac{2}{\log(1-p)}$, there exists at least one missing vertex in each grid with probability asymptotically approaching unity. Therefore, the Gibbs ensemble ρ_β is (R, ϵ) trivial for $R = \sqrt{c \log(L)}$ and $\epsilon = O(L^{-\delta})$ with some $\delta > 0$ in the thermodynamic limit. In other words, one can approximate the Gibbs ensemble via $\sqrt{c \log(L)}$ -range circuit.

By a similar argument, one is able to observe that topologically ordered models with pointlike excitations do not retain their topological order at any finite temperature. Such systems include the Levin-Wen model and the Walker-Wang model. In the case of the welded code [6], bifurcations of otherwise pointlike excitations are avoided in the presence of missing terms by the fact that the length of solids within the code grow as $O(L^{2/3})$, so by similar argument excitations can always be eliminated within a solid. See Appendix B for details.

C. Free energy

Let us take a short detour and present physical interpretations of the arguments in previous sections. Consider an isolated single excitation which is located inside some block of finite size in the system and randomly distributed sinks of excitations resulting from an imperfect Hamiltonian $H(p)$. If the single excitation can be eliminated by moving it to sinks via unitary transformation acting strictly inside a box, one can construct a local unitary transformation which brings $H(p)$ into a classical Hamiltonian. For simplicity of discussion, we assume that the sinks appear at some finite constant rate $p_0 > 0$. If excitations are particlelike, meaning that their propagation trajectories are deformable lines, then an isolated excitation can be absorbed into some sink after propagating $1/p_0$ sites on average. However, when the system has diverging energy barrier and an excitation may split into

multiple excitations, as in the cubic code and the welded code, it is not clear if excitations can be absorbed by randomly distributed sinks or not.

Consider a single isolated excitation and a surrounding box of finite size. Let \mathcal{R} be a distribution of sinks inside the box. We consider a set of all of the possible error configurations $\{\mathcal{C}_j\}$ into which the single isolated excitation may evolve. If there exists an excitation configuration \mathcal{C} such that $\mathcal{C} \subseteq \mathcal{R}$, then the isolated excitation can be absorbed via a unitary transformation acting strictly inside the box. Let us consider an error configuration \mathcal{C} such that the total number of excitations is given by $|\mathcal{C}| = E$. The probability for such a configuration that can be absorbed by \mathcal{R} is given by

$$p_0^E \sim \exp(-\beta_0 E), \quad e^{-\beta_0} \equiv p_0. \quad (21)$$

We would like to estimate a probability that such an excitation configuration exists from a list of all the possible excitations configurations $\{\mathcal{C}_j\}$. While an exact calculation of such a probability is difficult in general, we can crudely approximate it by assuming that the probabilities for excitation configurations to be absorbed by \mathcal{R} are statistically independent events. Let N_E be the total number of error configurations such that $|\mathcal{C}| = E$. Then, the total probability for excitation configurations of size E to be absorbed by sinks can be estimated as

$$P_{\text{total}}(E) \approx N_E \exp(-\beta_0 E) = \exp(-\beta_0 F), \quad (22)$$

where

$$F = E - T_0 S, \quad S = \log N_E. \quad (23)$$

See Fig. 2. Here we have approximated $P_{\text{total}}(E)$ by summing each probability $\exp(-\beta_0 E)$ which is valid when $P_{\text{total}}(E)$ is small. The exponent F can be viewed as a free energy of excitation configurations with energy E , and the approximation is valid for $F > 0$.

The observation above allows us to obtain an easy check for topological order at finite temperature by computing the free energy of excitations. If F is negative for a given E , such excitations are likely to be absorbed. If F is positive,

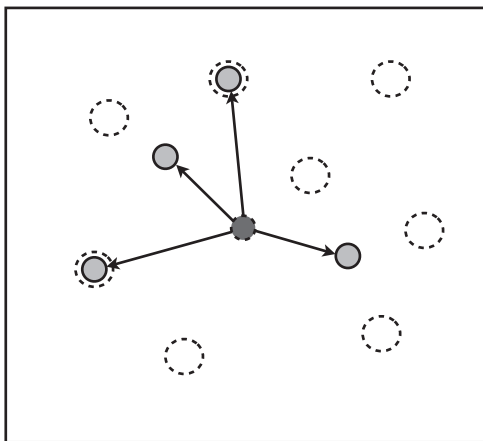


FIG. 2. An elimination of excitations and the Peierls argument. Grey circles represent excitations that evolved from the isolated single charge at the center. Dotted circles represent sinks of excitations. In this example, two excitations surrounded by sinks can be absorbed but the other two remain.

such excitations are not likely to be absorbed. This is indeed the essence of the Peierls argument [21]. Let us apply this argument to two-dimensional Ising model on a square lattice. The excitation energy E corresponds to the length of the domain wall of an excitation droplet. The total number of possible configurations can be estimated to be $N_E = 3^E$ by treating it as a random walk. (There are three directions to move on a square lattice). So the free energy of excitation at E is lower bounded by

$$F \geq E - \log 3^E T. \quad (24)$$

For T less than some constant, the free energy is positive, and we expect that excitations cannot be absorbed by sinks. However, for large T , the free energy becomes negative, implying that such excitations can readily be absorbed. By finding the temperature for which the right-hand side becomes zero, we are able to obtain an analytical lower bound the transition temperature: $1/\log 3 \sim 0.91 \leq T_c$.

III. CUBIC CODE

In this section, we show that the cubic code is not topologically ordered at nonzero temperature.

A. Fractal and polynomial

This subsection provides a quick review of the cubic code and its generalization, the quantum fractal codes. To treat stabilizer quantum codes with fractal-shaped logical operators, it is convenient to introduce the polynomial representation of Pauli operators. See Refs. [22,23] for detailed discussions. Consider a three-dimensional cubic lattice with $L \times L \times L$ qubits. Each site is labeled as (i, j, k) where $0 \leq i, j, k \leq L - 1$. Consider a polynomial over \mathbb{F}_2 fields,

$$f(x, y, z) = \sum_{i, j, k=0}^{L-1} c_{i, j, k} x^i y^j z^k, \quad c_{i, j, k} = 0, 1, \quad (25)$$

and corresponding Pauli operators,

$$X(f) = \prod_{i, j, k} (X_{i, j, k})^{c_{i, j, k}} \quad Z(f) = \prod_{i, j, k} (Z_{i, j, k})^{c_{i, j, k}}. \quad (26)$$

Polynomial representations are useful for dealing with translation invariant systems. For instance, $X(x^i f)$ represents a translation of $X(f)$ in the \hat{x} direction by i sites. We impose periodic boundary conditions by setting $x^L = y^L = z^L = 1$. To be more concrete, we have $x^{-1} = x^{L-1}$, which implies that the -1 th site and the $L - 1$ th site are identical. For simplicity of discussion, we assume $L = 2^m$ with positive integer m . (The system size L affects the number of logical qubits, but discussion on thermal topological order does not depend on choices of L .)

Quantum fractal codes are defined on a cubic lattice where *two qubits* live at each site of an $L \times L \times L$ cubic lattice. Pauli operators can be represented as

$$X \begin{pmatrix} f \\ g \end{pmatrix}$$

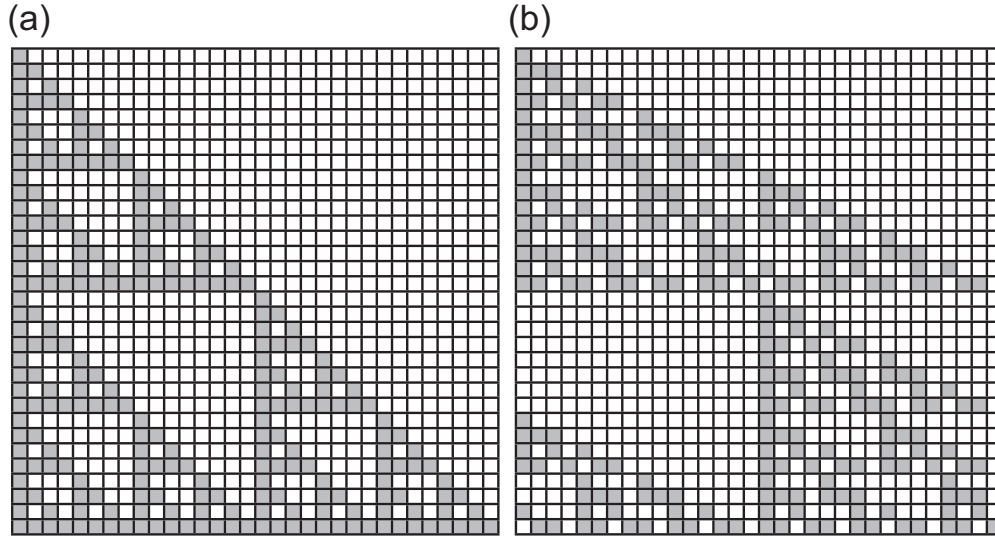


FIG. 3. (a) The Sierpinski triangle from $f = 1 + x$. (b) The model from $f = 1 + x + x^2$.

and

$$Z\left(\begin{array}{c} f \\ g \end{array}\right),$$

where polynomials on the upper (lower) row represent the first (second) qubit at each site.

Let us now specify the Hamiltonian of quantum fractal codes. For arbitrary polynomials $\alpha(x, y, z)$ and $\beta(x, y, z)$, consider

$$Z\left(\begin{array}{c} \alpha \\ \beta \end{array}\right), \quad X\left(\begin{array}{c} \bar{\beta} \\ \bar{\alpha} \end{array}\right), \quad (27)$$

where $\bar{\alpha}$ and $\bar{\beta}$ are duals of α and β obtained by taking $x \rightarrow x^{-1}$, $y \rightarrow y^{-1}$, $z \rightarrow z^{-1}$. The Hamiltonian is

$$H = - \sum_{ijk} Z\left(\begin{array}{c} x^i y^j z^k \alpha \\ x^i y^j z^k \beta \end{array}\right) - \sum_{ijk} X\left(\begin{array}{c} x^i y^j z^k \bar{\beta} \\ x^i y^j z^k \bar{\alpha} \end{array}\right). \quad (28)$$

One can verify that terms in the Hamiltonian commute with each other due to the use of dual polynomials. So this is a stabilizer Hamiltonian.

In Ref. [23], quantum codes with the following polynomials have been studied in detail:

$$\alpha = 1 + f(x)y, \quad \beta = 1 + g(x)z. \quad (29)$$

Logical operators have fractal shapes which are generated by polynomials $f(x)$ and $g(x)$:

$$\mathbf{f}(x, y) = 1 + fy + f^2y^2 + \dots, \quad (30)$$

$$\bar{\mathbf{f}}(x, y) = 1 + \bar{f}\bar{y} + \bar{f}^2\bar{y}^2 + \dots, \quad (31)$$

$$\mathbf{g}(x, z) = 1 + gz + g^2z^2 + \dots, \quad (32)$$

$$\bar{\mathbf{g}}(x, z) = 1 + \bar{g}\bar{z} + \bar{g}^2\bar{z}^2 + \dots. \quad (33)$$

Note that $\mathbf{f}(x, y)$ lives on a (\hat{x}, \hat{y}) plane while $\mathbf{g}(x, z)$ lives on a (\hat{x}, \hat{z}) plane. Here $\mathbf{f}(x, y)$ and $\mathbf{g}(x, y)$ have fractal shapes in general as shown in Fig. 3. Namely, for $f = 1 + x$, one obtains the well-known Sierpinski triangle while for

$f = 1 + x + x^2$, one obtains a fractal shape with Hausdorff dimension $\log(1 + \sqrt{5})/\log 2$.

Quantum fractal codes typically have $k = 2L$, and there are $2L$ of Z -type logical operators and $2L$ of X -type logical operators:

$$\ell_i^{(Z)} = Z\left(\begin{array}{c} 0 \\ x^i \mathbf{f}(x, y) \end{array}\right), \quad r_i^{(Z)} = Z\left(\begin{array}{c} x^i \mathbf{g}(x, z) \\ 0 \end{array}\right), \quad (34)$$

$$\ell_i^{(X)} = X\left(\begin{array}{c} x^i \bar{\mathbf{f}}(x, y) \\ 0 \end{array}\right), \quad r_i^{(X)} = X\left(\begin{array}{c} 0 \\ x^i \bar{\mathbf{g}}(x, z) \end{array}\right), \quad (35)$$

where $i = 0, \dots, L - 1$. Therefore, Z -type logical operators have geometric shapes of $\mathbf{f}(x, y)$ and $\mathbf{g}(x, z)$ while X -type logical operators have geometric shapes of $\bar{\mathbf{f}}(x, y)$ and $\bar{\mathbf{g}}(x, z)$. For a technical subtlety concerning the number of logical qubits, see Ref. [23].

The cubic code corresponds to the choices of

$$\alpha = 1 + (1 + x + x^2)y, \quad (36)$$

$$\beta = 1 + (1 + x)z + (1 + x + x^2)z^2.$$

Note that polynomials α, β generate fractal geometries. Loosely speaking, if α and β generate different fractal geometries, then the system is free from stringlike logical operators [23]. Thus there is no special meaning in choosing α, β to be as in Eq. (36), and our arguments apply to a larger class of translation symmetric stabilizer codes.

B. Thermal instability

The fact that fractal logical operators can be supported on two-dimensional planes implies that excitations may propagate according to fractal geometries confined on a plane. We shall concentrate on excitations associated with violations of X -type stabilizers. Excitations can be characterized by a polynomial E :

$$E(x, y, z) = \sum_{i,j,k} e_{i,j,k} x^i y^j z^k, \quad (37)$$

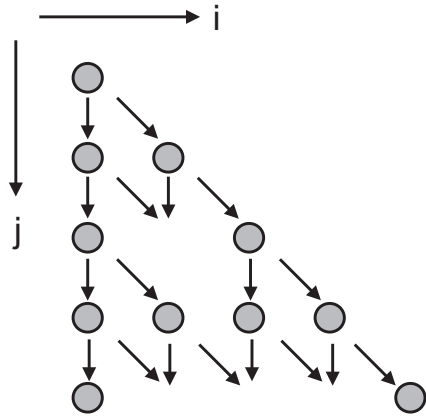


FIG. 4. Propagations of excitations according to a fractal geometry.

where $e_{i,j,k} = 1$ means that an excitation is present at (i, j, k) while $e_{i,j,k} = 0$ means that an excitation is absent.

Consider an excitation at $(0,0,0)$, which can be represented by $E = x^0 y^0 z^0 = 1$. If one applies a Pauli operator $Z_{0,0,0}^{(B)}$, then the excitation will propagate to $E = fy$ since $Z_{0,0,0}^{(B)}$ would create an excitation characterized by $E = 1 + fy$. By applying Pauli Z operators in an appropriate manner, one can move this single excitation at $(0,0,0)$ to

$$E = f^m y^m \quad (38)$$

in the \hat{y} direction by forming a fractal geometry. See an example for $f = 1 + x$ shown in Fig. 4.

Let us assume that a stabilizer term at $(0,0,0)$ is removed. Then, excitations $f^m y^m$ can be eliminated by bringing them to $(0,0,0)$. This is the key property we will use in our proof. Our strategy is to prove that an excitation $(0,0,0)$ can be eliminated with high probability by some local unitary operation when stabilizer terms are randomly removed.

We now prove that, for $f(x) = 1 + x$, the system is not topologically ordered at nonzero temperature. We need to introduce a few notations. Since one can move excitations inside a two-dimensional plane, we concentrate on a plane $(i, j, 0)$ by setting $k = 0$. We will also write $(i, j, 0)$ as (i, j) for brevity of notation. Consider the following sets of sites:

$$B = \{(i, j) : 0 \leq i, j \leq r-1, i \leq j\}. \quad (39)$$

We want to prove that a single excitation at $(0,0)$ can be almost surely eliminated inside a block B for $r = O(\log(L)^2)$. It is convenient to define

$$B_t = \{(i, j) \in B : i \geq r-t\}, \quad (40)$$

$$L_t = \{(i, j) \in B : i = r-t\}. \quad (41)$$

See Fig. 5 for graphical representations of B, B_t, L_t .

Assume that a stabilizer term at (i, j) is removed. Then excitations $x^i f^{r-1-j} y^{r-1}$ can be eliminated since a single excitation at (i, j) would propagate to $x^i f^{r-1-j} y^{r-1}$. Note that this excitation is located at the bottom of B . Let us consider a subset of B , denoted by B^* , whose elements are chosen randomly with probability $p > 0$. Also, define $L_t^* = L_t \cap B^*$ and $B_t^* = B_t \cap B^*$. We shall consider the following linear

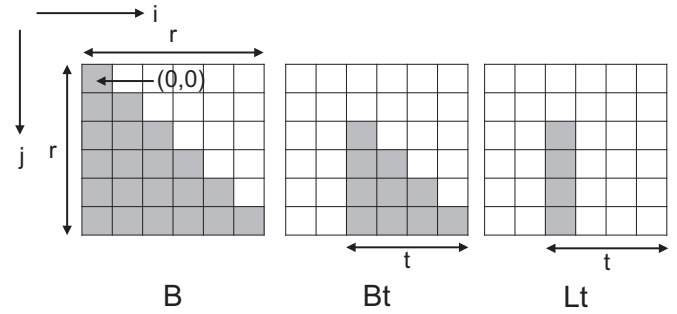


FIG. 5. Graphical representations of B, B_t, L_t .

spaces spanned by polynomials:

$$F_t = \langle x^i f^j : (i, j) \in B_t^* \rangle. \quad (42)$$

The dimension of F_t , denoted by d_t , is defined as the number of independent polynomials in F_t . Also, define

$$G_t = \langle x^i : i \geq r-t \rangle. \quad (43)$$

Then $F_t \subseteq G_t$, and $d_t \leq t$. If $d_r = r$, then any excitation inside B can be eliminated.

Lemma 3. For any $p > 0$ there exists a sufficiently large but finite constant c such that, for $r = \frac{c}{p^3} \cdot \log(L)$,

$$1 - \text{Prob}[d_r = r] \leq O\left(\frac{1}{L^{c'}}\right) \quad (44)$$

for some constant $c' > 3$ which is independent of L, p and can be made arbitrarily large by taking large c .

This lemma implies that a single excitation inside B can be eliminated with high probability. Note that there are $O(L^3)$ terms in the cubic code while Eq. (44) implies that the probability approaches to unity faster than $O(1/L^3)$. Thus, an imperfect Hamiltonian $H(p)$ is almost surely range- $\log(L)$ trivial, proving the thermal instability of the original Hamiltonian H .

Proof. Consider an arbitrary set of sites $\hat{B}_t \subseteq B_t$ and denote the dimension of a linear space spanned by polynomials generated from \hat{B}_t by \hat{d}_t . Let us choose sites in $L_{t+1}^* \subseteq L_{t+1}$ randomly with probability p . We define $\hat{B}_{t+1} = L_{t+1}^* \cup \hat{B}_t$ and denote the dimension of the linear space by \hat{d}_{t+1} . From the definition, one has $\hat{d}_{t+1} \geq \hat{d}_t$. Also note that $d_t \leq t$. We are interested in the probability distribution of the increase: $\hat{d}_{t+1} - \hat{d}_t$ for a given \hat{B}_t .

First, assume that \hat{d}_t is already at its maximal value: $\hat{d}_t = t$. If L_{t+1}^* is an empty set, then $\hat{d}_{t+1} = \hat{d}_t$. Since $x^i f^j \notin F_{t+1}$ for all $(i, j) \in L_{t+1}$, if L_{t+1}^* is a nonempty set, one has $\hat{d}_{t+1} = \hat{d}_t + 1$. So one has

$$\text{Prob}[\hat{d}_{t+1} = t] = (1-p)^{t+1}, \quad (45)$$

$$\text{Prob}[\hat{d}_{t+1} = t+1] = 1 - (1-p)^{t+1}.$$

Note that the above expressions are valid for any sets \hat{B}_t with $\hat{d}_t = t$. Thus, for sufficiently large t , the value of d_t will almost surely increase by 1. This implies that once d_t reaches its maximal value at some $t = t_0$, then we have $d_t = t$ for all $t > t_0$ with very high probability. Namely, if $t = \frac{c}{p} \log(L)$ for sufficiently large $c > 0$, then $\text{Prob}[\hat{d}_{t+1} = t]$ is

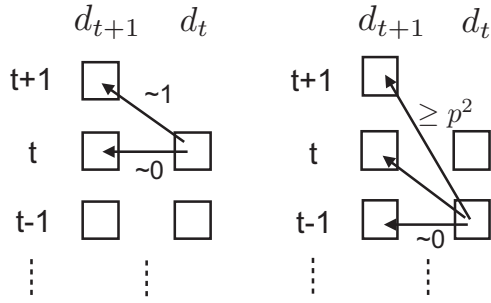


FIG. 6. Evolutions of probability distributions of d_t .

polynomially suppressed. Later, we will take $r = \frac{c'}{p^3} \cdot \log(L)$, so $\text{Prob}[\hat{d}_{t+1} = t]$ will be negligible for all $t \geq \frac{c}{p} \log(L)$.

Next, we consider sets \hat{B}_t with $\hat{d}_t < t$. We find a lower bound on the probability that \hat{d}_t increases by more than 1. Let $e_j = x^{r-t} f^{r-1-j}$. Note that e_j are independent for all $j = 0, \dots, t$. Also note that $e_j \notin F_t$ since e_j always contains x^{r-t} while polynomials in F_t do not contain it. For a given i ($0 \leq i \leq t$), there always exists $k(i)$ [$0 \leq k(i) \leq t, i \neq k(i)$] such that $e_i + e_{k(i)} \notin F_t$ since e_j are independent. To prove this statement, let us suppose that $e_i + e_j \in F_t$ for all $j \neq i$. Since $\hat{d}_t < t$ and $\langle e_0, \dots, e_t \rangle = G_{t+1}$, there must exist a pair of polynomials $a, b \in G_{t+1}$ such that

$$a, b, a + b \notin F_t. \quad (46)$$

Note that a, b can be expanded by e_0, \dots, e_t . Since adding or subtracting elements from F_t to a, b does not affect Eq. (46), we can choose a, b to be $a = e_i$ and $b = e_{k(i)}$ where $k(i) \neq i$ is some integer. But one has $a + b = e_i + e_{k(i)} \in F_t$, leading to a contradiction. So, if $(r-t, i), (r-t, k(i)) \in L_{t+1}^*$, then $\hat{d}_{t+1} \geq \hat{d}_t + 2$. Such a probability is lower bounded by p^2 :

$$\text{Prob}[\hat{d}_{t+1} \geq \hat{d}_t + 2] \geq p^2. \quad (47)$$

This lower bound is valid for any sets \hat{B}_t with $\hat{d}_t < t$. Also, recall that $\text{Prob}[\hat{d}_{t+1} = \hat{d}_t] = (1-p)^{t+1}$, which is very small for large t . Therefore, \hat{d}_t increases by at least 1 almost surely and increases by more than 1 with probability at least p^2 (see Fig. 6).

Intuitively, since the probability of \hat{d}_t increasing by more than 1 is constant while the probability of \hat{d}_t failing to increase is exponentially vanishing, a difference of $t_0 - \hat{d}_{t_0}$ (for $\hat{d}_{t_0} < t_0$) can be made up for in a large number of subsequent steps with low probability of introducing too many additional failures. More formally, let us consider an arbitrary set \hat{B}_{t_0} for $t_0 = \frac{c}{p} \log(L)$ with sufficiently large but finite $c > 0$ so that \hat{d}_t will almost surely increase by at least one for all $t \geq t_0$. In order for $d_r = r$, d_t needs to reach its maximal value for some t with $t_0 \leq t \leq r$. For each t , the probability of increase by more than 1 is lower bounded by p^2 . So the probability of d_r not reaching its maximal value is upper bounded by

$$\sum_{j=0}^{t_0-1} \binom{r-t_0}{j} (p^2)^j (1-p^2)^{r-t_0-j} \quad (48)$$

which corresponds to a lower tail of the binomial distribution. For $r = \frac{c'}{p^3} \log(L)$ with sufficiently large $c' > 0$, this

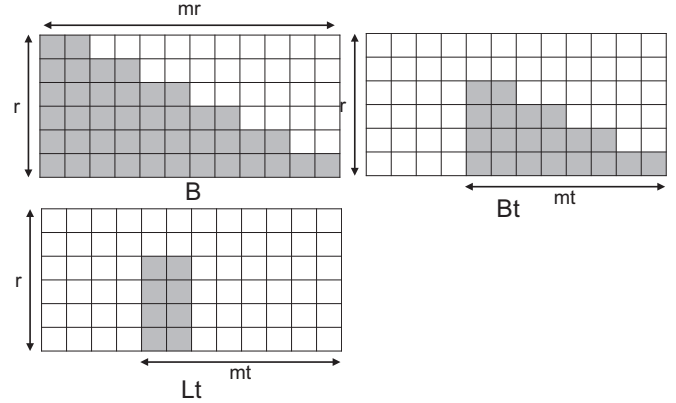


FIG. 7. Graphical representations of B, B_t, L_t for $m = 2$.

expression is polynomially suppressed with respect to L . Thus, d_r reaches its maximal value r with probability approaching to unity, and all the errors are polynomially suppressed with respect to L . This completes the proof.

This implies that there almost surely exists a unitary quantum circuit acting solely inside B that can transform all stabilizer operators so that they are diagonal in the same basis (X or Z). This probability goes to 1 in the thermodynamic limit ($L \rightarrow \infty$). Thus all the terms in $H(p)$ are almost surely $O(\log(L))$ removable. We have shown that quantum fractal codes with $f = 1 + x$ are not topologically ordered at nonzero temperature. The argument does not rely on the fact that the system consists of qubits, and thus can be generalized to quantum fractal codes with qudits.

Finally, we sketch the proof for generic $f(x)$. Without loss of generality, one can assume that $f(x)$ can be written as $f(x) = \sum_{i=0}^m c_i x^i$ and $c_0, c_m \neq 0$ where m is a positive integer. We consider a steplike region B , and then define $B^*, B_t, L_t, B_t^*, L_t^*$ in a similar manner as depicted in Fig. 7 where each step consists of m columns. We then define linear spaces by

$$F_t = \langle x^i f^j : (i, j) \in B_t^* \rangle \quad (49)$$

and denote its dimension by d_t . Note that $d_t \leq mt$. We want to prove that $d_r = mr$ with high probability.

When t is logarithmically large with respect to L , we will have $d_{t+1} - d_t < m$ with polynomially vanishing probability. So, at each step, d_t will almost surely increase by at least m . Also, one finds that the probability of having $d_{t+1} - d_t > m$ is lower bounded by p^{m+1} as long as d_t has not reached its maximal value by using an argument similar to the proof of Lemma 3. Note that the cubic code has $f = 1 + x + x^2$ with $m = 2$, so we should take $r = \frac{c}{p^3} \log(L)$ for some large but finite $c > 0$. Then the probability of d_r being at its maximal value approaches to unity polynomially. Thus, the cubic code is not topologically ordered at nonzero temperature. This proof strategy can be generalized to cases where α, β have the form of $1 + f_1(x)y + f_2(x)y^2 + \dots$ and $1 + g_1(x)z + g_2(x)z^2 + \dots$ by appropriately modifying the constructions of $B^*, B_t, L_t, B_t^*, L_t^*$. However, at this moment, we do not know how to generalize this proof to arbitrary α, β . It would be interesting to generalize our studies to arbitrary stabilizer codes with translation symmetries. Also it should be noted that

approximating the Gibbs ensemble of the toric code requires a $O(\sqrt{\log(L)})$ -range circuit only while approximating the Gibbs ensemble of the cubic code requires a $O(\log(L))$ -range circuit in our approach.

IV. MEMORY TIME

Having studied static properties of quantum memories, we now wish to switch gears to study dynamical properties of quantum memories by computing upper bounds on memory time, the expected lifetime of quantum information. In this section, we begin by reviewing a mathematical framework to treat thermalization of stabilizer quantum memory Hamiltonians based on Davies generators. We then present a method of choosing minimal coupling to the thermal bath, minimal in the sense that the system-bath interaction is as simple as possible but still yields an ergodic reduced dynamics.

A. Lindblad operators

We shall model the dynamics of the system-bath interaction using the Davies weak-coupling limit [24] and begin by reviewing this construction. The total Hamiltonian describing the system and the bath is given by

$$H = H^{\text{sys}} + H^{\text{bath}} + H^{\text{int}}, \quad H^{\text{int}} = \sum_{\alpha} S_{\alpha} \otimes f_{\alpha}, \quad (50)$$

where S_{α} are system operators, such as single-qubit Pauli operators, and f_{α} are bath operators. The density matrix of the system obeys

$$\dot{\rho}(t) = -i[H^{\text{sys}}, \rho(t)] + \mathcal{L}[\rho(t)] \quad (51)$$

with $\rho(0)$ the initialized ground state.

The Lindblad generator describing transfer of energy between the system and the bath is given by

$$\mathcal{L}[\rho(t)] = \sum_{\alpha, \omega} h(\alpha, \omega) \left[S_{\alpha}(\omega) \rho(t) S_{\alpha}(\omega)^{\dagger} - \frac{1}{2} \{ \rho, S_{\alpha}(\omega)^{\dagger} S_{\alpha}(\omega) \} \right]$$

with $S_{\alpha}(\omega)$ describing transitions from eigenvectors of H^{sys} with energy E to eigenvectors with energy $E - \omega$ with ω delivered to the bath and $h(\alpha, \omega)$ describing the rate of these transfers. To satisfy detailed balance, we demand

$$h(\alpha, -\omega) = e^{-\beta\omega} h(\alpha, \omega) \quad (52)$$

with $\beta = 1/T$, which ensures that the thermal density matrix $\rho_{\beta} \sim e^{-\beta H}$ is a fixed point of the dynamics ($\mathcal{L}[\rho_{\beta}] = 0$).

Moreover, if the action of \mathcal{L} is ergodic, meaning that all the eigenvectors can be visited, then ρ_{β} will be the *only* fixed point. So if a minimal interaction Hamiltonian is chosen, it is necessary to ensure that it is ergodic. The condition for this is given in Refs. [25,26] by

$$\{S_{\alpha}, H^{\text{sys}}\}' = \mathbb{C}1. \quad (53)$$

In other words, no system operator apart from multiples of the identity should commute with all of the S_{α} and the system Hamiltonian so that all energy levels of the system Hamiltonian are accessible. For a stabilizer Hamiltonian, let us assume that

S_{α} are Pauli operators. Then the condition for ergodicity is given by

$$\langle \{S_{\alpha}\}, H^{\text{sys}} \rangle = \mathcal{P}. \quad (54)$$

In other words, S_{α} and stabilizer generators in H^{sys} need to form the full Pauli operator group since otherwise there exists a nontrivial Pauli operator which commutes with S_{α} and H^{sys} . A natural choice of $\{S_{\alpha}\}$ is to pick Pauli X and Z operators acting on each qubit. Then the Davies generator has the form

$$\mathcal{L} = \sum_{j=1}^n \mathcal{L}_{X,j} + \sum_{j=1}^n \mathcal{L}_{Z,j} = \mathcal{L}_X + \mathcal{L}_Z. \quad (55)$$

The convergence time to the Gibbs ensemble is given by the inverse of the gap of the superoperator \mathcal{L} . When the Davies generator has the form of Eq. (55) and the Hamiltonian is described by a CSS stabilizer code, convergence time is given by $\tau_{\text{conv}}^{-1} = \min(\text{Gap}(\mathcal{L}_X), \text{Gap}(\mathcal{L}_Z))$. Our goal is to obtain lower bounds on quantum memory time of CSS codes. Since memory systems suffer from both X and Z errors, memory time is given by $\tau_{\text{memory}}^{-1} = \max(\text{Gap}(\mathcal{L}_X), \text{Gap}(\mathcal{L}_Z))$. Note that it is possible that memory time and convergence time are different. For instance, in the three-dimensional toric code, convergence time is $O(\exp(L))$ while quantum memory time is $O(1)$. See [16] for details.

To obtain an upper bound on quantum memory time, or equivalently lower bounds on gaps in \mathcal{L}_X and \mathcal{L}_Z , we pick a smaller, but ergodic $\{S_{\alpha}^{\text{min}}\}$. Observe that one can remove some of Pauli operators from $\{S_{\alpha}\}$ while retaining its ergodicity. The gap for $\{S_{\alpha}\}$ will be lower bounded by the one for $\{S_{\alpha}^{\text{min}}\}$ since choosing a smaller set always makes the gap smaller. In Ref. [16], Alicki *et al.* used this strategy to provide an upper bound on quantum memory time of the two-dimensional toric code, giving a rigorous proof that the model is not self-correcting. Namely, they choose $\{S_{\alpha}^{\text{min}}\}$ carefully such that the problem is reduced to computing classical memory time of one-dimensional Ising model.

B. Ergodic decomposition

In this section, we generalize an ergodic decomposition of the Davies generator applied by Alicki *et al.* to the toric code in Ref. [16] to all CSS codes.

Lemma 4. Given a CSS stabilizer code defined by stabilizer group \mathcal{S} with k logical qubits, there exists a tripartition of the full system into three nonoverlapping sets of qubits, denoted by L, R_X, R_Z such that

$$\begin{aligned} \mathcal{P}^{(X)} &= \langle S_X, \mathcal{P}^{(X)}(L \cup R_X) \rangle, \\ \mathcal{P}^{(Z)} &= \langle S_Z, \mathcal{P}^{(Z)}(L \cup R_Z) \rangle, \end{aligned} \quad (56)$$

where $\mathcal{P}^{(X,Z)}$ is a group of Pauli- X, Z operators, $\mathcal{P}^{(X,Z)}(R)$ is a group of Pauli- X, Z operators acting on region R , and

$$|L| = k \quad |R_X| = \log_2 |\mathcal{S}_Z| \quad |R_Z| = \log_2 |\mathcal{S}_X|.$$

Thus choosing

$$\{S_{\alpha}\} = \mathcal{P}^{(X)}(R_X \cup L) \cup \mathcal{P}^{(Z)}(R_Z \cup L) \quad (57)$$

for the noise model's Lindblad generator satisfies the ergodicity requirement in Eq. (53). Note that the choice of $\{S_{\alpha}^{\text{min}}\}$ is not unique, but the lemma teaches us the minimal number of

generators to ensure the ergodicity. The proof also provides an efficient algorithm (running in polynomial time) to find $\{S_\alpha^{\min}\}$, which is similar to the one presented in Ref. [27].

Proof. Denote independent generators of \mathcal{S}_Z as $S_j^{(Z)}$. We seek a set of operators $\{A_j^{(X)}\}$ such that $S_i^{(Z)}$ and $A_j^{(X)}$ anticommute if $i = j$ and commute otherwise. We shall therefore proceed by induction on j . For $S_1^{(Z)}$, we can find a single qubit Pauli- X operator $A_1^{(X)}$ such that $\{S_1^{(Z)}, A_1^{(X)}\} = 0$. However, $A_1^{(X)}$ may anticommute with other generators $S_j^{(Z)}$ for $j > 1$, so for all $j > 1$, if $\{S_j^{(Z)}, A_1^{(X)}\} = 0$, we update $S_j^{(Z)} \rightarrow S_1^{(Z)} S_j^{(Z)}$. This resolves all of the commutation relations for $A_1^{(X)}$, thereby completing the base case. The commutation relations so far can be concisely summarized in the following canonical form:

$$\left\langle \begin{array}{cccc} S_1^{(Z)} & S_2^{(Z)} & \dots & S_{n_z}^{(Z)} \\ A_1^{(X)} & & & \end{array} \right\rangle \quad (58)$$

with $n_z = |R_X| = \log_2 |\mathcal{S}_Z|$. Assume that there exist m operators $A_j^{(X)}$ with $1 \leq j \leq m$ satisfying the desired commutation relations,

$$\left\langle \begin{array}{cccccc} S_1^{(Z)} & \dots & S_m^{(Z)} & S_{m+1}^{(Z)} & \dots & S_{n_z}^{(Z)} \\ A_1^{(X)} & \dots & A_m^{(X)} & & & \end{array} \right\rangle. \quad (59)$$

The group $\langle A_1^{(X)} \dots A_m^{(X)} \rangle$ is a set of Pauli- X operators acting on m qubits. For $S_{m+1}^{(Z)}$, pick a single-qubit Pauli- X operator $A_{m+1}^{(X)}$ such that $\{S_{m+1}^{(Z)}, A_{m+1}^{(X)}\} = 0$. For $j \leq m$, if $\{S_j^{(Z)}, A_{m+1}^{(X)}\} = 0$ we update $A_{m+1}^{(X)} \rightarrow A_j^{(X)} A_{m+1}^{(X)}$. For $j > m + 1$ if $\{S_j^{(Z)}, A_{m+1}^{(X)}\} = 0$, we update $S_j^{(Z)} \rightarrow S_{m+1}^{(Z)} S_j^{(Z)}$. This ensures that the desired commutation relations are satisfied for $A_j^{(X)}$ for $1 \leq j \leq m + 1$:

$$\left\langle \begin{array}{cccccc} S_1^{(Z)} & \dots & S_m^{(Z)} & S_{m+1}^{(Z)} & \dots & S_{n_z}^{(Z)} \\ A_1^{(X)} & \dots & A_m^{(X)} & A_{m+1}^{(X)} & & \end{array} \right\rangle.$$

Further, the commutation relations and choice of operators for $A_j^{(X)}$ guarantees that $\langle A_1^{(X)} \dots A_{m+1}^{(X)} \rangle$ must be the group of Pauli- X operators acting on $m + 1$ qubits. This completes the inductive step to give, in all,

$$\left\langle \begin{array}{cccc} S_1^{(Z)} & \dots & S_{n_z}^{(Z)} \\ A_1^{(X)} & \dots & A_{n_z}^{(X)} \end{array} \right\rangle. \quad (60)$$

Define the group $\mathcal{P}^{(X)}(R_X)$ as

$$\mathcal{P}^{(X)}(R_X) = \langle A_1^{(X)} \dots A_{n_z}^{(X)} \rangle, \quad (61)$$

where R_X is a set of n_z qubits. Suppose that we now expand \mathcal{S}_Z by adding in k independent Z -type logical operators \bar{Z}_i :

$$\left\langle \begin{array}{cccccc} S_1^{(Z)} & \dots & S_{n_z}^{(Z)} & \bar{Z}_1 & \dots & \bar{Z}_k \\ A_1^{(X)} & \dots & A_{n_z}^{(X)} & & & \end{array} \right\rangle.$$

Then, via the algorithm described above, we can find k more operators $A_j^{(X)}$ defined on another region L disjoint from R_X with $|L| = k$. The full Pauli- X group $\mathcal{P}^{(X)}$ is thus given by

$$\mathcal{P}^{(X)} = \langle S^{(X)}, \mathcal{P}^{(X)}(R_X \cup L) \rangle. \quad (62)$$

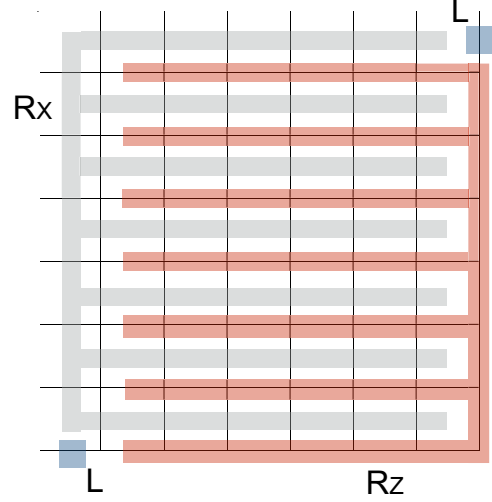


FIG. 8. Ergodic decomposition for the toric code. Qubits live on edges of the lattice.

Indeed, one can confirm this by observing that $k + n_z + n_x = n$, and $S_j^{(Z)} \notin \mathcal{P}^{(X)}(R_X \cup L)$ since $S_j^{(Z)}$ commutes with $S_j^{(Z)}$ and \bar{Z}_j .

We have shown that $\langle A_1^{(X)} \dots A_{n_z}^{(X)} \rangle$ is a set of Pauli- X operators acting on n_z qubits. Let us denote Pauli- X and Pauli- Z operators acting on the j th qubit in R_Z by X_j and Z_j ($j = 1, \dots, n_z$). Note that stabilizer generators $S_j^{(Z)}$ are independent and anticommute with at least one of Pauli- X operators X_i in R_Z . We can thus relabel stabilizer generators by taking their products such that $S_j^{(Z)} = Z_j U_j^{(Z)}$ where Z_j is Pauli- Z acting on the j th qubit and $U_j^{(Z)}$ is some product of Pauli- Z operators acting somewhere on \bar{R}_X . Therefore, if we want to generate $\mathcal{P}^{(Z)}$ using \mathcal{S}_Z , we only need $\mathcal{P}^{(Z)}(\bar{R}_X)$, noting that $L \subseteq \bar{R}_X$. The lemma is thus proved.

As an example of the decomposition, let us consider the two-dimensional toric code by following Alicki *et al.* Define a pair of comblike regions R_X, R_Z as depicted in Fig. 8. Also, define a set of two qubits at the corners as L . Then one can verify that a decomposition into R_X, R_Z, L satisfies the ergodicity condition. One can thus choose $\{S_\alpha\}$ so that Pauli X, Z operators act on $R_X \cup L, R_Z \cup L$ respectively. This reduces the problem to the mixing analysis of a classical Ising model defined on a comblike region $R_X \cup L$.

V. UPPER BOUND ON MEMORY TIME

In this section, we derive an upper bound on quantum memory time of the welded code.

A. Welded code

The three-dimensional toric code has pairs of one-dimensional Z -type logical operators and two-dimensional X -type logical operators which characterize propagations of pointlike electric charges and looplike magnetic fluxes. Since the energy barrier for propagations of magnetic loops is $O(L)$, the system is thermally protected from X -type errors. Yet, electric charges can propagate freely, so the system is not thermally stable against Z -type errors, and thus is not

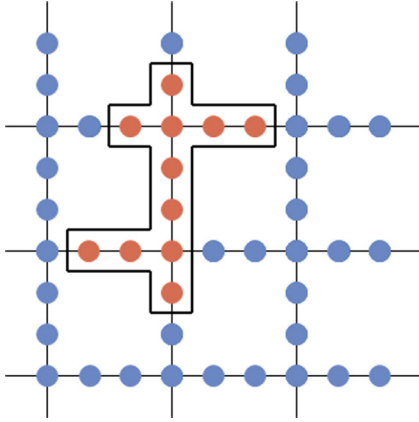


FIG. 9. Two-dimensional sparse lattice. The lattice has length L and each small square has dimensions $\sqrt{L} \times \sqrt{L}$. A Peierls droplet of frustration is shown in red, marking, for example, a cluster of spin down sites in a sea of spin up sites (blue). This contour has length $\ell = 6$, or energy cost $12J$.

self-correcting. The procedure of welding allows one to introduce energy an barrier to propagations of electric charges. Namely, by attaching multiple copies of three-dimensional toric code to form a larger stabilizer code, electric charges need to split in a “welded” surface, which leads to a polynomially diverging energy barrier. See Appendix B for a physical interpretation of the welding procedure.

By applying the aforementioned algorithm of finding an ergodic decomposition, we can show that it suffices to study classical memory time of the Ising model supported on a sparse graph, which is similar to the one shown in Fig. 9. By a sparse graph, we mean a two-dimensional skeletonlike graph where spins live on the vertices, and vertices with degree 4 are separated by $L^{1-\alpha} - 2$ vertices with degree 2. In Fig. 9, we show a two-dimensional example with $\alpha = 1/2$, though the analytical derivation below does not depend on the choice of α . A physical intuition for this reduction can be obtained by looking at trajectories of electric charges. At each welding surface, pointlike electric charges split into multiple charges, and their trajectories form a two-dimensional sparse graph.

We now bound the memory time of the welded code by bounding the memory time of a corresponding classical code, conceived by the Pauli-Z terms of the reduced \mathcal{L} , which can be found as per the algorithm described above. We henceforth show that the memory time of the welded solid code is maximized at each temperature T for a size $L_{\max}(T) \propto \exp(c\beta)$ for $c > 0$ and $\beta = T^{-1}$. Further, along the curve described by $L_{\max}(T)$, this bound follows $\tau_{\max} \propto \exp(\kappa'[\exp(\kappa\beta)])$ for positive constants κ and κ' .

The reduced Lindblad generator responsible for errors on X -type stabilizers [$S_\alpha \in \mathcal{P}^{(Z)}(R_Z \cup L)$] is geometrically equivalent to a sparse lattice and mathematically equivalent to a system of nearest-neighbor, ferromagnetically coupled bits as in Fig. 9. For this thermal noise, one could equivalently write the Hamiltonian as

$$H = -J \sum_{(ij)} Z_i \otimes Z_j, \quad J > 0, \quad (63)$$

where thermal noises are given by Pauli-X operators.

We shall develop the calculation of classical memory time via the Peierls argument as follows. Let us start with one of the ground states, say $|00\dots 0\rangle$, and couple the system to the thermal bath which would flip spins and create regions with $|1\rangle$ states. We can view these clusters of $|1\rangle$ states as excitation droplets whose contours cost energy as spins are not aligned. If the droplet grow too large or too numerous, then erroneous spin flips may proliferate, leading to the loss of encoded information. The probability of observing each excitation droplet of excitation energy E is roughly given by $e^{-\beta E}$. There are many different shapes of excitation droplets with the same energy E , and the probability of observing an excitation droplet of energy E is given by $\Omega \exp(-\beta E) = \exp(-\beta F)$ where $F = E - TS$ with $S = \log \Omega$. Thus, the probability of observing an erroneous spin flip can be estimated by computing free energies associated with excitation droplets.

We pause briefly to interpret this probability in terms of quantum circuit complexity and error correction. Excitation droplets represent errors which could be eliminated by active error correction. However, performing error correction instantaneously would require nonlocal feedback circuits to compute the paths of all quasiparticles in the system. Instead, as in Sec. II, we divide the system into local grids and are interested in the probability that there exists some circuit acting inside each grid to remove excitations without inducing a global error. The cost of employing local error correction, such as majority voting in finite neighborhoods, is a nonzero probability of failure. In particular, if errors grow too numerous, the correction circuits, decided with imperfect knowledge of the trajectories of the quasiparticles, may induce a global error on the system. The aforementioned probability in terms of free energy, then, can be viewed as the probability of such a failure. Generally, at high temperature, excitations proliferate and failure is likely, but at low temperature, these configurations are suppressed, so failure is unlikely. The lifetime of information stored initially is then limited by the time taken for this probability to rise. See Sec. VB for more applications of this connection.

We begin calculating the free-energy barrier by defining a Peierls contour on the dual lattice (Fig. 9). A contour is defined as a closed path on the dual lattice along which there is a domain wall (one side is spin up, the other side is spin down), and let ℓ be used to denote its length. Specifically, the length is the number of edges crossed by the contour. A domain of frustration (relative to the majority), or droplet, needs to grow to have at least macroscopic length $\ell = O(L^\alpha)$ with energy barrier is $O(JL^\alpha)$. This is the lowest energy barrier of an error pathway. The entropy of a contour is lower bounded as

$$S(\ell) \geq \log(\Omega(\ell)) \geq \log(L^{(1-\alpha)\ell} 2^\ell). \quad (64)$$

Here $L^{1-\alpha}$ results from the fact that, for each edge, there are $L^{1-\alpha}$ possible edges crossed by the contour. The 2^ℓ term lower bounds the entropy originating from the shape of the contour as a random walk. We estimate the rescaled time τ based on the Arrhenius law for a reaction rate as

$$\tau = \frac{\exp(\beta E_b)}{\Omega} = \exp(\beta F_b), \quad (65)$$

where E_b is the energy barrier associated with the domain and F_b is the corresponding free-energy barrier. In this

calculation to find an upper bound on τ , we have corrected the Arrhenius law to account for the different paths by which this macroscopic energy barrier can be overcome:

$$F_b(\beta, L) \leq 2JL^\alpha - \frac{1}{\beta} \log(L^{(1-\alpha)L^\alpha} 2^{L^\alpha}). \quad (66)$$

Observe that at a given β there is a system size L_{\max} such that the lower bound on F_b is maximized, given by

$$L_{\max} \propto \exp(c\beta) \quad (67)$$

with $c > 0$ and dimensions of energy or temperature. Substitute this into the expression for the energy cost,

$$E_{b,\max} \propto \exp(c'\beta) \quad (68)$$

with $c' > 0$. Finally, substitute it into the expression for the bound on the memory time,

$$\tau \leq \tau_{\max} \propto \exp[\kappa' \exp(\kappa\beta)] \quad (69)$$

with κ and κ' positive.

While we relied on the Arrhenius law to estimate an upper bound on quantum memory time, the above argument can be turned into a rigorous derivation of upper bound on quantum memory time and lower bounds on gaps in \mathcal{L} . Rigorous approaches of finding mixing time for the Glauber dynamics often relies on refinement of the Peierls argument; see Ref. [28] for instance. Namely, one considers canonical paths between two configurations in order to bound the mixing time under Markovian dynamics [29]. Our estimates of the entropy correspond to the logarithm of the number of possible canonical paths connecting two configurations in different ground-state sectors. While a rigorous mathematical proof is beyond the scope of this paper, we believe that our estimation of free energy will lead to rigorous derivation of the mixing time for the Ising model on a sparse graph.

In Ref. [7], Bravyi and Haah obtained a general formula for obtaining lower bounds on quantum memory time. Applying their bound to the welded code, we obtain a lower bound of the form $\exp[a \exp(a'\beta)]$, which was also predicted by Michnicki. Indeed, the optimization of the number of qubits in Ref. [7] can be viewed as maximizing the free-energy barrier at fixed T . Since our argument provides an upper bound of the double exponential form, we conclude that quantum memory time of the welded code is doubly exponential in the inverse temperature β .

B. Thermal topological order and self-correction

The Peierls argument, based on estimation of free energy, allowed us to obtain upper bound on quantum memory time, showing that the welded code is not a self-correcting quantum memory. This analysis relates the notion of self-correcting quantum memory to the notion of topological order at nonzero temperature. In Sec. II C, we showed that the problem of approximating the Gibbs ensemble ρ_β boils down to evaluations of free energy associated to excitation droplets. Namely, we divide the system into local regions and bound the probability that there exists some quantum error correction that will remove excitations inside local regions by sending excitations to sinks; this probability can be viewed as the rate of successful convergence to the thermal Gibbs ensemble.

To be concrete, let us return to the problem of thermal instability of topological order for the welded code. To show that the welded code is not topologically ordered at nonzero temperature, it suffices to prove that an isolated excitation can be eliminated with high probability inside some box of length R . We have already seen that the probability of eliminating excitations inside a box is closely related to free energy associated with excitation droplets. In the above Peierls argument, we have observed that free energy becomes negative at $L = L_0 \approx e^{c'\beta}$ for some positive constant c' . This implies that the excitation can be eliminated inside the box of length R if $R \geq L_0$. As such, excitations can be eliminated almost surely inside a box of length $L_0 \log(L)$ and thus the Gibbs ensemble can be approximated via a range $L_0 \log(L)$ -range circuit. Here the $\log(L)$ factor is necessary because one needs to successfully eliminate excitations inside each grid and there are $(\frac{L}{R})^3$ grids.

Having studied memory time of the welded code via the Peierls argument, let us revisit the cubic code. In Appendix A, we provide an ergodic decomposition for quantum fractal codes which enables reduction to problems concerning classical fractal spin models. This approach allows one to obtain upper bound on quantum memory time of quantum fractal codes via the Peierls argument. Here, we derive the upper bound from the perspective of thermal instability of topological order in the cubic code. We have shown that, for a fixed temperature, an isolated excitation can be removed inside a box of length $R \approx \frac{c}{p^2} \log(L)$ for large L where $p = \frac{2}{1+e^\beta}$. Let us write R as $R \approx e^{c'\beta} \log(L)$ with some constant $c' > 0$, and vary the system size L . Let $L_0 = e^{c'\beta}$ be the value of L such that $R = L$. Then, for $L \leq L_0$, an isolated excitation cannot be removed inside some finite box since $R \leq L$, and its elimination requires a quantum circuit of $O(L)$ range. We sketch the growth of complexity of the Gibbs ensemble, as measured by the range of quantum circuits, in Fig. 10 where the growth is linear for $L \leq L_0$ while it suffers from logarithmic slowdown for $L \geq L_0$. This reasoning also shows that in the thermodynamic limit the transition temperature T_c to a topologically ordered phase goes to 0.

It is also illuminating to assess the free-energy barrier separating different ground-state sectors of the cubic code. Observe that $R \leq L$ implies that entropic contributions dominate energetic contributions and the free energy becomes very small. For $R \geq L$, energetic contributions are dominant which is order of $\log(L)$. Thus, the free-energy barrier will grow logarithmically and then become small as shown in Fig. 10. A

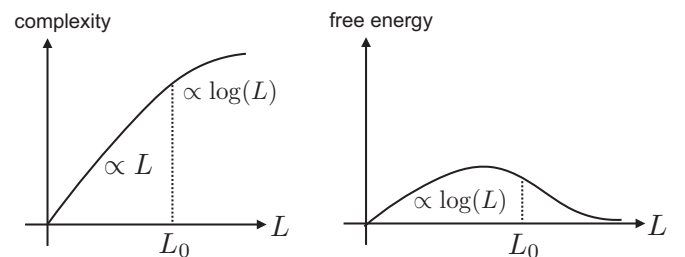


FIG. 10. Complexity of the Gibbs ensemble and free energy as a function of the system size L .

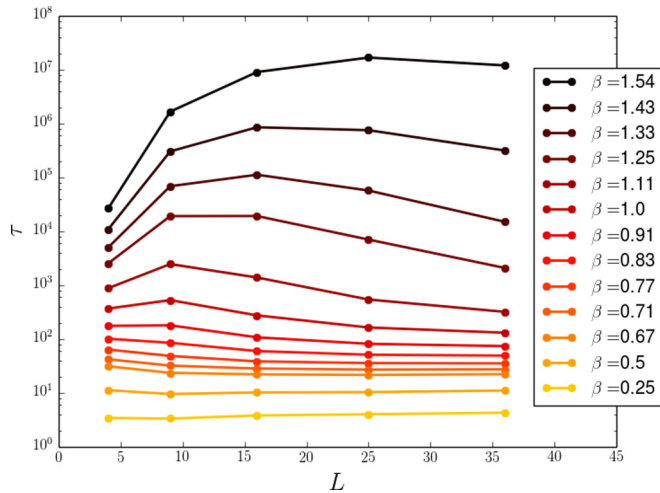


FIG. 11. Rescaled memory time τ as a function of system size L at different inverse temperatures β for the classical ferromagnetic sparse Ising model. The line is added to guide the eye along each temperature curve. At each temperature, there is an optimal system size, and the temperature at which L is the optimal size goes to zero as $L \rightarrow \infty$.

peak in the free-energy barrier is expected to be at $L \approx L_0$ which leads to the peak value of $O(\beta)$. This leads to the quantum memory time of $\exp(\kappa\beta^2)$ for some positive constant $\kappa > 0$.

C. Computational simulation

To verify the analytical result, the classical ferromagnetic sparse Ising model was simulated computationally. The Hamiltonian is $H = -\sum_{\langle i,j \rangle} Z_i Z_j$ where the linear size of the sparse graph is L while inner squares have linear length of \sqrt{L} ($\alpha = 1/2$). Numerical simulations were performed for values of L between 4 and 36, at β^{-1} from 0.65 to 2. In these numerics, the system is initialized in a completely ordered state, and transitions between successive states were governed by the Metropolis rule. The memory time is defined as the time required for the information stored initially to be compromised by at least 50%.

As discussed in the previous section, the memory time is expected to be doubly exponential in the inverse temperature which can be very long and poses difficulty in numerical simulations. For this reason, we relied on kinetic Monte Carlo methods using the Bortz-Kalos-Lebowitz (BKL) algorithm [30], a variation on the standard Metropolis algorithm. In this scheme, at each step the sites are divided into equivalence classes defined by the energy change that would be caused by flipping the spin at that site. The energy costs are calculated *a priori*, as well as the probabilities of once having flipped such a site that the new configuration is accepted. The BKL algorithm is particularly useful when the rates of spin flips are small.

The data from these simulations are reported in Fig. 11. At each temperature there is an optimal system size, as conjectured by Michnicki and derived in Eq. (67). Thus the welded code is only marginally self-correcting, as at no $T > 0$ can one grow the system arbitrarily large to achieve longer memory times. Qualitatively, this is because in large systems, it is more

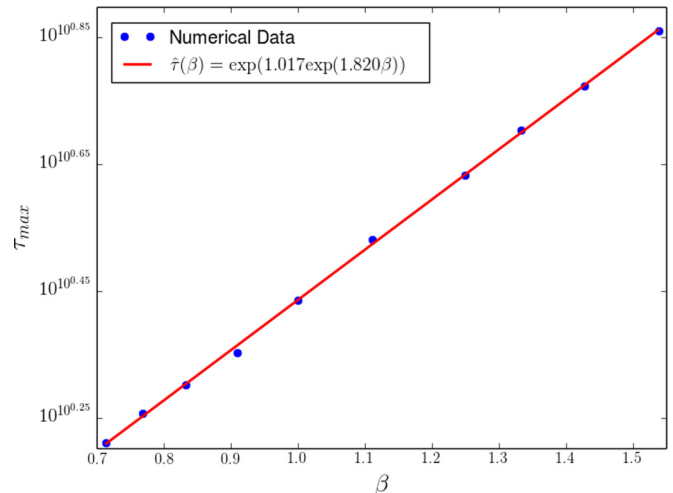


FIG. 12. Maximum memory time τ_{\max} as a function of β for the classical ferromagnetic sparse Ising model. Coefficient of determination $R^2 = 0.999567$ for this fit, which suggests that the double exponential prediction is accurate.

entropically favorable for droplets of frustration to proliferate and destroy the order. The other major crucial feature, though, is the scaling of the memory time with the inverse temperature β for each L_{\max} . Because L is not dense in the numerics, though, one cannot obtain to arbitrary precision the locations of the maxima at each β . The maxima were instead estimated by the highest data point available from each temperature. Nevertheless, it appears in Fig. 12 that these data points fit a doubly exponential law well (with coefficient of determination $R^2 = 0.999567$) with respect to β . The likely implication is that the memory time is not so strongly peaked that an off-peak estimate of the peak is not off by a large factor.

VI. DISCUSSION

In this paper, we proved the thermal instability of the cubic code and studied quantum memory properties of the welded code. In both arguments, we have utilized some variants of free-energy arguments, further establishing the connection between self-correction and topological order at nonzero temperature.

Our result raises a fundamental question concerning non-triviality of three-dimensional thermal Gibbs ensemble $\rho(\beta)$, or in other words, whether one can create an arbitrary thermal ensemble efficiently in three dimensions or not. The potential for CSS codes to achieve self-correction or topological order is particularly interesting, as these appear to be closely linked to models in classical physics. Perhaps, a reasonable step toward this question is to ask whether Brell's code is topologically ordered at nonzero temperature or not. The code can be viewed as a four-dimensional toric code supported on a fractal lattice with Hausdorff dimension less than 3. It has been argued that the code can be embedded into a three-dimensional space and the code has a finite transition temperature which is a strong indicative of the presence of some "order" at nonzero temperature. However such an embedding typically bring pairs of far-away sites to geometric proximity.

Another interesting question concerns the presence or absence of topological order in an intermediate Coulomb-like phase. Namely, for the four-dimensional qudit toric code, there exists an intermediate temperature range where quantum memory time diverges only polynomially with respect to the system size as opposed to exponential divergence at the low-temperature phase [31]. Whether such a phase is topologically ordered or not is an interesting question.

Our treatment, as well as the original result by Hastings, applies only to commuting frustration-free Hamiltonians. This limitation is due to technical difficulties of studying thermal ensembles of generic gapped Hamiltonians where perturbative analysis is not readily applicable in general. For instance, whether the four-dimensional toric code with small perturbation is topologically ordered at nonzero temperature is currently not known. This may be an interesting future problem.

An imperfect Hamiltonian may be interesting by its own. As a side result, we showed that an imperfect Hamiltonian $H(p)$ of a three-dimensional Hamiltonian H can be written in the form of “lattice gauge theories” where terms in $H(p)$ are either vertex and plaquette types after appropriate coarse-graining in Appendix D. Since lattice gauge theories often admit efficient characterizations via tensor product state (TPS) representations, our finding may hint that thermal Gibbs states for three-dimensional gapped spin systems can be efficiently studied by TPS formalism.

Generalization of an ergodic decomposition of the Davies generator to arbitrary stabilizer codes is an interesting future problem. Namely, studies of such decompositions in the presence of geometric locality of stabilizer generators may further strengthen no-go results for self-correcting quantum memory in two dimensions.

ACKNOWLEDGMENTS

We would like to thank J. Preskill and K. Temme for discussions and O. Motrunich for providing computational resources for numerics. This work was supported through Summer Undergraduate Research Fellowship (SURF) and NSF Grant No. PHY-1125565 at the California Institute of Technology. Research at Perimeter Institute was supported by the Government of Canada through Industry Canada and by the Province of Ontario through the Ministry of Research and Innovation.

APPENDIX A: DECOMPOSITION FOR FRACTALS

The ergodic decomposition method applies to quantum fractal codes, and the problem is reduced to finding classical memory time of classical fractal spin models [32,33]. Consider first the two-dimensional toric code represented by using polynomials: $\alpha = 1 + x$ and $\beta = 1 + y$. We show the decomposition in Fig. 13 where squares represent lattice sites which consist of two qubits. (Qubits live on faces, not on edges.) To construct the ergodic decomposition for quantum fractal codes, we repeat this decomposition in the \hat{x} direction. That is, for each layer of the \hat{y}, \hat{z} plane, we use the decomposition shown in Fig. 13. This reduces the problem to that of finding a classical memory time of a classical fractal spin model supported on

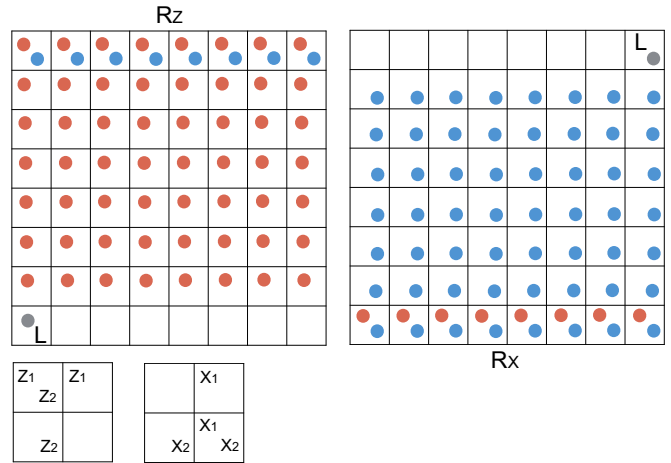


FIG. 13. Ergodic decomposition for the toric code in a polynomial representation.

multiple layers of two-dimensional planes which are attached by forming a comblike geometry.

APPENDIX B: WELDING AND GAPPED BOUNDARY

Welding is a procedure of combining multiple CSS codes into a larger CSS code. This procedure has an interesting interpretation in terms of anyon condensations in gapped boundaries. To be concrete, consider a welding of three copies of the two-dimensional toric code along a welding surface as shown in Fig. 14. Here, “welding” refers to a certain procedure of modifying stabilizer generators on the welding surface so that the resulting code has a set of commuting stabilizer generators.

There are two types of welding procedures, called X -type and Z -type weldings, which depend on how X -type and Z -type stabilizer generators are modified. In Fig. 14, we choose the X -type welding whose physical meaning is as follows. Let us denote three copies of the toric code by A, B, C , and associated charges and fluxes by $e_A, e_B, e_C, m_A, m_B, m_C$. In this setting, an electric charge from C bifurcates into two electric charges living on A and B respectively. On the other hand, a magnetic flux from each copy of the toric code can travel freely among A, B, C . More rigorously, an X -type weld of two CSS codes S_1 and S_2 is constructed as follows: pairs of qubits between S_1 and S_2 are identified and are each contracted to a single qubit and the new stabilizer group S of the welded code is generated by all Z -type operators from S_1 and S_2 and

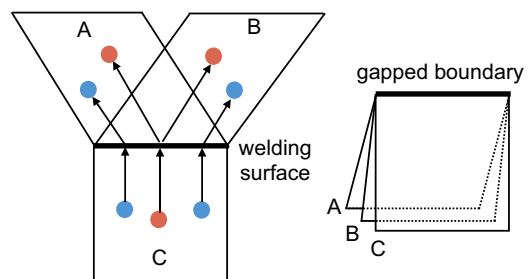


FIG. 14. Welding as a gapped boundary.

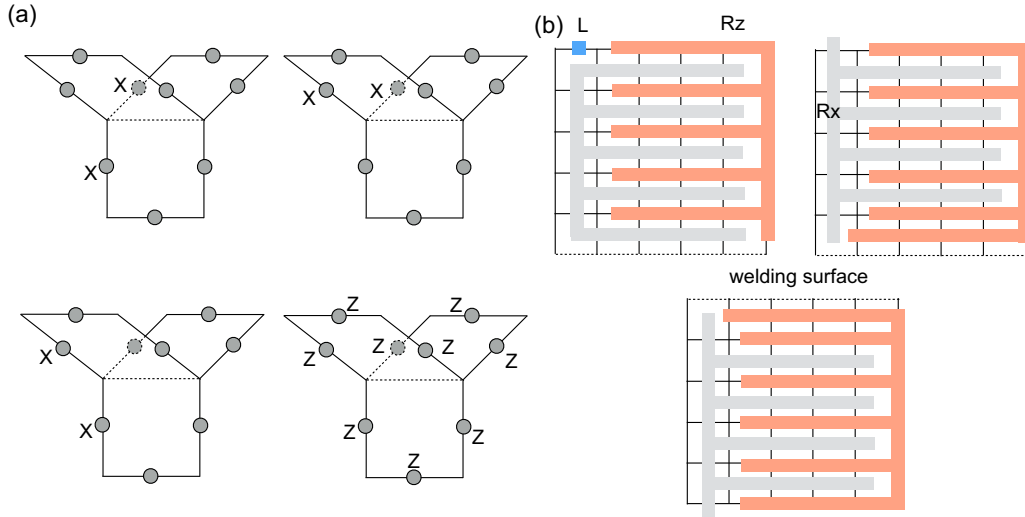


FIG. 15. (a) Terms near the welding surface. (b) The ergodic decomposition for the 2D welded code.

all X -type operators that commute with all of those Z -type operators. Further remarks on the code space can be found in Ref. [6]. By repeating this procedure, one can construct a larger CSS code, which is like a web of two-dimensional toric code. Since electric charges must bifurcate at welding surfaces, one can introduce an energy barrier for propagations of electric charges. Michnicki has applied this trick to multiple copies of the three-dimensional toric code to obtain a quantum memory with polynomially diverging energy barrier.

To discuss the welding procedure in the language of anyon condensations, it is convenient to fold the entire system along the welding surface as shown in Fig. 14. Then, the welding surface can be viewed as a gapped boundary defined for three copies of the toric code A, B, C . In this interpretation, condensing anyons are

$$e_A e_B e_C, m_A m_B, m_B m_C. \quad (\text{B1})$$

Figure 15(a) shows the terms near the welding surface explicitly. Note that the ergodic decomposition works in a similar manner as shown in Fig. 15(b). It has been shown that the three-dimensional topological color code with boundaries is equivalent to three copies of the toric code which are attached by a common boundary whose anyon condensations are given by Eq. (B1) [34]. In other words, the three-dimensional color code can be viewed as an outcome of welding three copies of the three-dimensional color code.

APPENDIX C: MORE ON WELDED CODE

In this appendix, we sketch a proof that the welded code is not topologically ordered at nonzero temperature. A formal construction of the welded code is rather involved, so interested readers are encouraged to read the original paper. The proof can be obtained straightforwardly by examining the patterns of anyon propagations, which was reviewed in Appendix B from the perspectives of gapped domain walls.

We begin by showing that the three-dimensional toric code is not topologically ordered at nonzero temperature. The proof was already given by Hastings. The model is defined as

follows:

$$H = - \sum_v A_v - \sum_p B_p, \quad (\text{C1})$$

where A_v is a Pauli- X vertex term and B_p is a Pauli- Z plaquette term. Violations of A_v correspond to pointlike excitations, and are referred to as electric charges. Violations of B_p correspond to looplike excitations, and are referred to as magnetic fluxes. We then consider an imperfect Hamiltonian $H(p)$ by removing terms of the original Hamiltonian. One technical subtlety is that lemma 1 does not apply to the three-dimensional toric code since there are too many redundant B_p operators. However, one can easily modify the lemma by using the CSS property of the toric code. As such, it suffices to consider an imperfect Hamiltonian $H(p)$ where only the vertex terms A_v are removed with probability p . We split the whole system into cubes of size $O(\log(L)^{1/3})$ which ensures that there is at least one sink of electric charge per cube with probability approaching to unity polynomially. Thus, the three-dimensional toric code is not topologically ordered at nonzero temperature. One can also prove that the three-dimensional toric code is not a self-correcting quantum memory by using the ergodic decomposition. Namely, Fig. 16 depicts the decomposition where Pauli Z operators (creating electric charges) form multiple comblike objects attached along a line, which have a finite energy barrier and are essentially described by one-dimensional Ising model, while Pauli X operators (creating magnetic fluxes) form sheetlike objects which are attached like a comb, which have $O(L)$ energy barrier.

The three-dimensional welded code can be constructed by gluing multiple copies of the three-dimensional toric code. As explained in the previous appendix, the welding procedure can be better understood from the paradigm of anyon condensations [35]. Consider six copies of the three-dimensional toric code. We label their anyonic excitations by e_1, \dots, e_6 and m_1, \dots, m_6 where e_j represent electric pointlike charges and m_j represent magnetic looplike fluxes. Let us label three coordinates of the lattice by x, y, z . We choose boundaries for surfaces perpendicular to x or y such that magnetic fluxes

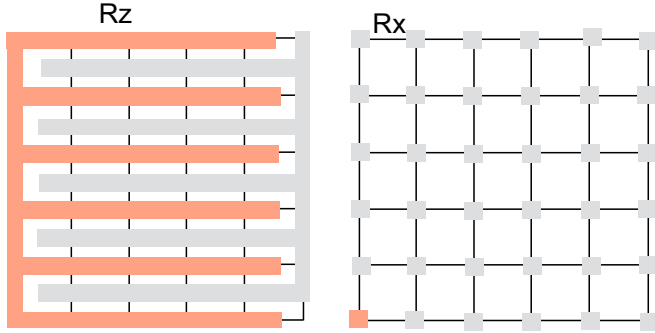


FIG. 16. The ergodic decomposition for the three-dimensional toric code. The figure shows only two layers of (x, y) planes in the three-dimensional lattice. For the full decomposition, one needs to stack these layers in the z direction.

condense. For the boundaries perpendicular to z , we construct joint boundaries involving six copies of the toric code. Namely, one can construct a gapped boundary for these six copies of the toric code such that the following anyonic excitations can condense into the boundary:

$$(e_1 \cdots e_6), \quad (m_1 m_j), \quad j = 2, \dots, 6. \quad (\text{C2})$$

One can easily see that braiding statistics of condensing anyons are mutually bosonic. Once the labels of condensing anyons are identified, it is easy to construct the corresponding gapped boundary explicitly; see [36,37] for instance. Note that a single electric charge, say e_1 , cannot condense into the boundary. Instead, if e_1 moves to the boundary, then it will be reflected back as a composite $e_2 e_3 e_4 e_5 e_6$. Since an electric charge splits into multiple excitations, it adds extra energy penalty for propagations of electric charges which are otherwise freely propagating. This joint boundary, involving six copies of the toric code, is indeed the welding surface which glues six copies of the toric code.

The three-dimensional welded code is essentially a web of multiple copies of the three-dimensional toric code which are glued together with boundaries described above. To be more specific, consider the three-dimensional cubic (sparse) lattice where its linear length is $O(L)$ and its edges have length $O(L^a)$ ($0 < a < 1$). [So, there are $O(L^{1-a})$ vertices in one direction of the sparse lattice.] We then place the three-dimensional toric code of linear size $O(L^a)$ on each edge of the lattice, and glue six copies of the toric code at boundaries which are facing to the same vertex of the sparse lattice. The gluing process requires deforming the shape of each of the toric code, but the point is that this process preserves locality of interaction terms as well as keeps the density of qubits finite [6]. By performing such gluing at each vertex, we obtain the three-dimensional welded code. The energy barrier for a magnetic looplike flux is $O(L^a)$ which is determined by the barrier for the original toric code of size $O(L^a)$. The energy barrier for an electric charge is $O(L^{2(1-a)})$ since electric charges split into multiple charges once they cross the welding surfaces. By setting $a = 2/3$, both excitations will have energy barrier $O(L^{2/3})$.

Let us finally show that the welded code is not topologically ordered at nonzero temperature. We shall split the system into blocks of $O(\log(L))$ qubits. If the box does not contain

the welding surface, then all the excitations can be brought to the sink, and can be eliminated with high probability. When the box contains the welding surface, an extra care is necessary. In such cases, we consider six blocks, attached at the welding surface, as a single region. Then, each of six types of excitations from six copies of the toric code can be eliminated in each block. Therefore, all the excitations can be eliminated in a box of size $O(\log(L))$. We also make a brief comment on the ergodic decomposition in the three-dimensional welded code. On the bulk, away from the welding surfaces, the decompositions look like those for the three-dimensional toric code, as depicted in Fig. 16 while, on the welding surfaces, Pauli X and Z operators are connected to other cubes in a way similar to Fig. 15(b). Here we are interested in the decomposition for Pauli- Z operators. Away from the welding surface, Pauli- Z operators form a comblike object whose thermodynamic properties can be essentially studied as a one-dimensional Ising model. On the welding surface, their trajectory splits into multiple combs which gives energy penalty for electric charges. As a whole, Pauli- Z operators form an object which is like a sparse Ising model.

APPENDIX D: REDUCTION TO LATTICE GAUGE THEORY

In this section, we show that an imperfect realization $H(p)$ of a three-dimensional local Hamiltonian H can be almost surely coarse-grained into the form of lattice gauge theories. By lattice gauge theories, we mean a class of interacting quantum spin systems where physical degrees of freedom live on edges of a graph and interaction terms are either vertex or plaquette types:

$$H_{LGT} = - \sum_v \sum_j A^{(v)}_j - \sum_p \sum_j B^{(p)}_j, \quad (\text{D1})$$

where $A^{(v)}_j$ acts only on edges radiating from a vertex v and $B^{(p)}_j$ acts on edges that form a plaquette p (see Fig. 17).

For a proof, one terminology needs to be introduced. Let $r > 0$ be a range of interaction terms, meaning that each term is contained in some ball of radius r . By assumption, r is finite. Consider a ball of finite radius r' ($r' \gg r$) such that all the terms in $H(p)$ act trivially on the ball. We call such a

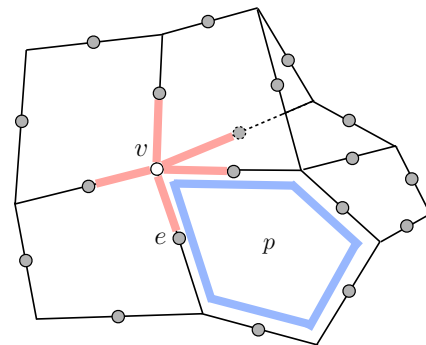


FIG. 17. A lattice gauge theory H defined on a homogeneous graph. Composite spins live on edges of the graph. Red lines represent a vertex operator on a vertex v and blue lines represent a plaquette operator on a plaquette p .

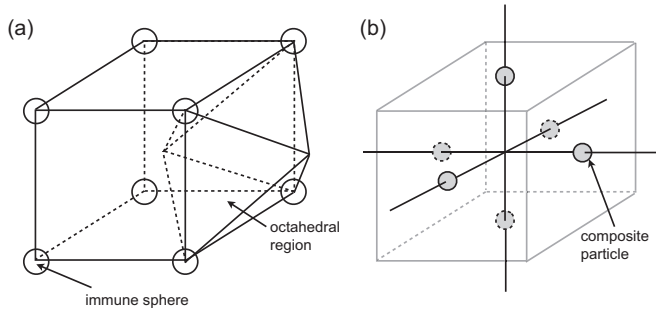


FIG. 18. (a) Immune spheres at vertices on a cubic grid. Shaded dots represent spheres at half-integer sites. We split the entire system into octahedral objects and view them as composite spins. (b) A construction of a lattice gauge theory. Composite spins live on faces of a cubic lattice.

ball of no interaction terms an immune sphere. Note that an immune sphere appear with some finite probability as long as r' is finite.

We begin by demonstrating a reduction to a lattice gauge theory for a certain simple realization of $H(p)$. Consider a cubic grid with spacing $d \gg r'$ on a lattice, denoted by Λ_{grid} . First, we study a special case where immune sphere are present at vertices of the grid. We draw lines connecting immune spheres and the centers of the cubes, then split the entire system into a collection of octahedral objects Fig. 18(a), each of which

can be identified with a face of the cubic lattice. The system is coarse-grained so that degrees of freedom in each octahedral object form single composite spins which live on faces of the cubic grid as in Fig. 18(b). We then construct a cubic graph Λ by connecting centers of unit cells in the grid so that composite particles live on edges of this graph.

We now show that interaction terms of a local Hamiltonian H are either vertex or plaquette type. One can classify interaction terms into three types: (a) terms which live near an edge of the cubic grid; (b) terms which live near a face of the cubic grid; (c) terms which live in the bulk of a cube. For (a), such terms can be considered as plaquette terms which are at most four-body. For (b), such terms are one-body operators. For (c), such terms can be considered as vertex terms which are at most six-body. Thus, the system can be viewed as a lattice gauge theory on a graph Λ .

Let us treat a generic case where interaction terms are removed randomly with probability p . We split the full three-dimensional system into unit cubes of $d \times d \times d$ qubits where $d > 0$ is some sufficiently large positive integer such that $d \gg r' \gg r$. For a given unit cube, the probability of finding at least one immune sphere of radius r' is lower bounded by some finite constant $q > 0$. By taking sufficiently large d while keeping r' , one can make q arbitrarily close to unity. One can connect immune spheres and construct a skewed cubic grid. We then use a coarse-graining strategy similar to the one used in the previous section. This reduces the system into a form of lattice gauge theories.

- [1] E. Dennis, A. Kitaev, A. Landahl, and J. Preskill, *J. Math. Phys.* **43**, 4452 (2002).
- [2] C. G. Brell, *New J. Phys.* **18**, 013050 (2016).
- [3] B. J. Brown, D. Loss, J. K. Pachos, C. N. Self, and J. R. Wootton, *Rev. Mod. Phys.* **88**, 045005 (2016).
- [4] O. Landon-Cardinal, B. Yoshida, D. Poulin, and J. Preskill, *Phys. Rev. A* **91**, 032303 (2015).
- [5] J. Haah, *Phys. Rev. A* **83**, 042330 (2011).
- [6] K. P. Michnicki, *Phys. Rev. Lett.* **113**, 130501 (2014).
- [7] S. Bravyi and J. Haah, *Phys. Rev. Lett.* **111**, 200501 (2013).
- [8] I. H. Kim, *Phys. Rev. A* **83**, 052308 (2011).
- [9] F. L. Pedrocchi, A. Hutter, J. R. Wootton, and D. Loss, *Phys. Rev. A* **88**, 062313 (2013).
- [10] C.-E. Bardyn and T. Karzig, *Phys. Rev. B* **94**, 094303 (2016).
- [11] S. Bravyi and B. Terhal, *New J. Phys.* **11**, 043029 (2009).
- [12] J. Haah and J. Preskill, *Phys. Rev. A* **86**, 032308 (2012).
- [13] O. Landon-Cardinal and D. Poulin, *Phys. Rev. Lett.* **110**, 090502 (2013).
- [14] F. Pastawski and B. Yoshida, *Phys. Rev. A* **91**, 012305 (2015).
- [15] B. Yoshida, *Ann. Phys.* **326**, 2566 (2011).
- [16] R. Alicki, M. Fannes, and M. Horodecki, *J. Phys. A: Math. Gen.* **42**, 065303 (2009).
- [17] A. Komar, O. Landon-Cardinal, and K. Temme, *Phys. Rev. A* **93**, 052337 (2016).
- [18] R. Alicki, M. Horodecki, P. Horodecki, and R. Horodecki, *Open Syst. Inf. Dyn.* **17**, 1 (2010).
- [19] M. B. Hastings, *Phys. Rev. Lett.* **107**, 210501 (2011).
- [20] X. Chen, Z.-C. Gu, and X.-G. Wen, *Phys. Rev. B* **82**, 155138 (2010).
- [21] R. Peierls, *Math. Proc. Cambridge Philos. Soc.* **32**, 477 (1936).
- [22] J. Haah, *Commun. Math. Phys.* **324**, 351 (2013).
- [23] B. Yoshida, *Phys. Rev. B* **88**, 125122 (2013).
- [24] E. B. Davies, *Commun. Math. Phys.* **39**, 91 (1974).
- [25] H. Spohn, *Lett. Math. Phys.* **2**, 33 (1977).
- [26] A. Frigerio, *Commun. Math. Phys.* **63**, 269 (1978).
- [27] B. Yoshida and I. L. Chuang, *Phys. Rev. A* **81**, 052302 (2010).
- [28] F. Martinelli, in *Lectures on Probability Theory and Statistics (Saint-Flour, 1997)* (Springer, Berlin, 1999), pp. 93–191.
- [29] K. Temme, *Commun. Math. Phys.* **350**, 603 (2017).
- [30] A. Bortz, M. Kalos, and J. Lebowitz, *J. Comput. Phys.* **17**, 10 (1975).
- [31] B. Yoshida, [arXiv:1404.0457](https://arxiv.org/abs/1404.0457).
- [32] M. E. J. Newman and C. Moore, *Phys. Rev. E* **60**, 5068 (1999).
- [33] B. Yoshida, *Ann. Phys. (N.Y.)* **338**, 134 (2013).
- [34] A. Kubica, B. Yoshida, and F. Pastawski, *New J. Phys.* **17**, 083026 (2015).
- [35] M. Levin and Z.-C. Gu, *Phys. Rev. B* **86**, 115109 (2012).
- [36] S. Beigi, P. W. Shor, and D. Whalen, *Commun. Math. Phys.* **306**, 663 (2011).
- [37] B. Yoshida, *Ann. Phys.* **377**, 387 (2017).