A little bit of self-correction

Michael Kastoryano[1,2], Lasse B. Kristensen[1], Chi-Fang Chen[3], and András Gilyén[4]

[1]Department of Computer Science, University of Copenhagen, Denmark
[2]AWS Center for Quantum Computing, Pasadena, CA
[3]Institute for Quantum Information and Matter, California Institute of Technology, Pasadena, CA, USA
[4]HUN-REN Alfréd Rényi Institute of Mathematics, Budapest, Hungary

arXiv:2408.14970v2 [quant-ph] 4 Jul 2025

We investigate the emergence of stable subspaces in the low-temperature quantum thermal dynamics of finite spin chains. Our analysis reveals the existence of effective decoherence-free qudit subspaces, persisting for timescales exponential in β. Surprisingly, the appearance of metastable subspaces is not directly related to the entanglement structure of the ground state(s). Rather, they arise from symmetry relations in low-lying excited states. Despite their stability within a 'phase', practical realization of stable qubits is hindered by susceptibility to symmetry-breaking perturbations. This work highlights that there can be non-trivial quantum behavior in the thermal dynamics of noncommuting many body models, and opens the door to more extensive studies of self-correction in such systems.

The pursuit of a self-correcting quantum memory, while still largely elusive, has been an exceptionally fertile ground for identifying new approaches to robustly process quantum information [1, 2, 3, 4, 5, 6, 7, 8]. Ideally, a self-correcting quantum memory would be a system that inherently stabilizes itself against thermal, stochastic and coherent noise. A less stringent, yet still valuable, form involves a qubit subspace that is robust against thermal fluctuations but may not withstand arbitrary non-thermal perturbations. At present, most of our understanding of thermal stability is based on commuting models [9, 10], mostly because, until recently, no reliable mathematical model existed for studying the thermalization of noncommuting many-body systems.

The interplay between quantum mechanics and statistical mechanics becomes particularly intriguing at low temperatures. While the high-temperature behavior of non-commuting systems tends to mimic classical Gibbs states [11, 12], low temperatures reveal a richer structure. The characteristics of the low-lying quantum states, particularly entanglement, intertwine with the statistical features of the Gibbs distribution. When the temperature is lower than the spectral gap, we expect the Gibbs state to approximate the ground state, suggesting that the quantum thermal dynamics should mirror the properties of the ground state. This leads to the hypothesis that the mixing time—the timescale over which the system reaches thermal equilibrium—might diverge at a quantum phase transition at low temperatures. Such a hypothesis would be consistent with the statics-dynamics connection that is well studied for classical spin systems [13] and commuting quantum spin systems [6, 14].

Building on the new quasi-local noncommutative Gibbs samplers constructed in Refs. [15, 16, 17, 18], we investigate the low-temperature quantum thermal dynamics of a specific class of one-parameter quantum spin chains: the $J_1 - J_2$ chain. Spin chains are notable as they lack a thermal phase transitions, yet they undergo quantum phase transitions [19]. Our analysis reveals that: (i) There exist qudit subspaces that remain decoherence-free for timescales on the order of $\Omega(N^{-1}e^{\beta c})$, where β is the inverse temperature, c is a constant related to gaps in the low lying spectrum of the Hamiltonian, and N is the chain length. (ii) The existence of metastable subspaces does not require ground state degeneracy of the model. Rather, it originates from symmetry constraints in the low lying spectrum, and depends on the system-bath coupling operators. (iii) The 'phase diagram' of these stable subspaces does not necessarily mirror the ground state quantum phase diagram. Instead, the metastable subspaces can appear and disappear at (avoided-

Michael Kastoryano: mika@di.ku.dk

)crossings of the low-lying excited states of the model (see Fig. 1).

Despite the intriguing nature of these findings, we temper expectations as for their practical utility in building noise resilient qubits. The primary limitation stems from the protection mechanism's reliance on discrete symmetries. Such symmetries are vulnerable to local perturbations, which can easily disrupt the stability afforded by these decoherence-free subspaces. Consequently, while the stable subspaces reveal interesting new physics in the thermal dynamics of quantum systems, they may not translate into practical, fault-tolerant quantum memories.

This work is significant for several reasons. First, it provides initial evidence of non-trivial quantum behavior of thermalization. This marks a departure from the classical understanding, showcasing how quantum effects can alter thermal dynamics in unexpected ways. Second, it highlights the critical influence of the choice of elementary jump operators on the system's dynamics. This observation suggests that subtle details in the model's implementation can have profound impacts on its behavior. Finally, the presence of approximate decoherence-free subspaces within the thermal dynamics of spin systems suggests that such phenomena might be more common than previously thought, opening new avenues for exploring the interplay between quantum coherence and thermal noise.

1 Metastability

While ground states of commuting Hamiltonians always have finite correlation length, noncommuting quantum many-body systems can have extensive ground state correlations, allowing for various forms of topological and symmetry-driven quantum phase transitions [19, 20]. While the quantum thermal dynamics of commuting systems is fairly well understood [6, 21], the quantum thermal dynamics of noncommuting many-body systems is essentially uncharted territory. This is due in large part to the lack of a go-to quantum thermal dynamics algorithm. Recently, a subset of the authors introduced a (quasi-)local quantum Lindbladian that satisfies detailed balance with respect to a noncommuting quantum Gibbs state [16], hence allowing for the analytic and numerical analysis of quantum Gibbs sampling. This new construction will be the starting point for our analysis.

We consider a quantum system described by a local many-body Hamiltonian H, on a regular lattice Λ. The thermal Markovian dynamics elaborated in Refs. [15, 16, 17, 18] is defined by a Lindbladian (the generator of a quantum dynamical semigroup)

$$\mathcal{L}(\rho) = -i[Q,\rho] + \sum_a L_a \rho L_a^\dagger - \frac{1}{2}(L_a^\dagger L_a \rho + \rho L_a^\dagger L_a). \tag{1}$$

L_a are the quantum jumps, or Lindblad operators, and are given explicitly by

$$L_a = \int_{-\infty}^{\infty} g(t) e^{-itH} S_a e^{itH} dt, \tag{2}$$

where $\{S_a\}$ are a complete set of elemental jump operators that guarantee ergodicity of the dynamics. In the case of a quantum spin model on a 1D lattice (a ring), it is customary to choose $\{S_a\}$ to be the set of local Pauli operators: $S_a \in \{X_n, Y_n, Z_n\}_{n\in\Lambda}$. $g(t) = \int_{-\infty}^{\infty} \sqrt{\gamma(\omega)} e^{it\omega} d\omega$, is a damping term, and $\gamma(\omega)$ is related to the bath auto-correlation function, and incorporates the system bath coupling strength. For concreteness, we assume that $\gamma(\omega) \propto 1/(1+e^{\beta\omega})$ for a specific temperature β. This choice is reminiscent of an Ohmic bath [22], but all of our conclusions hold more generally, as long as $\gamma(-\omega) = e^{\beta\omega}\gamma(\omega)$. We choose to work with the specific form of jump operators in Eqn. (2) for concreteness, however there is a wider class of constructions which lead to the same conclusions [17, 18].

Q is a Hermitian matrix, playing the role of the system Hamiltonian. The two natural candidates for Q are (i) $Q \propto H$, the Hamiltonian of the system, and (ii)

$$Q = \frac{i}{2} \sum_{ij} \tanh\left(\frac{\beta(e_i - e_j)}{4}\right) \langle e_i | \sum_a L_a^\dagger L_a |e_j\rangle |e_i\rangle\langle e_j|, \tag{3}$$

written out in the eigenbasis of $H = \sum_j e_j |e_j\rangle\langle e_j|$. The latter choice guarantees that the stationary state of the dissipative dynamics is the Gibbs state; i.e. $\mathcal{L}(\rho_\beta) = 0$ with $\rho_\beta \propto e^{-\beta H}$. The specific form in Eqn. (3) is necessary to enforce the quantum detailed balance condition [16, 17, 18]—the key mathematical relationship governing both classical and quantum Gibbs samplers. The first choice $Q = H$ is closer to the natural dynamics resulting from a Markovian weak-coupling limit, as shown in Ref. [23]. However, it

does not lead to a stationary state that is globally Gibbs at any temperature. We will see that at low enough temperatures both choices lead to metastable quantum subspaces, though $Q = H$ induces coherent dynamics in the metastable subspace.

It is worth pausing at this point to discuss how important the choice of bath couplers $\{S_a\}$ is for the thermalization process. By analogy with the classical Metropolis algorithm, it is generally believed that as long as the operators $\{S_a\}$ are local on the lattice, and span the full algebra of the space, then the dynamics will not depend sensitively on the specific choice of couplers $\{S_a\}$. This intuition builds upon a deep correspondence from Glauber dynamics of classical lattice spin systems, which shows a strong correspondence between static properties of the Gibbs states and the mixing time of the dynamics [13], at least in the case of finite-dimensional lattice systems [24]. We show, on the contrary, that the choice of couplers plays a crucial role in the quantum case in the large β limit, when the statics is dominated by the ground state.

Decoherence-free subspaces We say that a Lindbladian admits a decoherence-free subspace (DFS) $\mathcal{D}$ if for any state $\psi \in \mathcal{D}$, $\mathcal{L}(\psi) = 0$ [25, 26]. Equivalently [27], $\mathcal{D}$ is a DFS of $\mathcal{L}$ if $L_a|\psi\rangle = 0$ and $Q|\psi\rangle \propto |\psi\rangle$, for all Kraus operators L_a, and any pure state $|\psi\rangle \in \mathcal{D}$. We will consider an ϵ-approximate version of the DFS (ϵ-DFS):

$$||L_a|\psi\rangle|| \leq \epsilon||L_a|| \quad (4)$$
$$||Q|\psi\rangle|| \leq \epsilon^2||Q||. \quad (5)$$

In general, the Hamiltonian only needs $|\psi\rangle$ to be an approximate eigenstate for the subspace to be ϵ-DFS, but for our purpose, the more restrictive definition is sufficient. As mentioned above, this requirement can be further relaxed to allow for unitary dynamics within the subspace [28], as occurs in the case $Q = H$.

Definition 1.1 (Blind qubit subspace) *Consider a Hamiltonian $H = \sum_j e_j|e_j\rangle\langle e_j|$, with its eigenvalues in increasing order: $e_0 \leq e_1 \leq ...$, and a set of elemental jump operators $\{S_a\}$. If there exists an eigenvector $|e_k\rangle$, $k > 0$ such that $\langle e_j|S_a|e_k\rangle = 0$ for all $j \leq k$, and $\langle e_0|S_a|e_0\rangle = 0$, then $\mathcal{S} = \text{span}(|e_0\rangle, |e_k\rangle)$ is a blind qubit subspace of $\{S_a\}$.*

The extension to qudit subspaces is straightforward. We now show that a blind subspace of $\{S_a\}$ is metastable. Let $\mathcal{S}$ be a blind subspace of $\{S_a\}$, and consider $\psi \in \mathcal{S}$. We time evolve this state with respect to the thermal Lindbladian in Eqn. (1) for a time t. Then, note that [29]

$$||\psi(t) - \psi||_1 = ||\int_0^t \dot{\psi}(s)ds||_1 \leq t||\mathcal{L}(\psi)||_1. \quad (6)$$

Writing the Lindbladian as $\mathcal{L}(\rho) = T(\rho) + (\kappa\rho + \rho\kappa^\dagger)$, with $T(\rho) = \sum_a L_a\rho L_a^\dagger$ and $\kappa = -\frac{1}{2}\sum_a L_a^\dagger L_a - iQ$ we get

$$||\mathcal{L}(\psi)||_1 \leq 2||\kappa|\psi\rangle|| + ||T(\psi)||_1. \quad (7)$$

If ψ is in an ϵ-DFS subspace, then $||\kappa|\psi\rangle|| = O(\epsilon N)$, and $||T(\psi)||_1 = O(\epsilon^2 N)$, where $N = |\Lambda|$ is the volume, so the dominant contribution will come from $||\kappa|\psi\rangle||$ term.

Let us now look at the specifics of the thermal Lindbladian. Consider the action of L_a on an eigenstate of H:

$$L_a|e_j\rangle = \int_{-\infty}^{\infty} g(t)e^{-itH}S_a e^{itH}|e_j\rangle\, dt \quad (8)$$
$$= \sum_k \int_{-\infty}^{\infty} g(t)e^{-it(e_k-e_j)}\langle e_k|S_a|e_j\rangle|e_k\rangle\, dt$$
$$= 2\pi\sum_k \sqrt{\gamma(e_k-e_j)}\langle e_k|S_a|e_j\rangle|e_k\rangle. \quad (9)$$

By construction, note that $\gamma(x) \leq 1$, then combined with the property that $\gamma(\omega) = e^{-\beta\omega}\gamma(-\omega)$, we have that $\gamma(\omega) \leq e^{-\beta\omega}$. Then

$$||L_a|e_j\rangle|| \leq 2\pi\sum_k e^{\beta(e_j-e_k)/2}|\langle e_k|S_a|e_j\rangle| \quad (10)$$
$$\leq 2\pi\sum_{k=1}^{j} e^{\beta(e_j-e_k)/2}|\langle e_k|S_a|e_j\rangle| \quad (11)$$
$$+ O(e^{-\beta\Delta/2}), \quad (12)$$

where $\Delta = \min\{e_{j+1} - e_j, e_p - e_0\}$ with $|e_p\rangle$ the lowest-lying state not in the blind subspace, and we assume that $\sum_{k>j}|\langle e_k|S_a|e_j\rangle|e^{-\beta(e_k-e_{j+1})/2} = O(1)$. The latter assumption will be true whenever the bulk of the eigenvalues of H is sufficiently removed from the extremities of the spectrum, as is typically the case for local lattice models [1]. If $|e_0\rangle$ and $|e_j\rangle$ span a *blind qubit subspace* with respect to the elemental jump operators $\{S_a\}$, i.e.

$$|\langle e_k|S_a|e_j\rangle| = 0 \quad (13)$$

[1] See the appendix B.2 for a discussion of this assumption in the setting of the $J_1 - J_2$ model.

for all $k \leq j$, then we get that

$$||L_a|e_j\rangle|| \leq O(e^{-\beta\Delta/2}). \tag{14}$$

The 'Hamiltonian' term Q given by Eqn. (3) can be controlled similarly (see the appendix), leading to an overall scaling of $||\mathcal{L}(\psi)||_1 = O(Ne^{-\beta\Delta/2})$. Hence, from Eqn. (6), any state in the blind subspace survives for a time $\Omega(N^{-1}e^{\beta\Delta/2})$. The coefficient Δ is governed by the gaps above e_0 and e_j.

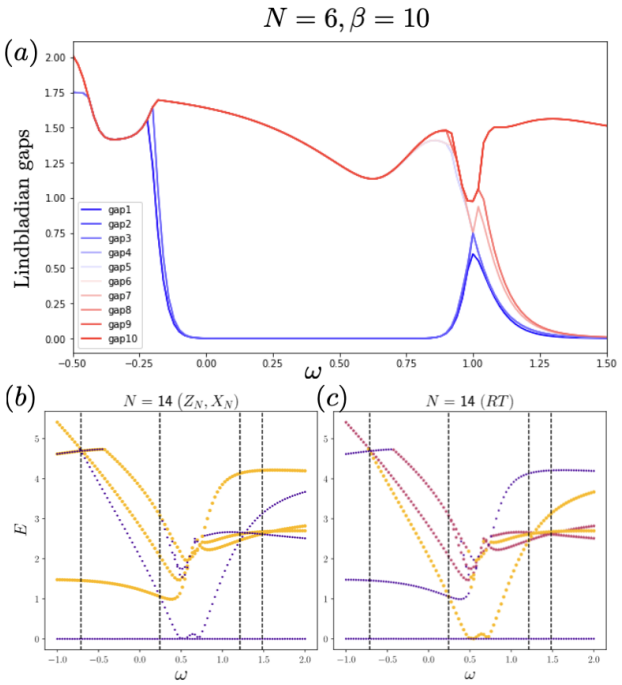


Figure 1: (a) Numerical simulation of the spectral gaps of the Lindbladian in Eqn. (1) for the J_1-J_2 model, with $N = 6$ sites, and $\beta = 10$. The zero eigenvalue of $\mathcal{L}$ is omitted. β is in the low temperature regime ($\beta \gg \text{gap}^{-1}(H)$) everywhere except around the Majumdar-Ghosh point ($\omega = .5$), where the Hamiltonian gap closes. Note that the operator $\mathcal{L}$ has real spectrum because of detailed balance [15]. The region $0 \leq \omega \leq .8$ has four nearly degenerate eigenvalues, while the region $\omega \geq 1.5$ has nine, reflecting qubit and qutrit metastable subspaces, respectively. (b) Low lying spectrum of H; i.e. $e_k - e_0$ for $k = 1, ..., 12$. The excited eigenstates are color coded to indicate which symmetry sector they belong to. The eigenstates $|e_k\rangle$, $k \neq 0$, are plotted in blue if $\langle e_k|Z_N|e_k\rangle = \langle e_0|Z_N|e_0\rangle$ and $\langle e_k|X_N|e_k\rangle = \langle e_0|X_N|e_0\rangle$, and plotted in yellow otherwise. (c) The eigenstates are plotted in blue if $\langle e_k|RT|e_k\rangle = \langle e_0|RT|e_0\rangle$, in yellow if $\langle e_k|RT|e_k\rangle \neq \langle e_0|RT|e_0\rangle$, and in pink if $|e_k\rangle$ is not a simultaneous eigenstate of R and T. Numerics were performed using the QuTiP library [30].

2 Symmetries

In order to illustrate the existence of metastable subspaces in the thermal dynamics, we will concentrate primarily on the $J_1 - J_2$ model, which exhibits a rich ground state and dynamical symmetry phase diagram (see Fig. 1). The 1D quantum J_1-J_2 model describes the behavior of spins on a one-dimensional lattice with competing nearest-neighbor (J_1) and next-nearest-neighbor (J_2) exchange interactions. The Hamiltonian is given by:

$$H = \sum_{n=1}^{N} J_1 \mathbf{S}_n \cdot \mathbf{S}_{n+1} + J_2 \mathbf{S}_n \cdot \mathbf{S}_{n+2}, \tag{15}$$

where $\mathbf{S}_n = (X_n, Y_n, Z_n)^T$ in the spin-1/2 case. We will assume periodic boundary conditions and N even. J_1 and J_2 are the (real) exchange coupling constants. Without loss of generality, we will fix $J_1 = 1$ and write $J_2/J_1 \equiv \omega$. The system has a gappeless unique ground state when $\omega < \omega^*$, and a gapped phase for $\omega > \omega^*$ with degenerate ground state in the thermodynamic limit. The system undergoes a Kosterlitz-Thouless transition at $\omega^* \approx 0.2411$ [31]. The point $\omega = 0.5$—the Majumdar-Ghosh point—has a known exact solution. Specifically, the degenerate ground states at $\omega = 0.5$ have an exact representation as matrix product states. The spin-1/2 model enjoys a number of discrete symmetries: total 'charge' $X_N = \bigotimes_n X_n$, total 'phase' $Z_N = \bigotimes_n Z_n$, translation symmetry $T : S_n^\alpha \to S_{n+1}^\alpha$, and reflection symmetry $R : S_n^\alpha \to S_{L-n+1}^\alpha$, along any line cutting the ring in two equal-size pieces (see Fig. 2).

We start our analysis with the spin-1/2 case at the Majumdar-Ghosh point: $\omega = 0.5$. At this special point, the model has a degenerate ground subspace spanned by the states: $|L\rangle = \bigotimes_{e \in E_{\text{even}}} |-_e\rangle$ and $|R\rangle = \bigotimes_{o \in E_{\text{odd}}} |-_o\rangle$, where $|-_n\rangle = (|00\rangle_{n,n+1} - |11\rangle)_{n,n+1}/\sqrt{2}$. The tensor product is taken over even edges of the ring for the $|L\rangle$ state and over odd edges for the $|R\rangle$ state. These are known as valence bond solid (VBS) states. The model is believed to remain in the VBS phase all the way down to the phase transition point ω^*. A straightforward calculation shows that the states $|L\rangle, |R\rangle$ are protected against local Pauli errors:

$$\langle L|P_n|R\rangle = \langle L|P_n|L\rangle = \langle R|P_n|R\rangle = 0, \tag{16}$$

where $P_n \in \{X_n, Y_n, Z_n\}$. To see this, note that when N is even, X_N, Z_N and H all commute with each other, therefore the eigenvectors of H are also eigenvectors of X_N and Z_N with eigenvalues ± 1. In particular, $Z_N|L,R\rangle = |L,R\rangle$ and $X_N|L,R\rangle = |L,R\rangle$ for $N/2$ even, and $X_N|L,R\rangle = -|L,R\rangle$ for $N/2$ odd [2]. Then, for any site $n = 1, 2, ...$, we note that

$$\langle L|Z_n|L\rangle = \langle L|X_N Z_n X_N|L\rangle \quad (17)$$
$$= -\langle L|Z_n|L\rangle, \quad (18)$$

where the second equality follows from anti-commutation of X_N and Z_n. It follows that the expectation value $\langle L|Z_n|L\rangle$ has to be zero. The same argument holds for all combinations of $|L\rangle, |R\rangle$ and all local Pauli operators $\{X_n, Y_n, Z_n\}$. All we need for the argument to hold is for $|L\rangle$ and $|R\rangle$ to have the same eigenvalue under Z_N and X_N. It turns out [32] that the third eigenstate does not satisfy this property, hence the protection is restricted to a two-dimensional subspace. By the arguments above, the subspace spanned by $|L\rangle, |R\rangle$ is a blind subspace with respect to local Paulis, and hence an approximate DFS under this noise. Note that the constraints arising from discrete symmetries are somewhat reminiscent of the selection rules governing allowed and forbidden transitions in nuclear, atomic, and molecular systems [33, 34, 35].

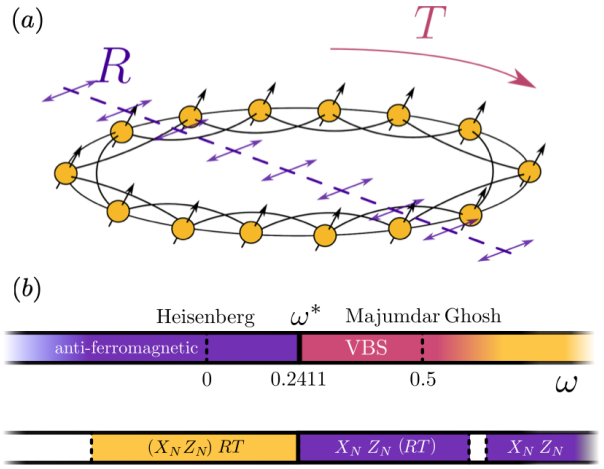


Figure 2: (a) Illustration of the $J_1 - J_2$ spin chain, with reflection R and translation T symmetries. (b) Ground state phase diagram of the model with $\omega \equiv J_1/J_2$, and low temperature metastability diagram indexed by the protecting symmetries for $N = 14$. The locations of the metastability diagram depend on the chain length N.

[2] $|L,R\rangle$ is shorthand notation to mean $|L\rangle$ or $|R\rangle$.

Away from the Majumdar-Ghosh point, the model can only be explored numerically [3]. In Fig. 1b, we plot the low lying spectrum of H for $\omega \in [-1, 2]$ and $N = 14$. We color code each excited state according to whether it has the same symmetry as the ground state (blue), it has a different symmetry (yellow), or the state is not an eigenstate of the symmetry (pink). We plot this for the X_N, Z_N symmetries and for the RT symmetry. What we observe in the figure is that for $\omega^* < \omega \lesssim 1.2$, the two lowest states are protected by the X_N, Z_N symmetry. For $\omega \gtrsim 1.5$, the model's first excited states are degenerate, but satisfy the same symmetry conditions as the ground state, hence yielding a qutrit blind subspace. Below $\omega_L \sim -.8$, there is no blind subspace, as all of the symmetry conditions are violated.

For $-.8 \lesssim \omega < \omega^*$, the excited states violate the X_N, Z_N symmetry conditions, suggesting no ϵ-DFS. However, a closer inspection shows that a new subspace emerges, protected by both the X_N, Z_N and the R, T symmetries. Note first that, in general, translations and reflections do not commute, $[R, T] \neq 0$. This means that even though they each commute with the Hamiltonian, we are not guaranteed the existence of a full set of states that are simultaneous eigenstates of all three operators $\{H, T, R\}$. For simplicity, we define our states as eigenstates of H and T. This has the benefit that in the cases where a state is also symmetric with respect to a reflection, translation symmetry guarantees that it will respect *all* of the possible reflections of the ring, with identical eigenvalue. Now, observe that for any given single site Pauli P_n, we can choose a reflection R_n such that $R_n P_n R_n = P_{n-1}$. Then, the TR_n transformation yields $T^\dagger R_n P_n R_n T = P_n$. The same argument holds for any single-site Pauli operator. Assuming that low-lying states $\{|e_j\rangle\}$ are eigenstates of both T and reflections, and denoting the eigenvalues as $T|e_j\rangle = \lambda_j^T|e_j\rangle$ and $R_n|e_j\rangle = \lambda_j^{R_n}|e_j\rangle$, we get

$$\langle e_j|P_n|e_k\rangle = \lambda_j^{R_n}\lambda_k^{R_n}\bar{\lambda}_j^T\lambda_k^T\langle e_j|P_n|e_k\rangle. \quad (19)$$

Hence, $\langle e_j|P_n|e_k\rangle$ must be zero, unless $\lambda_j^{R_n}\lambda_k^{R_n}\bar{\lambda}_j^T\lambda_k^T = 1$. Recall, because of translation invariance, λ^{R_n} is independent of the site n.

[3] Some analytic expressions exist for excited states [32, 36], but there is no guarantee that these states are the first few excited states

Consider now the table of symmetries at the value $\omega = 0$ (i.e. the Heisenberg model) for the five lowest eigenvectors:

$\omega = 0$	λ^{X_N}	λ^{Z_N}	λ^{R}	λ^{T}
$\|e_0\rangle$	-1	-1	-1	-1
$\|e_1\rangle$	1	1	1	1
$\|e_2\rangle$	-1	1	1	1
$\|e_3\rangle$	1	-1	1	1
$\|e_4\rangle$	-1	-1	-1	1

From the table of symmetries, which was obtained numerically, and the argument above, it follows that $\langle e_j|P_n|e_4\rangle = 0$ for $j = 0, 1, 2, 3, 4$ at the Heisenberg point. Given that there is a gap to the fifth excited eigenstate, there therefore exists a blind qubit subspace spanned by $|e_0\rangle$ and $|e_4\rangle$. The blind subspace is protected by all of the unitary symmetries of the system, but predominantly by the R and T symmetries. We expect the state $|e_4\rangle$ to be related to the anti-ferromagnetic ground state via the Bethe ansatz with total magnetization zero, as a sum of string operators, as argued in Ref. [37, 38]. The metastable subspace is also found in the Heisenberg XXZ model throughout the paramagnetic region. This is relevant, as physical implementation of next-nearest-neighbor models are much more challenging on present-day hardware than nearest-neighbor models.

Topological stabilizer Hamiltonians. Another class of models which very naturally exhibits the conditions for an ϵ-DFS are topological stabilizer models. Indeed, consider the 1D Wen model [39] $H_{\text{wen}} = \sum_{n=1}^{N} h_n$, with $h_n = X_n Z_{n+1} Z_{n+2} X_{n+3}$, and again assume periodic boundary conditions. There are two stabilizer/ground states of the model, both satisfying $h_n|\psi\rangle = 0$ for all n. Then, we note that for any local Pauli operator P_m, we can find a stabilizer h_n such that $h_n P_m h_n = -P_m$. This implies that $\langle\psi|P_m|\psi'\rangle = 0$, for any $|\psi\rangle, |\psi'\rangle$ in the stabilizer subspace. In fact this even holds of every string of Pauli operators that are not stabilizer-equivalent to the X_N and Z_N operators [4]; i.e. the logical operators of the code when N is odd. In this case, the protection is due to local symmetries of the stabilizer group.

[4] A and B are stabilizer equivalent if there exists a product of stabilizers $\Omega = \prod_\alpha h_{j_\alpha}$ such that $A = \Omega B$.

2.1 Robustness

The type of robustness afforded by the ϵ-DFSs depends crucially on the symmetries protecting the subspace. Going over the models covered here, the VBS-type phase, protected by global charge X_N and global phase Z_N, will be stable to any weak perturbation of the Hamiltonian preserving the symmetries. This includes any term of the form XX, ZZ on any two sites, and arbitrary combinations thereof. However, any 'magnetic field' (i.e. single or odd Pauli-weight terms) will in general break the metastability. One exception is translationally invariant combinations of such perturbations, since the RT symmetries ensure metastability in this phase even when the X_N and Z_N symmetries are broken. The anti-ferromagnetic phase is partly protected by X_N, and Z_N, but only fully protected by RT. Hence, only transitionally invariant perturbations are allowed, which is a much stronger restriction. For more details on the relationship between symmetries and metastability in these phases, see App. C.

The topological models are in turn only stable to multiplicative prefactors in front of the stabilizer terms, as well as perturbations that are in the stabilizer group. It is important to note that even for the toric code, whose ground subspace is topologically protected, non-stabilizer perturbations will kill the ϵ-DFS, leaving only classical information as metastable states of the thermal dynamics.

Stable qubits It is tempting to ask whether the metastable subspaces can be used to protect and manipulate logical information. At a first glance, the metastability described in this article looks pretty fragile. However, certain features lend themselves to optimism. First, because the effect described here is not very sensitively dependent on the precise form of the thermal dynamics, only on the system bath couplings, it is reasonable to expect that several of the recently proposed cooling protocols [40, 41, 42] might reveal the existence of low-temperature metastable states. Furthermore, in these synthetic situations, it might be possible to enforce the symmetries robustly. The error correction properties of the VBS phase have been described extensively

in Ref. [43, 44, 45], though explicit constructions are only known for the specific Majumdar-Ghosh point. In particular, the logical operators are only known in the large-N limit. It is unclear at this point how to engineer logical operations at other points in the phase diagram.

3 Discussion

The presented results show that there can be metastable states in the low-temperature thermal dynamics of non-commuting spin chain models. Importantly, the existence of the metastable subspaces does not require ground state degeneracy. Conversely, ground state degeneracy is not a sufficient condition either. For instance, the transverse field Ising model in the ferromagnetic phase has exponentially small splitting between the ground state and first excited state but does not possess a metastable subspace, since the ground state expectation value of local Pauli operators is non-zero. As a consequence, there is dephasing within the ground subspace, leading to a two-dimensional stationary space of the Lindbladian; in other words, a stable classical bit.

One important question which we leave open is whether some weaker forms of protection can be seen in these systems as the chain length grows. So far, our results do not really depend upon chain length within a given phase. Rather, the effects discovered here are most pronounced for small system sizes ($N < 12$). In the appendix, we provide some additional analysis of the behavior of metastable subspaces as the chain length increases. Finally, we ask whether similar effects can be seen in systems with additional structure; including, but not limited to, fermionic systems, lattice gauge theories, and subsystem codes.

Acknowledgments

We acknowledge support from the AWS Center for Quantum Computing, and from the Carlsberg foundation.

References

[1] Benjamin J Brown, Daniel Loss, Jiannis K Pachos, Chris N Self, and James R Wootton. "Quantum memories at finite temperature". Reviews of Modern Physics **88**, 045005 (2016).

[2] A Yu Kitaev. "Fault-tolerant quantum computation by anyons". Annals of physics **303**, 2–30 (2003).

[3] Héctor Bombín. "Gauge color codes: optimal transversal gates and gauge fixing in topological stabilizer codes". New Journal of Physics **17**, 083002 (2015).

[4] Dave Bacon. "Operator quantum error-correcting subsystems for self-correcting quantum memories". Physical Review A **73**, 012340 (2006).

[5] Jeongwan Haah. "Local stabilizer codes in three dimensions without string logical operators". Physical Review A **83**, 042330 (2011).

[6] Michael J Kastoryano and Fernando GSL Brandao. "Quantum Gibbs samplers: The commuting case". Communications in Mathematical Physics **344**, 915–957 (2016).

[7] Héctor Bombín. "Single-shot fault-tolerant quantum error correction". Physical Review X **5**, 031043 (2015).

[8] Frederik Nathan, Liam O'Brien, Kyungjoo Noh, Matthew H Matheny, Arne L Grimsmo, Liang Jiang, and Gil Refael. "Self-correcting GKP qubit and gates in a driven-dissipative circuit" (2024). arXiv:2405.05671.

[9] Robert Alicki, Mark Fannes, and Michal Horodecki. "On thermalization in Kitaev's 2d model". Journal of Physics A: Mathematical and Theoretical **42**, 065303 (2009).

[10] Robert Alicki, Michal Horodecki, Pawel Horodecki, and Ryszard Horodecki. "On thermal stability of topological qubit in Kitaev's 4d model". Open Systems & Information Dynamics **17**, 1–20 (2010).

[11] Ainesh Bakshi, Allen Liu, Ankur Moitra, and Ewin Tang. "High-temperature Gibbs states are unentangled and efficiently preparable". In 2024 IEEE 65th Annual Symposium on Foundations of Computer Science (FOCS). Pages 1027–1036. IEEE (2024).

[12] Cambyse Rouzé, Daniel Stilck França, and Álvaro M Alhambra. "Efficient thermalization and universal quantum computing with quantum Gibbs samplers" (2024). arXiv:2403.12691.

[13] Fabio Martinelli. "Lectures on Glauber dynamics for discrete spin models". In Lectures on probability theory and statistics. Pages 93–191. Springer (1999).

[14] Ángela Capel, Cambyse Rouzé, and Daniel Stilck França. "The modified logarithmic Sobolev inequality for quantum spin systems: classical and commuting nearest neighbour interactions" (2021) arXiv:2009.11817.

[15] Chi-Fang Chen, Michael J Kastoryano, Fernando GSL Brandão, and András Gilyén. "Quantum thermal state preparation" (2023). arXiv:2303.18224.

[16] Chi-Fang Chen, Michael J Kastoryano, and András Gilyén. "An efficient and exact noncommutative quantum Gibbs sampler" (2023). arXiv:2311.09207.

[17] András Gilyén, Chi-Fang Chen, Joao F Doriguello, and Michael J Kastoryano. "Quantum generalizations of Glauber and Metropolis dynamics" (2024). arXiv:2405.20322.

[18] Zhiyan Ding, Bowen Li, and Lin Lin. "Efficient quantum Gibbs samplers with Kubo–Martin–Schwinger detailed balance condition". Communications in Mathematical Physics **406**, 67 (2025).

[19] Subir Sachdev. "Quantum phase transitions". Physics world **12**, 33 (1999).

[20] Bei Zeng, Xie Chen, Duan-Lu Zhou, Xiao-Gang Wen, et al. "Quantum information meets quantum matter". Springer. (2019).

[21] Ivan Bardet, Ángela Capel, Li Gao, Angelo Lucia, David Pérez-García, and Cambyse Rouzé. "Rapid thermalization of spin chain commuting Hamiltonians". Physical Review Letters **130**, 060401 (2023).

[22] Crispin Gardiner and Peter Zoller. "Quantum noise: a handbook of Markovian and non-Markovian quantum stochastic methods with applications to quantum optics". Springer Science & Business Media. (2004).

[23] Frederik Nathan and Mark S Rudner. "Universal Lindblad equation for open quantum systems". Physical Review B **102**, 115109 (2020).

[24] Noam Berger, Claire Kenyon, Elchanan Mossel, and Yuval Peres. "Glauber dynamics on trees and hyperbolic graphs". Probability Theory and Related Fields **131**, 311–340 (2005).

[25] Emanuel Knill, Raymond Laflamme, and Lorenza Viola. "Theory of quantum error correction for general noise". Physical Review Letters **84**, 2525 (2000).

[26] Paolo Zanardi. "Stabilizing quantum information". Physical Review A **63**, 012301 (2000).

[27] Barbara Kraus, Hans P Büchler, Sebastian Diehl, Adrian Kantian, Andrea Micheli, and Peter Zoller. "Preparation of entangled states by quantum Markov processes". Physical Review A **78**, 042307 (2008).

[28] Daniel A Lidar and K Birgitta Whaley. "Decoherence-free subspaces and subsystems". In Irreversible quantum dynamics. Pages 83–120. Springer (2003).

[29] Stefano Chesi, Daniel Loss, Sergey Bravyi, and Barbara M Terhal. "Thermodynamic stability criteria for a quantum memory based on stabilizer and subsystem codes". New Journal of Physics **12**, 025013 (2010).

[30] J Robert Johansson, Paul D Nation, and Franco Nori. "QuTiP: An open-source Python framework for the dynamics of open quantum systems". Computer physics communications **183**, 1760–1772 (2012).

[31] Kiyomi Okamoto and Kiyohide Nomura. "Fluid-dimer critical point in S = 12 antiferromagnetic Heisenberg chain with next nearest neighbor interactions". Physics Letters A **169**, 433–437 (1992).

[32] WJ Caspers and W Magnus. "Some exact excited states in a linear antiferromagnetic spin system". Physics letters A **88**, 103–105 (1982).

[33] M. Hamermesh. "Group theory and its application to physical problems". Dover Books on Physics. Dover Publications. (2012).

[34] B.H. Bransden and C.J. Joachain. "Physics of atoms and molecules". Pearson Education. Prentice Hall. (2003). url: https://books.google.dk/books?id=i5IPWXDQlcIC.

[35] Jorge A Campos-Gonzalez-Angulo, Raphael F Ribeiro, and Joel Yuen-Zhou. "Generalization of the Tavis–Cummings model for multi-level anharmonic systems". New Journal of Physics **23**, 063081 (2021).

[36] Daniel P Arovas. "Two exact excited states

for the S = 1 AKLT chain". Physics Letters A **137**, 431–433 (1989).

[37] Lyudvig Dmitrievich Faddeev and Leon Armenovich Takhtadzhyan. "Spectrum and scattering of excitations in the one-dimensional isotropic Heisenberg model". Journal of Soviet Mathematics **24**, 241–267 (1984).

[38] Anastasia Doikou and Rafael I Nepomechie. "Parity and charge conjugation symmetries and S matrix of the XXZ chain". Statistical Physics on the Eve of the Twenty-First Century, M. Batchelor and L. Wille, edsPages 391–411 (1999). arXiv:hep-th/9810034.

[39] Xiao-Gang Wen. "Quantum orders in an exact soluble model". Physical review letters **90**, 016803 (2003).

[40] Anne Matthies, Mark Rudner, Achim Rosch, and Erez Berg. "Programmable adiabatic demagnetization for systems with trivial and topological excitations". Quantum **8**, 1505 (2024).

[41] Gilad Kishony, Mark S Rudner, Achim Rosch, and Erez Berg. "Gauged cooling of topological excitations and emergent fermions on quantum simulators". Physical Review Letters **134**, 086503 (2025).

[42] Trond I Andersen, Nikita Astrakhantsev, Amir H Karamlou, Julia Berndtsson, Johannes Motruk, Aaron Szasz, Jonathan A Gross, Alexander Schuckert, Tom Westerhout, Yaxing Zhang, et al. "Thermalization and criticality on an analogue–digital quantum simulator". Nature **638**, 79–85 (2025).

[43] Jerome Lloyd, Alexios A Michailidis, Xiao Mi, Vadim Smelyanskiy, and Dmitry A Abanin. "Quasiparticle cooling algorithms for quantum many-body state preparation". PRX Quantum **6**, 010361 (2025).

[44] Masaaki Nakamura and Synge Todo. "Order parameter to characterize valence-bond-solid states in quantum spin chains". Physical review letters **89**, 077204 (2002).

[45] Dong-Sheng Wang. "Classes of topological qubits from low-dimensional quantum spin systems". Annals of Physics **412**, 168015 (2020).

A Contribution from Q

We derive the bound on $||\kappa|\psi\rangle||$. To start with, we break up the norm as

$$||\kappa|\psi\rangle|| \leq \sum_a ||Q_a|\psi\rangle|| + \frac{1}{2}||L_a^\dagger L_a|\psi\rangle||, \tag{20}$$

where we have defined

$$Q_a = \frac{i}{2}\sum_{ij} \tanh\left(\frac{\beta(e_i - e_j)}{4}\right) \langle e_i|L_a^\dagger L_a|e_j\rangle |e_i\rangle\langle e_j|. \tag{21}$$

Now lets consider the blind qubit subspace span$\{|e_0\rangle, |e_l\rangle\}$, for some eigenstate $|e_l\rangle$ of H; i.e. $\langle e_l|S_a|e_j\rangle = 0$ for all Paulis S_a and all $j \leq l$. We showed in the main text,

$$||L_a|e_l\rangle|| \leq O(e^{-\beta\Delta/2}). \tag{22}$$

The second term in Eqn. (20) can therefore be bounded simply as

$$\frac{1}{2}||L_a^\dagger L_a|e_l\rangle|| \leq \frac{1}{2}||L_a^\dagger||\ ||L_a|e_l\rangle|| \leq O(e^{-\beta\Delta/2}), \tag{23}$$

since $||L_a^\dagger|| = O(1)$ [18]. The first term can similarly be bounded by noting that

$$|\langle e_j|L_a^\dagger L_a|e_l\rangle| \leq O(e^{-\beta\Delta/2}), \tag{24}$$

for the same reasons as above, and thus

$$||Q_a|e_l\rangle|| \leq \frac{1}{2}\sum_j \tanh\left(\frac{\beta(e_j - e_l)}{4}\right) |\langle e_j|L_a^\dagger L_a|e_l\rangle| \tag{25}$$

$$\leq \frac{1}{2}\sum_j |\langle e_j|L_a^\dagger L_a|e_l\rangle| \leq O(e^{-\beta\Delta/2}). \tag{26}$$

By summing up all Pauli operators, we get a scaling $||\kappa|e_l\rangle|| = O(Ne^{-\beta\Delta/2})$.

B The large system limit

In the main text, we showed that if a system has a blind subspace, then this subspace will be metastable for a time scaling as $\Omega(N^{-1}e^{\beta\Delta/2})$. Our arguments relied on numerical evidence on small systems ($N < 16$). Below, we explore to what extend some of our conclusions might break down for larger system sizes.

B.1 System size dependence of Δ.

Leakage out of the blind subspace is bounded by the quantity $e^{-\beta\Delta/2}$, where $\Delta = \min\{e_{j+1}-e_j, e_p-e_0\}$ is the smallest energy gap from a state in the blind subspace to a higher-energy state outside of it. To ensure slow leakage out of the subspace, we must therefore require $\beta\Delta \gg 1$. Thus, the size of the gap is important for practical applications, since Δ^{-1} sets a lower bound for the required inverse temperature β, as well as determines the strength of the suppression at low temperatures. To investigate how this quantity evolves as the system size grows, Lanczos-based numerical diagonalization was used to determine the low-energy spectrum and calculate the gap for system-sizes up to $N = 20$. In Fig. 3, the

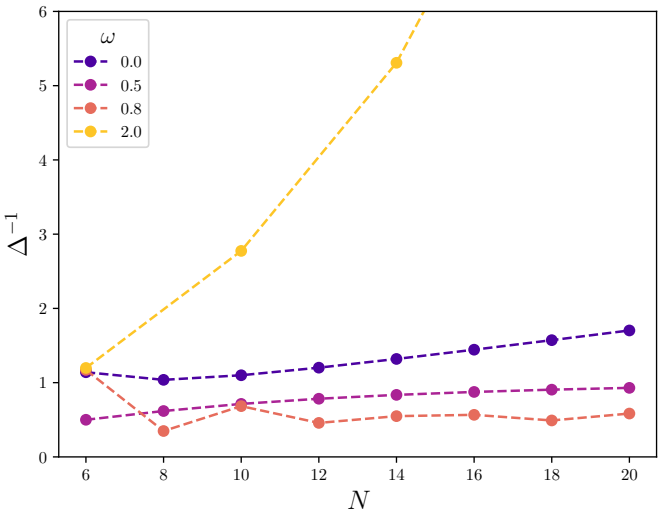


Figure 3: Magnitude of the gap above the blind subspace (Δ) as a function of the system size N for different values of ω. Dotted lines have been added between data points to guide the eyes.

inverse gap above the blind subspace is plotted for representative points in the qubit VBS subspace ($\omega = 0.8$), in the 'Heisenberg' subspace ($\omega = 0$), and in the qutrit VBS subspace ($\omega = 2$). The gap for the Majumdar-Gosh point ($\omega = 0.5$) is also plotted. As can be seen from this figure, the gaps for values of ω in the 'Heisenberg' and qubit VBS regions are relatively constant, with some indications of slowly decreasing gaps as system size is scaled up. However, in the qutrit-region $\omega \geq 1.5$, the gap closes rapidly, corresponding to the disappearance of the protected subspace. Note that the qutrit blind subspace is only observed for $N \equiv 2 \,(\mathrm{mod}\, 4)$, meaning the observed vanishing gap is consistent with an expectation that $N \equiv 2 \,(\mathrm{mod}\, 4)$ and $N \equiv 0 \,(\mathrm{mod}\, 4)$ should behave equivalently in the thermodynamic limit.

B.2 Contributions from the bulk of the spectrum

In the main text, it was conjectured that for low enough temperatures, the following holds as the system size is increased:

$$\mathcal{K}_a(N,\beta,\omega) := \sum_{k>j+1} e^{-\beta(e_k-e_{j+1})/2} |\langle e_k|S_a|e_j\rangle| \sim O(1)\ , \tag{27}$$

where $\{e_k\}$ are the eigenenergies of the system numbered in increasing order, and j is the index of the high-energy state of the blind qubit subspace. Here, we provide numerical evidence to support this claim, and explore what constitutes sufficiently low temperatures.

First, note that

$$\mathcal{K}_a(N,\beta,\omega) \leq \mathcal{C}(N,\beta,\omega) := \sum_{k>j+1} e^{-\beta(e_k - e_{j+1})/2} \,. \tag{28}$$

If the degeneracy for the energy level that the $(j+1)$th state belongs to is $D(N,\omega)$, the sum takes the form

$$\mathcal{C}(N,\beta,\omega) = D(N,\omega) + \sum_{k>(j+D)} e^{-\beta(e_k - e_{j+1})/2} \tag{29}$$

where now the energies in the exponent of the sum are all strictly positive by definition. Thus, in the limit of $\beta \to \infty$, $\mathcal{C}(N,\beta,\omega)$ and $D(N,\omega)$ coincide. Arguing that $\mathcal{C} \sim O(1)$ therefore boils down to arguing that a) $D(N,\omega) \sim O(1)$, i.e., that the degeneracies of the low-energy spectrum does not grow with system size, and b) the temperature scale where $\mathcal{C} \simeq D$ remains constant, or at the very least only shrinks at a relatively slow rate. To get an intuition for both of these questions, the evolution of $\mathcal{C}$ towards D as β is increased is depicted in Fig. 4 for $\omega = 0$, i.e., in the Heisenberg phase. As can be

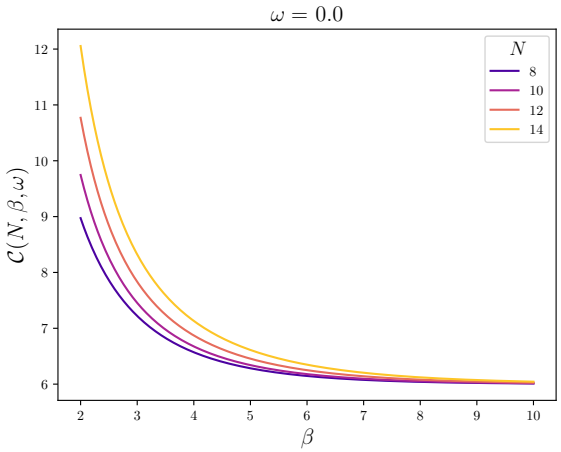


Figure 4: Evaluation of $\mathcal{C}(N,\beta,\omega)$ as a function of β for $\omega = 0$ and various system sizes.

seen from this figure, the limit of $\mathcal{C}$ as β is increased is unchanged as N is increased, indicating that the degeneracy D does not depend on the system size for even-size systems. This was additionally numerically verified up to a system size of $N = 22$ using the Lanczos algorithm for finding the low-energy spectrum. A similar independence of D to system size was observed across a range of ω in the relevant regimes, including in the qutrit-regime $\omega \gtrsim 1.5$ for systems of the size where the protected qutrit subspace is present, i.e., systems where $N \equiv 2 \,(\mathrm{mod}\, 4)$.

The results above provide support to the first of the two claims, namely that the degeneracies of the low-energy spectrum is independent of system size, meaning the low-temperature limit of $\mathcal{C}(N,\beta,\omega)$ is also independent of system size. However, as can be seen in Fig. 4, the rate at which the asymptotic value is reached does depend on the system size, with larger systems typically displaying a slower decay. One way to quantify this slowdown is to investigate how large a value of β is required to achieve a certain value of $\mathcal{C}$,

$$\beta^*_{\mathcal{C}}(N,\omega,K) = \inf\{\beta : \mathcal{C}(N,\beta,\omega) \leq K\} \,. \tag{30}$$

However, numerics indicate that the bound in Eqn. (28) is loose in some regimes. Specifically, picking values of ω that lie deep in the two meta-stable phases, while avoiding the special points ($\omega = 0$ and

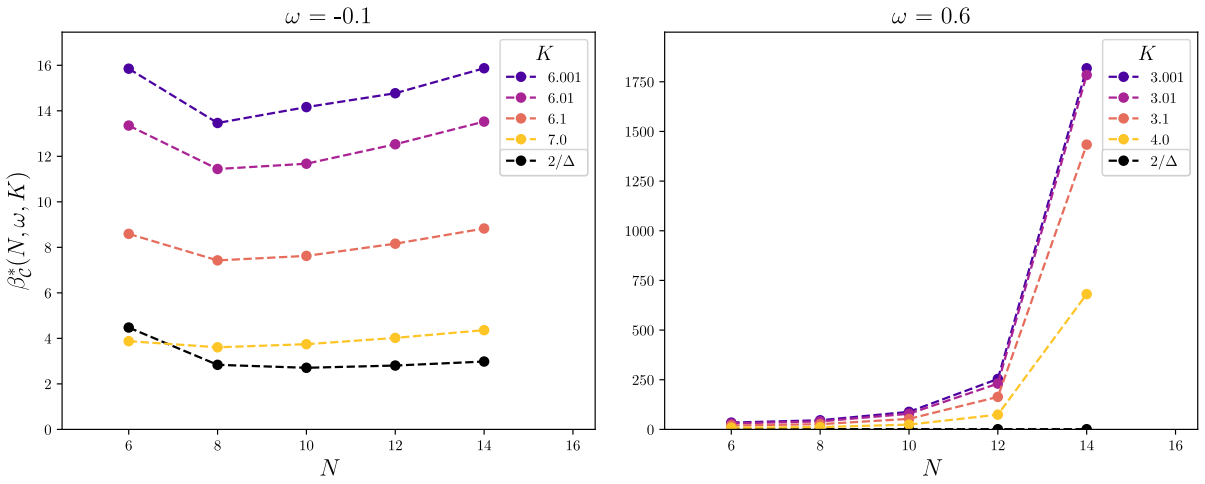


Figure 5: Numerical calculations of $\beta^*_{\mathcal{C}}(N,\omega,K)$, the value of β required to achieve a given value K of the upper bound $\mathcal{C}$ from Eqn. (28). The corresponding values for the inverse gap have been added for reference, and dotted lines added to guide the eyes.

$\omega = 0.5$), yields the results in Fig. 5 for $\beta^*_{\mathcal{C}}$. This highlights that achieving $\mathcal{C} \simeq D$ requires rapidly increasing β with system size in some regimes of ω. On the other hand, defining a similar quantity for the full sum to be bounded,

$$\beta^*_{\mathcal{K}}(N,\omega,K) = \inf\{\beta : \mathcal{K}(N,\beta,\omega) \leq K\} \tag{31}$$

$$\mathcal{K}(N,\beta,\omega) = \sum_{S_a \in \{X_0,Y_0,Z_0\}} \mathcal{K}_a(N,\beta,\omega)\ , \tag{32}$$

yields the results in Fig. 6, showing a slower increase as the system size is increased. Note that for

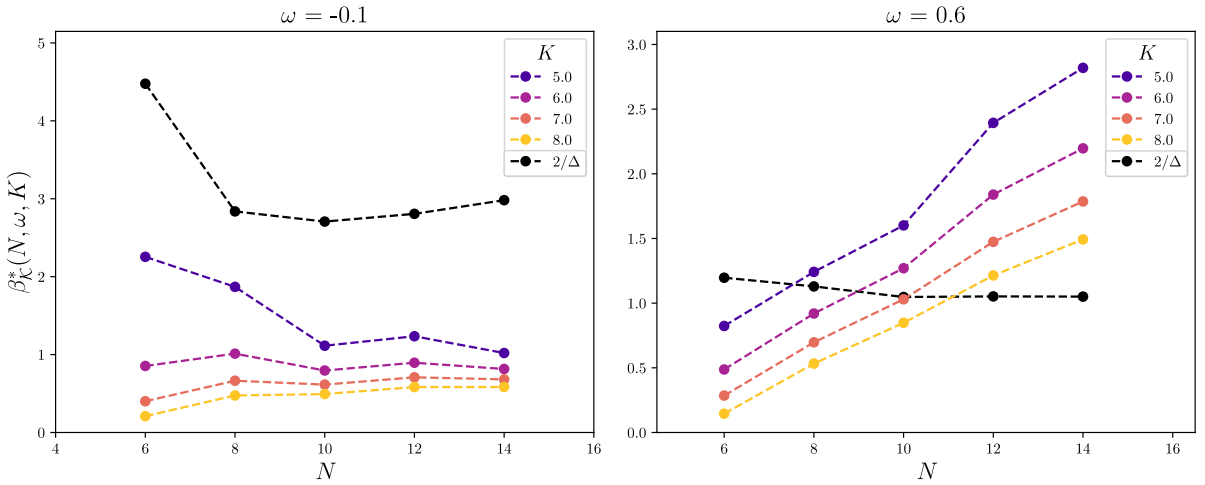


Figure 6: Numerical calculations of $\beta^*_{\mathcal{K}}(N,\omega,K)$, the value of β required to achieve a given value K of the quantity $\mathcal{K}$ in Eqn. (27). The corresponding values for the inverse gap have been added for reference, and dotted lines added to guide the eyes.

simplicity, the effect of X, Y and Z-type operators has been combined, and translational invariance implicitly used to allow a focus on operators on the first qubit. The exact way that $\beta^*_{\mathcal{K}}$ evolves as N grows depends on ω and K, but generally follows similar patterns of approximately linear growth with N for the investigated system sizes. This supports the claim that $\mathcal{K}$ can be kept constant as the

system size is increased using at most a linear increase in β, also beyond the regimes where it follows from a bound on the simpler quantity $\mathcal{C}$.

Beyond supporting the scaling behaviour, similar numerics can be used to investigate the full value of the sum in Eqn. (10). Specifically, we can define the inverse temperature needed to reduce the combined operator-norm below a certain threshold,

$$\beta^*(N, \omega, K) = \inf\{\beta : 2\pi N e^{-\beta\Delta/2}\mathcal{K}(N, \beta, \omega) \leq K\} \,, \tag{33}$$

which indicates more directly the temperature required to reach a certain level of protection of the

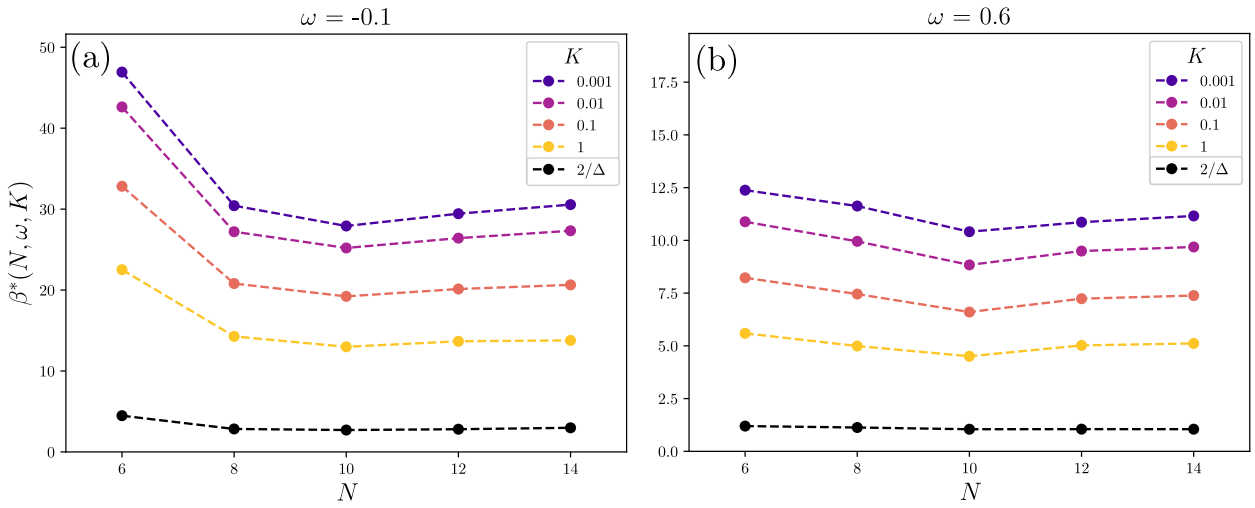


Figure 7: Numerical calculations of $\beta^*(N, \omega, K)$, the value of β required to achieve a given suppression K in Eqn. (10). The corresponding values for the inverse gap have been added for reference, and dotted lines added to guide the eyes.

blind subspace. This quantity is depicted on Fig. 7 for $\omega = -0.1$ and $\omega = 0.6$. While it is difficult to make any strong claims based on the figure, the value of β necessary to reach a given protection does appear to increase slowly beyond $N \approx 10$. A similar behavior is seen for other choices of ω. The exact way that β^* evolves as N grows depends on ω and K, but often follows a pattern where stricter (lower) values of K have a positive curvature, while more lenient (higher) values of K have a negative curvature. One exception is near $\omega \simeq 0.8$, where oscillatory finite-size effects are observed, similarly to the behaviour of the gaps in Fig. 3.

In conclusion, numerics suggest that if $\beta \geq \beta^*$ for a β^* that increases sub-linearly with system size, the scaling $\mathcal{K} = O(1)$ is achieved, and all of the decoherence-inducing contributions are therefore bounded by $O(e^{-\beta\Delta/2})$. Since the inverse gap also scales approximately linearly (see Fig. 3), both of these effects enforce an approximately linearly increasing β with system size.

B.3 Phase transition points as the system scales up

A compelling feature of the results presented here is that the transitions into and out of metastable subspaces do not appear to coincide with the ground state phase diagram for finite N. To test whether this feature survives as the chain length grows, we plot the (approximate) locations of the transition points in Fig. 8. The numerics suggest that the locations of transitions to the qubit and qutrit metastable subspaces grow linearly with N, strongly suggesting that the qubit metastable subspace will cover the whole phase in the thermodynamic limit. The 'Heisenberg' subspace is also moving left as the system size increases, but it is difficult to know whether it will continue expanding as the system size grows further.

If the two phases extend out to infinity, this would suggest that the metastability 'phase diagram' coincides with the ground state phase diagram for the J_-J_2 model. In parallel experiments, that are

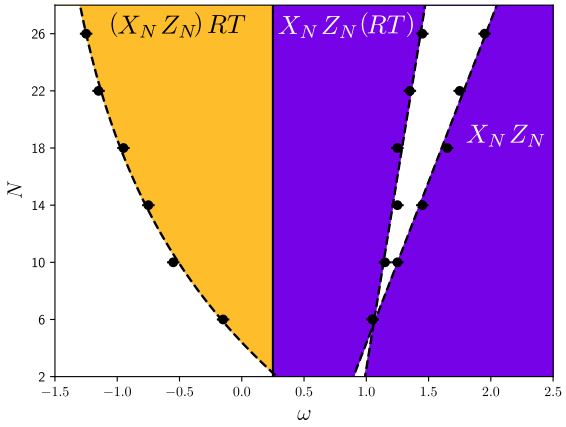


Figure 8: Location of the transitions between metastable phases. From left to right, the second region hosts a protected qubit, while the fourth hosts a protected qutrit. Fits are either linear or exponential in the system size, depending on the observed behaviour. For the exponential fit, a limit of $\omega_L^* \simeq -1.647$ is predicted, with the characteristic scale of $k^{-1} \simeq 15$ in the exponential decay $\omega(N) = \omega_L^* + (\omega_L^{(0)} - \omega_L^*)e^{-kN}$.

not reported here, we found that this is clearly not the case for the XXZ Heisenberg model, which undergoes a metastable transition at zero ZZ field strength, which does not coincide with the ground state phase transition of the model.

Finally, in view of the observation above, that the critical temperature β^* above which there is exponential survival of the blind subspace grows with system size, the effects that we report are most pronounced for finite chains.

C Sensitivity to symmetry-breaking perturbations

As outlined in the main text, the blind subspaces are formed due to protection caused by the symmetries of the system. Specifically, the symmetries R and T protect the fourth excited state of the $\omega < \omega^*$ Heisenberg-like phase, while the first excited state of the $\omega^* < \omega \lesssim 1$ VBS-like phase is independently and redundantly protected by both the symmetries Z_N and X_N and the symmetries R and T (See, e.g., Fig. 2 of the main text).

As mentioned in the main text, this protection is expected to break down if perturbations breaking the symmetries are added to the system Hamiltonian. Below, we will elucidate this effect for a small selection of local perturbations, and provide numerical results on the performance of the protection when the system is subjected to disordered local fields. The central figure of merit will be the overlaps induced by the symmetry breaking:

$$\tau_n^A = \sum_{i<j} |\langle e_i| P_n |e_j\rangle|^2 \qquad P \in \{X, Y, Z\},\ n \in \{1, 2 \ldots N\}\ . \tag{34}$$

In the case of no perturbations, the symmetries enforce these matrix elements to be zero. On the other hand, non-zero values implies non-zero contributions to the sum in Eq. (10) in the main text, and thus a breakdown of metastability. The results below are based on analyzing these quantities numerically for system sizes of $N = 8$ to $N = 14$ at $\omega = -0.1$ (Heinsenberg-like phase) and $\omega = 0.6$ (VBS-like phase), though the conclusions are not expected to depend sensitively on these parameters.

C.1 Global homogeneous field

Consider first a homogeneous local field applied identically to every site. Due to the symmetry of the un-perturbed model, we assume without loss of generality that the field points in the Z direction, i.e.,

we consider the perturbation

$$H_p(x) = x \sum_{n=1}^{N} Z_n \ .$$

This perturbation breaks the X_N symmetry, but leaves the rest of the symmetries intact. This means that the R and T-based protection of both the Heisenberg phase and VBS phase remains intact, with the relevant blind subspaces remaining blind and fully protected. The breakdown of the protection thus only occurs once the perturbation becomes sufficiently strong to induce (avoided) crossings in the spectrum of the Hamiltonian, which for $N = 12$ occurs at $x \sim 0.4$ for the Heisenberg phase and $x \sim 0.9$ for the VBS phase.

C.2 Local single-site perturbation

Consider now a perturbation of the form:

$$H_p(x) = x\, Z_{N/2} \ . \tag{35}$$

This perturbation breaks both the X_N symmetry, the translation symmetry T, and most of the reflection symmetries. However, it does preserve symmetry under reflection about the axis through qubits 0 and $N/2$. Denote this symmetry by R'_0. Using arguments similar to those of the main text, it is then straightforward to show that the matrix elements $\langle e_i|\, P_0\, |e_j\rangle$ and $\langle e_i|\, P_{N/2}\, |e_j\rangle$ related to single-qubit Paulis at the two qubits preserved by the reflection R'_0 can only be non-zero if $\lambda_i^{R'_0} = \lambda_j^{R'_0}$. Thus, R'_0 can provide blindness with respect to Paulis at those two qubits whenever $\lambda_i^{R'_0} \neq \lambda_j^{R'_0}$. In fact, looking at the $N = 12$ symmetries of the low-energy states, the states $|e_4\rangle$ (Heisenberg) and $|e_0\rangle$ (VBS) indeed do have the requisite properties for this protection (See Table 1 below). In addition, protection against some local Paulis is implied by the remaining Z_N symmetry by similar arguments. A depiction of the sensitivities (quantified by τ_n^A) as the strength of the perturbation is scaled up is shown on Fig. 9, showing clearly the special protection of qubits 0 and $N/2$ implied by the R'_0 symmetry. Note also that the preserved Z_N symmetry causes intact protection against X and Y-noise in the VBS phase, thus causing the observed yellow tint in the VBS-phase from the combination of preserved X (red) and Y (green) robustness.

$\omega = -0.1$	λ^{Z_N}	$\lambda^{R'_0}$	$\omega = 0.6$	λ^{Z_N}	$\lambda^{R'_0}$
$\|e_0\rangle$	1	1	$\|e_0\rangle$	1	1
$\|e_1\rangle$	−1	1	$\|e_1\rangle$	1	−1
$\|e_2\rangle$	1	1			
$\|e_3\rangle$	−1	1			
$\|e_4\rangle$	1	−1			

Table 1: Symmetry properties of the low-energy states of a system of size $N = 12$ for both the Heisenberg-like phase (left) and the VBS-like phase (right).

C.3 Two-site perturbation

Consider next a perturbation of the form:

$$H_p(x) = x\, Z_{N/2-1} Z_{N/2} \ . \tag{36}$$

This perturbation breaks only the T and R symmetries, while preserving Z_N and X_N. As a result, the blind subspace in the VBS-like phase is preserved under this perturbation. On the other hand, the Heisenberg-like phase relies on T and R for the protection of the blind subspace, and thus no longer

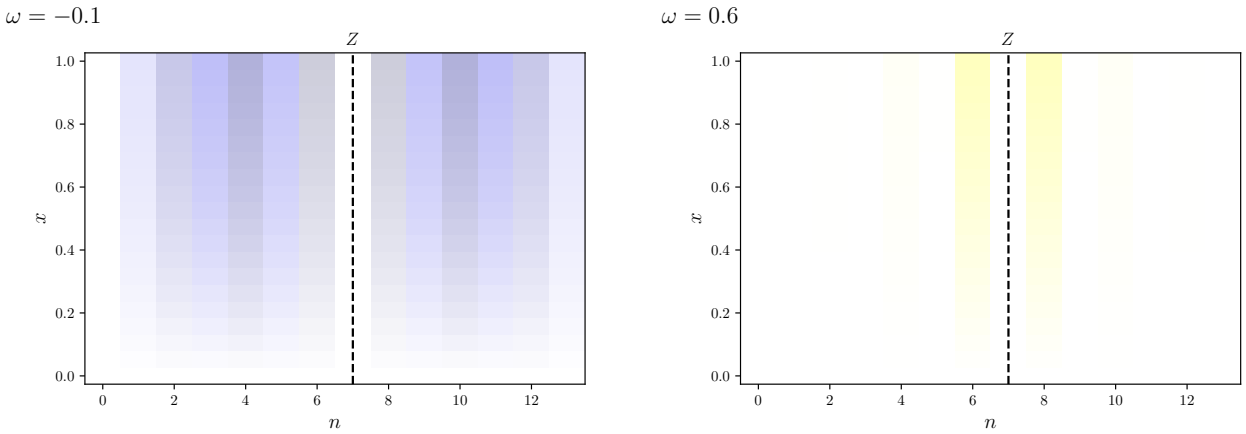


Figure 9: Sensitivities caused by a Hamiltonian perturbation $x\,Z_7$ on a chain of length $N = 14$ in both the Heisenberg-like phase (left) and VBS-like phase (right). The x-axis labels the 14 qubits, and the y-axis represents the strength of the perturbation. The color of the patch at location (n, x) is given by $(R, G, B)_{n,x} = \left(1 - \sqrt{\tau_n^X(x)}, 1 - \sqrt{\tau_n^Y(x)}, 1 - \sqrt{\tau_n^Z(x)}\right)$, i.e., each of the primary colors indicate robustness with respect to one of the three Paulis. Thus, white corresponds to full protection and vanishing sensitivities, $\tau_n^A = 0$, while darker colors indicate larger sensitivities.

exhibits protection. The induced sensitivities in the Heisenberg phase is depicted on Fig. 10. A few interesting observations from this figure is that the induced sensitivity is in general weaker than the one induced by the single-site perturbation, and that the induced sensitivity is predominantly localized around the site of the perturbation.

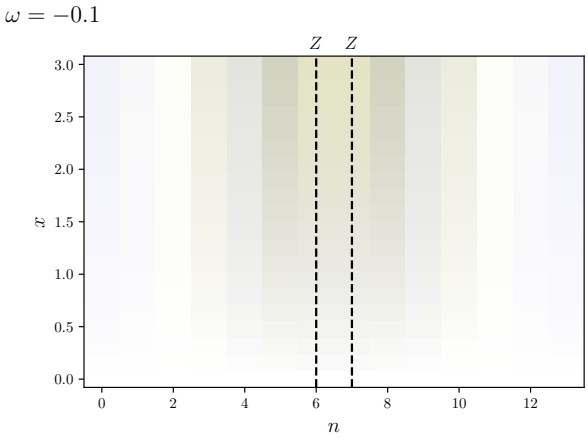


Figure 10: Sensitivities caused by a Hamiltonian perturbation $x\,Z_6\,Z_7$ on a chain of length $N = 14$ in the Heisenberg-like phase. The x-axis labels the 14 qubits, and the y-axis represents the strength of the perturbation. The color of the patch at location (n, x) is given by $(R, G, B)_{n,x} = \left(1 - \sqrt{\tau_n^X(x)}, 1 - \sqrt{\tau_n^Y(x)}, 1 - \sqrt{\tau_n^Z(x)}\right)$, i.e., each of the primary colors indicate robustness with respect to one of the three Paulis. Note the increased scale of the perturbation compared to Fig. 9.

C.4 Inhomogeneous field

The preceding examples provide evidence of the effects of specific perturbations breaking specific symmetries. A different but equally important question is the effect of less structured perturbations. Indeed, small random deviations of the implemented Hamiltonian from the fully symmetric theoretical model is likely to occur in any physical experiment investigating the observed effects. As a small preliminary investigation of the effect of such disordered perturbations on the observed protection, we

consider now perturbations of the form:

$$H_p(x) = x \sum_n s_n Z_n \ , \tag{37}$$

with each $s_n \in \mathbb{R}$ drawn independently from a normal distribution of mean zero and variance one. Tracking the evolution of the robustness proxy $1 - \sum_{n,A} \tau_n^A(x)$ as the overall magnitude x of this in-homogeneous field is scaled up provides some evidence of the degree of robustness of the protection for different phases and system sizes. Numerical results from 8 disorder realizations is shown on Fig. 11. The main qualitative observations is that the VBS-like phase exhibits larger general robustness of the blind subspace (note the different axes), and that the Heisenberg-like phase seems to exhibit less robustness as system size increases.

An additional qualitative observation worth making is the presence of oscillatory revivals of the robustness for the VBS-like phase. By the definition of the robustness proxy, this indicates an oscillation in the magnitude of the matrix elements between the two states $|e_0\rangle$ and $|e_1\rangle$ for some noise realizations. In fact, similar oscillations is also observed in the large-strength limit of the simple $x\, Z_{N/2}$ perturbation for the VBS-like phase. For the Heisenberg-like phase, the averaging over multiple overlaps between multiple pairs of states may be the reason why similar state-structure effects are not observed, including in Fig. 11. However, some evidence does exist for non-monotonic behavior also in the Heisenberg-like phase for $x\, Z_{N/2-1} Z_{N/2}$ and $x\, Z_{N/2-1} Z_{N/2} Z_{N/2+1}$ perturbations, particularly for perturbation strengths just below values at which the perturbation induces (avoided) crossings involving states of the blind subspace.

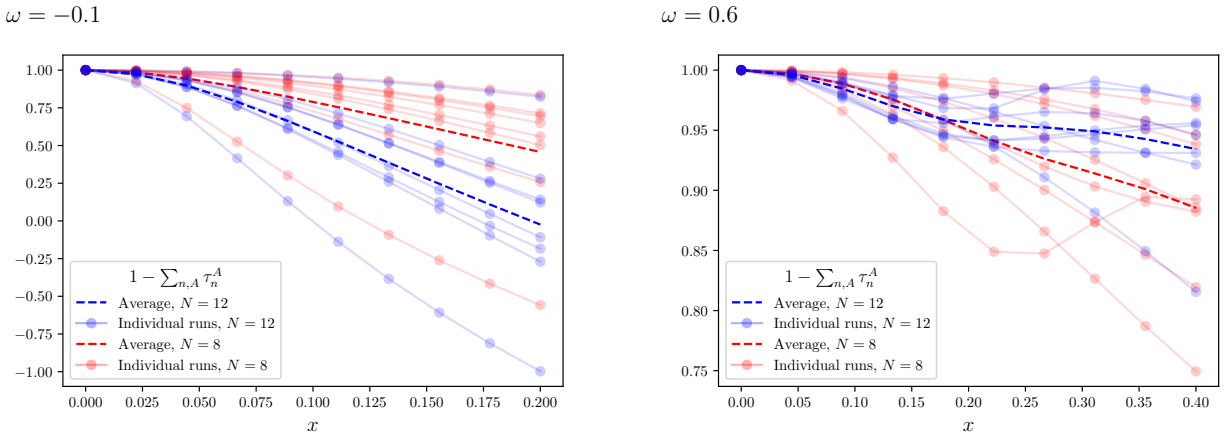


Figure 11: Robustness-proxy $1 - \sum_{n,A} \tau_n^A$ for different realizations of the noise disorder as the overall magnitude of the noise is increased. Identical noise realizations was used across both system sizes and in both phases to facilitate easier comparison. Note the different axes for the two plots.