Prethermal Strong Zero Modes and Topological Qubits

Dominic V. Else,[1] Paul Fendley,[2,3] Jack Kemp,[3] and Chetan Nayak[4,1]

[1]*Physics Department, University of California, Santa Barbara, California 93106, USA*
[2]*All Souls College, University of Oxford, Oxford OX1 4AL, United Kingdom*
[3]*Rudolf Peierls Centre for Theoretical Physics, 1 Keble Road, Oxford OX1 3NP, United Kingdom*
[4]*Station Q, Microsoft Research, Santa Barbara, California 93106, USA*

We prove that quantum information encoded in some topological excitations, including certain Majorana zero modes, is protected in closed systems for a time scale exponentially long in system parameters. This protection holds even at infinite temperature. At lower temperatures, the decay time becomes even longer, with a temperature dependence controlled by an effective gap that is parametrically larger than the actual energy gap of the system. This nonequilibrium dynamical phenomenon is a form of prethermalization and occurs because of obstructions to the equilibration of edge or defect degrees of freedom with the bulk. We analyze the ramifications for ordered and topological phases in one, two, and three dimensions, with examples including Majorana and parafermionic zero modes in interacting spin chains. Our results are based on a nonperturbative analysis valid in any dimension, and they are illustrated by numerical simulations in one dimension. We discuss the implications for experiments on quantum-dot chains tuned into a regime supporting end Majorana zero modes and on trapped ion chains.

I. INTRODUCTION

Solid-state systems supporting non-Abelian anyons, such as Majorana zero modes (MZMs), are the focus of considerable research aimed at exploiting them for quantum information processing [1–4]. In the limit of zero temperature, the quantum information stored in a collection of non-Abelian anyons is protected, up to corrections exponentially small in the separation between anyons. At nonzero temperatures, however, thermally excited bulk quasiparticles can be absorbed or emitted by a zero mode, thereby corrupting the quantum information contained therein. Thus a separation-independent failure of protection is expected to increase with temperature as $e^{-\Delta/T}$, where Δ is an energy gap. These processes increase the width and reduce the height of the predicted [5–7] zero-bias peak that appears to have been observed in tunneling experiments [8–14].

However, the decay of a zero mode and the quantum information encoded in it is a nonequilibrium dynamical process, and it is not clear if thermodynamic reasoning can describe it properly. In the absence of electron-electron interactions—for instance, in the Kitaev chain Hamiltonian [15] or the transverse-field Ising chain, to which it is related by a Jordan-Wigner transformation—thermally excited

quasiparticles do not affect the zero modes at all. This can be reconciled with the previous paragraph by noting that, in the absence of interactions, the coefficient in front of $e^{-\Delta/T}$ vanishes. While the absence of interactions is a fine-tuned special case, a similar conclusion holds in systems with strong disorder in which many-body localization [16] occurs. Here disorder-induced localization prevents bulk excitations from carrying quantum information away from a zero mode [18–20].

Disorder, however, is not necessary to have zero modes in interacting systems. In at least one integrable system, the XYZ spin chain, exact edge zero modes survive the presence of interactions [21]. This edge "strong zero mode" is an operator that commutes with the Hamiltonian up to exponentially small corrections in the finite size of the system [22–24]. Moreover, as with the transverse-field Ising chain, thermally excited quasiparticles do not cause the edge degrees of freedom to equilibrate with the bulk. Rather, the edge spin coherence lasts forever in a semi-infinite chain, even at infinite temperature [25]. Even more strikingly, similar behavior was found in several nonintegrable deformations of the Ising chain. Here the coherence time is not infinite, but extremely long lived [25].

The purpose of this paper is to show that such long-lived edge modes are a more general phenomenon and to give a direct and rigorous method for understanding them. We demonstrate that "prethermalization," the exponentially slow approach to thermal equilibrium that occurs in some closed quantum systems [26–31], can protect edge zero modes and, in fact, topological degrees of freedom in higher-dimensional systems as well. (Prethermalization can also occur in periodically driven systems [31–39], but this

is not our focus here.) Our analysis gives a clear meaning to the notion of an "almost" strong zero mode: it is an operator that commutes with the full Hamiltonian of a system up to corrections that are a nearly exponentially small function of a ratio of energy scales. Here by "nearly" exponentially small, we mean with a logarithmic correction to the exponent, as set out in Eq. (3) below. We call such an operator a "prethermal strong zero mode." Its lifetime is bounded below by a nearly exponentially growing function of this ratio of energy scales because prethermalization delays equilibration of a prethermal strong zero mode until this late time.

By relating the protection of quantum information to prethermalization, we reveal the limits of such protection. We elucidate the nature of this protection in one-dimensional, two-dimensional, and three-dimensional closed systems. However, solid-state devices are not closed systems and prethermalization in such devices is eventually superseded by thermalization driven by electron-phonon interactions. Thus, the applicability of these ideas to Majorana zero modes in semiconductor-superconductor devices depends on the particular device considered (since the prethermal limit is not accessible in all devices) and, even then, is only in some temperature range over which the electron-phonon interaction does not dominate. We quantitatively analyze the quantum-dot chain proposed in Refs. [40,41]; although it has not been realized in experiments yet, it is a useful case study. We show that prethermalization can occur over a range of time scales. In this prethermal regime, arguments relying on thermal equilibrium are ultimately correct, but with a (nearly) exponentially small prefactor that reflects the slow thermalization of the system. Moreover, the naive energy gap Δ is replaced with a much larger effective energy gap Δ_{eff}. This suggests that the $T > 0$ protection of quantum information may be optimized by entering the prethermal regime, in addition to—or even rather than—maximizing the energy gap. We demonstrate this tradeoff in explicit models.

The type of prethermalization we describe is not special to one-dimensional or topological systems. We describe explicitly how analogous phenomena occur in some two- and three-dimensional systems and how prethermalization protects edge modes for long times in systems not topologically ordered. One particular example we describe in detail is the transverse-field Ising chain perturbed by integrability-breaking interactions. While this chain is related to the quantum-dot chain via a Jordan-Wigner transformation, the nonlocality of the map means that topological order in the latter is simply ordinary ferromagnetic order in the former. Nonetheless, we show how prethermalization means the edge spin coherence lasts for very long times here as well under *any* perturbation preserving the $\mathbb{Z}_2$ spin-flip symmetry. Namely, the spin coherence lasts for a time that is (nearly) exponentially large in terms of the couplings.

We begin with a conceptual overview in Sec. II. In Sec. III, we explain the prethermal regime and the theorem of Abanin *et al.* [31] that guarantees its existence for certain Hamiltonians. In the first part of Sec. IV, we show how this theorem can be used to provide a lower bound on the lifetime of edge zero modes and apply it explicitly to a model of interacting Majorana fermions. In the remainder of Sec. IV, we discuss the lifetime of the Majorana zero modes at nonzero temperatures and present numerical simulations supporting our arguments. We then generalize this analysis in the following Sec. V, describing the conditions needed to observe prethermally protected zero modes and giving details of several examples that illustrate these conditions. In Sec. VI, we apply this general strategy to analyze systems in two and three dimensions. Section VII explores possible practical applications of our results, in quantum-dot and ion chains. Finally, in Sec. VIII, we consider integrable systems, where the zero modes may survive much longer (possibly even infinitely longer) than the lower bound.

II. TOPOLOGICAL ZERO MODES AT FINITE TEMPERATURE

A common feature of systems exhibiting topological order at zero temperature is topological degeneracy, where there are several nearly degenerate ground states. Their energy splitting scales as $e^{-L/\xi}$ for some correlation length ξ, where L is the system size. Moreover, these degenerate states are locally indistinguishable. This means that, at zero temperature, quantum information can be stored in the degenerate ground state subspace in a topologically protected way. However, this topological protection usually does not extend to finite temperature. Here we review, in a schematic way, the standard arguments for this. We then explain why, in an *isolated* system, prethermalization can improve the situation considerably, both in topologically and conventionally ordered systems. A rigorous argument is given in later sections.

Our prototypical example is a one-dimensional topological superconductor, exemplified by the Kitaev chain [15]. In such systems with open boundary conditions, there is a pair of Majorana zero modes on the two ends of the chain, represented by Majorana operators γ and γ' [see Fig. 1(a)]. At zero temperature, these can be used to encode quantum information in the qubit. The qubit can be decohered by randomly acting on it with the logical operators $\sigma^z = i\gamma\gamma'$ or $\sigma^x = \gamma$. Since these are both nonlocal (the latter because it is a fermionic operator), the qubit is therefore immune to decoherence from any local noise process.

At finite temperature, on the other hand, there is a finite density of fermionic quasiparticles in the bulk of the system. If such a quasiparticle is near one of the Majorana zero modes (say, the one corresponding to γ), then it can annihilate on it [Fig. 1(b)], which has the effect of acting
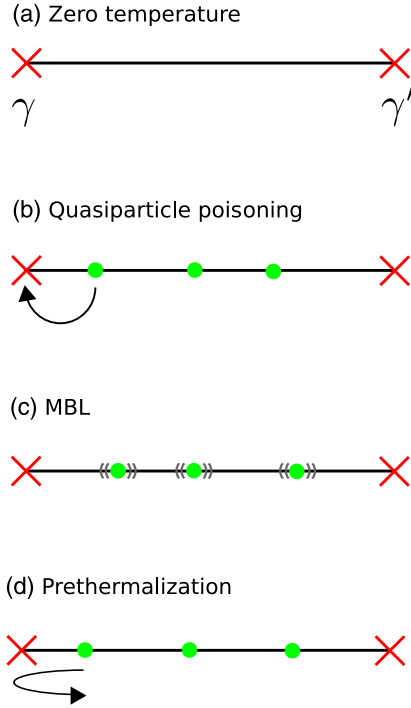
FIG. 1. (a) A 1D chain of a topological superconductor, with Majorana zero modes at the edges. (b) At finite temperature, mobile quasiparticles can annihilate on the Majorana zero modes, decohering the quantum information stored. (c) With strong disorder, the chain can be made to be MBL, such that the localized quasiparticles are not able to annihilate on the boundaries. (d) In a suitable "prethermal" regime, the quasiparticles are mobile but are prevented from annihilating on the boundary by an approximate conservation law.

on the encoded qubit with the logical operator $\sigma^x = \gamma$. At finite temperature, such processes happen continuously, and so the encoded qubit quickly decoheres. This is known as "quasiparticle poisoning."

How can quasiparticle poisoning be overcome? One way is by many-body localization (MBL) [19,42–50]. This causes the quasiparticles to become immobile and thus prevents them from moving onto the boundary and annihilating [50], as shown in Fig. 1(c). Going to a MBL phase, however, requires strong disorder, and it is not at all clear if such phases exist outside one dimension. Moreover, MBL systems may not be suitable for topological quantum computation with Majorana zero modes because of nonlocal rearrangements that occur when varying the Hamiltonian adiabatically [51,52].

We show in this work that there is another way of avoiding quasiparticle poisoning. Our approach exploits the fact that if the number of quasiparticles in the bulk is conserved (in fact, it is sufficient for the number to be conserved modulo 2), then they cannot be annihilated on the boundaries [Fig. 1(d)]. Of course, outside integrable systems, there is no reason why such a conservation law should hold exactly, but often parameters of the

Hamiltonian can be tuned such that it holds in an approximate way, leading to a long decoherence time.

One might imagine that obtaining a long decoherence time in this way would require significant fine-tuning. Remarkably, this turns out not to be the case. We show that many Hamiltonians possess a significant parameter regime with an approximately conserved quantity that we identify with quasiparticle number. Terms violating the conservation law are *exponentially* small in parameters of the Hamiltonian. This is based on the mechanism of prethermalization, as discussed further in later sections.

Our approach is not necessarily limited to one-dimensional systems. In higher dimensions, decoherence of quantum information is also generally related to processes involving quasiparticles. For example, in the 2D toric code defined on a torus, decoherence is caused by quasiparticles moving around a noncontractible loop on the torus. On the other hand, the 4D toric code is known to be immune from decoherence at low temperatures, precisely because this system has no topologically nontrivial particle-like excitations [53]. Unfortunately, no such system is known in dimension less than four.

We show that prethermalization can be used to suppress decoherence in an isolated system arising from creation or annihilation of quasiparticles. Thus, it is no help in the case of the toric code on a torus, since quasiparticles can move around noncontractible loops without changing the total quasiparticle number. However, in the *planar* version of the toric code, the decoherence mechanism involves quasiparticles annihilating on the edges. Prethermalization therefore is useful here. We analyze this and other higher-dimensional examples in more detail in Sec. VI.

III. PRETHERMAL REGIME

A closed quantum system is said to be "prethermal" if, en route to thermalization, it is in an exponentially long-lived quasisteady state. One cause of prethermalization is an approximate conservation law: over intermediate time scales—known as the prethermal regime—the system maximizes its entropy, subject to the constraint that the conserved quantity takes a fixed value. Over sufficiently long time scales, the entropy is maximized without any constraint.

A theorem of Abanin, De Roeck, Huveneers, and Ho (henceforth ADHH) [31] guarantees the existence of such a prethermal regime for Hamiltonians of the form

$$\hat{H} = -J\hat{N} + \hat{Y}, \qquad (1)$$

where $\hat{N}$ is a sum of finite-range commuting terms, such that $\hat{N}$ has integer eigenvalues, i.e., $e^{2\pi i \hat{N}} = 1$. The proof in Ref. [31] assumes that each term in $\hat{N}$ acts only on a single site, but in Appendix A we show that this assumption can be relaxed. We define a parameter J_0 that is essentially the

largest operator norm of any local term in Y; more explicit forms are given in specific examples below. The theorem says that for J_0/J sufficiently small, there exists a local unitary transformation $\mathcal{U}$ such that

$$\mathcal{U}\hat{H}\mathcal{U}^\dagger = -J\hat{N} + \hat{D} + \hat{E}, \qquad (2)$$

where $[\hat{N}, \hat{D}] = 0$ and $\|\hat{E}\| = O(e^{-cn_*})$, where

$$n_* = \left\lceil \frac{J/J_0}{[1 + \log(J/J_0)]^3} \right\rceil \qquad (3)$$

and c is a constant. Since we focus on the $J \gg J_0$ limit, we often drop the 1 in the denominator of Eq. (3) for the sake of uncluttering the equations. As a result, the dynamics of the system conserves $\hat{N}$ until a time $t_* \propto e^{cn_*}$. A more precise statement of the ADHH theorem can be found in Ref. [31].

For our purposes, the essential point of the ADHH theorem is that a Hamiltonian, Eq. (1), has an emergent approximate $U(1)$ symmetry generated by $\hat{N}$. The symmetry violations come from terms in the transformed Hamiltonian that are nearly *exponentially small* in the large-J limit. We apply the ADHH theorem to Hamiltonians with edge zero modes and show that the approximate $U(1)$ symmetry can protect these zero modes—even far from the ground state of the system, where we do not ordinarily expect topological protection. It also protects them in systems with no topological order whatsoever, thus providing a completely different mechanism for preserving quantum coherence distinct from topological considerations.

There is a nice heuristic picture for why a Hamiltonian of the form in Eq. (1) has an approximate $U(1)$ symmetry. The transformation to the form of Eq. (2) means that it is very difficult for the terms in $\hat{Y}$ to cause transitions between different eigenspaces of $\hat{N}$: in order to conserve energy, many excitations of $\hat{Y}$ must be created or annihilated. Such a process occurs slowly, so violation of the approximate $U(1)$ symmetry takes a very long time.

However, in applying these ideas to real systems, it is important to keep in mind that the ADHH theorem requires that $\hat{Y}$ be a sum of local terms, each of which has a bounded (indeed, small) norm. This condition cannot be satisfied in any real solid since phonons and photons do not have finite-dimensional local Hilbert spaces. The energy associated with a transition between different eigenspaces of $\hat{N}$ need not wait for many excitations of $\hat{Y}$ to be created or annihilated; it can, instead, be supplied or carried away by a phonon or photon. However, if the couplings of the electronic degrees of freedom to phonons and photons are sufficiently small, they may not play a role during the prethermal regime. We show that this is the case in semiconductor devices in Sec. VII A. We note also that the presence of gapless excitations does not automatically destroy all symmetry-protected edge modes [54].

IV. PRETHERMAL PROTECTION OF A TOPOLOGICAL QUBIT

The importance of the ADHH theorem for the present work is that in many cases $\mathcal{U}\hat{N}\mathcal{U}^\dagger$ can be identified with an effective "quasiparticle number," and its approximate conservation can suppress the decoherence mechanisms of a topological qubit, as outlined in Sec. II.

In this section we introduce the idea by describing a particular example in depth: a topological superconducting chain. We show that it is possible to construct a prethermal edge Majorana zero mode for arbitrary interactions, as long as the topological superconducting ordering term is the dominant coupling. We present evidence from numerical simulations that the prethermal regime persists over a surprisingly large range of couplings, including values of the dominant coupling that are not so very large. Later, in Sec. VII A, we discuss the relevance of these ideas to quantum-dot chains in a semiconductor-superconductor heterostructure, where the electron-phonon coupling cuts off prethermalization. As we see, there are circumstances under which the effect of the electron-phonon coupling is weaker than electron-electron interactions, so that prethermalization acts to suppress the dominant MZM decay channel, leading to relatively long-lived MZMs. Although this first example is one dimensional, the basic idea and the theorem on which it relies work in any dimension, as we discuss in later sections.

A. Prethermalization in the interacting Kitaev chain

The Kitaev chain [15] is a simplified model of a one-dimensional topological superconductor of spinless fermions. Its Hamiltonian is

$$H = -\sum_j [t(c_j^\dagger c_{j+1} + c_j c_{j+1}^\dagger) + \mu c_j^\dagger c_j$$
$$+ \Delta(c_j c_{j+1} + c_{j+1}^\dagger c_j^\dagger)]. \qquad (4)$$

The topological superconducting phase occurs for $2|t| > |\mu|$, provided $\Delta > 0$.

Moreover, we need to include interactions in our description. They can be and usually are dropped for the purposes of demonstrating the existence of the topological superconducting phase and its concomitant MZMs, but they are necessary for any discussion of nonzero temperature dynamics. With the addition of such terms, Eq. (4) takes the form

$$H = \sum_j \left[-\frac{t+\Delta}{2}(c_j^\dagger c_{j+1} + c_j c_{j+1}^\dagger + c_j c_{j+1} + c_{j+1}^\dagger c_j^\dagger) \right.$$
$$- \frac{t-\Delta}{2}(c_j^\dagger c_{j+1} + c_j c_{j+1}^\dagger - c_j c_{j+1} - c_{j+1}^\dagger c_j^\dagger)$$
$$\left. - \mu c_j^\dagger c_j + V c_j^\dagger c_j c_{j+1}^\dagger c_{j+1} + \cdots \right]. \qquad (5)$$

We have only written the simplest interaction term explicitly (the V term) and denoted the rest implicitly with the ellipses. In writing the Hamiltonian in this way, we have explicitly separated the presumed largest term, written on the first line, from the smaller terms, written on the second and third lines.

This model can be written in terms of the Majorana fermion operators γ_j^A and γ_J^B defined according to $c_j = (\gamma_j^A + i\gamma_J^B)/2$. This rewriting gives $\hat{H} = -J\hat{N} + \hat{Y}$, where

$$\hat{N} = i\sum_{j=1}^{L-1}\gamma_j^B\gamma_{j+1}^A = \sum_{j=1}^{L-1}\sigma_i^z\sigma_{i+1}^z, \qquad (6)$$

and $J \equiv (t + \Delta)/2$. On the right-hand side of the second equal sign, we have written the Jordan-Wigner transformed representation of this operator according to the definitions

$$\gamma_j^A \equiv \sigma_j^z\prod_{k=1}^{j-1}\sigma_k^x, \qquad \gamma_j^B = i\sigma_j^x\gamma_j^A. \qquad (7)$$

Note that we have taken open boundary conditions in order to focus on the physics of the edge. One important thing to note is that with open boundary conditions, the two Majorana fermions at the edges, γ_1^A and γ_L^B, do not appear in $\hat{N}$. Each therefore commutes with it. We also have

$$\hat{Y} = -ih\sum_{j=1}^{L}\gamma_j^A\gamma_j^B - \sum_{\alpha,\beta}J_{\alpha\beta}\sum_{j=1}^{L-1}i\gamma_j^\alpha\gamma_{j+1}^\beta$$
$$- h_2\sum_{j=1}^{L-1}\gamma_j^A\gamma_j^B\gamma_{j+1}^A\gamma_{j+1}^B - J_2\sum_{j=1}^{L-2}\gamma_j^B\gamma_{j+1}^A\gamma_{j+1}^B\gamma_{j+2}^A$$
$$+ \cdots, \qquad (8)$$

where … denotes other third-neighbor and more-distant hopping and interaction terms. The transverse field h is the chemical potential of the topological superconductor, Eq. (5), according to the identification $h \equiv \mu$. The two four-Fermi terms we single out are the two simplest, and in spin language are $h_2\sigma_j^x\sigma_{j+1}^x$ and $J_2\sigma_j^z\sigma_{j+2}^z$, respectively; in terms of the original topological superconductor, $h_2 \equiv V$, while J_2 includes slightly longer-ranged interactions and Cooper pair-hopping terms not explicitly included in Eq. (5). For simplicity, we take couplings in $\hat{Y}$ to be spatially uniform, but this is not necessary for the approach to work; in fact, adding disorder to $\hat{Y}$ typically enhances the effects that we describe. We thus include a J_{AB} term in Eq. (8); because of the integer-eigenvalue restriction, it cannot be absorbed into the $J\hat{N}$ term if it is disordered. We assume that the hopping and interaction terms have finite range but make no further assumptions.

We now apply the ADHH theorem to the perturbed Ising-Kitaev Hamiltonian $-J\hat{N} + \hat{Y}$, with open boundary conditions and the operators defined by Eqs. (6) and (8). The theorem is applicable for J/J_0 sufficiently large, where the energy scale J_0 is given in this case by

$$J_0 = \frac{1}{\kappa_0^2}[e^{\kappa_0}h + e^{2\kappa_0}(h_2 + J_{AB} + J_{BA} + J_{AA} + J_{BB})$$
$$+ e^{3\kappa_0}(\text{three-site terms}) + \cdots]. \qquad (9)$$

The J_{AB} term is included in the presence of disorder, where it is the deviation of J from its mean value. In the disordered case, J_0 is defined as the maximum possible value of the right-hand side of Eq. (9) for any site in the system, stipulating that the interactions on the right-hand side of Eq. (9) touch that site. The number κ_0 is chosen so that this sum is finite. By choosing e^{κ_0} as large as possible while satisfying this requirement, we can maximize the range of J over which the theorem applies. Then the theorem guarantees that there exists a local unitary transformation $\mathcal{U}$ such that

$$\mathcal{U}(-J\hat{N} + \hat{Y})\mathcal{U}^\dagger = -J\hat{N} + \hat{D} + O(e^{-cn_*}), \qquad (10)$$

where $[\hat{N}, \hat{D}] = 0$ and n_* is given by Eq. (3). Another way to say this is that the original Hamiltonian $\hat{H} = -J\hat{N} + \hat{Y}$ has an approximately conserved quantity $\mathcal{U}^\dagger\hat{N}\mathcal{U}$, which is conserved until times $t_* = O(e^{cn_*})$.

To illustrate the significance of this approximately conserved quantity, let us consider the limit $J_0 \to 0$, which implies $\hat{Y} = 0$ so the Hamiltonian is simply $\hat{H} = -J\hat{N}$. In that case, the local unitary rotation $\mathcal{U} = \mathbb{1}$, so the conserved quantity is simply $\mathcal{U}^\dagger\hat{N}\mathcal{U} = \hat{N}$. In this limit, excitations (henceforth, "quasiparticles") correspond to violations of one of the terms in Eq. (6) and have no dynamics. The conserved quantity $\hat{N}$ simply counts the number of quasiparticles. Precisely, the number of quasiparticles is the eigenvalue of $(L - 1 - \hat{N})/2$. In fact, for our purposes, it is sufficient to use the fact that the number of quasiparticles is conserved modulo 2 in this limit; in other words, here there is a conserved $\mathbb{Z}_2$ charge,

$$\tilde{\mathcal{F}} \equiv e^{i\pi\hat{N}/2} = i\gamma_1^B\left(\prod_{j=2}^{L-1}i\gamma_j^A\gamma_j^B\right)\gamma_L^A. \qquad (11)$$

We emphasize that this is *not* the same as the fermion parity $\mathcal{F}$, which is also a symmetry of any local fermionic Hamiltonian; the latter is instead given by

$$\mathcal{F} = \prod_{j=1}^{L}(i\gamma_j^A\gamma_j^B), \qquad (12)$$

which differs from $\tilde{\mathcal{F}}$ in that the boundary Majorana operators γ_1^A and γ_L^B are also included (thus, $\tilde{\mathcal{F}}$ can in a sense be interpreted as the "fermion parity in the bulk").

The important result is that as perturbations are added, moving J_0 away from zero, there is a continuous deformation of $\hat{N}$, namely, $\mathcal{U}^\dagger \hat{N} \mathcal{U}$, which we continue to identify as the effective quasiparticle number, and it is approximately conserved. By the intuitive discussion in Sec. II, this suggests that topological information stored in the Majorana zero modes should have a long decoherence time. We are now in a position to rigorously establish this. Indeed, a sufficient condition to allow for quantum information to be stored with infinite (respectively, very large) decoherence time is that there exist Majorana operators Ψ^A and Ψ^B that square to the identity; anticommute with each other; and commute (respectively, almost commute) with the Hamiltonian. This means that they form an "edge strong MZM" in the language of Ref. [24]. We show that this is an implication of the ADHH result.

The presence of the strong MZMs has important ramifications. In the fermionic picture, both Ψ^A and Ψ^B toggle between *all* states in the sectors, even the highly excited ones. Their presence means that the full Hilbert space of the system can be decomposed into the tensor product of a two-state quantum system, i.e., a topological qubit and a nontopological "bulk" Hilbert space, of dimension 2^{L-1}, such that the Hamiltonian vanishes upon projection onto the topological qubit, up to finite-size corrections $\sim e^{-L/\xi}$. Consequently, the topological qubit is protected at any temperature: regardless of how the dynamics of the system affects the projection of the state of the system into the bulk Hilbert space, the topological qubit is unaffected. In the spin language, this means that the autocorrelator of the boundary spin operator $\sigma_1^z = \gamma_1^A$ for any temperature or initial state is nonvanishing up to exponentially long times of order $(J/h)^L$ [25]. The same goes for the other edge spin, $\sigma_L^z = \mathcal{F}\gamma_L^B$.

To see why ADHH implies a strong MZM, let us first return to the limit $J_0 = 0$, where the Hamiltonian is simply $-J\hat{N}$. In this limit, the Majorana operators γ_1^A and γ_L^B already form a strong MZM. Now turn on any or all of the other terms in $\hat{Y}$, while keeping $0 < J_0 \ll J$. The ADHH theorem states that there exists a local unitary change of basis $\mathcal{U}$ that transforms the problem into one in which $\hat{N}$ is conserved, up to nearly exponentially small corrections. The locality properties of $\mathcal{U}$, along with the fact that $[\hat{N}, \hat{D}] = 0$, require that the approximate transformed Hamiltonian $-J\hat{N} + \hat{D}$ commutes with the edge Majorana fermions:

$$[(-J\hat{N} + \hat{D}), \gamma_1^A] = [(-J\hat{N} + \hat{D}), \gamma_L^B] = 0. \quad (13)$$

To see this, observe that $-J\hat{N} + \hat{D}$ commutes with both the fermion parity $\mathcal{F}$ and the "bulk fermion parity" $\tilde{\mathcal{F}}$ defined in Eq. (11). Therefore, it also commutes with their product, $\mathcal{F}\tilde{\mathcal{F}} = i\gamma_1^A\gamma_1^L$. The fact that $\mathcal{U}$ is a local unitary ensures that the norm of any term coupling both γ_1^A and γ_L^B must be

exponentially small in L. We assume that L is large enough that such terms can be ignored. It then follows that all terms must commute with γ_1^A and γ_1^B individually, proving Eq. (13).

Thus, in the presence of interactions, we define

$$\Psi^l = \mathcal{U}^\dagger \gamma_1^A \mathcal{U}, \qquad \Psi^r = \mathcal{U}^\dagger \gamma_L^B \mathcal{U}. \quad (14)$$

The vanishing commutator, Eq. (13), and the theorem, Eq. (10), show that these commute with H up to order $cn_* < L$:

$$[\hat{H}, \Psi^r] = O(e^{-cn_*}), \qquad [\hat{H}, \Psi^r] = O(e^{-cn_*}). \quad (15)$$

By construction, Ψ^l and Ψ^r square to the identity operator and anticommute with $\mathcal{F}$. We call such operators, satisfying Eq. (15), "prethermal strong Majorana zero modes." The authors of Ref. [25] found evidence for "almost" strong zero modes in the special case where $\hat{Y}$ contains only nonzero h, h_2, J_2. Here we have established that they are an example of a prethermal strong MZM, confirming the claims made there.

B. Temperature dependence of the lifetime

Now suppose that the system is at temperature T. The dynamics of the prethermal strong MZM Ψ^l are visible in the retarded Green function

$$G(\omega) \equiv \int_0^\infty dt e^{i\omega t} \text{Tr}\{e^{\beta H}\gamma_1^A(t)\gamma_1^A(0)\}$$
$$= \int_0^\infty dt e^{i\omega t} \text{Tr}\{e^{\beta H}\sigma_1^z(t)\sigma_1^z(0)\}. \quad (16)$$

This is not the Green function of Ψ^l but rather that of the "bare" operator γ_1^A. However, there is nonzero overlap between these two operators, so the Green function of the latter can be used to probe the former. In the limit of large system size, we can ignore the interaction between Ψ^l and Ψ^r, so this Green function takes the form

$$G(\omega) = Z(\omega + i\Gamma)^{-1}. \quad (17)$$

The "wave-function renormalization" Z is a measure of the overlap between γ_1^A and the MZM operator Ψ^l, as determined by the unitary transformation $\mathcal{U}$. The decay rate Γ is our primary figure of merit in judging prethermal strong zero modes. It is directly reflected in the width of a zero-bias peak observed in tunneling into the end of a Majorana chain (see Sec. VII A) and determines the error rate for topological qubits encoded in the MZMs Ψ^l, Ψ^r. The decay rate Γ is determined by $\hat{E}$, the correction term in the transformed basis, and would vanish if $\hat{E}$ were to vanish, with the decay time becoming infinite. From Eq. (15), which expresses the fact that $\|\hat{E}\| = O(e^{-cn_*})$, we see that

$$\Gamma = A(T)e^{-cn_*} \approx A(T)e^{-[(cJ/J_0)/\ln^3(J/J_0)]}. \quad (18)$$

Let us now examine the temperature dependence in more detail. After a local unitary rotation by $\mathcal{U}$, the Hamiltonian can be shifted to Eq. (2) as $\hat{H} = -J\hat{M} + \hat{D} + \hat{E}$, where $\hat{M} = -\hat{N} + c$, with the constant c chosen so that $\hat{M}$ has the smallest eigenvalue 0. The term $\hat{D}$ commutes with $\hat{M}$ and can be jointly diagonalized with it, so the decay is attributed to resonant transitions between different $\hat{M}$ sectors induced by $\hat{E}$. These transitions can only happen between states with approximately the same energy with respect to $J\hat{M} + \hat{D}$. The corresponding energy bandwidth of the sector with $\hat{M} = m$ is bounded by $[E_0 + mJ - mCJ_0, E_0 + mJ + mCJ_0]$, where E_0 is the ground-state energy, and C is some dimensionless constant. Roughly, this is because a state in this sector differs from the ground state only in at most m spots, and only those spots can contribute to the energy difference from the ground state. For a rigorous proof, see Appendix B. States in the $\hat{M} = m$ sector can have the same energy as those in the $\hat{M} = m + 1$ sector when $E_0 + mJ + mCJ_0 \geq E_0 + (m+1)J - (m+1)CJ_0$ or, in other words, when $m \geq m_{\min} \sim J/J_0$. A transition can only occur when the resonance condition is satisfied. In an excited state with a nonzero density of excitations, at least $m_{\min}$ of them must be close to the MZM in order for a transition to occur. Hence,

$$\Gamma_{\text{el-el}} = A_0 e^{-cn_*} \rho^{m_{\min}}, \quad (19)$$

where ρ is the density of excited quasiparticles.

We thus have shown that the decay time of a MZM of Eq. (19) is determined by two factors. The important consequence is that decay of a MZM is suppressed by a high power of the quasiparticle density, in addition to the state-independent exponential suppression e^{-cn_*}.

At very low temperatures, the density of excited quasiparticles in superconductors is generally higher than the expected thermal (or prethermal) equilibrium value $\rho \sim e^{-\Delta/T}$. (The reasons for this lie outside the purview of our discussion; for a recent theoretical analysis, see Ref. [55]. References [1–7] in this paper contain experimental measurements of the quasiparticle density.) At temperatures that are not too low, the quasiparticle density exhibits equilibrium behavior, and we have $\rho \sim e^{-\Delta/T}$. When this is the case, the decay rate is controlled by $e^{-\Delta_{\text{eff}}/T}$, where we have defined the "effective gap" $\Delta_{\text{eff}} \sim J(J/J_0)$. The effective gap is much larger than the actual gap, which is $\Delta \sim J$, up to corrections of order J_0. Therefore, for $J \gg J_0$ and $T \ll \Delta_{\text{eff}}$, the decay rate is of the form

$$\Gamma_{\text{el-el}} = A_0 e^{-cn_*} e^{-\Delta_{\text{eff}}/T}. \quad (20)$$

This shows that the finite lifetime of a MZM is exponentially large in $n_* \sim (J/J_0)/\ln^3(J/J_0)$ and also exponentially

large in $1/T$. Equation (20) suggests that MZM qubits can be optimized by maximizing J/J_0 even at the cost of reducing Δ, since in any case $\Delta_{\text{eff}} \gg \Delta$ when J/J_0 is large.

C. Numerical results for prethermal MZMs

The preceding general arguments can be substantiated by computations in finite-size systems. For chains of length $N = 8$–14, we study prethermal strong MZMs all the way up to infinite temperature by exact diagonalization. As may be seen in the top panel of Fig. 2, the MZM survives to very long times at infinite temperature. The lifetimes are consistent with an exponential dependence on the ratio of scales J/J_0 until the lifetime becomes so long that finite-size effects become important. In a model in which the only terms in $\hat{Y}$ are h and h_2, we can write $J_0 = h_2/f(h/h_2)$ for some function $f(x)$. As may be seen from Fig. 2, the data collapse onto this form.

Using time-evolving block decimation (TEBD) [56], we can study much larger systems, where finite-size effects are less severe. This approach is applicable to low-energy initial states for which the entanglement does not grow too much, allowing accurate simulation with bond dimension $\chi = 100$. Our results confirm that, at least for low-energy states, the Majorana lifetime remains large for larger system sizes (see Fig. 3). Note that the decay time shown in Fig. 3 should not be directly compared to Fig. 2 because Fig. 2 is at infinite temperature, whereas Fig. 3 is at very low temperature.
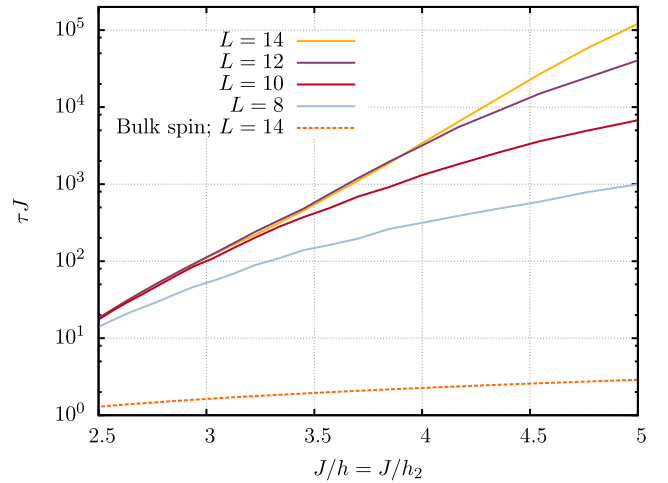

FIG. 2. The decay time (on a logarithmic scale) of a MZM at $T = \infty$ for $L = 8$–14 sites, as a function of h and h_2, as obtained by exact diagonalization. The data collapse onto the form $\ln \tau J = J/h_2 f(h/h_2) + \text{constant}$, as expected, with deviations attributable to finite-size effects when the lifetime is long. The decay time of the spin at the center of the $L = 14$ chain (dashed line) is not prethermalization protected and is much shorter.
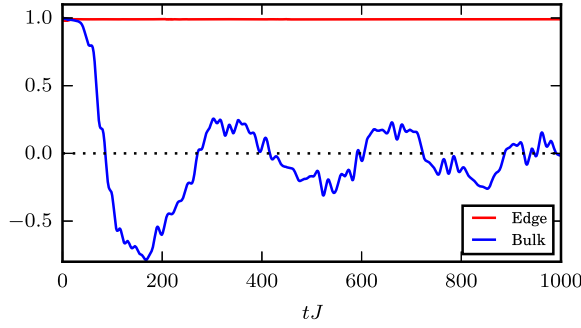
FIG. 3. The decay of a MZM at low energies, as obtained by TEBD for systems of length $L = 40$, which shows that a long-lived MZM persists even in a much larger system. Red: Decay of the strong zero mode Green's function $\langle\Psi|\gamma_1^A(t)\gamma_1^A(0)|\Psi\rangle$ for an initial state $|\Psi\rangle$ containing a single quasiparticle close to the left edge (i.e., $-\gamma_3^B\gamma_4^A = -1$ and $\gamma_i^B\gamma_{i+1}^A = 1$ for $i \neq 3$), for $L = 40$ and $h/J = h_2/J = 0.075$. This is equivalent in the corresponding transverse-field Ising chain to the spin correlation $\langle\Psi|\sigma_1^z(t)\sigma_1^z(0)|\Psi\rangle$. Blue: Decay of the bulk spin correlation $\langle\Psi|\sigma_{L/2}^z(t)\sigma_{L/2}^z(t)|\Psi\rangle$ in the transverse-field Ising chain.

V. GENERAL CRITERIA AND EXAMPLES FOR PRETHERMALIZATION-PROTECTED TOPOLOGICAL DEGREES OF FREEDOM

The procedure described in preceding sections uses the ADHH theorem to find prethermal strong MZMs. In this section, we give a systematic approach to applying the theorem to determine when topological degrees of freedom are protected from thermal fluctuations by prethermalization. This approach involves two key observations:

(i) The ADHH theorem guarantees the presence of a single long-lived local $U(1)$ charge for J/J_0 large enough.

(ii) In turn, this nearly conserved $U(1)$ prevents local bulk excitations from violating the conservation of some topological charge, up to exponentially small terms.

More prosaically, the idea is that the nearly conserved quantity guarantees that a topological charge localized at edges or defects cannot be changed by the absorption or emission of small numbers of bulk excitations. These general criteria can be applied in any dimension, and so we give some examples in one, two, and three dimensions. We illustrate them through several one-dimensional examples in the remainder of this section and with higher-dimensional examples in the next.

A. Three-state Potts

It is instructive to analyze an example where the ADHH theorem, although applicable, does *not* guarantee any prethermal strong zero modes. As shown in Refs. [22,23], the three-state Potts chain does not have any sort of edge zero mode for finite J. The easy edge-spin-flip process that kills

the putative zero mode can be seen easily in perturbation theory for J large. Here we rephrase this result in the more general setup of this paper.

The quantum-chain analog of the three-state Potts model has a three-state quantum system on each of L sites. The basic operators σ and τ acting nontrivially on a single site generalize the Pauli matrices σ^x and σ^z. Instead of squaring to the identity and anticommuting, they obey

$$\sigma^3 = \tau^3 = 1, \qquad \sigma^\dagger = \sigma^2, \qquad \tau^\dagger = \tau^2, \qquad \sigma\tau = \omega\tau\sigma,$$

where $\omega \equiv e^{2\pi i/3}$. Matrices satisfying this algebra are

$$\sigma = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \omega & 0 \\ 0 & 0 & \omega^2 \end{pmatrix}, \qquad \tau = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}. \qquad (21)$$

Here σ generalizes the Pauli matrix σ^z to measure the value of a clock variable, while τ generalizes σ^x to shifting the value. The operators σ_j, τ_j are defined analogously to the Ising case, where they act nontrivially on the jth site of the chain and trivially elsewhere.

The three-state Potts chain is invariant under global S_3 permutations of the three states and has nearest-neighbor interactions. This fixes the Hamiltonian with open boundary conditions to be $H = -J\hat{N}_P + Y_P$, where

$$\hat{N}_P = \sum_{j=1}^{L-1}(\sigma_j^\dagger\sigma_{j+1} + \sigma_j\sigma_{j+1}^\dagger), \qquad \hat{Y}_P = -\sum_{j=1}^{L}(\tau_j + \tau_j^\dagger).$$

$$(22)$$

If desired, these operators can be rewritten in terms of $\mathbb{Z}_3$ parafermionic operators akin to Majorana fermions [24].

The operator $\hat{N}_P$ has integer eigenvalues, since $\sigma_j^\dagger\sigma_{j+1} + \sigma_j\sigma_{j+1}^\dagger$ has eigenvalue 2 if the $\mathbb{Z}_3$ spins at j and $j+1$ are identical and -1 if they are not. Thus, $\hat{N}_P$ is related to counting kinks, just as in Ising. However, an important difference with the Ising case is that here there are two types of kinks. We label the three states at each site as A, B, C, with the $\mathbb{Z}_3$ symmetry cyclically permuting them. When the states AB (or BC or CA, their cyclic permutations) occur on sites j and $j+1$, respectively, we call the configuration a kink, and when BA appears (or CB or AC), we call the configuration an antikink. Thus, $[2(L-1) - \hat{N}_P]/3$ counts the number of kinks *plus* the number of antikinks.

Since $\hat{N}_P$ is an integer, the ADHH theorem says there is an emergent $U(1)$ symmetry in the prethermal regime, conserving the total number of quasiparticles. However, this symmetry does not prevent a zero mode from decaying, because a kink can scatter off the edge and turn into an antikink without changing $\hat{N}_P$. In perturbation theory, this results from the easy-spin-flip process described in Ref. [23] (see [57]). Namely, consider, say, the AB kink

on sites 1 and 2. The Y_P term allows the spin on site 1 to be shifted from A to C, and so converts the AB kink to the CB antikink. This process conserves $\hat{N}_P$ while flipping the edge spin. Clearly there can be no long edge-spin relaxation time, and no edge strong zero mode, even in the prethermal regime.

This fact appears nicely in the language of the unitary transformations used in this paper. The point is that the transformation guarantees only that the resulting $\hat{D}$ commutes with $\hat{N}$. It says nothing directly about the edge spin, which is why in the Majorana case we needed to argue that no terms involving edge-spin flips could appear in $\hat{D}$. Here they can. Such operators are easy to write out using the projectors

$$P^{(r)}_{j,j+1} = (1 + \omega^r \sigma^\dagger_j \sigma_{j+1} + \omega^{2r} \sigma_j \sigma^\dagger_{j+1})/3, \qquad (23)$$

satisfying $P^{(r)}_{j,j+1} P^{(s)}_{j,j+1} = \delta_{rs} P^{(r)}_{j,j+1}$ and $\tau_j P^{(r)}_{j,j+1} = P^{(r-1)}_{j,j+1}$. Since $\hat{N}_P$ can be written as a sum over $P^{(0)}$, it follows that

$$[\hat{N}_P, \tau_1 P^{(2)}_{1,2}] = 0,$$

so there is no obstacle to including $\tau_1 P^{(2)}_{1,2}$ into $\hat{D}$. Indeed, it does appear for Eq. (22) and for a generic Hamiltonian with the same dominant term $\hat{N}_P$. Since it shifts the edge spin, it rules out any strong zero mode. The Q-state Potts model with S_Q permutation symmetry is the obvious generalization of this to any integer $Q \geq 2$ and, except for the Ising case $Q = 2$, the same arguments apply: there is no edge strong zero mode for any finite J.

B. $\mathbb{Z}_3$ parafermions

Despite the results for the three-state Potts model, it is still possible to have prethermal edge zero modes of $\mathbb{Z}_3$ parafermions. One way of doing this is to follow Refs. [22,23] and to deform the dominant term $\hat{N}_P$ from Eq. (22) to

$$\hat{N}_\theta = \frac{1}{2\cos\theta} \sum_{j=1}^{L-1} (e^{-i\theta} \sigma^\dagger_j \sigma_{j+1} + e^{i\theta} \sigma_j \sigma^\dagger_{j+1}). \quad (24)$$

For θ not a multiple of $\pi/3$, this explicitly breaks spatial parity and time-reversal symmetries. It also breaks the S_3 permutation symmetry to $\mathbb{Z}_3$ and thereby breaks the symmetry between kinks and antikinks.

To utilize the ADHH theorem, we need to chose θ so that $\hat{N}_\theta$ has integer eigenvalues. The simplest nonzero value, $\theta = \pi/3$, is the Potts antiferromagnet. Here, the edge zero mode does not exist for similar reasons as described above for the ferromagnet. Thus, we choose $\theta = \pi/6$, halfway between ferromagnet and antiferromagnet. In this case, the energy of an antikink is twice that of a kink when $h = 0$,

and $L + \hat{N}_{\pi/6}$ counts the number of kinks plus *twice* the number of antikinks. Turning on h and applying the theorem means that the resulting $\hat{D}$ does not conserve the number of kinks individually, but allows scattering processes that convert one antikink to two kinks. The nice fact is that no such simple process flips the edge spin. For example, an antikink near the edge scatters into two kinks via the process

$$BAAAA... \Rightarrow BCAAA \cdots$$

that does not flip the edge spin. More formally, there is no local operator involving τ_1 inside $\hat{D}$ here, just as there is none involving σ^x_1 in the Majorana case. To prove this, we utilize the identity [59]

$$\omega^{L+\hat{N}_{\pi/6}} = \sigma^\dagger_1 \sigma_L. \qquad (25)$$

Hence, conservation of $\hat{N}_{\pi/6}$ implies that any term that does not couple the two ends of the chain must commute with σ_1 and σ_L individually.

The physics for other values of θ seems akin to the J_2 large case discussed next. For certain values of θ, $\hat{N}_\theta$ can be rescaled to have integer eigenvalues, but then edge spin-shifting terms can occur at some order [59].

C. Interacting Majorana chains with multiple large couplings

We now return to the interacting Majorana chain, but we allow $\hat{Y}$ to include terms that are not very small. We cannot apply the ADHH theorem to ensure a long decay time for the zero mode. We may be tempted to fix this by moving the offending terms from $\hat{Y}$ to $\hat{N}$. Naively, this fix would be valid as long as the extra terms commute with our putative zero mode and the eigenvalues of $\hat{N}$ remain integers. We show that similar considerations as for the Potts and clock models arise: the prethermalization theorem holds, but the resulting $U(1)$ symmetry need not protect the zero mode.

As a simple example, returning to our original Ising model in (8), let us allow the next-nearest neighbor coupling J_2 to be of the same magnitude as J. In order to apply the ADHH theorem, we must then include it in the dominant term $J\hat{N}$. This presents a potential problem, since the theorem requires integer eigenvalues of $\hat{N}$. However, if we take rational $J_2/J = p/q$ for coprime integers p and q, then

$$\hat{N}_{nn} = \sum_j (q \sigma^z_j \sigma^z_{j+1} + p \sigma^z_j \sigma^z_{j+2}) \qquad (26)$$

still has integer eigenvalues. The Hamiltonian becomes $\hat{H}_{nn} = -(J/q)\hat{N}_{nn} + \hat{Y}_{nn}$.

We may now use the ADHH theorem to obtain an approximate conservation law for $\hat{N}_{nn}$. However, $\hat{N}_{nn}$ no longer counts kink number; instead, it is the sum of the

number of broken nearest and next-nearest neighbor bonds, weighted appropriately by q and p. This means it is possible to flip the edge spin while conserving $\hat{N}_{nn}$ by converting broken bonds of one type to the other.

Equivalently, terms appear in $\hat{D}_{nn}$ that do not commute with the MZM γ_1^A. For example if $J = J_2$, then $\sigma_1^x(\sigma_2^z - \sigma_3^z)$ commutes with $\hat{N}_{nn}$ but flips the edge spin, ruining the conservation of the MZM. These are the resonances described in detail in Ref. [25]. They allow easy edge-spin flips by exchanging energy between different types of bonds. For example, consider the process that swaps between these two spin configurations, identical but for the edge spin:

$$\uparrow\uparrow\downarrow\cdots\Leftrightarrow\downarrow\uparrow\downarrow\cdots. \qquad (27)$$

The energy contribution from the three spins on the left is $2(J + J_2)$, while on the right it is $2(2J)$, so when $J = J_2$ the edge spin can be flipped for no energy cost.

This easy spin flip is analogous to a kink scattering off the edge into an antikink in the Potts model, but there is an important difference between the two. Here, there is an energy cost associated with either of the domain walls in the right configuration in Eq. (27) moving. Thus once a kink has moved to the edge and transformed from the left configuration to the right via an edge-spin flip, it is trapped at the edge, as opposed to the Potts case, where the newly produced antikink is not confined. Thus the only way the kink can move away from the edge is to transform back into the left configuration, reversing the edge-spin flip. This implies that, at low energy densities where there are few kinks, we should expect the MZM to retain a long lifetime, despite the resonances. We have confirmed this via the TEBD. It should be noted that an exception to this survival at low energy densities occurs at the critical point $J_2 = -J/2$, which, for any nonvanishing $\hat{Y}_{nn}$, becomes a paramagnetic regime between competing ferromagnetic and antiferromagnetic orders. This allows kinks to move freely from the edge in either configuration.

These results and the ADHH theorem shows that it is possible for the MZM to survive for long times even for large J_2/J, because the term in D with the edge-spin flip may occur only with some high power of $1/(Jq)$. In other words, the order in perturbation theory in which the resonance occurs may be some large value n_r. The time to decay is then of order e^{cn_r}. For general p and q, we expect that n_r is roughly $\max(p, q)$, in accordance with the analysis of Ref. [25]. It is worth noting that at $J = 0$ but nonzero J_2, the ensuing Hamiltonian model is equivalent to two copies of the Ising-Kitaev chain, and so $n_* \to \infty$ as $J_2/J \to \infty$ as well.

VI. TWO- AND THREE-DIMENSIONAL SYSTEMS

As we have emphasized throughout, the ADHH theorem applies in any dimension. In this section, we apply the general criteria developed in Sec. V to show how the resulting almost-conservation law can result in topological protection analogous to the one-dimensional examples we have analyzed.

A. Two dimensions

We study a perturbed toric code Hamiltonian [60] on a finite square lattice with sides of lengths L_1 and L_2. The spins live on the links i, with Hamiltonian

$$H = -u\left(\sum_v A_v + \sum_p B_p\right) + h^z\sum_i \sigma_i^z + h^x\sum_i \sigma_i^x + \dots,$$

$$(28)$$

where $A_v = \prod_{i\in\mathcal{N}(v)}\sigma_i^z$ and $B_p = \prod_{i\in p}\sigma_i^x$ for vertices v and plaquettes p. In the bulk, there are four spins on the links entering each vertex v and four spins on the links in each plaquette p. We put "rough" and "smooth" boundary conditions [61,62] on, respectively, the horizontal and vertical sides. At a rough edge, there are only three links around each edge plaquette; at a smooth edge, there are only three links entering each edge vertex. At the rough boundaries, B_p is modified so that it is the product of the three σ_i^x operators around a rough boundary plaquette; the vertex terms are unchanged since there are still four links attached to each vertex (we include no term for the "dangling" vertices touching only one link). At the smooth boundaries, A_v is modified so that it is the product of the three σ_i^z operators around each smooth boundary vertex; the nearby B_p are unchanged since each plaquette still contains four links. The … represents all other possible local terms, which are assumed to be small, including a term $\propto (\sum_v A_v - \sum_p B_p)$ which would give electric and magnetic charges different energies.

Such a system has a doubly degenerate ground state, which can be used as a qubit. One basis for this qubit is given by the eigenstates of the electric charge (modulo 2) on a rough edge. To make this more precise, consider the unperturbed toric code Hamiltonian [Eq. (28) with all couplings other than u set to zero]. Each term in the Hamiltonian is a projector plus a constant, and so ground states are annihilated by each individually. Now consider a path $\mathcal{P}$ of length P stretching along the lattice from one of the dangling vertices on one rough edge to a dangling vertex on the other edge labeled by consecutive links l_1, $l_2 \dots l_P$ and the operator $M_\mathcal{P} = \prod_{k=1}^P \sigma_{l_k}^x$. Then the eigenstates of the unperturbed toric code Hamiltonian can be grouped into eigenstates of $M_\mathcal{P}$ with eigenvalues ± 1. It is easy to check that this eigenvalue is independent of the choice of the path $\mathcal{P}$. This eigenvalue is the magnetic charge on either smooth boundary, with which $\mathcal{P}$ is roughly parallel. Likewise, the smooth edge corresponds to a rough edge on the dual lattice, and so we can define a path $\hat{\mathcal{P}}$ on

the dual lattice stretching from one smooth edge to the other. The electric charge operator is then defined as $E_{\hat{P}} = \prod_{k=1}^{P} \sigma_{l_k}^x$, and analogously to the magnetic charge, any eigenstate of the Hamiltonian can be grouped into eigenstates of $E_{\hat{P}}$ with eigenvalue ± 1. However, $E_{\hat{P}}$ and $M_{\mathcal{P}}$ anticommute, since $\mathcal{P}$ and $\hat{\mathcal{P}}$ always intersect. There are thus two ground states, not four, and the operators acting on this qubit can be identified as $Z \equiv E_{\hat{P}}$ and $X \equiv M_{\mathcal{P}}$.

Now consider the perturbed model, in which the other couplings are allowed to be nonzero. If h^x or h^z is large, the system undergoes a zero-temperature phase transition to a trivial phase; see Refs. [63,64] and references therein. However, if $h^x, h^z, \ldots$ are not too large, then the system remains in the zero-temperature topological phase, and the ground state is still doubly degenerate, up to corrections that are exponentially small in $\min(L_1, L_2)$. We can still associate the Z and X eigenstates with the eigenstates of electric and magnetic charge (modulo 2) on, respectively, the rough and smooth edges, but the Z and X Wilson lines, $E_{\hat{P}}$ and $M_{\mathcal{P}}$, need to be thickened since the ground state has fluctuations in which virtual pairs of e particles straddle $\hat{\mathcal{P}}$, and similarly with m particles and $\mathcal{P}$. The resulting thickened operators are not washed out by such quantum fluctuations in the ground state because virtual pairs never get too widely separated before recombining, so long as the system is in the topological phase. (They have an opportunity to become more and more widely separated from each other as the system approaches the quantum phase transition.)

However, conventional wisdom holds that this degeneracy is only a feature of the ground states and the excited states are not degenerate. Thermally excited pairs of quasiparticles can wander far from each other since they are real, not virtual, excitations. In particular, the naive expectation is that the qubit—meaning the boundary electric and magnetic charges—has a decay rate Γ of order $(L_1 + L_2)e^{-u/T}$ at nonzero temperature. Suppose, however, that $u \gg u_0$, where u_0 is built analogously to Eq. (9) from h^x, h^z, and all other terms lumped into the … remainder. The interesting question is then whether the prethermal conservation law implied by the ADHH theorem results in the qubit living longer than this naive expectation, analogous to the edge modes in 1D.

Because all A_v and B_p operators mutually commute, their sum has integer eigenvalues. The ADHH theorem thus implies that $\hat{N}_{TC} \equiv \sum_v A_v + \sum_p B_p$ is conserved, up to exponentially small corrections. This integer is zero in the ground state, and otherwise is simply the number of bulk excitations, not just their parity. To see if the approximate conservation of the number of bulk excitations protects the qubit, we need to define X and Z operators. The bulk Wilson line operators described above do not work because, even if the number of bulk excitations is

conserved, a bulk excitation can still cross a Wilson line and so flip its charge. However, we can define boundary operators that measure the rough-boundary electric charge and the smooth-boundary magnetic charge by simply pushing $\hat{\mathcal{P}}$ to either rough boundary and $\mathcal{P}$ to either smooth boundary,

$$Z = \mathcal{U}^{\dagger} \left(\prod_{i \in HB} \sigma_i^z \right) \mathcal{U}, \qquad X = \mathcal{U}^{\dagger} \left(\prod_{i \in VB} \sigma_i^x \right) \mathcal{U}, \quad (29)$$

where $\mathcal{U}$ is the unitary given by ADHH, HB is the set of vertical links that belong to the horizontal rough boundary at the top of the rectangle (this choice is arbitrary—the bottom would work equally well), and VB is the set of vertical links that belong to the vertical smooth boundary at the left side of the rectangle.

The transformed edge Wilson-line operators in Eq. (29) commute with the Hamiltonian, up to $O(e^{-cn_*})$ corrections, except at the corners. Away from the corners, D contains only terms commuting with Z, because a noncommuting term would necessarily create or annihilate an e particle in the bulk, thereby violating the conservation of $\hat{N}_{TC}$. Similarly, terms in D that do not commute with X would create an m particle. However, at the corners, D can contain terms that cause the rough edge to absorb an e particle and the smooth edge to emit an m particle or vice versa. One example is $\sigma_4^z \sigma_3^z \sigma_1^x (1 + B_{p_{2345}})(1 + A_{v_{123}})$, where the dangling link is labeled by 1; v_{123} is the corner vertex connected to links 1,2,3; and p_{2345} is the corner plaquette that overlaps it and contains links 2,3,4,5. An error-causing process associated with this term is depicted in Fig. 4.

We thus arrive at the interesting result that the only possible violations of the long-lived conservation law occur at the corners. Prethermalization therefore suppresses the low-temperature error rate from $\Gamma \propto (L_1 + L_2)e^{-u/T}$ to

$$\Gamma \propto e^{-cn_*}(L_1 + L_2)e^{-(u/u_0)u/T} + e^{-u/T}$$

for some constant c. It would be interesting to see how this argument generalizes to weakly perturbed Levin-Wen models [65].

B. Three dimensions

There are four dimensional topological phases with no pointlike excitations, and they protect quantum information at nonzero temperatures below the phase transition into the topological phase, as noted in Sec. II. In three dimensions, we are halfway there, since magnetic excitations are looplike; their line tension prevents them from causing errors. However, there are still pointlike excitations that cause errors at nonzero temperature. Prethermalization can suppress them.

It was recently shown [66] that Abelian topological phases in $3 + 1$ dimensions have loop excitations carrying
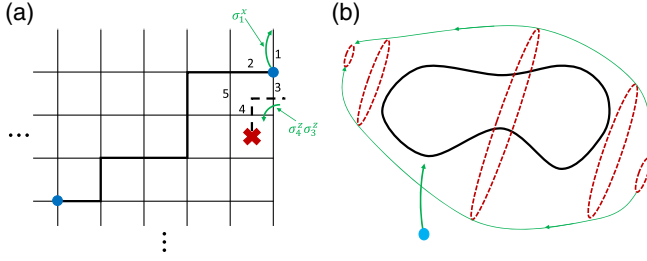
(a)

(b)

FIG. 4.   (a) Prethermalization does not protect against qubit errors at the boundaries between "rough" and "smooth" edges of the toric code on a surface with boundary. Here, an e particle (blue dot) is absorbed by the rough boundary (top edge), while an m particle (red cross) is emitted by the neighboring smooth boundary (right edge), as described in the text. (b) When a qubit is encoded in a loop that carries "Cheshire charge" in 3D, errors associated with a magnetic flux loop encircling (sequence of dotted red lines, whose growth and subsequent shrinkage are indicated by the green arrows) the Cheshire loop are suppressed by the size of the Cheshire loop, since the magnetic flux loops have a line tension at temperatures below the bulk phase transition temperature. There is no such protection against the emission or absorption of a pointlike electric charge (blue circles), even below the phase transition temperature into the topological phase. However, prethermalization can suppress such processes exponentially.

Cheshire charge. The simplest example is the 3D toric code, which has a Hamiltonian of the same form as Eq. (28) generalized to the cubic lattice, with one modification discussed below. There are six links attached to each vertex, so A_v is a product of six σ_i^z operators while B_p is still the product of four σ_i^x operators around each plaquette. There are loop excitations on which e particles condense in a manner analogous to the "rough" boundaries considered above in the case of the 2D toric code. To analyze such loop excitations, we can modify the Hamiltonian so that the ground state has one at a specified loop K, as was done in Ref. [66]. We label the loop K, of length L, in terms of vertices v_k and links l_k such that the vertices v_{k+1} and v_k are connected by the link l_k for all $k = 1 \dots L$ with $v_{K+1} = v_1$. We now modify the Hamiltonian along this loop according to

$$\sum_{k=1}^{K} A_{v_k} \rightarrow -\sum_{k=1}^{K} \sigma_{l_k}^x. \tag{30}$$

This transverse field commutes with the B_p terms, so eigenstates of the Hamiltonian are eigenstates of $\sigma_{l_k}^x$ for every link l_{k+1} on the loop. Thus, they resonate between a state with two e particles at either end of the link and a state without e particles at the ends of the link. Alternatively, we can insert such a loop excitation in a manner that emphasizes the similarity with the "rough" edge in 2D: we remove all v_k, $l_k \in K$ from the lattice. The Hamiltonian is unchanged at all of the remaining vertices and the plaquette

operator is modified to a product of three spins for the plaquettes that previously contained $l_k \in K$. In this alternative construction, it is again clear that the states of the system resonate between having zero and two e particles at either end of each link on K. The spectrum is degenerate in the limit of a large loop, provided that there are at least two such loops in the system, since the total charge on a loop can be either 1 or e, subject to the constraint that the total topological charge of the system is fixed. In the simplest case, in which there are two such loops in the system, states are doubly degenerate and the degenerate subspace at each energy forms a qubit spanned by states with charge 1 or e on both loops. This charge is locally unobservable; hence it is "Cheshire charge," which explains why the two states are degenerate in the limit of a large loop.

A straightforward generalization of the arguments applied in two dimensions in Sec. VI A shows that Cheshire-charge-carrying loops cannot emit or absorb an e particle in the analogous prethermal regime in 3D. Hence, the Z operator acting on the qubit commutes with the prethermal Hamiltonian $\hat{D}$. A dressed version of the X operator, meanwhile, is expected to be conserved in the low-temperature phase $T < T_c$ below the phase transition at which long flux loops unbind and proliferate [67]. Thus, the qubit is partially protected by the dynamics of the low-temperature phase (as was already known [53]) and partially protected by prethermalization. Unlike in the 2D case discussed above, our 3D topological qubit has error rate $\Gamma < e^{-cn_*}e^{-(u/u_0)u/T}$, where u_0 is the appropriate energy scale derived from the couplings in $\hat{Y}$. The corner error processes that caused trouble in 2D are not present here because an electric charge cannot become a magnetic charge, since one is pointlike and the other looplike, unlike in 2D, where both are pointlike and necessarily have the same energy in the prethermal limit. In other words, in the prethermal regime, the 3D toric code is, up to exponentially small corrections, a self-correcting nonzero temperature quantum memory; the only other known examples are 4D topological phases [53]. Thus, in this case, prethermalization buys us an extra dimension, which might be rather difficult to otherwise acquire.

VII. EXPERIMENTAL REALIZATIONS

Although the primary focus of this paper is the conceptual advance in the theory of topological phases of matter that results when the theory of prethermalization is brought to bear on it, it is important to note that this advance may have implications for near-term devices and experiments. For illustrative purposes, we discuss two examples: Majorana zero modes in semiconductor-superconductor devices and Ising spins in a trapped ion chain. The former example allows us to introduce an important point, which is that prethermalization is generically terminated by a coupling to a heat bath; when the coupling is

weak, this occurs at very late times, and the prethermal regime extends until then. Since the electron-phonon interaction is suppressed by the ratio of the electron mass to the ion mass (it must vanish in the limit that the latter approaches infinity), this can be weak, thus giving an example of this scenario. Meanwhile, the ion chain has interactions that are long ranged, so it is not immediately obvious that our analysis applies. However, the calculations that we present do, in fact, indicate that prethermalization occurs in this case as well. Although this system does not have a topological phase, the edge spins are governed by similar dynamics. We present a quantitative discussion of the experimental requirements for observing edge spins protected by prethermalization.

A. Quantum-dot chains and the fate of prethermalization in the presence of electron-phonon interactions

The authors of Refs. [40,41] have proposed devices in which quantum dots (possibly defined by gates acting on MZM-supporting nanowires) are the basic building blocks for superlattices that have an effective low-energy description as models of Majorana fermions hopping on a lattice. They therefore realize literal versions of the Hamiltonian of Eqs. (6) and (8). Moreover, Ref. [41] gives a procedure for tuning a system of quantum dots to the limit in which the coupling J is much larger than all the other coupling constants from Eq. (8). In the standard notation for the Kitaev chain [15], this translates into making $t + \Delta$ much larger than μ and $t - \Delta$. In other words, Ref. [41] gives a procedure for tuning into the prethermal regime of large J/J_0. These devices are a bit futuristic, so we are not, in this section, proposing immediate experimental tests of prethermalization-protected MZMs. Rather, our purpose here is to illustrate the considerations that must enter into any effort to exploit prethermalization in a solid.

As we discuss momentarily, prethermalization leads us to predict extremely small decay rates for the MZMs at the ends of such a chain. However, these extremely small decay rates are not observed since the electron-phonon coupling ends prethermalization before these exponentially small effects do. In other words, the electron-phonon coupling can violate the conservation of N before the time t_*. But if the electron-phonon contribution to thermalization is smaller than the electron-electron contribution over some range of temperatures, device geometries, and device parameters, then prethermalization still plays an important role in extending the lifetime of a topological qubit. We show that the naive MZM decay rate attributable to electron-electron interactions, which is what would be seen outside the prethermal regime, is $\Gamma_{\text{el-el}}^{\text{non-pt}} \sim n_{\text{qp}}^2 \cdot 0.6$ GHz. Meanwhile, the decay rate attributable to electron-phonon interactions is $\Gamma_{\text{el-ph}} < n_{\text{qp}} \cdot 10$ MHz, which is almost certainly a gross overestimate since it does not take into account the screening of piezoelectric interactions by the superconductor or the effect of the device geometry, which can be designed to have a phonon band gap. Since the former is quadratic in n_{qp}, while the latter is linear, we conclude that, even for this overestimate of phonon effects, the electron-electron interactions are the dominant decay channel for $n_{\text{qp}} > 0.01$, and prethermalization suppresses this dominant decay channel. If the bulk quasiparticle density is thermal (which may or may not be a valid assumption, for the reasons discussed in Ref. [55] and Refs. [1–7] of that paper), this translates into a range of temperatures, $(t + \Delta)/T < 4.6$. Again, this range could be greatly increased, depending on the effects of device geometry on the phonon spectrum. In the rest of this subsection, we explain these estimates in more detail.

The initial motivation for considering such a system was that it may be possible to tune parts of the system more reliably into and out of the topological superconducting phase. The potential drawback of such systems is that the energy gap of the coupled-MZM system is significantly smaller than the energy gap of a single nanowire. For instance, Ref. [41] finds $t \approx 9$ μV and $\Delta \approx 6$ μV. One might consequently fear that much lower temperatures would be necessary in order to protect the end MZMs of such a quantum-dot chain.

Indeed, suppose that the hopping parameter were twice as large, in which case the system would be outside the prethermal regime, with $(t + \Delta)/2 = 7.5$ μV and $(t - \Delta)/2 = 1.5$ μV, and assume a nearest-neighbor repulsion $V = 3$ μV. Then, a rough naive estimate for the decay rate due to electron-electron interactions would be $\Gamma_{\text{el-el}}^{\text{non-pt}} \sim n_{\text{qp}}^2 \cdot V^2/(t - \Delta)$, since $(t - \Delta)$ is the bandwidth of excited quasiparticles (so its inverse is the density of states). This gives $\Gamma_{\text{el-el}}^{\text{non-pt}} \sim n_{\text{qp}}^2 \cdot 0.6$ GHz. Thus, the zero mode decays in a few nanoseconds unless the quasiparticle density is very low, which is unlikely to be the case since the gap is relatively small.

However, the results of this paper show that the small gap may not doom the MZM: prethermalization protects quantum information until a time that is nearly exponentially long in the ratio between $t + \Delta$ and some combination of the other couplings, such as $t - \Delta$, V, etc., according to Eq. (20). Moreover, the temperature dependence has a characteristic energy scale $n_*(t + \Delta)$, rather than $(t + \Delta)$ itself, so relatively small $(t + \Delta)$ is not as detrimental as one might fear. As a result of prethermalization, the decay rate due to electron-electron interactions should, instead, be $\Gamma_{\text{el-el}}^{\text{pt}} \sim (n_{\text{qp}})^{m_{\min}} \cdot e^{-cn_*} \cdot V^2/(t - \Delta)$, as in Eq. (19). In other words, prethermalization suppresses the decay rate attributable to electron-electron interactions by a factor $\Gamma_{\text{el-el}}^{\text{pt}}/\Gamma_{\text{el-el}}^{\text{non-pt}} \sim (n_{\text{qp}})^{m_{\min}-2} e^{-cn_*}$. For the values of t, Δ given above, $(t + \Delta)/(t - \Delta) \approx 5$, and assuming that all other couplings are smaller than $t - \Delta$, we find $n_* \approx m_{\min} \approx 5$. If we replace the $O(1)$ constant c by 1, then we can use $e^{-n_*} \approx 6 \times 10^{-3}$. Thus, we have $\Gamma_{\text{el-el}}^{\text{pt}} \sim n_{\text{qp}}^5 \cdot 3.6$ MHz.

If we assume that the system is at a temperature $T = 50$ mK, and that the quasiparticle density is given by the equilibrium value, then $n_{qp} \approx e^{-(t+\Delta)/T} \approx 0.05$, while $(n_{qp})^5 \approx e^{-\Delta_{eff}/T} = e^{-m_{min}(t+\Delta)/T} \approx 3 \times 10^{-7}$. (We are assuming that, in the initial state, the quasiparticle density is equal to its equilibrium value but its subsequent evolution is impeded by prethermalization, especially its equilibration with the edge modes.) Thus, the enhancement of $(t + \Delta)$ to $\Delta_{eff} = m_{min}(t + \Delta)$ is the larger effect at these temperatures, assuming that the system equilibrates rapidly with respect to $\hat{D}$, i.e., is in the prethermal state. The resulting decay rate is $\Gamma_{el\text{-}el}^{pt} \sim n_{qp}^5 \cdot 3.6$ MHz ~ 1 Hz. However, if the system is out of equilibrium, the factor of e^{-cn_*} may be more important since it protects the zero modes even if there are nonequilibrium excitations in the bulk that would render the equilibrium estimate $(n_{qp})^{m_{min}} e^{-\Delta_{eff}/T}$ moot.

We must compare the prethermal decay rate to the phonon-assisted decay rate $\Gamma_{el\text{-}ph}$. A bulk fermionic excitation, which has energy $E = t + \Delta$, can be absorbed by a zero mode and its energy can be emitted as a phonon of momentum $q = E/v$, where v is the speed of sound. The Hamiltonian governing the electron-phonon interaction is

$$H_{el\text{-}ph} = \int d^3x \, d^3x' \rho_{el}(x) V_{ij}(x - x') \partial_i u_j(x'). \quad (31)$$

Here, $\rho_{el}(x) = i\gamma_1^A \gamma_1^B |\psi_1(x)|^2$, where $\psi_1(x)$ is the wave function of the zero-energy fermionic level of a single dot coupled to a superconductor [41], $u_j(x)$ is the displacement in the j-direction of the ion whose equilibrium position is x, and $V_{ij}(x - x') = D\delta(x - x')\delta_{ij} + eh_{14} w_{ij}(x - x')$. The electron-phonon coupling has two parts, the deformation potential D and the piezoelectric coupling h_{14}. The piezoelectric potential satisfies $q_i w_{ij}(q) = \sum_\lambda i M_\lambda(q)(\epsilon_q^\lambda)_j$, where λ are the phonon polarizations, ϵ_q^λ are the corresponding polarization unit vectors, and $M_\lambda(q)$ depend on the direction of q but not its overall scale. Hence, the decay rate for a MZM attributable to the deformation potential electron-phonon coupling is

$$\Gamma_{el\text{-}ph}^{DP} = \int \frac{d^3q}{(2\pi)^3} |Q(q)|^2 (Dq)^2 \frac{1}{\rho} \delta(E^2 - v_l^2 q^2) n_{qp}$$
$$< \frac{1}{4\pi^2 \rho v_l E} \left(D\frac{E}{v_l}\right)^2 \left(\frac{E}{v_l}\right)^2 n_{qp}. \quad (32)$$

Here, ρ is the density of the solid and, as before, n_{qp} is the probability of a bulk fermionic excitation on dot 1. In going to the second line, we have bounded $Q(q) \equiv \int d^3x e^{iq \cdot x} |\psi_1(x)|^2$ by $|Q(q)|^2 < 1$. The first two factors in the second line are the matrix element for such a process; the third factor in the second line is the density of states for the phonon, which is $\propto q^2$; and the final factor is the

probability for a bulk quasiparticle excitation to be near enough to the MZM for absorption to occur. (We emphasize that the relevant gap for phonon-assisted decay of a MZM is $E = t + \Delta$, not Δ_{eff}, which is the relevant scale for electron-electron interactions in the prethermal regime.) The reverse process, in which a bulk quasiparticle excitation is emitted and a phonon is absorbed, has the same amplitude at low temperature. Prethermalization occurs if $\Gamma_{el\text{-}ph} < \Gamma_{el\text{-}el}$. For InAs we take the following values [68]: $D = 5.1$ eV; the speed of longitudinal sound waves is $v_l \approx 4.7$ km/s; the density is $\rho \approx 5.67$ g/cm^3. We take $t + \Delta \approx 15$ μV estimated in Ref. [41], as in our discussion of the prethermal decay rate attributable to electron-electron interactions. This gives $\Gamma_{el\text{-}ph}^{DP} < n_{qp} \cdot 300$ kHz. For an equilibrium distribution of excited quasiparticles at $T = 50$ mK, this gives $\Gamma_{el\text{-}ph}^{DP} < 15$ kHz. Turning now to the piezoelectric coupling, we first note that, in the presence of strong coupling to superconducting leads, this effect may be suppressed by screening. However, if we neglect this screening effect and compute, as an upper bound, the decay rate due to an unscreened piezoelectric coupling, we find

$$\Gamma_{el\text{-}ph}^{PE} < \int \frac{d^3q}{(2\pi)^3} |Q(q)|^2 (eh_{14})^2 \frac{1}{\rho} \delta(E^2 - v^2 q^2) n_{qp}$$
$$\sim \frac{1}{4\pi^2 \rho v E} (eh_{14})^2 \left(\frac{E}{v}\right)^2 n_{qp}. \quad (33)$$

Here, we have made the approximation of ignoring the difference between the longitudinal and transverse sound velocities $v_l \approx 4.7$ km/s and $v_t \approx 3.3$ km/s and simply set them both to $v \equiv 4.2$ km/s. We have also made the simplification of replacing $M_\lambda(q)$ by an upper bound $M_\lambda(q) < 1$. Using $h_{14} = 3.5 \times 10^6$ V/cm, given in Ref. [69], we find $\Gamma_{el\text{-}ph}^{PE} \sim n_{qp} \cdot 10$ MHz. For an equilibrium distribution of excited quasiparticles at $T = 50$ mK, this gives $\Gamma_{el\text{-}ph}^{PE} \sim 500$ kHz. We note that this estimate does not take into account the effect of phonons in the superconductor, although we expect this to be a smaller contribution to the decay rate since the electronic wave function is concentrated primarily in the semiconductor; it has also not taken into account the effect of the device geometry on the phonon spectrum at the wavelengths of interest. In this regard, we note that it may be possible to pattern a material in order to engineer a phonon band gap at the wavelength $2\pi k^{-1} = hv/E \approx 1$ μm, potentially strongly suppressing the effect of phonons. Thus, even at relatively high temperatures, 1 μs is a conservative estimate of the potential lifetime of a MZM in a quantum-dot chain in the prethermal regime, but the lifetime may be as long as 1 ms.

We briefly note that much of our discussion of quantum-dot chains in Sec. VII A also applies to the model of Ref. [70], which uses short topological nanowires to

construct a two-dimensional model of Ising anyons on the honeycomb lattice [71]. The Ising anyons emerge as low-energy excitations of a superlattice of Coulomb-blockaded islands, each containing two nanowires. Our results do not apply to the physics of the nanowires themselves, but instead to the effective model of the low-energy degrees of freedom, which has a prethermal regime in the limit that the bulk Majorana fermion operators have a flat band. The existence of such a prethermal regime would facilitate universal topological quantum computation using the strategy of Ref. [70], since it ameliorates the drawback of a reduced energy gap.

B. Trapped atomic chains

Another possible experimental realization would be a trapped ion or neutral atom chain governed by a perturbed transverse-field Ising model [72,73]. Here, coupling to an external heat bath would be less of a concern, although the effective system size might be smaller than in the quantum-dot case.

For example, in Ref. [74], the authors use chains of up to 22 ^{171}Yb$^+$ ions in linear radiofrequency (Paul) traps, encoding effective two-state systems in their $^2S_{1/2}$ hyperfine ground states. Long-range spin-spin interactions are generated using laser-mediated spin-phonon interactions. In particular, using the beatnote between two overlapped laser beams to drive stimulated Raman transitions, they generate the effective Hamiltonian

$$\hat{H} = \sum_{i<j} J_{i,j}\sigma_i^z\sigma_j^z + B\sum_i \sigma_i^x, \qquad (34)$$

where interaction is long ranged and antiferromagnetic:

$$J_{i,j} = \frac{J_I}{|i-j|^\alpha}, \qquad (35)$$

with $J_I > 0$. For nearest-neighbor interactions in Ising, ferromagnet and antiferromagnet are unitarily equivalent. Here, the distinction is important, because the ferromagnet has a phase with long-ranged order for $\alpha < 2$ and nonzero temperatures less than some critical temperature T_c [75], while the antiferromagnet does not have such an ordered phase for $T > 0$, like the nearest-neighbor model [76]. Consequently, for initial states that are near the *top* of the spectrum here, the end and bulk spin lifetimes are infinite since these are low-energy states of the ferromagnetic Hamiltonian $-H$.

In the setup of Ref. [74], the experimentally realizable range of α is $0.5 < \alpha < 2$, while $J_I/2\pi \leq 1$ kHz, achieved by changing the trap voltages and the detuning of the beatnote from resonance. The B term is generated by driving further resonant stimulated Raman transitions out of phase with the beatnote. It can range from negligible to a
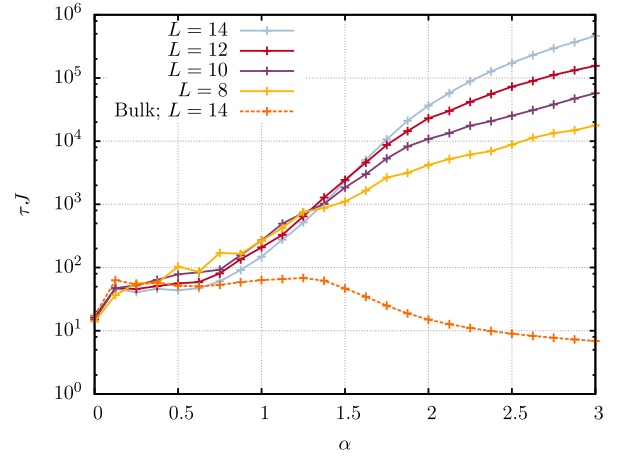

FIG. 5. The decay time (on a logarithmic scale) of a MZM in the transverse-field Ising model with long-ranged interactions at $T = \infty$ for $L = 8$–14 sites and $h = 0.2$, as obtained by exact diagonalization. It is plotted as a function of α, the power of the decay of the long-ranged interaction. As in Fig. 2, the decay time saturates in L for larger values of the perturbing couplings. The decay times for ferromagnetic and antiferromagnetic interactions are the same at $T = \infty$; it is only at lower temperatures where the effect of the long-ranged order for $\alpha < 2$ in the ferromagnetic case becomes important.

maximum of $B/2\pi = 10$ kHz. So it should certainly be feasible to enter the prethermal regime $B \ll J_I$.

The autocorrelators of individual spins $\langle\sigma_j^z(t)\sigma_j^z(0)\rangle$ may be measured up to times of order $100/J_I$. Thus, it should be possible to observe the prethermal protection of the edge spin through the survival of its autocorrelator, in contrast with the bulk spins, which would decay over experimentally accessible time scales. The one caveat is that, although the edge spin is long lived for any system temperature or any energy initial state, for very high-energy initial states near the top of the spectrum, or for negative temperatures, the protection comes from the long-ranged ferromagnetic order when $\alpha < 2$, rather than prethermalization. This case is easily distinguishable because the bulk spins are also long lived.

Simulations with exact diagonalization (ED) at infinite temperature on similar system sizes confirm this picture, at least for $\alpha \gtrsim 1.25$; see Fig. 5. The main theoretical concern for prethermalization is the long-ranged nature of the interaction, because for the ADHH theorem to hold we require small J_0/J. In fact, the ADHH theorem has not yet been proven for power-law decay interactions. In our discussion above, $\hat{N}$ consists of just the nearest-neighbor interaction magnitude J_I. For $\alpha = 2$, the shortest-possible range in the experiments, the next-largest term is a quarter of the size, so we might be justified in putting it and longer-range terms in $\hat{Y}$. However, for smaller α, we should include at least the next-nearest-neighbor term in $\hat{N}$ as well. This case, including the possibility of resonances, is discussed in detail in Sec. V C below. Crucially, for there to be any

chance for the ADHH theorem to hold, $\hat{N}$ must have integer eigenvalues, which is of course impossible to tune exactly experimentally for more than one coupling. However, for the system sizes $L \leq 22$ that are experimentally accessible, we do not expect this to be of practical issue in observing the prethermal protection of the edge spin in contrast to the bulk, and indeed the ED results support this.

Of course, once α becomes small, the arguments above based on locality break down completely. The two ends of the chain come into contact, so that terms that flip the edge spin may be immediately added to $\hat{D}$, and the edge spin is no longer protected, regardless of the applicability of the ADHH theorem. It is also worth noting that the smaller α is, the less localized at the edge the zero mode is, and so the greater its overlap with the bulk spins. This means that there may be some small but observable part of the bulk spin correlator that survives to long times, albeit exponentially suppressed in magnitude compared to the edge.

VIII. INTEGRABLE SYSTEMS

Turning now to more formal considerations, we emphasize that the number n_* following from the ADHH theorem only defines a lower bound on how long a zero mode lives. Indeed, in a free-fermion system, the recursive procedure defined in Ref. [31] converges, so that $n_* \to \infty$ as $L \to \infty$, as we now show explicitly. In fact, one might go so far as to say that the ADHH recursive procedure gives us an alternate formulation of Onsager's solution of the Ising model.

The strong MZMs and corresponding unitary operators in the free-fermion Ising-Kitaev chain can be found directly, with no need for the full-blown ADHH procedure. The Hamiltonian here is $H = -J\hat{N} + \hat{Y}$ with $\hat{N}$ given by (6) and $\hat{Y}$ given by (8) with all couplings other than h and J set to zero. The strong MZMs at the left and right edges, respectively, are then given by [15]

$$\Psi = \mathcal{N} \sum_{j=1}^{L} \left(\frac{h}{J}\right)^j \gamma_j^A, \qquad \Psi' = \mathcal{N} \sum_{j=1}^{L} \left(\frac{h}{J}\right)^j \gamma_{L+1-j}^B.$$

where the normalization $\mathcal{N}$ is chosen to make $\Psi^2 = (\Psi')^2 = 1$. Each of these commutes with the Hamiltonian, up to terms of order $(h/J)^L$. Thus, n_* for the edge modes is L throughout the ordered phase $h < J$. The unitary operator $\mathcal{U}_e$ relates the strong edge MZM at arbitrary h to that at $h = 0$ (the latter commuting with $\hat{N}$). We have

$$\Psi = \mathcal{U}_e^{\dagger} \gamma_1^A \mathcal{U}_e, \qquad (36)$$

and we write

$$\mathcal{U}_e = \exp(i[U_0 + U_1 + U_2 + \ldots]), \qquad (37)$$

where U_n is Hermitian and of order $(h/J)^n$. If we insist that Ψ commute with the Hamiltonian $\hat{H}$ up to order $(h/J)^n$, as per Sec. III, then we may calculate U_n by inverting the equation:

$$[\hat{H}, [U_n, \gamma_1^A]] = i[\hat{H}, e^{-i\sum_{j=0}^{n-1} U_j} \gamma_1^A e^{i\sum_{j=0}^{n-1} U_j}]. \quad (38)$$

For the Ising-Kitaev chain, we may calculate this boundary unitary transform exactly:

$$\mathcal{U}_e = \cos\frac{\theta}{2} + \cos\theta \sin\frac{\theta}{2} \gamma_1^A \sum_{j=2}^{L} \left(\frac{h}{J}\right)^{j-2} \gamma_j^A, \quad (39)$$

where $\sin\theta = h/J$.

Of course the $\mathcal{U}_e$ constructed in this manner is not the full unitary transform shown to exist by the ADHH theorem, and the correspondingly transformed Hamiltonian would not display the emergent $U(1)$ symmetry. Nevertheless this "boundary" unitary transform is sufficient to show conservation of the edge mode in the prethermal regime. Namely, if we combine the boundary unitaries from both ends, then the transformed Hamiltonian conserves the $\mathbb{Z}_2$ bulk fermion parity $(\mathcal{U}_e \mathcal{U}_e')^{\dagger} \prod_{j=1}^{L-1} \gamma_j^B \gamma_{j+1}^A \mathcal{U}_e \mathcal{U}_e' = (\mathcal{U}_e \mathcal{U}_e')^{\dagger} \sigma_1^z \sigma_L^z \mathcal{U}_e \mathcal{U}_e'$ up to order $n_* = L$. As discussed in Sec. IV A, this approximate conservation law is all that is needed.

Breaking the integrability by including nonzero h_2 and/or J_2 in Eq. (8), we have calculated $\mathcal{U}_e$ up to eleventh order using a computer-aided algebra program. The resulting edge zero modes agree with those calculated explicitly in Ref. [25]. For example, when the only perturbing terms in $\hat{Y}$ are h and h_2, we find to second order that

$$\mathcal{U}_e = \exp\left(\frac{1}{2}[h\sigma_1^y\sigma_2^z + h_2\sigma_1^x\sigma_2^y\sigma_3^z + h^2\sigma_1^y\sigma_2^x\sigma_3^z \right.$$
$$\left. + hh_2\sigma_1^y(\sigma_3^z - \sigma_2^y\sigma_3^y\sigma_4^z) - h_2^2\sigma_1^x\sigma_2^y\sigma_3^y\sigma_4^y\sigma_5^z]\right).$$

Furthermore, this method of calculating the edge zero modes is preferable and more efficient than the method outlined in Ref. [25], because the edge zero modes are automatically normalized at each order by construction.

It is also illuminating to implement the ADHH procedure directly and explicitly on the full bulk Hamiltonian. We show that for the free-fermion Ising-Kitaev case, the unitary transformation can be computed exactly. For simplicity, here we take $L \to \infty$ so the Hamiltonian of the Ising-Kitaev chain is

$$H_{\text{IK}} = -\sum_j (J\sigma_j^z\sigma_{j+1}^z + h\sigma_j^x). \qquad (40)$$

We assume that J is sufficiently large compared to h so that we can apply ADHH with

$$\hat{N}_{\text{IK}} = \sum_j \sigma_i^z \sigma_{i+1}^z, \qquad \hat{Y}_{\text{IK}} = -h \sum_j \sigma_j^x. \qquad (41)$$

We rewrite the Hamiltonian in the form

$$H_{\text{IK}} = -J\hat{N}_{\text{IK}} - h\hat{D}_0 - i\frac{h}{2}\sum_j (\gamma_j^A \gamma_j^B + \gamma_{j-1}^B \gamma_{j+1}^A),$$

so that the last term is the "error" term $\hat{E}$, which does not commute with $\hat{N}_{\text{IK}}$, and

$$\hat{D}_0 \equiv \frac{1}{2}\sum_j \sigma_j^x (1 - \sigma_{j-1}^z \sigma_{j+1}^z) = \frac{i}{2}\sum_j (\gamma_j^A \gamma_j^B - \gamma_{j-1}^B \gamma_{j+1}^A)$$

commutes with $\hat{N}_{\text{IK}}$. This can be checked explicitly; the key observation is that if there is a single kink at either $j - 1/2$ or $j + 1/2$, flipping the spin at site j hops the kink, while with no kink or two kinks adjacent to j, flipping the spin creates or annihilates two kinks, respectively. The former process conserves $\hat{N}_{\text{IK}}$, while the latter does not, so the former is allowed in $\hat{D}_{\text{IK}}$. Acting with the operator $(1 - \sigma_{j-1}^z \sigma_{j+1}^z)/2$ annihilates any configuration with zero or two kinks adjacent to j, and gives the identity if there is a single one. Thus we arrive at the expression given above for $\hat{D}_0$.

At each step of the ADHH recursive procedure, new terms are included in $\hat{D}$ and the coefficient of the error term is reduced by a power of h/J. At the very first recursive step, the error term can be canceled by using the relation

$$[\hat{N}_{\text{IK}}, \gamma_j^A \gamma_{j+1}^A] = -[\hat{N}_{\text{IK}}, \gamma_{j-1}^B \gamma_j^B] = 2(\gamma_j^A \gamma_j^B + \gamma_{j-1}^B \gamma_{j+1}^A).$$

For simplicity, we impose periodic boundary conditions and define

$$G_n = \frac{1}{4}\sum_j (\gamma_j^A \gamma_{j+n}^A - \gamma_j^B \gamma_{j+n}^B). \qquad (42)$$

Then, after the first step in the recursive procedure, the transformed Hamiltonian has the form given on the right-hand side of the following equation:

$$\left(1 - \frac{h}{2J}G_1\right)H_{\text{IK}}\left(1 + \frac{h}{2J}G_1\right)$$

$$= -J\hat{N}_{\text{IK}} - h\hat{D}_0 + O\left(\frac{h^2}{J}\right).$$

Thus all terms in the transformed Hamiltonian that do not commute with $\hat{N}_{\text{IK}}$ are at least of order h^2/J. The terms of this order can be split into pieces that do commute with $\hat{N}_{\text{IK}}$, which then comprise $\hat{D}_1$, and those that do not, which comprise another error term. This procedure can be repeated, so that an order $(h/J)^2$ term can be added to $\mathcal{U}$ to yield an error term of order $h(h/J)^2$.

The ADHH theorem guarantees that this procedure can be repeated, at least up to order n_*. Not surprisingly, the procedure can be implemented to all orders in a free-fermion system and so n_* and hence t_* become infinite as $L \to \infty$ when $h < J$ in the Ising-Kitaev chain. This method was, in essence, how Onsager originally computed the free energy of the two-dimensional Ising model. His original calculations [77] are manipulations of fermion bilinears, just as is required to find the unitary transformation $\mathcal{U}$. We find

$$\mathcal{U} = \exp\left(-\sum_{n=1}^{\infty} \frac{1}{2n}\left(\frac{h}{J}\right)^n G_n\right), \qquad (43)$$

where the fermion bilinears G_n are defined in Eq. (42).

The key to deriving Eq. (43) is to utilize the *Onsager algebra*, as derived in the original paper [77]. This is simply the algebra of fermion bilinears, and a quick glance at the paper shows that Onsager defines them in terms of the conventional Jordan-Wigner expressions (given the clarity of his expression, one wonders why he missed defining individual fermions). The generators of the algebra are then given by the G_n in Eq. (42) and

$$A_m = -i\sum_j \gamma_j^B \gamma_{j+1-m}^A. \qquad (44)$$

The Hamiltonian of the Ising-Kitaev chain $H_{\text{IK}} = -J\hat{N} + \hat{Y}_{\text{IK}}$ is then written in terms of these generators as $\hat{N}_{\text{IK}} = -A_0$ and $\hat{Y}_{\text{IK}} = hA_1$. Onsager carefully works out the effect of periodic and antiperiodic boundary conditions, but we simplify matters here by taking $L \to \infty$. It is then easy to work out the Onsager algebra

$$[G_n, G_m] = 0$$
$$[G_n, A_m] = A_{m+n} - A_{m-n}$$
$$[A_n, A_m] = 8G_{n-m} \qquad (45)$$

for all integer n and m. These are exactly Eqs. (61a), (61), and (60) of Ref. [77], with a rescaling of the generators G_n.

A key property of the Onsager algebra is that when G_n is commuted with any linear combination of the A_m, the result remains linear in the A_m. Moreover, because by definition $G_{-n} = -G_n$, a series of quantities preserving the $U(1)$ symmetry is given by $A_m + A_{-m}$. This suggests then building the unitary transformation out of the G_n, a task made even easier by the fact that they commute among themselves. Such a construction is done straightforwardly by Fourier-transforming the A_m as

$$\tilde{A}(k) \equiv \sum_{m=-\infty}^{\infty} e^{-ikm} A_m, \qquad (46)$$

where k takes values in the Brillouin zone $-\pi < k \le \pi$. Commuting any G_n with $\tilde{A}(k)$ is then diagonal in k:

$$[G_n, \tilde{A}(k)] = (e^{-ink} - e^{ink})\tilde{A}(k). \quad (47)$$

Because $G_n^\dagger = -G_n$, then

$$\mathcal{U} = e^{\sum_{n=1}^\infty u_n G_n}$$

is indeed unitary when u_n is real. Then

$$\mathcal{U}\tilde{A}(k)\mathcal{U}^\dagger = \exp\left(\sum_{n=1}^\infty u_n(e^{-ink} - e^{ink})\right)\tilde{A}(k). \quad (48)$$

Finding the appropriate coefficients u_n for the ADHH transformation then becomes easy by rewriting the Hamiltonian using the inverse Fourier transformation as

$$H_{\text{IK}} = \int_{-\pi}^{\pi} \frac{dk}{2\pi}(J - he^{ik})\tilde{A}(k).$$

Note then if we take $u_n = -(h/J)^n/(2n)$ as in (43), then (48) gives

$$\mathcal{U}\tilde{A}(k)\mathcal{U}^\dagger = \left(\frac{J - he^{-ik}}{J - he^{ik}}\right)^{1/2}\tilde{A}(k).$$

Thus, by making this choice, we find

$$\mathcal{U}H_{\text{IK}}\mathcal{U}^\dagger = \int_{-\pi}^{\pi} \frac{dk}{2\pi}(J^2 + h^2 - 2hJ\cos(k))^{1/2}\tilde{A}(k). \quad (49)$$

This transformed Hamiltonian commutes with $\hat{N}_{\text{IK}}$, since $\tilde{A}(k) + \tilde{A}(-k)$ does, and the function in the integrand of Eq. (49) is even in k. The unitary transformation in Eq. (48) thus indeed does the job required of the ADHH theorem. It also shows that $n_* \to \infty$ for $L \to \infty$, since Eq. (49) shows that the error goes to zero in this limit.

In an integrable model such as the XYZ model, there is no free-fermion representation, so the zero mode can interact with bulk excitations. Nonetheless, it is possible to show by rewriting the exact strong zero mode of Ref. [21] as a matrix-product operator that n_* is indeed infinite here as well [78], at least for the edge mode. That tantalizingly suggests that n_* is infinite for all integrable systems.

IX. DISCUSSION

We show that prethermalization can extend topological protection into regimes where it might have been expected to fail. Perhaps the simplest experimental realization of the prethermal protection of edge modes, albeit in a system without topological order, would be in a trapped ion or neutral atom chain governed by a perturbed transverse-field Ising model. In a solid state system, prethermalization is ultimately interrupted by the thermalization driven by the electron-phonon interaction. However, over intermediate

time scales, a chain of quantum dots tuned to a Kitaev-type Hamiltonian may have prethermal strong zero modes that survive until the late time at which the electron-phonon interaction causes thermalization. We emphasize that prethermalization can occur in any dimension, and the long-lived zero modes can, as a result, occur in two or three dimensions. Our work is therefore relevant to the proposal of Ref. [70] for a universal topological quantum computer.

In a more theoretical direction, we note that the ADHH theorem applies to much more general systems than those with topological order. It provides a precise and rigorous approach to finding an approximately conserved charge. Integrable models are so because of the presence of nontrivial conserved charges, and the resemblance of the Onsager analysis of the Ising model to the results the unitary transform used in this paper hints at a more general connection between the ADHH procedure and integrability. For example, the $U(1)$ quantum number conservation implied by the theorem is strongly reminiscent of the conservation of quasiparticle number in integrable field theories. Moreover, the general connection between integrability breaking and prethermalization suggests that the extremely long lifetime of the prethermal strong zero modes is a consequence of the structure of integrability still affecting the perturbed system, at least in one dimension. In higher dimensions, the role of integrability seemingly is being played by the requirement of integer eigenvalues of $\hat{N}$, but why this is so is somewhat mysterious. Clearly more research in both more formal and experimental directions is warranted.

APPENDIX A: EXTENDING THE ADHH THEOREM TO NON-SINGLE-SITE $\hat{N}$

Here, we explain how the proof of ADHH can extended to the case where the unperturbed Hamiltonian $\hat{N}$ can be written as a sum $\hat{N}_\Gamma = \sum_\Gamma \hat{N}_\Gamma$, where each $\hat{N}_\Gamma$ is supported on a set of sites Γ of finite radius. Recall that the original proof of ADHH assumed that each $\hat{N}_\Gamma$ acts on just a single site. The only place where this assumption was used is Sec. 5.4 of Ref. [31], in which it was implicitly assumed that

$$\|e^{is\hat{N}}Ve^{-is\hat{N}}\|_\kappa = \|V\|_\kappa, \quad (A1)$$

where $\| \cdot \|_\kappa$ is the norm for potentials introduced in Sec. 2.2 of Ref. [31]. This assumption no longer holds if the terms of $\hat{N}$ do not act on one site each, because then evolution by $\hat{N}$ does not preserve the support of local terms.

We can get around this problem quite straightforwardly. We say that an operator V_Z is "strongly supported" on a set Z if V_Z is supported on the set Z, *and* V_Z also commutes with all $\hat{N}_\Gamma$'s such that $\Gamma \not\subseteq Z$. For $\hat{N}$ finite range, this just requires increasing the size of Z slightly. We then redefine the norm $\| \cdot \|_\kappa$ so that it is based on the strong support of local terms rather than the support. With this modification, we see that Eq. (A1) is recovered, since the "strong support" of an operator does not grow under $\hat{N}$. We observe that if A_Z and $B_{Z'}$ are strongly supported on sets Z, Z', then $[A_Z, B_{Z'}] = 0$ if $Z \cap Z' = \varnothing$, and moreover in general, $[A_Z, B_{Z'}]$ is strongly supported on $Z \cup Z'$. This allows the rest of the proof in Ref. [31] to carry over without change.

APPENDIX B: ENERGY BANDWIDTH

Here we prove the claim we made about the energy bandwidth of the $\hat{M} = m$ sector under the Hamiltonian $J\hat{M} + \hat{D}$. The starting point is the observation that a consequence of Ref. [31] is a bound on the "local norm" of $\hat{D}$. Namely,

$$\|\hat{D}\|_0 \le CJ_0, \qquad (B1)$$

where we define $J_0 \equiv (1/\kappa_0^2)\|\hat{Y}\|_{\kappa_0}$ and fix κ_0 such that $J_0 < \infty$. The dimensionless constant C is proportional to κ_0^2. We define the local norm $\|\mathcal{H}\|_\kappa$ of a Hamiltonian $\mathcal{H} = \sum_\Gamma \mathcal{H}_\Gamma$ (where the Γ are subsets of the lattice Λ, and $\mathcal{H}_\Gamma$ is an operator supported on Γ), as

$$\|\mathcal{H}\|_\kappa = \sup_{x \in \Lambda} \sum_{\Gamma \ni x} e^{\kappa|\Gamma|}\|\mathcal{H}_\Gamma\|. \qquad (B2)$$

We derive Eq. (B1) by applying the bound in the unnumbered equation just after Eq. (4.10) of Ref. [31], invoke the fact that $\| \cdot \|_0 \le \| \cdot \|_\kappa$ for $\kappa > 0$, and then sum over n.

Now, we state our results in some degree of generality. Rather than considering any particular form of $\hat{M}$, we just assume that it can be written as

$$\hat{M} = \sum_{x \in \mathcal{X}} P_x, \qquad (B3)$$

where the P_x's are commuting projectors, and $\mathcal{X}$ is some set to index the projectors. We assume that each projector P_x is supported on a set $B_x \subseteq \Lambda$ (where Λ is the set of all sites in the lattice).

We can label eigenstates of $\hat{M}$ by their simultaneous eigenvalue under the projectors P_x, which we refer to as a "syndrome." More precisely, a syndrome is a subset $S \subseteq \mathcal{X}$, whose corresponding projectors are not satisfied.

Moreover, for each syndrome s, we can construct a corresponding projector $P_s \equiv (\prod_{x \in S} P_x)(\prod_{x \notin s}(1 - P_x))$. We let P denote the projector corresponding to the trivial syndrome, $P \equiv P_\varnothing$ (i.e., the projector onto the ground state subspace of $\hat{N}$).

We also introduce the notion of a "partial syndrome" $(\mathcal{Y}, s_\mathcal{Y})$, where $\mathcal{Y} \subseteq \mathcal{X}$ and $s_\mathcal{Y} \subseteq \mathcal{Y}$. We say that $(\mathcal{Y}, s_\mathcal{Y})$ is the "restriction" of a syndrome s if $s_\mathcal{Y} = \mathcal{Y} \cap s$. A partial syndrome specifies the eigenvalue of only those projectors indexed by $x \in \mathcal{Y}$. The projector corresponding to a partial syndrome is

$$Q_{\mathcal{Y}, s_\mathcal{Y}} = \left(\prod_{x \in s_\mathcal{Y}} P_x\right)\left(\prod_{x \in \mathcal{Y} \setminus s_\mathcal{Y}}(1 - P_x)\right). \qquad (B4)$$

We now state the following condition under which we prove our results.

Local-TQO. There exists a constant K such that, for any syndrome s, there exists a region $\mathcal{R}_s$, of size at most $K|s|$, such that $B_x \subseteq \mathcal{R}_s$ for all $x \in s$, and furthermore for any region Γ with $\Gamma \cap \mathcal{R}_s = \varnothing$, we have

$$Q_{\mathcal{A}, s_\mathcal{A}} V_\Gamma Q_{\mathcal{A}, s_\mathcal{A}} = c(V_\Gamma) Q_{\mathcal{A}, s_\mathcal{A}}, \qquad (B5)$$

where $c(V_\Gamma) = \mathrm{Tr}(PV_\Gamma P)/\mathrm{Tr}(P)$, and $\mathcal{A} = \mathcal{X} \setminus s$, which implies that the restriction $s_\mathcal{A} = s \cap \mathcal{A} = \varnothing$.

Roughly, this is saying that there is a topologically protected degeneracy in the ground state of $\hat{M}$ (by applying it to the trivial syndrome), and moreover, that an excited state with eigenvalue m is localized to a region of size at most Km, and looks like the ground state elsewhere. The condition "Local-TQO" is closely related to the conditions under which stability of the topological order in the ground state subspace of $\hat{M}$ was proven in Refs. [79–81] (though the result we want to prove here is somewhat different). We observe that "Local-TQO" is indeed satisfied for the $\hat{N}$ of the Kitaev chain [Eq. (6)], with $K = 2$.

Now we can prove the following theorem:

Theorem 1. If the condition "Local-TQO" is satisfied, then for any Hamiltonian $V = \sum_\Gamma V_\Gamma$ that commutes with $\hat{M}$, the spectrum of $J\hat{M} + V$ in the eigenspace of $\hat{M}$ with eigenvalue m lies within the interval

$$[c(V) + m(J - K\|V\|_0), c(V) + m(J + K\|V\|_0)]. \qquad (B6)$$

Proof.—Consider an operator V_Γ supported on a set $\Gamma \subseteq \Lambda$. We want to consider the circumstances under which $P_s V_\Gamma P_t - \delta_{s,t} c(V_\Gamma) P_s$ can be nonzero for two syndromes $s, t \in \mathcal{S}$ such that $|s| = |t|$. Let us partition the set $\mathcal{X}$ into $\mathcal{X}_\Gamma$ and $\mathcal{X}_\Gamma^c$, where $\mathcal{X}_\Gamma = \{x \in \mathcal{X} : B_x \cap \Gamma \ne \varnothing\}$. We consider the following cases:

(i) $s \ne t$ and $s \cap \mathcal{X}_\Gamma^c \ne t \cap \mathcal{X}_\Gamma^c$.

Without loss of generality, we can say that there exists $x \in s \cap \mathcal{X}_\Gamma^c$ such that $x \notin t$. Then we note that $P_s P_x = 1$ and $P_x P_t = 0$. Furthermore, $x \in \mathcal{X}_\Gamma^c$

implies that $[P_x, V_\Gamma] = 0$. Hence, we can write $P_s V_\Gamma P_t = P_s P_x V_\Gamma P_t = P_s V_\Gamma P_x P_t = 0$.

(ii) $s \neq t$, $s \cap \mathcal{X}_\Gamma^c = t \cap \mathcal{X}_\Gamma^c$ and $\mathcal{R}_s \cap \Gamma = \varnothing$.

$\mathcal{R}_s \cap \Gamma = \varnothing$ implies that for all $x \in s$, $B_x \cap \Gamma = \varnothing$ (since $B_x \subseteq R_s$). This implies that $s \subseteq \mathcal{X}_\Gamma^c$, and hence that $s = s \cap \mathcal{X}_\Gamma^c = t \cap \mathcal{X}_\Gamma^c$. Since $|s| = |t|$, this implies that $s \neq t$, which contradicts our assumption.

(iii) $s = t$ and $\mathcal{R}_s \cap \Gamma = \varnothing$.

We decompose s into partial syndromes $(\mathcal{A}, s_\mathcal{A})$ and $(\mathcal{B}, s_\mathcal{B})$, where $\mathcal{A} = s$ and $\mathcal{B} = \mathcal{X} \backslash \mathcal{A}$. Then we can write $P_s = Q_{\mathcal{A}, s_\mathcal{A}} Q_{\mathcal{B}, s_\mathcal{B}}$. We observe that $\mathcal{R}_s \cap \Gamma \neq \varnothing$ ensures that $Q_{\mathcal{B}, s_\mathcal{B}}$ commutes with V_Γ. Hence, we find that

$$P_s V_\Gamma P_s = Q_{\mathcal{B}, s_\mathcal{B}} (Q_{\mathcal{A}, s_\mathcal{A}} V_\Gamma Q_{\mathcal{A}, s_\mathcal{A}}) Q_{\mathcal{B}, s_\mathcal{B}}$$
$$= c(V_\Gamma) P_s, \tag{B7}$$

by Eq. (B5).

In conclusion, we find that $P_s V_\Gamma P_t - \delta_{s,t} c(V_\Gamma) P_s = 0$ except when $\mathcal{R}_s \cap \Gamma \neq \varnothing$ and $\mathcal{R}_t \cap \Gamma \neq \varnothing$.

Now consider a Hamiltonian $V = \sum_\Gamma V_\Gamma$. Let $\mathbb{P}_m = \sum_{s \in \mathcal{S}:|s|=E} P_s$ (that is, $\mathbb{P}_m$ is the projector onto the subspace with eigenvalue m under $\hat{M}$). Furthermore, let $\mathbb{P}_m^\Gamma = \sum_{s \in \mathcal{S},|s|=m, \mathcal{R}_s \cap \Gamma \neq \varnothing} P_m$. Then we see that

$$\mathbb{P}_m V \mathbb{P}_m - c(V) \mathbb{P}_m = \sum_\Gamma \mathbb{P}_m V_\Gamma \mathbb{P}_m - c(V_\Gamma) \mathbb{P}_n$$
$$= \sum_\Gamma \mathbb{P}_m^\Gamma V_\Gamma \mathbb{P}_m^\Gamma - c(V_\Gamma) \mathbb{P}_m^\Gamma$$
$$\leq \sum_\Gamma \|V_\Gamma\| \mathbb{P}_m^\Gamma$$
$$= \sum_{s:|s|=m} \left(\sum_{\Gamma: \mathcal{R}_s \cap \Gamma \neq 0} \|V_\Gamma\| \right) P_s$$
$$\leq \sum_{s:|s|=m} |\mathcal{R}_s| \|V\|_0 P_s$$
$$\leq K m \|V\|_0 \sum_{s:|s|=m} P_s$$
$$= K m \|V\|_0 \mathbb{P}_m.$$

Hence,

$$\|\mathbb{P}_m V \mathbb{P}_m - c(V) \mathbb{P}_m\| \leq K m \|V\|_0. \tag{B8}$$

The theorem immediately follows. ∎

[1] C. Nayak, S. H. Simon, A. Stern, M. Freedman, and S. Das Sarma, *Non-Abelian Anyons and Topological Quantum Computation*, Rev. Mod. Phys. **80**, 1083 (2008).

[2] J. Alicea, *New Directions in the Pursuit of Majorana Fermions in Solid State Systems*, Rep. Prog. Phys. **75**, 076501 (2012).

[3] C. W. J. Beenakker, *Search for Majorana Fermions in Superconductors*, Annu. Rev. Condens. Matter Phys. **4**, 113 (2013).

[4] S. Das Sarma, M. Freedman, and C. Nayak, *Majorana Zero Modes and Topological Quantum Computation*, npj Quantum Inf. **1**, 15001 (2015).

[5] J. D. Sau, R. M. Lutchyn, S. Tewari, and S. Das Sarma, *Generic New Platform for Topological Quantum Computation Using Semiconductor Heterostructures*, Phys. Rev. Lett. **104**, 040502 (2010).

[6] R. M. Lutchyn, J. D. Sau, and S. Das Sarma, *Majorana Fermions and a Topological Phase Transition in Semiconductor-Superconductor Heterostructures*, Phys. Rev. Lett. **105**, 077001 (2010).

[7] Y. Oreg, G. Refael, and F. von Oppen, *Helical Liquids and Majorana Bound States in Quantum Wires*, Phys. Rev. Lett. **105**, 177002 (2010).

[8] V. Mourik, K. Zuo, S. M. Frolov, S. R. Plissard, E. P. A. M. Bakkers, and L. P. Kouwenhoven, *Signatures of Majorana Fermions in Hybrid Superconductor-Semiconductor Nanowire Devices*, Science **336**, 1003 (2012).

[9] L. P. Rokhinson, X. Liu, and J. K. Furdyna, *The Fractional a.c. Josephson Effect in a Semiconductor-Superconductor Nanowire as a Signature of Majorana Particles*, Nat. Phys. **8**, 795 (2012).

[10] M. T. Deng, C. L. Yu, G. Y. Huang, M. Larsson, P. Caroff, and H. Q. Xu, *Anomalous Zero-Bias Conductance Peak in a Nb–InSb Nanowire–Nb Hybrid Device*, Nano Lett. **12**, 6414 (2012).

[11] H. O. H. Churchill, V. Fatemi, K. Grove-Rasmussen, M. T. Deng, P. Caroff, H. Q. Xu, and C. M. Marcus, *Superconductor-Nanowire Devices from Tunneling to the Multichannel Regime: Zero-Bias Oscillations and Magnetoconductance Crossover*, Phys. Rev. B **87**, 241401 (2013).

[12] A. Das, Y. Ronen, Y. Most, Y. Oreg, M. Heiblum, and H. Shtrikman, *Zero-Bias Peaks and Splitting in an Al-InAs Nanowire Topological Superconductor as a Signature of Majorana Fermions*, Nat. Phys. **8**, 887 (2012).

[13] A. D. K. Finck, D. J. Van Harlingen, P. K. Mohseni, K. Jung, and X. Li, *Anomalous Modulation of a Zero-Bias Peak in a Hybrid Nanowire-Superconductor Device*, Phys. Rev. Lett. **110**, 126406 (2013).

[14] S. M. Albrecht, A. P. Higginbotham, M. Madsen, F. Kuemmeth, T. S. Jespersen, J. Nygård, P. Krogstrup, and C. M. Marcus, *Exponential Protection of Zero Modes in Majorana Islands*, Nature (London) **531**, 206 (2016).

[15] A. Y. Kitaev, *Unpaired Majorana Fermions in Quantum Wires*, Phys. Usp. **44**, 131 (2001).

[16] For more on many-body localization, see Ref. [17] and references therein.

[17] R. Nandkishore, S. Gopalakrishnan, and D. A. Huse, *Spectral Features of a Many-Body-Localized System Weakly Coupled to a Bath*, Phys. Rev. B **90**, 064203 (2014).

[18] D. A. Huse, R. Nandkishore, V. Oganesyan, A. Pal, and S. L. Sondhi, *Localization-Protected Quantum Order*, Phys. Rev. B **88**, 014206 (2013).

[19] B. Bauer and C. Nayak, *Area Laws in a Many-Body Localized State and Its Implications for Topological Order*, J. Stat. Mech. 9, 09005 (2013).

[20] Y. Bahri, R. Vosk, E. Altman, and A. Vishwanath, *Localization and Topology Protected Quantum Coherence at the Edge of Hot Matter*, Nat. Commun. **6,** 7341 (2015).

[21] P. Fendley, *Strong Zero Modes and Eigenstate Phase Transitions in the XYZ/Interacting Majorana Chain*, J. Phys. A **49,** 30LT01 (2016).

[22] P. Fendley, *Parafermionic Edge Zero Modes in Z_n-Invariant Spin Chains*, J. Stat. Mech. 11, 20 (2012).

[23] A. S. Jermyn, R. S. K. Mong, J. Alicea, and P. Fendley, *Stability of Zero Modes in Parafermion Chains*, Phys. Rev. B **90,** 165106 (2014).

[24] J. Alicea and P. Fendley, *Topological Phases with Parafermions: Theory and Blueprints*, Annu. Rev. Condens. Matter Phys. **7,** 119 (2016).

[25] J. Kemp, N. Y. Yao, C. R. Laumann, and P. Fendley, *Long Coherence Times for Edge Spins*, J. Stat. Mech. (2017) 063105.

[26] J. Berges, Sz. Borsányi, and C. Wetterich, *Prethermalization*, Phys. Rev. Lett. **93,** 142002 (2004).

[27] M. Moeckel and S. Kehrein, *Interaction Quench in the Hubbard Model*, Phys. Rev. Lett. **100,** 175702 (2008).

[28] M. Kollar, F. A. Wolf, and M. Eckstein, *Generalized Gibbs Ensemble Prediction of Prethermalization Plateaus and Their Relation to Nonthermal Steady States in Integrable Systems*, Phys. Rev. B **84,** 054304 (2011).

[29] M. Gring, M. Kuhnert, T. Langen, T. Kitagawa, B. Rauer, M. Schreitl, I. Mazets, D. A. Smith, E. Demler, and J. Schmiedmayer, *Relaxation and Prethermalization in an Isolated Quantum System*, Science **337,** 1318 (2012).

[30] F. H. L. Essler, S. Kehrein, S. R. Manmana, and N. J. Robinson, *Quench Dynamics in a Model with Tuneable Integrability Breaking*, Phys. Rev. B **89,** 165104 (2014).

[31] D. Abanin, W. De Roeck, F. Huveneers, and W. W. Ho, *A Rigorous Theory of Many-Body Prethermalization for Periodically Driven and Closed Quantum Systems*, Commun. Math. Phys. **354,** 809 (2017).

[32] M. Bukov, S. Gopalakrishnan, M. Knap, and E. Demler, *Prethermal Floquet Steady States and Instabilities in the Periodically Driven, Weakly Interacting Bose-Hubbard Model*, Phys. Rev. Lett. **115,** 205301 (2015).

[33] D. A. Abanin, W. De Roeck, and F. Huveneers, *Exponentially Slow Heating in Periodically Driven Many-Body Systems*, Phys. Rev. Lett. **115,** 256803 (2015).

[34] E. Canovi, M. Kollar, and M. Eckstein, *Stroboscopic Prethermalization in Weakly Interacting Periodically Driven Systems*, Phys. Rev. E **93,** 012130 (2016).

[35] D. A. Abanin, W. De Roeck, and W. W. Ho, *Effective Hamiltonians, Prethermalization and Slow Energy Absorption in Periodically Driven Many-Body Systems*, Phys. Rev. B **95,** 014112 (2017).

[36] T. Kuwahara, T. Mori, and K. Saito, *Floquet-Magnus Theory and Generic Transient Dynamics in Periodically Driven Many-Body Quantum Systems*, Ann. Phys. (Amsterdam) **367,** 96 (2016).

[37] T. Mori, T. Kuwahara, and K. Saito, *Rigorous Bound on Energy Absorption and Generic Relaxation in Periodically*

[38] M. Bukov, M. Heyl, D. A. Huse, and A. Polkovnikov, *Heating and Many-Body Resonances in a Periodically Driven Two-Band System*, Phys. Rev. B **93,** 155132 (2016).

[39] D. V. Else, B. Bauer, and C. Nayak, *Prethermal Phases of Matter Protected by Time-Translation Symmetry*, Phys. Rev. X **7,** 011026 (2017).

[40] J. D. Sau and S. D. Sarma, *Realizing a Robust Practical Majorana Chain in a Quantum-Dot-Superconductor Linear Array*, Nat. Commun. **3,** 964 (2012).

[41] I. C. Fulga, A. Haim, A. R. Akhmerov, and Y. Oreg, *Adaptive Tuning of Majorana Fermions in a Quantum Dot Chain*, New J. Phys. **15,** 045020 (2013).

[42] D. M. Basko, I. L. Aleiner, and B. L. Altshuler, *Metal Insulator Transition in a Weakly Interacting Many-Electron System with Localized Single-Particle States*, Ann. Phys. (Amsterdam) **321,** 1126 (2006).

[43] D. M. Basko, I. L. Aleiner, and B. L. Altshuler, *On the Problem of Many-Body Localization*, Phys. Rev. B **77,** 041409 (2008).

[44] V. Oganesyan and D. A. Huse, *Localization of Interacting Fermions at High Temperature*, Phys. Rev. B **75,** 155111 (2007).

[45] A. Pal and D. A. Huse, *Many-Body Localization Phase Transition*, Phys. Rev. B **82,** 174411 (2010).

[46] J. H. Bardarson, F. Pollmann, and J. E. Moore, *Unbounded Growth of Entanglement in Models of Many-Body Localization*, Phys. Rev. Lett. **109,** 017202 (2012).

[47] M. Serbyn, Z. Papić, and D. A. Abanin, *Universal Slow Growth of Entanglement in Interacting Strongly Disordered Systems*, Phys. Rev. Lett. **110,** 260601 (2013).

[48] M. Serbyn, Z. Papić, and D. A. Abanin, *Local Conservation Laws and the Structure of the Many-Body Localized States*, Phys. Rev. Lett. **111,** 127201 (2013).

[49] J. Z. Imbrie, *On Many-Body Localization for Quantum Spin Chains*, J. Stat. Phys. **163,** 998 (2016).

[50] D. A. Huse, R. Nandkishore, and V. Oganesyan, *Phenomenology of Fully Many-Body-Localized Systems*, Phys. Rev. B **90,** 174202 (2014).

[51] V. Khemani, R. Nandkishore, and S. L. Sondhi, *Nonlocal Adiabatic Response of a Localized System to Local Manipulations*, Nat. Phys. **11,** 560 (2015).

[52] S. Johri and R. Nandkishore, *Non-Local Response in the Disordered Majorana Chain*, arXiv:1608.00022.

[53] E. Dennis, A. Kitaev, A. Landahl, and J. Preskill, *Topological Quantum Memory*, J. Math. Phys. (N.Y.) **43,** 4452 (2002).

[54] T. Scaffidi, D. E. Parker, and R. Vasseur, *Gapless Symmetry Protected Topological Order*, arXiv:1705.01557.

[55] A. Bespalov, M. Houzet, J. S. Meyer, and Y. V. Nazarov, *Theoretical Model to Explain Excess of Quasiparticles in Superconductors*, Phys. Rev. Lett. **117,** 117002 (2016).

[56] G. Vidal, *Efficient Simulation of One-Dimensional Quantum Many-Body Systems*, Phys. Rev. Lett. **93,** 040502 (2004).

[57] A field-theory approach gives a consistent picture. The Hamiltonian (22) is not integrable except at the critical point $t = J$, but the corresponding field theory describing the scaling limit is. The resulting exact boundary scattering matrix turns out to always scatter kinks to antikinks [58].

[58] L. Chim, *Boundary S Matrix for the Integrable q-Potts Model*, J. Phys. A **28**, 7039 (1995).

[59] N. Moran, D. Pelligrino, J. Slingerland, and G. Kells, *Parafermionic Clock Models and Quantum Resonance*, Phys. Rev. B **95**, 235127 (2017).

[60] A. Y. Kitaev, *Fault-Tolerant Quantum Computation by Anyons*, Ann. Phys. (Amsterdam) **303**, 2 (2003).

[61] S. B. Bravyi and A. Y. Kitaev, *Quantum Codes on a Lattice with Boundary*, arXiv:quant-ph/9811052.

[62] A. Kitaev and L. Kong, *Models for Gapped Boundaries and Domain Walls*, Commun. Math. Phys. **313**, 351 (2012).

[63] S. Trebst, P. Werner, M. Troyer, K. Shtengel, and C. Nayak, *Breakdown of a Topological Phase: Quantum Phase Transition in a Loop Gas Model with Tension*, Phys. Rev. Lett. **98**, 070602 (2007).

[64] S. Dusuel, M. Kamfor, R. Orús, K. P. Schmidt, and J. Vidal, *Robustness of a Perturbed Topological Phase*, Phys. Rev. Lett. **106**, 107203 (2011).

[65] M. A. Levin and X.-G. Wen, *String-Net Condensation: A Physical Mechanism for Topological Phases*, Phys. Rev. B **71**, 045110 (2005).

[66] D. V. Else and C. Nayak, *Cheshire Charge in $(3 + 1)$-D Topological Phases*, Phys. Rev. B **96**, 045136 (2017).

[67] R. Alicki, M. Horodecki, P. Horodecki, and R. Horodecki, Open Syst. Inf. Dyn. **17**, 1 (2010).

[68] I. Vurgaftman, J. R. Meyer, and L. R. Ram-Mohan, *Band Parameters for III–V Compound Semiconductors and Their Alloys*, J. Appl. Phys. **89**, 5815 (2001).

[69] O. Madelung, *Semiconductors: Data Handbook* (Springer, Berlin, 2004).

[70] M. Barkeshli and J. D. Sau, *Physical Architecture for a Universal Topological Quantum Computer Based on a Network of Majorana Nanowires*, arXiv:1509.07135.

[71] A. Y. Kitaev, *Anyons in an Exactly Solved Model and Beyond*, Ann. Phys. (Amsterdam) **321**, 2 (2006).

[72] B. Neyenhuis, J. Smith, A. C. Lee, J. Zhang, P. Richerme, P. W. Hess, Z.-X. Gong, A. V. Gorshkov, and C. Monroe, *Observation of Prethermalization in Long-Range Interacting Spin Chains*, arXiv:1608.00681.

[73] S. Hild, T. Fukuhara, P. Schau, J. Zeiher, M. Knap, E. Demler, I. Bloch, and C. Gross, *Far-from-Equilibrium Spin Transport in Heisenberg Quantum Magnets*, Phys. Rev. Lett. **113**, 147205 (2014).

[74] P. W. Hess, P. Becker, H. B. Kaplan, A. Kyprianidis, A. C. Lee, B. Neyenhuis, G. Pagano, P. Richerme, C. Senko, J. Smith, W. L. Tan, J. Zhang, and C. Monroe, *Non-Thermalization in Trapped Atomic Ion Spin Chains*, arXiv:1704.02439.

[75] A. Dutta and J. K. Bhattacharjee, *Phase Transitions in the Quantum Ising and Rotor Models with a Long-Range Interaction*, Phys. Rev. B **64**, 184106 (2001).

[76] A. Kerimov, *Uniqueness of Gibbs States in One-Dimensional Antiferromagnetic Model with Long-Range Interaction*, J. Math. Phys. (N.Y.) **40**, 4956 (1999).

[77] L. Onsager, *Crystal Statistics. 1. A Two-Dimensional Model with an Order Disorder Transition*, Phys. Rev. **65**, 117 (1944).

[78] F. Verstraete (private communication).

[79] S. Bravyi, M. B. Hastings, and S. Michalakis, *Topological Quantum Order: Stability under Local Perturbations*, J. Math. Phys. (N.Y.) **51**, 093512 (2010).

[80] S. Bravyi and M. B. Hastings, *A Short Proof of Stability of Topological Order under Local Perturbations*, Commun. Math. Phys. **307**, 609 (2011).

[81] S. Michalakis and J. P. Zwolak, *Stability of Frustration-Free Hamiltonians*, Commun. Math. Phys. **322**, 277 (2013).