


Harnessing symmetry-protected topological order for quantum memories

M. Goihl,¹ N. Walk,¹ J. Eisert,^{1,2} and N. Tarantino¹

¹*Dahlem Center for Complex Quantum Systems, Freie Universität Berlin, 14195 Berlin, Germany*

²*Helmholtz-Zentrum Berlin für Materialien und Energie, 14109 Berlin, Germany*

 (Received 16 September 2019; revised manuscript received 2 December 2019; accepted 6 December 2019; published 4 February 2020)

Spin chains with symmetry-protected edge modes are promising candidates to realize intrinsically robust physical qubits that can be used for the storage and processing of quantum information. In any experimental realization of such physical systems, weak perturbations in the form of induced interactions and disorder are unavoidable and can be detrimental to the stored information. At the same time, the latter may in fact be beneficial; for instance, by deliberately inducing disorder which causes the system to localize. We explore the potential of using an XZX cluster Hamiltonian to encode quantum information into the local edge modes and comprehensively investigate the influence of both many-body interactions and disorder on their stability over time, adding substance to the narrative that many-body localization may stabilize quantum information. We recover the edge state at each time step, allowing us to reconstruct the quantum channel that captures the locally constrained out-of-equilibrium time evolution. With this representation in hand, we analyze how well classical and quantum information are preserved over time as a function of disorder and interactions. We find that the performance of the edge qubits varies dramatically between disorder realizations. Whereas some show a smooth decoherence over time, a sizable fraction is rapidly rendered unusable as memories. We also find that the stability of the classical information—a precursor for the usefulness of the chain as a quantum memory—depends strongly on the direction in which the bit is encoded. When employing the chain as a genuine quantum memory, encoded qubits are most faithfully recovered for low interaction and high disorder.

DOI: [10.1103/PhysRevResearch.2.013120](https://doi.org/10.1103/PhysRevResearch.2.013120)

I. INTRODUCTION

The prospect of highly controllable quantum simulators [1–4] allows for the simulation of complex Hamiltonians, potentially realizing exotic phases of matter. A celebrated class of such phases is topological phases of matter, which offer the promise of robust encodings of quantum information with a reduced need for active error correction [5–7]. Simple instances of these exist in one-dimensional spin chains, where one finds *symmetry-protected topological phases* (SPTs) which host emergent edge zero modes that enforce approximate degeneracies in the spectrum and define edge states which are dynamically decoupled from the bulk [8–12]. The latter feature makes these systems interesting candidates for physical qubits used in quantum memories, as well as in implementations of measurement-based quantum computation [13–15], which exploit the entanglement properties of the SPT to effectuate gate operations. However, for a practical quantum memory architecture, the topologically facilitated isolation of the edge states from the bulk must preserve the encoded data for a sufficiently long time, ideally several thousand multiples of the time taken to implement logical gates.

The ability of an SPT edge mode to retain information depends largely on the strength of interactions which perturb the SPT order. These drive two related phenomena: the splitting of the ground-state degeneracy and the broadening of its *edge zero modes*. While the former controls the phase transition, the latter can have a significant impact on our ability to practically use the edge modes for storage purposes. We should typically expect to encode information into only a finite number of spins located at the boundary of the chain. The edge mode has no such restriction and usually decays exponentially into the chain (see Fig. 1). This mismatch between the finite number of sites we can encode into and the long tails of the edge mode operators means that we can never expect our information to be totally preserved even when we take the system to the thermodynamic limit to make the edge modes exact constants of motion. With this in mind, one naturally turns to search for features which can further localize the edge modes without engaging in fine tuning.

One readily available and promising candidate for this role is *disorder*. While usually considered an obstacle to analytical and numerical study, this feature is both generically present in physical systems and also enables a variety of localization phenomena, which may conceivably enhance the localization of the emergent edge modes. These go by different names depending on whether the underlying physics is describable using free fermions (*Anderson localization* [16]) or is intrinsically interacting (*many-body localization*, or MBL [17,18]). While the exact phenomenology of these two differs, both feature local constants of motion. In the SPT context, this

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

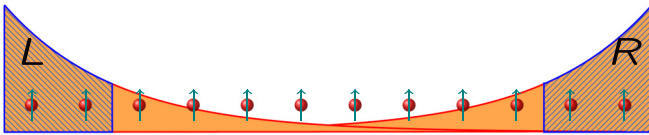


FIG. 1. This is a sketch of the one-dimensional spin chain investigated in this work. The support of the edge zero mode operators under the influence of perturbations is depicted in orange. The blue regions are the local edge qubits to which an experimenter would have access.

has been proposed as a method to attenuate the effective interaction between the two edge modes since any coupling between them must be mediated by bulk eigenstates which, if only locally supported, can only connect the two via a prohibitively small number of virtual processes [19–21].

While this heuristic is appealing, precious little confirmation of it has been forthcoming since initial investigations established that the edge modes are not destabilized by the presence of disorder [22]. In fact, a recent work by some of the authors [23]—built on the earlier Ref. [24]—suggests that the shape of the edge mode when subjected to disorder and interaction changes in a more intricate way than expected. There, the support of the edge modes was investigated and, surprisingly, found to depend heavily on the specific edge mode that was analyzed. In light of this, a thorough examination of disorder effects on the temporal stability of the edge mode appears to be urgently required. Here, we achieve this goal: We explicitly determine the dynamics of the encoded subspace and quantify the extent to which information can be recovered at a given time.

Before proceeding, it is critical to specify what is meant by a good quantum memory. Any imperfect memory device can be viewed as a partially decohering channel acting upon the encoded quantum states. For a comprehensive investigation, the first step is to compute the operators that define this channel, from which all potentially relevant measures can then be calculated. It is important to carefully select measures that meaningfully quantify the notion of information preservation in the context of quantum computation and communication applications. Ideally, a measure should be operationally well motivated, straightforward to calculate, and amenable to experimental verification. In this work, we perform full reconstruction of the time-evolved state on the edge mode subspace. This allows us to fully reconstruct the effective quantum channel applied at each time step. With this reconstruction at hand, we compute bounds on two measures of information preservation when using the system as a quantum memory: the retrievability of a *logical bit*, which is a prerequisite for a quantum memory, and the *coherent information*, which bounds the *quantum capacity* that is the ultimate limit to quantum information transmission and particularly relevant for quantum communication applications, e.g., quantum repeaters [25,26]. For the classical information, we find a multifaceted behavior consistent with Ref. [23] supporting the observation that encoding with different edge state bases results in different stability behavior. In the quantum information setting, we find that disorder helps the recovery of the encoded qubits, but low interaction strength is mandatory. Our results extend and enrich the narrative of disorder helping localize edge

modes, as we find that there are encodings and Hamiltonian parameters for which this paradigm holds and others where it does not. In practice, this either makes an optimization over all possible parameters necessary to find the “sweet spot” of information coherence or demands an educated choice of physical realization. One can equally see our work as shedding light on notions of locally constrained nonequilibrium dynamics that interplays with disorder [27].

II. SETTING

In this work, we analyze the interplay of disorder, interactions, and symmetry-protected topological (SPT) order. We are specifically interested in the stability of the protected subspace spanned by the unperturbed edge modes when the system is subjected to many-body interactions and disorder. To measure the extent to which the locally encoded state gets entangled with the bulk, we time evolve the system and reconstruct the unperturbed edge mode subspace. We study a spin chain of length N hosting a disordered XZX cluster Hamiltonian,

$$H_0(\{h_j\}) = - \sum_{j=2}^{N-1} (1 + h_j) X_{j-1} Z_j X_{j+1}, \quad (1)$$

where the h_j are drawn uniformly from the interval $[-\frac{\Delta}{2}, \frac{\Delta}{2}]$, and X_j, Y_j, Z_j are the Pauli operators acting at site j . Here, we restrict the disorder to act only on the cluster terms, disordering around a mean bulk gap of 1. We include this offset to distinguish the ability of disorder to help localize the edge modes from the bulk energy gap’s power to do the same. The choice to have disorder appear in the coefficient of the cluster terms is a pragmatic one; any other model of disorder, e.g., a random local magnetic field, competes with the SPT order and thereby drives a transition to a topologically trivial phase [19,22]. Disordering the cluster terms themselves is an ideal version of disorder in the sense that this version splits degenerate levels for excited states, while preserving the ground-state manifold. Therefore, we expect any other model of disorder to cause less localization in the system.

Even disordered, the Hamiltonian in Eq. (1) is a solvable representative of an SPT phase protected by the time-reversal operator,

$$\mathcal{T} = \prod_{i=1}^N Z_i \mathcal{K}, \quad (2)$$

where \mathcal{K} is complex conjugation. Thus, this system supports protected edge zero modes, i.e., operators localized on the edge which commute with the Hamiltonian, at least up to an error that is exponentially suppressed in system size. At this solvable point, these operators commute exactly and can be identified by inspection. There are six such operators,

$$\mathcal{O}_{\text{Edge}}^i = \left\{ \begin{array}{ll} \mathcal{O}_L^x = X_1, & \mathcal{O}_R^x = X_N, \\ \mathcal{O}_L^y = Y_1 X_2, & \mathcal{O}_R^y = X_{N-1} Y_N, \\ \mathcal{O}_L^z = Z_1 X_2, & \mathcal{O}_R^z = X_{N-1} Z_N \end{array} \right\}, \quad (3)$$

that split into two independent Pauli algebras, i.e., one for the left and right edge, respectively. Each of these Pauli algebras implies the existence of a special two-dimensional

Hilbert space hosted on these edges, corresponding to the fundamental representation of these algebras.

The physical situation underlying this algebraic picture is reflected by a subtle decomposition of the Hilbert space. The *logical* left (right) edge qubit is only encoded into a two-dimensional subspace \mathcal{H}_L (\mathcal{H}_R) of the four-dimensional Hilbert space described by the physical qubits on the two left-most (right-most) sites. It is important to stress that the bulk degrees of freedom, associated with $\mathcal{H}_{\text{Bulk}}$, involve physical sites $3, \dots, N-2$ as well as the remaining two-dimensional subspaces of each of the two edges. Initial preparations are product state vectors of the form

$$|\psi\rangle = |\psi_L\rangle \otimes |\psi_{\text{Bulk}}\rangle \otimes |\psi_R\rangle. \quad (4)$$

In the ideal limit, without any interactions, these edge states $|\psi_L\rangle, |\psi_R\rangle$ are independent of the global unitary time evolution generated by H_0 —never becoming entangled with the bulk—and thus do not decohere. This changes, however, as soon as coupling terms are added to the Hamiltonian. In this case, the actual edge modes broaden (sketched in Fig. 1) and no longer coincide with their compact form in Eq. (3). The corresponding edge states hybridize, splitting the degeneracy by an amount which is suppressed exponentially with system size [8,9,28,29]. Our logical qubits are no longer time independent and become entangled with the rest of the chain. By restricting ourselves to measurements of the edge operators in Eq. (3), we will be partially tracing an entangled state leading to decoherence over time.

To model this effect, we introduce an Ising-type interaction,

$$H_{ZZ}(J) = -J \sum_{j=1}^{N-1} Z_j Z_{j+1}, \quad (5)$$

resulting in the following Hamiltonian:

$$H(\{h_j\}, J) = H_0(\{h_j\}) + H_{ZZ}(J), \quad (6)$$

where we vary both J and Δ . Note that the resulting Hamiltonian is no longer free and hence requires a treatment in the full Hilbert space.

As we are interested in studying the evolution of the states of the edge Hilbert space, we need a way of constructing the reduced dynamics on what is effectively an open quantum system. This can be done by time evolving an initial pure product state on the entire system of the form in Eq. (4), and then evaluating the time-evolved reduced state of the edge, ρ_t , as

$$\begin{aligned} \rho_t &= \text{tr}_{\text{Bulk}}(e^{-iHt} |\psi\rangle \langle \psi| e^{iHt}) \\ &= \mathcal{E}_t(|\psi_{L,R}\rangle \langle \psi_{L,R}|), \end{aligned} \quad (7)$$

with $|\psi_{L,R}\rangle = |\psi_L\rangle \otimes |\psi_R\rangle$. These dynamics, for fixed bulk preparations, gives rise to a family of quantum channels \mathcal{E}_t , acting on the state space over $\mathcal{H}_{L,R}$. This time-parametrized family is also referred to as the *dynamical map* that captures time evolution of the edge. We now investigate the effects of these dynamics on the viability of the edge modes as a quantum memory.

III. METHODS

A. Characterizing quantum memories

The approximate conservation of the observable quantities is often used as a proxy for the information storage capabilities of a physical system, topological or otherwise. In fact, disordered models in a many-body localized phase have been shown to preserve spin orientation along a magnetization direction both theoretically [30,31] and experimentally [32] in one dimension, with similar results in two dimensions [33]. While these phenomena are indeed hallmarks of dynamical decoupling, they are only a prerequisite for a useful quantum memory. Specifically, the information encoded into a polarized spin state in an MBL system is only a classical bit string, even if the system itself is quantum. Given the abundance of efficient classical memories, using a quantum system for bit storage seems wasteful. A true quantum memory must be able to robustly store an arbitrary—typically unknown—qubit. This necessitates the existence of a protected two-dimensional Hilbert space, which is not guaranteed by the presence of even an arbitrary number of conserved quantities. Rather, preservation of any fixed basis can be seen as a necessary condition, but to promote the system in question to a quantum memory, all other possible bases must be preserved too.

To assess the usefulness of a system as a quantum memory, we will now introduce measures which quantify the amount of both classical and quantum information that is retrievable from our system. As mentioned above, classical information storage is a necessary condition for realizing a quantum memory. The storage capabilities of quantum states are relevant for quantum computations, where one reads in a classical bit string into a quantum system, then performs some (possibly) quantum gates on it and reads out the classical bit string encoding the result of the computation again. Moreover, many quantum communication protocols require a quantum memory, e.g., for quantum networks, it is often necessary to store an entangled state between some nodes while preparing further entanglement between others. As such, quantum memories are a crucial building block for many quantum technologies [34].

Since we are interested in the stability of information (whether classical or quantum) stored in edge states, care should be taken in defining measures thereof. The robustness of a logical bit, whose value is encoded via orthogonal state vectors $|\psi\rangle$ and $|\phi\rangle$ in an imperfect quantum memory described in its time evolution by \mathcal{E}_t , is essentially given by the distinguishability of the logical states at the output [i.e., between $\rho = \mathcal{E}_t(|\psi\rangle \langle \psi|)$ and $\sigma = \mathcal{E}_t(|\phi\rangle \langle \phi|)$]. This can be captured by the trace distance,

$$\mathcal{D}(\rho, \sigma) = \frac{1}{2} \|\rho - \sigma\|_1. \quad (8)$$

Operationally, this is directly related to the probability of successfully recovering an encoded logical bit. For example, the maximum probability p of distinguishing between quantum states ρ and σ on a single-shot level is

$$p = \frac{1}{2} + \frac{1}{2} \mathcal{D}(\rho, \sigma), \quad (9)$$

assuming each state occurs with initial identical probability of 1/2. This, then, is precisely the optimal probability that

an unknown equiprobable logical bit could be successfully determined from the output of the memory.

As outlined above, these notions can be extended to quantify the performance of a quantum memory by considering all possible encodings. This insight was used in Ref. [35] to propose a figure of merit called the *integrity*, which is essentially the output distinguishability minimized over all possible orthogonal state encodings. Intuitively, this measure tracks how well the entire logical subspace is preserved under the time evolution. It can also be related to other important metrics in quantum information processing such as the logical fidelity and the error-correction pseudothreshold [36,37].

For our system, the integrity at time t will be given by

$$\mathcal{I}(t) = \min_{|\psi_{L,R}\rangle \perp |\phi_{L,R}\rangle} \mathcal{D}[\mathcal{E}_t(|\psi_{L,R}\rangle\langle\psi_{L,R}|), \mathcal{E}_t(|\phi_{L,R}\rangle\langle\phi_{L,R}|)], \quad (10)$$

where $|\psi_{L,R}\rangle$ is a quantum state vector of $\mathcal{H}_{L,R}$ and $|\phi_{L,R}\rangle$ is an orthogonal state vector, evolved in time. Note that this is a slight generalization with respect to [35] where only single-qubit systems were considered, for which the orthogonal complement of a pure state is uniquely defined. The generalization here to two qubit states will thus require simultaneous minimization over pairs of orthogonal input states, which can still be captured as an efficiently solvable semidefinite problem. Nevertheless, the primary minimization in the integrity is highly nontrivial for general \mathcal{E}_t so we will resort to upper bounding this quantity by evaluating the output distinguishability for particular orthogonal encodings.

Given the above definition of the integrity, there is an immediate connection to the *classical information capacity* of the quantum channel \mathcal{E}_t . This quantifies the asymptotic rate at which classical bits encoded in quantum states can be retrieved, once the quantum systems undergo an evolution dictated by the quantum channel \mathcal{E}_t . This is the most meaningful definition of capturing the capability of storing classical information. On the formal level, it is defined as the *regularized Holevo- χ* of the quantum channel [38]. Since the integrity gives rise to a specific product encoding, we have

$$p_{\mathcal{I}}(t) \log_2 p_{\mathcal{I}}(t) + [1 - p_{\mathcal{I}}(t)] \log_2 [1 - p_{\mathcal{I}}(t)] \leq C_{\text{Cl}}(\mathcal{E}_t), \quad (11)$$

where $p_{\mathcal{I}}(t)$ is the distinguishing probability associated with $\mathcal{I}(t)$ via Eq. (9).

In what follows, we restrict ourselves to eigenstates of the unperturbed edge mode operators in Eq. (3) as inputs. That is, the initial orthogonal state vectors $|\psi_{L,R}\rangle$ and $|\phi_{L,R}\rangle$ entering the definition of the integrity are chosen to be orthogonal eigenvectors of $X_L \otimes X_R$, $Y_L \otimes Y_R$, and $Z_L \otimes Z_R$. We make this choice because it should be relevant in the context of possible experimental realizations, and also because similar input states have been shown to saturate the minimum in Eq. (10) for some classes of channels [35]. We refer to these physically motivated variants as *directed integrities* and will denote them as $\mathcal{I}_X(t)$, $\mathcal{I}_Y(t)$, $\mathcal{I}_Z(t)$, respectively.

For quantum communication, arguably the most meaningful quantity is the *quantum channel capacity* Q , which is the maximum number of qubits that can be reliably transmitted per use of the channel in the asymptotic limit of many uses.

More precisely, this means that $Q(\mathcal{E}_t)$ is the maximum rate at which one can send qubits through a channel \mathcal{E} , such that the probability of error goes to zero in the limit of many uses of the channel. To date, this quantity has only been evaluated for certain classes of channels. This is because, in general, it requires a complicated optimization over input encodings (possibly entangled across channel uses). Typically, one writes Q using a “regularization” of the single-shot capacity Q_1 ,

$$Q(\mathcal{E}_t) = \lim_{n \rightarrow \infty} \frac{1}{n} Q_1(\mathcal{E}_t^{\otimes n}), \quad (12)$$

where $\mathcal{E}_t^{\otimes n}$ represents n parallel uses of the channel. To calculate Q_1 , one only optimizes over encodings for a single input round, and thus

$$\begin{aligned} Q_1(\mathcal{E}_t) &= \sup_{\rho} \{S[\mathcal{E}_t(\rho)] - S[(\mathcal{E}_t \otimes \mathbb{I})(|\psi\rangle\langle\psi|)]\} \\ &= \sup_{\rho} C(\mathcal{E}_t)_{\rho}, \end{aligned} \quad (13)$$

where $|\psi\rangle$ is a purification of ρ and $S(\rho) = -\text{tr}(\rho \log_2 \rho)$ is the von Neumann entropy. The quantity on the right, $C(\mathcal{E}_t)_{\rho}$, is called the *coherent information* [38].

Even this optimization can be quite challenging, so we instead choose a particular input state which yields a lower bound to the channel capacity. We choose to input one-half of a maximally entangled state, such that the state going through the channel, ρ , will be the maximally mixed state. This state is completely isotropic and should hence be a good candidate to assess the (possibly asymmetric) induced noise. In fact, it also can be shown via the Choi-Jamiołkowski isomorphism that the resulting bipartite output state is a unique alternative representation of the map \mathcal{E} . In any event, this procedure yields a convenient lower bound to the quantum capacity of the memory.

B. Computing figures of merit

Having identified meaningful quantities for determining whether information is protected, we turn to the messy business of actually computing them. The first step is to compute the quantum channel \mathcal{E}_t which captures the edge state dynamics. There are several mathematically equivalent ways to represent a quantum channel, such as its Choi-Jamiołkowski isomorphic quantum state or its Kraus operators [39]. We will employ an alternative approach, better suited to the particularities of the problem, known as the *matrix form* of the channel. Obtaining this representation will essentially entail computing the final edge states given a set of well-chosen initializations. While this is formally done by evaluating the partial trace in Eq. (7), doing so directly is very inefficient. The difficulty comes from attempting to trace over $\mathcal{H}_{\text{Bulk}}$, which contains information that lives on the same *physical* sites as the edge qubits. This means that we cannot use the physical site index to perform the partial trace, forcing us to change to a basis where the decomposition of our full Hilbert space into $\mathcal{H}_{\text{Bulk}} \otimes \mathcal{H}_{L,R}$ is made explicit.

To avoid this complication, we compute the density matrix by instead evaluating the expectation values of a tomographically complete set of observables, built using the edge mode operators defined in Eq. (3). With these, we reconstruct the

density matrix on the two-qubit system residing on the edge using

$$\rho(t) = \frac{1}{4} \sum_{\substack{i, j \in \\ \{0, x, y, z\}}} \langle \mathcal{O}_L^i \mathcal{O}_R^j(t) \rangle \sigma_L^i \otimes \sigma_R^j, \quad (14)$$

where a superscript of 0 indicates that we take an identity operator in the relevant Hilbert space. Beyond providing a computational advantage in the following analysis, this method of evaluating the density matrix also provides an intuitive picture for how one would characterize a memory experimentally. The experimenter would measure a tomographically complete set of observables until sufficient statistics are obtained to accurately determine the expectation values in Eq. (14) and hence reconstruct the density matrix [39,40]. This procedure, known as *quantum state tomography*, has been implemented in a wide variety of quantum information experiments [41–44]. In general, at least $d^2 - 1$ different measurement settings are required to fully characterize a d -dimensional state.

With the edge state determined, only the computational challenge of obtaining the effective channel \mathcal{E}_t remains. Since \mathcal{E}_t is linear, we can always think of it as acting on a vector space spanned by the Hilbert-Schmidt scalar products of basis elements, usually written as $|m\rangle\langle n|$, and find its matrix elements in that basis. In our case, the matrix form can be obtained from the numerical values of $\mathcal{E}_t(|m\rangle\langle n|)$, using a Z -spin basis, letting $|m\rangle, |n\rangle \in \{|\uparrow, \uparrow\rangle, |\downarrow, \downarrow\rangle, |\uparrow, \downarrow\rangle, |\downarrow, \uparrow\rangle\}$. However, our procedure in Eq. (14) only allows us to evaluate the map on elements of the form $|m\rangle\langle m|$, and so producing all the operator basis elements seems out of reach. Thankfully, there exists an algorithm for producing the off-diagonal elements, namely,

$$\begin{aligned} \mathcal{E}_t(|m\rangle\langle n|) &= \mathcal{E}_t(|+\rangle\langle +|) + i\mathcal{E}_t(|-\rangle\langle -|) - \frac{1+i}{2}\mathcal{E}_t(|n\rangle\langle n|) \\ &\quad - \frac{1+i}{2}\mathcal{E}_t(|m\rangle\langle m|), \end{aligned} \quad (15)$$

where $|+\rangle = 1/\sqrt{2}(|m\rangle + |n\rangle)$ and $|-\rangle = 1/\sqrt{2}(|m\rangle + i|n\rangle)$. This entire process of characterizing a map by performing state tomography on the output of a set of judiciously chosen inputs is often referred to as quantum *process* tomography [39]. Performing this calculation at every time step allows us to track the progressive degradation of the memory and calculate any of the single-shot quantities defined in Sec. III A. We do so for a number of disorder configurations and interaction strengths, obtaining a comprehensive picture of the edge's response to both.

IV. RESULTS

In this section, we present the numerical results obtained for the disordered SPT model with Ising-type perturbations defined in Eq. (6). We carried out simulations for a system size $N = 14$ and 100 disorder realizations. The Hamiltonian parameters used are $J \in \{0.1, 0.075, 0.05, 0.025\}$ and $\Delta \in \{0.5, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5\}$, where we deliberately chose

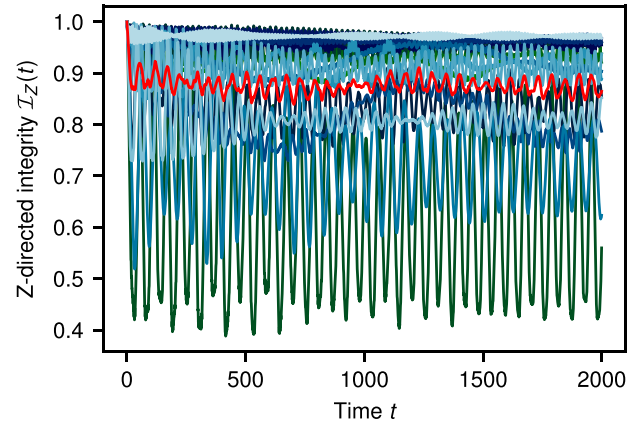


FIG. 2. This plot shows the Z -directed integrity $\mathcal{I}_Z(t)$ as a function of time for 15 random samples (blue/green shades) and their average (red). The Hamiltonian parameters in this plot are $J = 0.075$ and $\Delta = 1.0$. The system is of size $N = 14$ and we simulated the time evolution until $t = 2000$ using step sizes of $\Delta t = 0.1$.

some values $\Delta \geq 2$ to study the regime where disorder competes with the bulk gap.

All shown time traces were simulated from $t = 0$ to 2000 (in units of inverse energy of the XZX coupling) by integrating Schrödinger's equation with a step size of $\Delta t = 0.1$. For each disorder realization, we use 20 different initial states to obtain full tomographic information of the two-qubit subspace. At every integer time slice, we calculate the expectation value of all 15 combinations of the unperturbed edge mode operators given in Eq. (3). Tracking these quantities over time allows us to reconstruct the full channel that represents the time evolution reduced to the edges as laid out in Eq. (7).

As discussed in Sec. III A, we will first investigate directed integrities which, individually, are only a measure of the quality of a classical memory. When all three Pauli directions are taken together, however, they serve as an upper bound for the integrity of the quantum channel. We find that the behavior of each instance varies considerably between different disorder realizations. To demonstrate this issue, in Fig. 2 we show the time traces of the Z -directed integrity $\mathcal{I}_Z(t)$ for 15 randomly chosen instances of disorder with Hamiltonian parameters $J = 0.1$ and $\Delta = 1.5$. The red curve is an average over the 15 traces; the traces themselves are shown in shades of green and blue.

Though this average looks well behaved, the obtained traces show a wide variety of behavior from one realization to the next. It is useful to split the traces into two groups: those which decrease exponentially with gentle fluctuations and those which oscillate wildly within an exponentially decaying envelope. The former of these should dominate any average, and so we expect that our disorder-averaged quantities should softly decay, a finding consistent with previous results in related work [22].

While this would usually be the end of a story featuring disorder, the appearance of rapidly fluctuating traces is very concerning. Since a low value in the integrity and coherent information both indicate the formation of entanglement, we can only conclude that the edge state is oscillating between states of high and low bulk entanglement. This scenario constitutes a

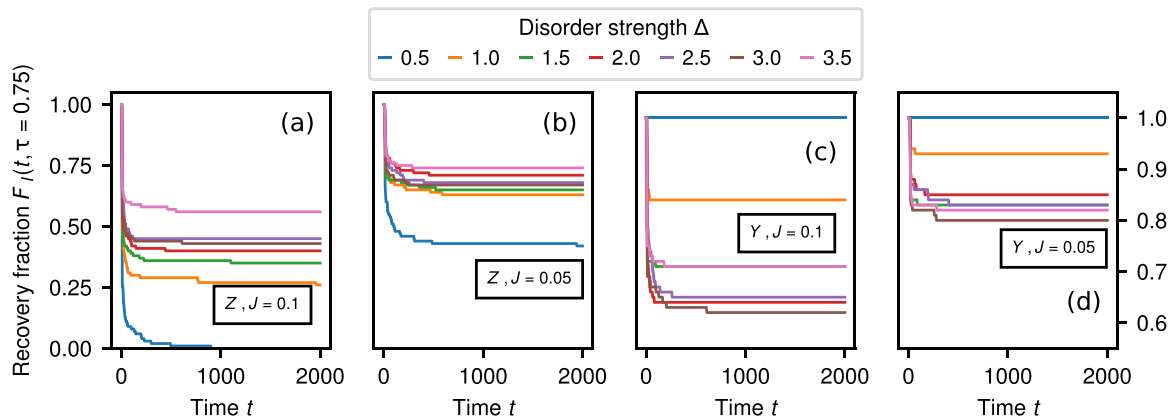


FIG. 3. Recovery fractions of directed integrities for a fixed threshold as a function of time. Colors encode the disorder strength Δ . The system size is $N = 14$ and we simulated the time evolution until $t = 2000$ using step sizes of $\Delta t = 0.1$. (a) Data for the Z -directed integrity $\mathcal{F}_{\mathcal{I}_Z}(t, \tau = 0.7)$ with $J = 0.1$. (b) Data for $\mathcal{F}_{\mathcal{I}_Z}(t, \tau = 0.7)$ with $J = 0.05$. (c) Data for the Y -directed integrity $\mathcal{F}_{\mathcal{I}_Y}(t, \tau = 0.7)$ with $J = 0.1$. (d) Data for $\mathcal{F}_{\mathcal{I}_Y}(t, \tau = 0.7)$ with $J = 0.05$.

fiendish obstacle for any experimental implementation. Even if one could, in principle, decode the stored data by picking a moment of low entanglement, the experimenter would need detailed knowledge about the system dynamics to do so reliably. One might hope that these wild fluctuations are limited to this single integrity plot, indicating a pathology with limited scope. But, as we shall see, they appear in a sizable fraction of disorder realizations and so are intrinsic to the dynamics of this system.

Given the extreme variation between disorder realizations, we find ourselves in a circumstance where disorder averaging would impede our ability to determine the suitability of this system to act as a quantum memory. Instead, we will define a phenomenological quantity appropriate for the given problem, which we call the recovery fraction,

$$\mathcal{F}_A(t, \tau) := \frac{N_{\text{rec}}}{N_{\text{tot}}}, \quad (16)$$

where N_{rec} is defined as the number of instances for which the measured quantity A is larger than τ , i.e., $\{|A(t)|A(t') > \tau \forall t' \leq t\}$. The coefficient τ captures a minimum threshold that a hypothetical experimenter would need in order to recover the encoded information. In the limit of infinite sample size, the recovery fraction converges to a probability. This recovery probability gives the relative proportion of systems that *reliably*, meaning for all times smaller than t , stays above the recovery threshold τ .

A. Directed integrities

We first investigate the directed integrities $\mathcal{I}_X(t)$, $\mathcal{I}_Y(t)$, $\mathcal{I}_Z(t)$, which quantify how well the two initially orthogonal states in the X , Y , or Z logical basis remain so. Each of these gives an upper bound to the true integrity, but there are significant differences between these bases which require comment.

As described above, we will look at the recovery fraction to identify which realizations would be unusable due to rapid oscillatory behavior. Figure 3 shows the time evolution of the recovery fraction for each directed integrity. The left panels show the recovery fraction for the Z -directed integrity

$\mathcal{F}_{\mathcal{I}_Z}(t, \tau = 0.7)$ for interaction strengths $J \in \{0.1, 0.05\}$. Colors encode the different disorder strengths. For any combination of Hamiltonian parameters, we find a steep initial drop of the recovery probability followed by a long plateau. When tuning the strengths of the interactions and disorder, we find that decreasing interactions and increasing disorder increases the apparent saturation value of the recovery probability. This is plausibly explained by the sharper localization behavior in such conditions, a mechanism that conceivably improves the storage of classical information. The sharp initial drop-off is both intriguing and highly disturbing, as it implies that even for relatively weak interaction strengths (e.g., $J = 0.05$), a huge fraction of disorder realizations fall below threshold immediately. The loss of 25% to 50% of these shows that while the overall decay profile may be exponential, the fine structure of the evolution would preclude us from saying that the information is protected in any practical sense. We would like to point out that the X -directed integrity shows the same qualitative behavior and thus is not displayed separately here.

When analyzing the recovery fraction for the Y -directed integrity $\mathcal{F}_{\mathcal{I}_Y}(t, \tau = 0.7)$ shown in the two panels on the right, we find a sizable quantitative increase of the recovery fraction accompanied by a similar qualitative effect of interactions, namely, the recovery probability decreases with interaction strength. However, its disorder dependence differs significantly from the $\mathcal{I}_Z(t)$ case. Here, increasing the disorder strength initially *decreases* the integrity of the Y basis, before rising again past $\Delta = 3$. This indicates that the Y -directed integrity is controlled by the likelihood of having cluster terms with coefficients $1 + h_j$ close to zero. This begins to occur when $\Delta \sim 2$ and is most likely in the parameter region of $\Delta \in [2, 3]$. The resurgence of the recovery fraction past this regime can be understood as the disorder parameter dominating both the mean gap and interaction strength. Thus we expect the system to behave identically to one with weaker interactions. While initially shocking, the large quantitative and qualitative difference between bases actually mirrors results from previous numerical studies performed by some of the authors [23], where the perturbed edge zero mode operators showed a similar disorder and interaction dependence.

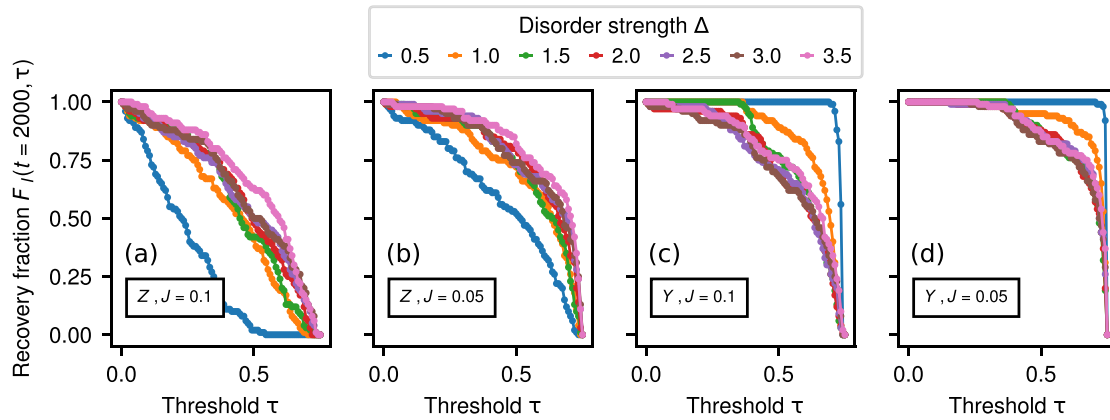


FIG. 4. Final recovery fraction of directed integrities at $t = 2000$ as a function of the threshold τ . Colors encode the disorder strength Δ . The system size is $N = 14$. (a) Data for the Z -directed integrity $\mathcal{F}_{I_Z}(t = 2000, \tau)$ with $J = 0.1$. (b) Data for $\mathcal{F}_{I_Z}(t = 2000, \tau)$ with $J = 0.05$. (c) Data for the Y -directed integrity $\mathcal{F}_{I_Y}(t = 2000, \tau)$ with $J = 0.1$. (d) Data for $\mathcal{F}_{I_Y}(t = 2000, \tau)$ with $J = 0.05$.

While the time dependence of \mathcal{F}_I reveals curious behavior, in practice, an experimenter will be interested in how well the bit can be recovered after time evolution. It is thus instructive to consider the final value of our simulation at $t = 2000$ and plot this value as a function of the threshold τ , ensuring that we have not artificially depressed the recovery fraction by choosing an overly optimistic threshold. The resulting plots are shown in Fig. 4. Again, the color encodes disorder and the two panels on the left show two interaction strengths $J \in \{0.1, 0.05\}$ for the final recovery fraction of the Z -directed integrity $\mathcal{F}_{I_Z}(t = 2000, \tau)$. Here, we find that the recovery fraction starts at 1 and smoothly goes to zero when the threshold reaches 1. Again, the interaction plays a crucial role here, depressing the recovery fraction uniformly as it increases. The disorder blunts this effect as before, but the effect saturates when it is of the order of the mean bulk gap. When turning to the final recovery fraction of the Y -directed integrity $\mathcal{F}_{I_Y}(t = 2000, \tau)$ shown in the two panels on the right, we again see an extreme sensitivity to the disorder strength, with a final recovery fraction which is almost identically 1 for all but the most aggressive values of τ , while disorder is very low (< 0.5),

with an immediate drop-off for higher disorder strengths. What becomes more apparent is disorder’s homogenization of the bases, improving the integrity of the X and Z encodings while simultaneously degrading the Y encoding so that the final recovery fractions are similar.

These results are nicely summarized in Figs. 5 and 6, with heat maps of $\mathcal{F}_{I_Z}(t, \tau = 0.7)$ and $\mathcal{F}_{I_Y}(t, \tau = 0.7)$, respectively. Here we can clearly see the monotonic improvement of the Z -directed integrity with increasing disorder and decreasing interaction strength, while the Y -directed integrity shows a clear dip as Δ approaches 2. We do see a resurgence in both plots when Δ moves beyond 4, but this corresponds to the disorder scale becoming the dominant energy scale in the system, playing a similar role to the mean gap in the original problem.

Taken together, the directed integrities paint a curious picture. Classical bits encoded using any basis will be susceptible to degradation from interactions. Bizarrely, our ability to attenuate this with disorder depends more on the chosen edge basis than the interaction strength, with the Y basis being clearly superior for weak disorder. We would like to

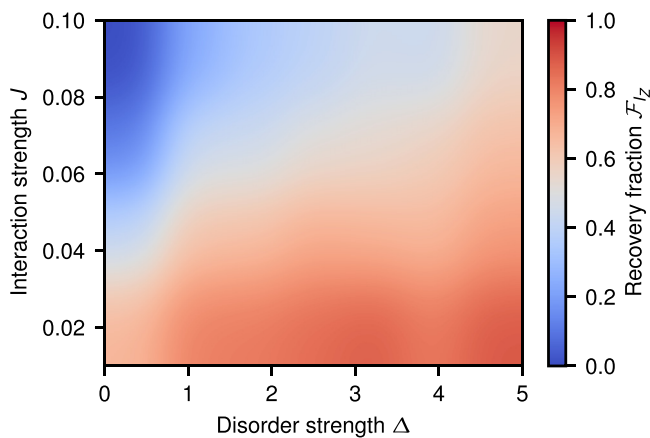


FIG. 5. In this heat map, we show the disorder and interaction dependence of the final recovery fraction $\mathcal{F}_{I_Z}(t = 2000, \tau = 0.7)$ for a fixed threshold for the Z -directed integrity.

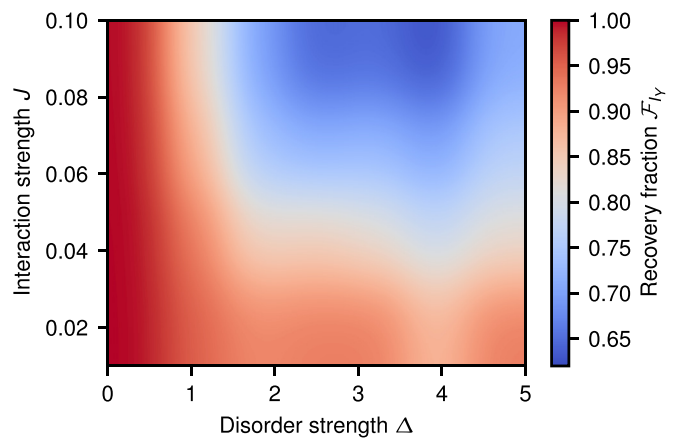


FIG. 6. In this heat map, we show the disorder and interaction dependence of the final recovery fraction $\mathcal{F}_{I_Y}(t = 2000, \tau = 0.7)$ for a fixed threshold for the Y -directed integrity.

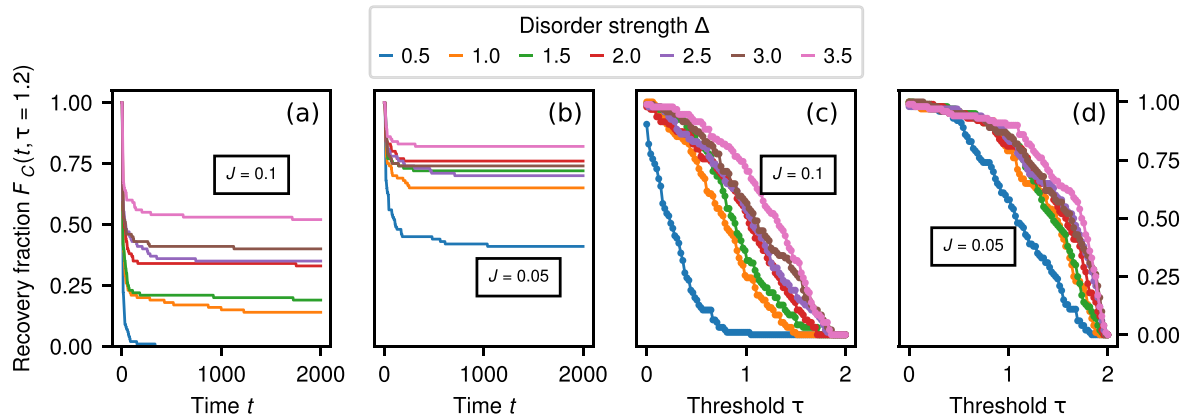


FIG. 7. This figure shows data for the coherent information for various interaction strengths. Colors encode the disorder strength Δ . The system is of size $N = 14$. Simulations run up to 2000 tunneling times with a step size of $\Delta t = 0.1$. $J \in \{0.1, 0.05\}$. (a) Time evolution of the recovery fraction of the coherent information $\mathcal{F}_C(t, \tau = 1.2)$ with $J = 0.1$. (b) The coherent information $\mathcal{F}_C(t, \tau = 1.2)$ with $J = 0.05$. (c) The final recovery fraction of the coherent information $\mathcal{F}_C(t = 2000, \tau)$ as a function of the threshold τ with $J = 0.1$. (d) The final recovery fraction $\mathcal{F}_C(t = 2000, \tau)$ with $J = 0.05$.

stress again that considering single directed integrities only allows conclusions about classical information storage. The actual integrity is the minimum of all possible encodings and is thus upper bounded by the worst of our directed integrity results, e.g., the Z -directed integrity. Since this is improved by introducing disorder, this suggests that disorder does in fact help its ability to store quantum information, albeit at the cost of reducing its classical information storage capabilities, which are captured by the Y -directed integrity. To confirm this, we will now calculate the coherent information of the dynamical map.

B. Coherent information

In this part, we carry out the very same analysis as for the directed integrities, but now we use the coherent information $C(t)$ which gives a lower bound to the single-shot capacity $Q_1(t)$ of our system. $Q_1(t)$ quantifies the number of qubits which could be reliably extracted from our system after our time evolution using some optimal single-shot encoding

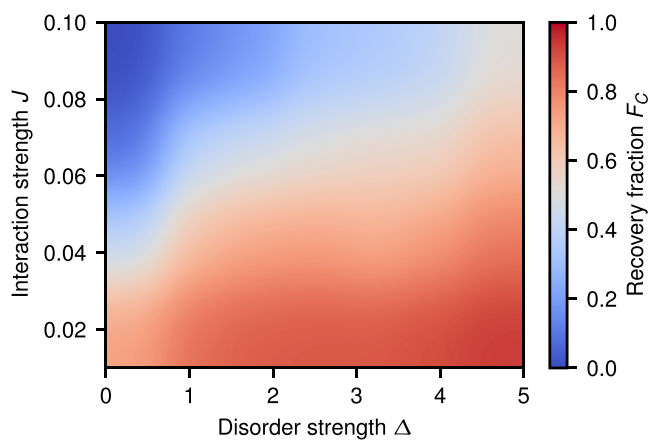


FIG. 8. In this heat map, we show the disorder and interaction dependence of the final recovery fraction of the coherent information $\mathcal{F}_C(t = 2000, \tau = 1.2)$.

strategy. The coherent information then corresponds to fixing an encoding strategy, and thus both range from zero to two.

The time dependence of the recovery fraction of the coherent information $\mathcal{F}_C(t, \tau = 1.2)$ for a fixed threshold and different interaction strengths is shown in the left panels of Fig. 7. The behavior shown is strikingly similar to the Z -directed integrity, showing the same immediate decline in recovery fraction, but now for a quantity which is genuinely quantum. Decreasing the interaction strength and increasing the disorder strength both improve the coherent information monotonically, a feature also shared with the Z -directed integrity. These similarities continue in the final recovery probability of the coherent information $\mathcal{F}_C(t = 2000, \tau)$ as a function of cutoff τ as shown in the panels on the right in Fig. 7, showing that most of the improvement comes from reducing the interaction strength.

If we now again consider the final recovery probability of the coherent information $\mathcal{F}_C(t = 2000, \tau = 1.2)$ with fixed threshold as a function of disorder and interaction strength as shown in Fig. 8, we again see that high disorder and low interaction lead to the highest final recovery fraction of the coherent information. The figure looks very similar to what we have obtained for the Z -directed integrities, confirming that the ability of these to store information is the limiting factor in the edge's ability to preserve quantum information reliably.

V. CONCLUSIONS

Using edge modes of topologically protected systems is a promising avenue for quantum information tasks. It is, however, largely unclear how stable information encoded into these systems will actually be when the system is perturbed away from its solvable point by spurious interactions and disorder, both of which are omnipresent in realistic experimental realizations. Here, we present numerical results for the dynamical recoverability of information—both classical and quantum—encoded into the edge mode subspace of an XZX cluster Hamiltonian with both tunable many-body interactions

and disorder as a perturbation. We reconstruct the edge state density matrix to obtain the channel which represents the time evolution. With these data, we are in position to calculate measures of classical information, namely, directed integrities, and quantum information, namely, the coherent information.

By studying the directed integrities, we find that not only do disorder strength and interaction strength change the amount of recoverable information, but also that the specific encoding direction plays a crucial role in determining the coherence of the encoded bits. The best encoding is found to be in the Y basis of the unperturbed Hamiltonian, yielding very stable bits for a large range of interaction strengths. The narrative of disorder aiding the encoding is wrong for this setting, as we find an intermediate regime when disorder strength is comparable to gap size ($\Delta \sim 2$) which has minimal recovery fraction, before rising again monotonically beyond that. It holds, however, for an encoding in the Z or X bases, where the detrimental effect of interaction strength can be counteracted by introducing disorder. Since the cluster Hamiltonian is not rotationally symmetric to begin with, we do not expect the bases to behave uniformly. However, the level of asymmetry we have observed is startling and provides urgent motivation to comprehensively understand the relationship between the microscopic model and the resulting preferred classical bit-encoding basis.

All of this said, a single encoding basis does not a qubit make. For quantum information, one needs to characterize the channel generated by the time evolution reduced to the edge qubits. We did so by calculating the time-dependent coherent information stored in the edge states, a basis-independent quantity. In doing so, we recovered and provided further substance to the narrative that disorder increases the ability of the system to store quantum information. Interactions decrease the quality of the memory in a similar way to the classical encoding in the Z or X direction.

To distinguish well- and ill-behaved disorder realizations, we set a minimum performance threshold and evaluated the fraction of runs which stayed above it over the course of the evolution. Quite surprisingly, this fraction failed to decay smoothly over time, as one would expect if there was gradual decoherence of the edge states. Instead, the recover fraction dropped precipitously, indicating that a large fraction of disorder realizations describe systems that would be worthless for information storage purposes. Furthermore, these may not have an appreciable effect on disorder averages, as they are rapidly oscillatory instead of simply dropping to zero immediately. It seems fair to say that the interplay of disorder, interactions, and SPT order is less understood than previously anticipated. Our findings are compatible with those in Ref. [23], where a similarly multifaceted picture for the edge mode support was found using a completely different method.

Numerical simulations	
Total Kernel Hours [h]	308000
Thermal Design Power per Kernel [W]	5.75
Total Energy Consumption of Simulations [kWh]	1771
Average Emission of CO ₂ in Germany [kg/kWh]	0.56
Total CO ₂ -Emission from Numerical Simulations [kg]	992
Were the Emissions Offset?	Yes
Transportation	
Total CO ₂ -Emission from Transportation [kg]	0
Were the Emissions Offset?	Yes
Total CO ₂ -Emission [kg]	992

FIG. 9. Estimated climate footprint of the computations presented in this paper. Prototyping is not included in these calculations. Estimations have been calculated using the examples of Scientific CO₂nduct [50] and are correct to the best of our knowledge.

These results demand more foundational work, possibly including an extension of perturbation theory to systems with topologically protected degeneracies.

Our work covers information loss due to unitary dynamics under perturbed Hamiltonians, but real-world experiments are also coupled to the outside world leading to further decoherence. Since localization effects are also expected to be diminished in the presence of dissipation [45,46], the amount of information protection induced by disorder seems questionable. Given that open system dynamics are also capable of corrupting topological data directly, even without the presence of many-body interactions [47–49], investigating the circumstances where disorder, interactions, and a particle reservoir are all present is likely to yield fascinating results.

VI. CO₂-EMISSION TABLE

Figure 9 summarizes the climate expenses of the simulations presented in this work.

ACKNOWLEDGMENTS

We would like to thank Ingo Roth for discussions on optimization schemes. N.W. acknowledges funding support from the European Unions Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant Agreement No. 750905. This work has been supported by the ERC (TAQ), the DFG (CRC 183, FOR 2724, EI 519/7-1), and the Templeton Foundation. This work has also received funding from the European Union’s Horizon 2020 research and innovation programme under Grant Agreement No. 817482 (PASQuaS).

- [1] C. Gross and I. Bloch, *Science* **357**, 995 (2017).
 [2] R. Blatt and C. F. Roos, *Nat. Phys.* **8**, 277 (2012).
 [3] A. Aspuru-Guzik and P. Walther, *Nat. Phys.* **8**, 285 (2012).
 [4] J. Eisert, M. Friesdorf, and C. Gogolin, *Nat. Phys.* **11**, 124 (2015).

- [5] E. Dennis, A. Kitaev, A. Landahl, and J. Preskill, *J. Math. Phys.* **43**, 4452 (2002).
 [6] A. Y. Kitaev, *Ann. Phys.* **303**, 2 (2003).
 [7] A. Kitaev and J. Preskill, *Phys. Rev. Lett.* **96**, 110404 (2006).

- [8] F. D. M. Haldane, *Phys. Rev. Lett.* **50**, 1153 (1983).
- [9] A. Y. Kitaev, *Phys. Usp.* **44**, 131 (2001).
- [10] R. Verresen, R. Moessner, and F. Pollmann, *Phys. Rev. B* **96**, 165124 (2017).
- [11] X.-G. Wen, *Phys. Rev. D* **88**, 045013 (2013).
- [12] X. Chen, Z.-C. Gu, Z.-X. Liu, and X.-G. Wen, *Phys. Rev. B* **87**, 155114 (2013).
- [13] J. Miller and A. Miyake, *Phys. Rev. Lett.* **114**, 120506 (2015).
- [14] D. T. Stephen, D.-S. Wang, A. Prakash, T.-C. Wei, and R. Raussendorf, *Phys. Rev. Lett.* **119**, 010504 (2017).
- [15] R. Raussendorf, D.-S. Wang, A. Prakash, T.-C. Wei, and D. T. Stephen, *Phys. Rev. A* **96**, 012302 (2017).
- [16] P. W. Anderson, *Phys. Rev.* **109**, 1492 (1958).
- [17] D. A. Abanin, E. Altman, I. Bloch, and M. Serbyn, *Rev. Mod. Phys.* **91**, 021001 (2019).
- [18] C. Gogolin and J. Eisert, *Rep. Prog. Phys.* **79**, 56001 (2016).
- [19] D. A. Huse, R. Nandkishore, V. Oganesyan, A. Pal, and S. L. Sondhi, *Phys. Rev. B* **88**, 014206 (2013).
- [20] A. Chandran, V. Khemani, C. R. Laumann, and S. L. Sondhi, *Phys. Rev. B* **89**, 144201 (2014).
- [21] J. A. Kjäll, J. H. Bardarson, and F. Pollmann, *Phys. Rev. Lett.* **113**, 107204 (2014).
- [22] Y. Bahri, R. Vosk, E. Altman, and A. Vishwanath, *Nat. Commun.* **6**, 7341 (2015).
- [23] M. Goihl, C. Krumnow, M. Gluza, J. Eisert, and N. Tarantino, *SciPost Phys.* **6**, 72 (2019).
- [24] M. Goihl, M. Gluza, C. Krumnow, and J. Eisert, *Phys. Rev. B* **97**, 134202 (2018).
- [25] H.-J. Briegel, W. Dür, J. I. Cirac, and P. Zoller, *Phys. Rev. Lett.* **81**, 5932 (1998).
- [26] S. Muralidharan, L. Li, J. Kim, N. Lütkenhaus, M. D. Lukin, and L. Jiang, *Sci. Rep.* **6**, 20463 (2016).
- [27] C. Chen, F. Burnell, and A. Chandran, *Phys. Rev. Lett.* **121**, 085701 (2018).
- [28] I. Affleck, T. Kennedy, E. H. Lieb, and H. Tasaki, *Commun. Math. Phys.* **115**, 477 (1988).
- [29] I. Affleck, T. Kennedy, E. H. Lieb, and H. Tasaki, *Phys. Rev. Lett.* **59**, 799 (1987).
- [30] M. Serbyn, Z. Papić, and D. A. Abanin, *Phys. Rev. Lett.* **111**, 127201 (2013).
- [31] D. A. Huse, R. Nandkishore, and V. Oganesyan, *Phys. Rev. B* **90**, 174202 (2014).
- [32] M. Schreiber, S. S. Hodgman, P. Bordia, H. P. Lüschen, M. H. Fischer, R. Vosk, E. Altman, U. Schneider, and I. Bloch, *Science* **349**, 842 (2015).
- [33] J.-Y. Choi, S. Hild, J. Zeiher, P. Schauß, A. Rubio-Abadal, T. Yefsah, V. Khemani, D. A. Huse, I. Bloch, and C. Gross, *Science* **352**, 1547 (2016).
- [34] A. Acin, I. Bloch, H. Buhrman, T. Calarco, C. Eichler, J. Eisert, D. Esteve, N. Gisin, S. J. Glaser, F. Jelezko, S. Kuhr, M. Lewenstein, M. F. Riedel, P. O. Schmidt, R. Thew, A. Wallraff, I. Walmsley, and F. K. Wilhelm, *New J. Phys.* **20**, 080201 (2018).
- [35] X. Xu, N. de Beaudrap, J. O’Gorman, and S. C. Benjamin, *New J. Phys.* **20**, 023009 (2018).
- [36] K. M. Svore, A. W. Cross, I. L. Chuang, and A. V. Aho, *Quantum Inf. Comput.* **6**, 193 (2006).
- [37] A. W. Cross, D. P. Divincenzo, and B. M. Terhal, *Quantum Inf. Comput.* **9**, 541 (2009).
- [38] M. M. Wilde, *Quantum Information Theory* (Cambridge University Press, Cambridge, 2013).
- [39] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).
- [40] Z. Hradil, *Phys. Rev. A* **55**, R1561 (1997).
- [41] A. G. White, D. F. V. James, P. H. Eberhard, and P. G. Kwiat, *Phys. Rev. Lett.* **83**, 3103 (1999).
- [42] C. F. Roos, G. P. T. Lancaster, M. Riebe, H. Häffner, W. Hänsel, S. Gulde, C. Becher, J. Eschner, F. Schmidt-Kaler, and R. Blatt, *Phys. Rev. Lett.* **92**, 220402 (2004).
- [43] M. Steffen, M. Ansmann, R. C. Bialczak, N. Katz, E. Lucero, R. McDermott, M. Neeley, E. M. Weig, A. N. Cleland, and J. M. Martinis, *Science* **313**, 1423 (2006).
- [44] A. I. Lvovsky and M. G. Raymer, *Rev. Mod. Phys.* **81**, 299 (2009).
- [45] E. van Nieuwenburg, J. Y. Malo, A. Daley, and M. Fischer, *Quantum Sci. Tech.* **3**, 01LT02 (2017).
- [46] H. P. Lüschen, P. Bordia, S. S. Hodgman, M. Schreiber, S. Sarkar, A. J. Daley, M. H. Fischer, E. Altman, I. Bloch, and U. Schneider, *Phys. Rev. X* **7**, 011034 (2017).
- [47] J. C. Budich, S. Walter, and B. Trauzettel, *Phys. Rev. B* **85**, 121405(R) (2012).
- [48] D. Rainis and D. Loss, *Phys. Rev. B* **85**, 174533 (2012).
- [49] L. Mazza, M. Rizzi, M. D. Lukin, and J. I. Cirac, *Phys. Rev. B* **88**, 205142 (2013).
- [50] M. Goihl and R. Sweke (unpublished).