

Explicit construction of local conserved operators in disordered many-body systemsT. E. O'Brien,¹ Dmitry A. Abanin,² Guifre Vidal,³ and Z. Papić⁴¹*Instituut-Lorentz, Universiteit Leiden, P.O. Box 9506, 2300 RA Leiden, The Netherlands*²*Department of Theoretical Physics, University of Geneva, 24 quai Ernest-Ansermet, 1211 Geneva, Switzerland*³*Perimeter Institute for Theoretical Physics, Waterloo, Ontario N2L 2Y5, Canada*⁴*School of Physics and Astronomy, University of Leeds, Leeds LS2 9JT, United Kingdom*

(Received 18 August 2016; revised manuscript received 4 October 2016; published 24 October 2016)

The presence and character of local integrals of motion—quasilocal operators that commute with the Hamiltonian—encode valuable information about the dynamics of a quantum system. In particular, strongly disordered many-body systems can generically avoid thermalization when there are extensively many such operators. In this work, we explicitly construct local conserved operators in one-dimensional spin chains by directly minimizing their commutator with the Hamiltonian. We demonstrate the existence of an extensively large set of local integrals of motion in the many-body localized phase of the disordered XXZ spin chain. These operators are shown to have exponentially decaying tails, in contrast to the ergodic phase where the decay is (at best) polynomial in the size of the subsystem. We study the algebraic properties of localized operators and confirm that in the many-body localized phase, they are well described by “dressed” spin operators.

DOI: [10.1103/PhysRevB.94.144208](https://doi.org/10.1103/PhysRevB.94.144208)**I. INTRODUCTION**

Isolated quantum systems (i.e., those disconnected from any thermal bath) evolve according to unitary evolution. Recently there has been much interest in understanding and classifying the possible outcomes of such evolution for generic systems, e.g., those containing many particles, subject to local potentials and interactions, and in the presence of quenched disorder degrees of freedom. For simplicity, we consider systems defined on a lattice, and therefore with a local Hilbert space of finite dimension (e.g., two in the case of spin-1/2 qubits). The total Hilbert space, being a tensor product of local Hilbert spaces, is exponentially large in lattice size L , which makes the problem still very difficult to treat in general.

One possible outcome that is relatively well understood is ergodic evolution: a system, prepared in an arbitrary initial state, evolves towards local thermal equilibrium at long times. This process is a result of the “eigenstate thermalization hypothesis” (ETH), which governs the underlying structure of individual (many-body) eigenvectors of ergodic systems [1–3]. More recently, there has been a surge of activity focused on understanding a distinct class of systems which undergo *nonergodic* dynamics. A well-known example of such systems is the Anderson insulator [4]. Anderson localization is a generic property of low-dimensional systems which is not sensitive to a particular type of lattice or disorder, but only applies to noninteracting systems. When interactions are comparable in strength to hopping and disorder energy scales, the system may exhibit *many-body localization*, a type of localization in the many-body Hilbert space [5–7].

The interest in many-body localized (MBL) systems is partly fundamental, as they provide a new paradigm for nonergodic systems that violate the basic premise of equilibrium statistical mechanics. On the other hand, MBL systems have measurable properties that distinguish them from both Anderson insulators and ergodic systems. For example, their dynamics is governed by slow dephasing between different parts of the system [8–11]. On the other hand, it was shown that quantum information can be recovered in MBL systems

using spin-echo techniques [12], and there are proposals that exotic types of order can be stabilized using the MBL mechanism [13–16], which may be applicable to designing quantum information processing schemes.

In order to explain the basic phenomenology of MBL systems, including their failure to thermalize, a picture of *local integrals of motion* (LIOMs) has been put forward [17,18]. According to this picture, the basic mechanism of MBL systems is similar to integrable models [19]: there emerges an extensive number of operators (“conserved charges”) τ_i , which commute amongst themselves $[\tau_i, \tau_j] = 0$ as well as with the Hamiltonian, $[\hat{H}, \tau_i] = 0$. A special property of MBL systems is that τ_i have eigenvalues ± 1 , thus they resemble the bare spin-1/2 operators, and generically there are L such operators in a lattice system of size L . This means that any Hamiltonian eigenvector can be specified by the conserved quantum numbers corresponding to operators τ_i , $i = 1, \dots, L$. Because of this extensive number of emergent quantum numbers (that by definition do not change during unitary evolution), the thermalization of the system is prevented as the MBL state retains the memory of its initial condition. The difference between integrable models and MBL systems is in the form of individual τ_i : in the integrable case, each τ_i is an extended sum of local operators, while in the MBL case, each τ_i is a single local operator, up to corrections that vanish exponentially with distance to the core. The subleading (exponentially suppressed) corrections are important, as they cause the distinction between Anderson and MBL insulators. For example, the presence of tails in LIOMs is responsible for the dephasing dynamics and the spreading of entanglement in MBL systems [9,10], which does not occur in Anderson insulators.

The extensive number of LIOMs can therefore be viewed as a defining characteristic of MBL systems. For this reason, several previous works have explored ways of constructing the LIOMs for MBL systems. In Refs. [20,21], a local operator is time evolved with the Hamiltonian and averaged over time. Other works [22,23] have perturbatively constructed LIOMs of the form written in Refs. [10,18]. In particular, Ref. [22]

rigorously proves the existence of LIOMs near the classical limit of a one-dimensional (1D) spin model. On the other hand, Ref. [23] follows the perturbative treatment laid out in Ref. [6] near Anderson localized free fermions. Most recently, in Ref. [24] an iterative construction for LIOMs was proposed based on a similarity transformation (see also Refs. [25] and [26]), while there have also been attempts at constructing the LIOMs via quantum Monte Carlo [27].

In this work, we perform an explicit construction of local conserved operators, focusing on the many-body localized phase in 1D spin chains. We follow the method introduced in Ref. [28] where a slow-moving mode was identified in ergodic Hamiltonian and Floquet systems. Our method directly minimizes the commutator, with the system's Hamiltonian, of an operator restricted to a small subregion of an infinite chain. The main result is an unambiguous signal of the exponentially decaying tails of localized operators in the MBL phase of the disordered XXZ spin chain. This is in contrast to the polynomial scaling of the analogous operator in the ergodic phase. Further, we demonstrate that an *extensively large set* of localized operators can be found in the MBL phase, in contrast to the ergodic phase. We also study the algebraic properties of localized operators, finding that when the system is localized, they resemble spin operators with exponentially decaying tails.

As a consequence, our results allow one to directly visualize the LIOMs that continue to play a vital role in the theoretical studies of MBL systems; see Fig. 1. By plotting the disorder-averaged logarithm of the operator density, defined in Eq. (14) below, we map out the spatial profile of a typical integral of motion τ_i in the MBL phase of the disordered XXZ model. This indeed confirms that τ_i decays exponentially in the distance from the site i .

The remainder of this paper is organized as follows. In Sec. II, we introduce the notation and definitions. In Sec. III, we briefly review the properties of the disordered XXZ chain—the prototype model of the MBL phase and the localization-delocalization transition. In Sec. IV, we present

our method for explicitly constructing operators supported on a finite region which minimize the commutator with the given Hamiltonian. The method can be formulated as finding the approximate zero modes of a certain matrix, which we call the “commutant matrix.” In Sec. V, we study the scaling of the smallest eigenvalue of the commutant matrix with the size of the subsystem. We confirm the polynomial decay of this eigenvalue in the ergodic phase as previously found in Ref. [28], and show that it decays exponentially in the MBL phase. In Sec. VI, we show that the difference between ergodic, integrable, and MBL phases can also be characterized by the *counting* of approximate zero modes in the commutant matrix. In particular, the MBL phase is distinguished by having an extensively large manifold of such zero modes in contrast to other phases. In Secs. VII and VIII, we examine in more detail the properties of the constructed LIOMs, such as the “operator participation ratio” and their spectral measures. Our conclusions are presented in Sec. IX, where we also discuss some future directions.

II. BACKGROUND AND DEFINITIONS

In the following, for simplicity, we restrict ourselves to one-dimensional spin chains with local Hamiltonians (sums of terms with local support). Our methods can be directly extended to higher-dimensional systems, although this comes with considerable computational cost.

Consider a spin chain \mathcal{S} . Around any spin i in \mathcal{S} , a series of concentric one-dimensional spheres of radius l (connected line segments and their endpoints) can be formed:

$$\mathcal{S}_{i,l} = \{i-l, i-l+1, \dots, i+l-1, i+l\}, \quad (1)$$

where the boundary of this segment is denoted by

$$\partial\mathcal{S}_{i,l} = \{i-l, i+l\}. \quad (2)$$

We are interested in the behavior of operators within the areas $\mathcal{S}_{i,l}$. We can focus on traceless operators \hat{O} satisfying $\text{Tr}[\hat{O}] = 0$, as any operator can be written as a basis-independent linear combination of a traceless part and the identity (which has no local character).

A traceless operator \hat{O} acting on the Hilbert space of \mathcal{S} ($\mathcal{H}_{\mathcal{S}}$) may then be decomposed into a limit of operators $\hat{O} = \lim_{l \rightarrow \infty} \hat{O}_l$,

$$\hat{O}_l = \text{Tr}_{\overline{\mathcal{S}_{i,l}}} \hat{O}, \quad (3)$$

where the trace is taken over the complement of $\mathcal{S}_{i,l}$. We write the difference $\hat{O}_l - \hat{O}_{l-1} =: \partial\hat{O}_{i,l}$. We say \hat{O} is *centered at i with exponential tails* if there exists an exponential function $A \exp(-l/\xi)$ that bounds $\|\partial\hat{O}_{i,l}\|$. Here, ξ is a characteristic length scale that governs some of the correlations involving \hat{O} . It is tempting to identify ξ as the “many-body localization length,” which should diverge at the MBL transition. Note, however, that the decay of correlations within the MBL phase could be governed by multiple lengths [18]. Some of those may diverge at the phase transition, while others may not, e.g., the length scale ξ_c defined below, which controls the spreading of entanglement, is expected to remain finite at the phase transition. Our analytic arguments below are restricted to systems deep inside the MBL phase, where ξ is finite.

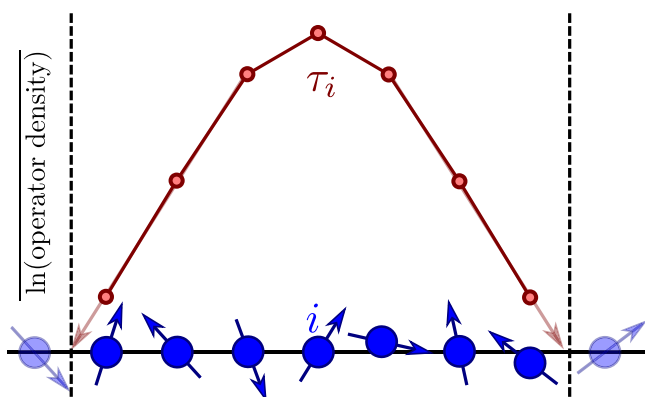


FIG. 1. A typical integral of motion τ_i in the many-body localized phase of the disordered XXZ model. τ_i decays exponentially away from the central site i . The operator τ_i is determined using the method described in Sec. IV, and its spatial profile is obtained by plotting the disorder-averaged logarithm of the operator density, given by Eq. (14). The system is an XXZ spin chain with interaction $V = 0.5t$ and disorder $\Delta = 5t$ in Eq. (5).

Given a Hamiltonian \hat{H} acting on \mathcal{S} , an *integral of motion* (IOM) is an operator \hat{O} that commutes with \hat{H} . In contrast to this generic operator, the MBL phase [17,18] supports extensively many integrals of motion, each centered around some site i in \mathcal{S} , with exponentially decaying tails. We call such an operator a *local integral of motion* (LIOM) and denote it by τ_i . In Sec. IV, we will outline the procedure of how such τ_i can be directly constructed. The result of such a calculation gives a typical τ_i , illustrated in Fig. 1. In this figure, we plot the logarithm of the operator density [defined in Eq. (14)]. The important sign of the MBL phase is the straight-line fit to the tails on either side, as this corresponds to exponential decay.

As \hat{H} is a sum of terms with local support, its commutator with the restricted $\tau_{i,l}$ must live in the area around $\partial\mathcal{S}_{i,l}$ and be canceled by operators that extend to this area. Thus, we can bound it as

$$\|[\hat{H}, \tau_{i,l}]\| \leq C \|\partial\tau_{i,l}\| \leq C' \exp(-l/\xi). \quad (4)$$

We intend to approximate $\tau_{i,l}$ in each $\mathcal{S}_{i,l}$ by finding the *approximate local integral of motion* (ALIOM) $t_{i,l}$ in $\mathcal{S}_{i,l}$ that minimizes the commutator with \hat{H} .

III. MODEL: THE DISORDERED XXZ SPIN CHAIN

A prototype model of a system with an MBL phase is the disordered XXZ spin chain [29]. This consists of a one-dimensional spin-1/2 chain with the Hamiltonian

$$\begin{aligned} \hat{H} = & \frac{t}{4} \sum_{j=1}^{N-1} (\sigma_j^x \sigma_{j+1}^x + \sigma_j^y \sigma_{j+1}^y) \\ & + \frac{V}{4} \sum_{j=1}^{N-1} \sigma_j^z \sigma_{j+1}^z + \frac{1}{2} \sum_{j=1}^N h_j \sigma_j^z, \end{aligned} \quad (5)$$

with σ_j^a , $a = \{0, x, y, z\}$ being the Pauli operator acting on site j and satisfying $(\sigma_j^a)^2 = 1$, and h_j selected randomly from a uniform distribution over $[-\Delta, \Delta]$. We consider open boundary conditions. The model has two independent parameters (V and Δ), assuming the overall energy scale is fixed by t . Via the Jordan-Wigner transformation, this system is equivalent to the (spinless) 1D Hubbard model, where Δ is random on-site disorder.

The disordered XXZ model was found to capture several interesting physical regimes. Let us assume $V = t$. At zero disorder $\Delta = 0$, the model is integrable and thus fails to thermalize. For disorder $\Delta < \Delta_c$ with $\Delta_c \approx 3.5-4$ the critical value of disorder, the system is ergodic, in the sense that nearly all eigenstates obey the ETH [30]. For small values of $\Delta \lesssim 0.55$, entanglement spreads ballistically [31] and obeys the ‘‘volume law’’ in the system’s eigenstates [32]. This means that the von Neumann entropy S_{ent} , evaluated for a system bipartitioned into A and B subsystems (such that $L_A + L_B = L$), scales proportionally to the number of degrees of freedom in A : $S_{\text{ent}} \propto L_A$. This is a typical property of ergodic systems, and a similar relation holds even for completely random vectors in the Hilbert space. At slightly larger disorder $0.55 \lesssim \Delta \lesssim \Delta_c$, the system exhibits subballistic entanglement spreading and subdiffusive charge transport [32].

Studies of the inverse participation ratio (IPR) and eigenvalue statistics [29] have demonstrated the existence of a finite- V phase transition in the range $3 \lesssim \Delta \lesssim 4$. There is evidence that this transition may be complicated by the presence of a mobility edge [33,34], making the position of the phase transition dependent on energy density, but it has been suggested that this may be due to finite-size effects [35].

Finally, at strong disorder $\Delta \gtrsim 4$, the system enters the MBL phase where nearly all eigenstates are localized (apart from rare resonant states). In the MBL phase, the entropy of most eigenstates scales as the area of the subsystem, i.e., $S_{\text{ent}} \propto \text{const}$ in 1D. This means, on average, that there is far less entanglement in MBL states compared to ergodic systems. This ‘‘area law’’ for entropy [14,18,36] is similar to what is found in the ground states of gapped systems [37]. The special property of the MBL phase is that the area law holds throughout the spectrum, even at arbitrarily high excited eigenstates. This also has consequences for the dynamics in the MBL phase; for example, during the global quench from an initial product state, the entanglement spreads logarithmically in time [8–10]: $S_{\text{ent}}(t) \propto \xi_c \ln(Vt/\hbar)$, where ξ_c is an effective length controlling the spread of entanglement (note that ξ_c is expected to remain finite at the MBL transition [18]).

As mentioned in Sec. I, the fundamental reason why entanglement and dynamics are so constrained in the MBL phase compared to the ergodic phase is the appearance of extensively many LIOMs. In order to fully understand the properties of MBL systems, it is therefore desirable to have a method that has direct access to such operators. In the following section, we describe our method for constructing such operators directly by solving an eigenvalue problem.

IV. METHOD

We now introduce the method for explicitly constructing operators that ‘‘best commute’’ with the Hamiltonian, such as the one in Eq. (5) (see also Ref. [28]). The method is general and can be applied to any model. In subsequent sections, we will apply the method to construct local operators and study their scaling with system size in the MBL and ergodic phases of the disordered XXZ model.

Given a spin chain \mathcal{S} , we wish to minimize the commutator [28]

$$\rho = \frac{\text{Tr}([\hat{H}, \hat{O}]^\dagger [\hat{H}, \hat{O}])}{\text{Tr}(\hat{O}^\dagger \hat{O})}, \quad (6)$$

for operators \hat{O} that are sums of terms with support on a subsystem S of \mathcal{S} . We exclude the identity from this sum, as it trivially commutes with \hat{H} . This requirement is equivalent to requiring \hat{O} be traceless.

We now expand \hat{O} as a linear combination of operators from a basis for the set of traceless operators with support in S . A natural basis to take (for spin-1/2 systems) is the set \mathcal{P}_S of local products of Pauli operators. Each operator in \mathcal{P}_S is defined by choosing a Pauli operator from $\{\sigma_i^0, \sigma_i^x, \sigma_i^y, \sigma_i^z\}$ for each site i , and taking the tensor product of the combination. \mathcal{P}_S contains all possible such operators excluding the identity (which corresponds to the choice of σ^0 on all sites). Note that this

gives $4^{|S|} - 1$ elements in \mathcal{P}_S , making it a complete basis for operators with support on S when combined with the identity.

With this choice, we write

$$\hat{O} = \sum_{\Sigma \in \mathcal{P}_S} a_{\Sigma} \Sigma. \quad (7)$$

If we substitute this into (6), and write \vec{a} as the vector of individual a_{Σ} terms, we can rewrite [using the fact that operators in \mathcal{P}_S satisfy $\text{Tr}(\Sigma_i^{\dagger} \Sigma_j) = \delta_{i,j} 4^{|S|}$]

$$\rho = \frac{\vec{a}^{\dagger} \mathcal{C} \vec{a}}{\vec{a}^{\dagger} \vec{a}}, \quad (8)$$

where we introduced a ‘‘commutant matrix’’ \mathcal{C} , defined as

$$C_{\Sigma_i, \Sigma_j} = \text{Tr}([\hat{H}, \Sigma_i]^{\dagger} [\hat{H}, \Sigma_j]) 4^{-|S|}. \quad (9)$$

In order to minimize ρ , we need $\frac{\partial \rho}{\partial \vec{a}^{\dagger}} = 0$. Multiplying both sides of Eq. (8) by $\vec{a}^{\dagger} \vec{a}$, and differentiating by \vec{a}^{\dagger} , gives

$$\rho \vec{a} = \mathcal{C} \vec{a}. \quad (10)$$

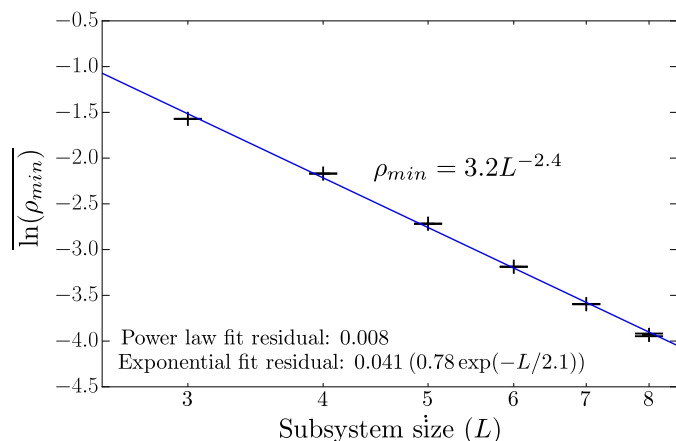
Thus, the problem reduces to finding the eigenvector \vec{v} of the commutant matrix \mathcal{C} corresponding to the lowest eigenvalue ρ . From this, we can construct the *best-localized* operator $\hat{O}_{\rho} = \sum_{\Sigma \in \mathcal{P}_S} \vec{v}_{\Sigma} \Sigma$.

Note that \mathcal{C} is positive semidefinite, so all eigenvalues are ≥ 0 . To show this, let us define $S + \hat{H}$ as the set of sites connected to S by terms in \hat{H} (i.e., site r lies in $S + \hat{H}$ if and only if there is some site s in S and some $\Sigma_{rs} \in \mathcal{P}_S$ that contributes to \hat{H} and acts nontrivially on both sites r and s). Then, if $\Sigma_i \in \mathcal{P}_S$, $[\hat{H}, \Sigma_i]$ acts nontrivially only on $S + \hat{H}$, and we can write

$$[\hat{H}, \Sigma_i] = \sum_{\Omega \in \mathcal{P}_{S+\hat{H}}} c_{\Omega, i} \Omega, \quad c_{\Omega, i} = \text{Tr}(\Omega^{\dagger} [\hat{H}, \Sigma_i]) 4^{-|S|}. \quad (11)$$

Then, if $\Sigma_j \in \mathcal{P}_S$ also, we have

$$\text{Tr}([\hat{H}, \Sigma_j]^{\dagger} [\hat{H}, \Sigma_i]) 4^{-|S|} = \sum_{\Omega} c_{\Omega, j}^{\dagger} c_{\Omega, i}. \quad (12)$$



Then, from the definition of \mathcal{C} , we can immediately write $C = C^{\dagger} C$ with

$$C_{\Omega, \Sigma_i} = c_{\Omega, i} = \text{Tr}(\Omega^{\dagger} [\hat{H}, \Sigma_i]) 2^{-|S|}. \quad (13)$$

This result implies that \mathcal{C} does not have any eigenvalues other than what we are searching for; ρ does indeed correspond to the eigenvalue closest to 0. Furthermore, it is computationally far easier to construct the matrix C and then calculate $C = C^{\dagger} C$ than directly calculate \mathcal{C} from its definition.

We briefly describe here the method used to construct Fig. 1. Given a fixed subsystem S of seven sites, we find the best-localized operator \hat{O}_{ρ} that corresponds to the lowest eigenvalue of the localization matrix. For each site $i \in S$, we calculate

$$1 - \|\text{Tr}_{S/i} \hat{O}_{\rho}\|^2, \quad (14)$$

where S/i is the set of points with i excluded. This is equivalent to finding the trace norm of the piece of \hat{O}_{ρ} with nontrivial support on i . This is then averaged for each site over 10000 disorder realizations.

V. EIGENVALUE DECAY IN THE DISORDERED XXZ SPIN CHAIN

To begin our study of the disordered XXZ spin chain via the method explained in Sec. IV, we demonstrate the difference in operator decay between the MBL and ergodic phases. In the Appendix A we argue that this corresponds to the exponential decay of the lowest eigenvalue of the commutant matrix.

We fix a subsystem \mathcal{S}_L of L spins (labeled $i = 1, \dots, L$) and take all combinations of Pauli matrices acting on these spins as a basis of operators from which we construct the commutant matrix in Eq. (9). This matrix is then independent of Hamiltonian terms that do not act on these spins, and so for our XXZ model, we need only embed this in a larger chain of $L + 2$ spins. It is then possible to obtain the lowest eigenvalue of the commutant matrix for $L = 8$ via iterative diagonalization, e.g., using ARPACK. This method is notably different in spirit from the diagonalization of the Hamiltonian, as the total size of the system can be assumed infinite and the boundary conditions play no role.

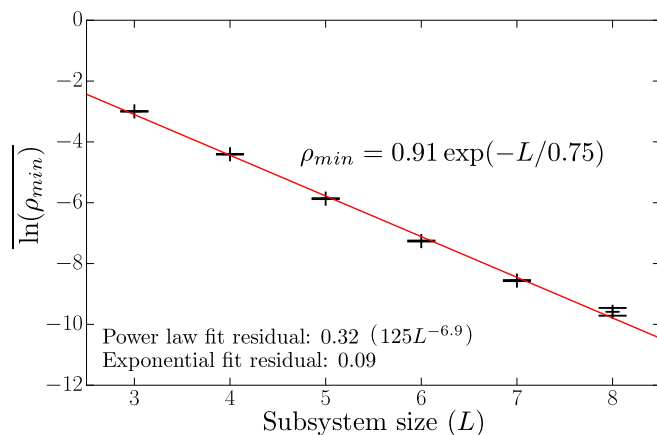


FIG. 2. Minimum eigenvalue of the commutant matrix for the disordered XXZ spin chain (5) as a function of the subsystem size L . Parameters are $V = 0.5t$ and $\Delta = 0.5t$ (left) or $\Delta = 5t$ (right). Points are averaged over either 10000 ($L = 3, 4, 5, 6, 7$) or 200 ($L = 8$) disorder realizations. Lines are polynomial (left) and exponential (right) fits to the points. Fit parameters are given in the figures.

In Fig. 2, we plot the disorder-averaged lowest eigenvalue of the commutant matrix for the range $L = 3$ to $L = 8$, in the ergodic ($V/t = 0.5, \Delta/t = 0.5$) and MBL ($V/t = 0.5, \Delta/t = 5$) regimes. The data shows *typical* lowest eigenvalues of the commutant matrix (i.e., we average the logarithm of the lowest eigenvalue). Similar results are obtained when averaging the eigenvalue itself. In both cases, we attempt linear fits of $\ln(y) = ax + b$ (exponential decay) and $\ln(y) = a \ln(x) + b$ (power-law decay). Found values of a and b are given in the figure. In the ergodic phase, the power-law fit has a three times smaller residual than the exponential, while in the MBL phase, the reverse occurs. Beyond the difference in the fit discrepancies, the power-law fit to the results of the MBL system has almost L^{-7} decay, which appears implausible. The power-law decay in the ergodic phase ($L^{-2.4}$) matches similar trends in other ergodic systems (Ref. [28] finds a $L^{-2.6}$ decay in a different spin chain within an ergodic phase).

VI. EIGENVALUE COUNTING IN THE COMMUTANT MATRIX

In an ergodic system, Ref. [28] found a single slow operator with an eigenvalue that scales as a power-law function of localization area L . By contrast, the MBL phase is distinguished by the presence of not one, but extensively many, local integrals of motion. This suggests that one should not only focus on the “ground state” of the commutant matrix, but on the entire low-eigenvalue manifold of eigenvectors. In this section, we analyze the structure of this manifold and provide counting rules for the complete set of localized operators that span this subspace.

In Fig. 3, we plot the averaged low-lying spectrum of the commutant matrix for a subsystem of length $L = 7$, in the MBL ($\Delta = 8t$), ergodic ($\Delta = 1t$), and integrable ($\Delta = 0$) regimes (all with $V = 0.5t$). We see that at the lowest

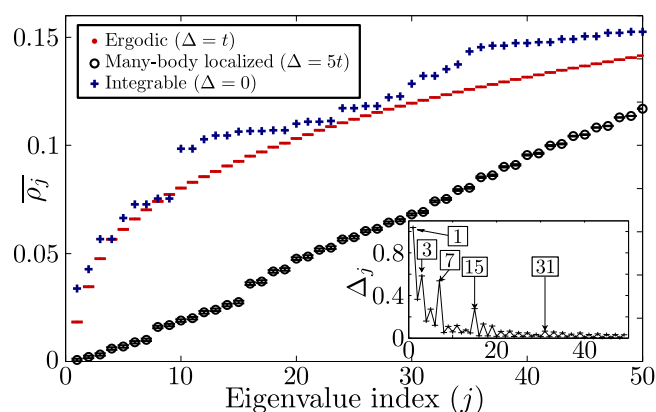


FIG. 3. The lowest 50 eigenvalues of the commutant matrix for systems in the integrable (red), ergodic (blue), and MBL (black) phases, with a subsystem of seven sites. The results for the integrable system are exact; those for the ergodic and MBL systems are averaged over 5000 disorder realizations (error bars are shown in figure). Inset: Relative difference between the eigenvalues of the commutant matrix in the MBL phase, given by Eq. (15). Points that correspond to changes between the distance d_l in Fig. 4 have been marked, and a local maximum is visible at each point.

eigenvalues, the trend of the points in the MBL phase is slightly concave up; a large number of eigenvalues are clustered near zero. By comparison, the ergodic and integrable spectra are strongly concave down, and the lowest eigenvalue is atypically small.

Of interest is the clear presence of gaps in the MBL phase, which are not present when the system is ergodic. In order to display these more clearly, in the inset of Fig. 3 we plot the relative difference,

$$\Delta_j = 2 \frac{\rho_{j+1} - \rho_j}{\rho_{j+1} + \rho_j} \quad (15)$$

(the specific labeling of the peaks in this plot is explained below). Interestingly, gaps are also present in the integrable limit, where new conserved operators have recently been constructed [38–40]. Pairs of degenerate eigenvalues in the integrable system are due to the $\sigma_x \leftrightarrow \sigma_y$ symmetry in the commutant matrix (these operators exist in separated blocks of the commutant matrix that are coupled by the disorder terms σ_z).

To explain the presence of gaps in the MBL case, let us consider the L τ_i operators present in an L -site subsystem \mathcal{S} in the MBL phase. If we select a subset of sites $S \subset \mathcal{S}$, the product $P_S = \prod_{i \in S} \tau_i$ will have exponentially decaying tails, as we can bound

$$\|\partial P_{S,l}\| < |S|A \exp(-l/\xi), \quad (16)$$

with $A \exp(-l/\xi)$ the bounding function for the exponential decay of the LIOMs (which is system dependent), and l defined by the distance from the edge of S . This gives a total of $2^L - 1$ LIOMs centered in \mathcal{S} . Each of these LIOMs should be approximated by a different eigenvector of the commutant matrix [as they satisfy $\text{Tr}(P_S^\dagger P_{S'}) = \delta_{S,S'} 4^{|S|}$]. If the approximation is good, the size of the corresponding eigenvalue will be proportional to the size of the tail of the LIOM that lies outside \mathcal{S} , as per Eq. (4). The size of this tail grows exponentially with the minimum distance d_S between any $i \in S$ and the boundary. Thus, we would expect groups of eigenvalues corresponding to P_S operators with different d_S .

On the small system sizes studied, the situation is slightly complicated by disorder, which breaks the reflection symmetry within \mathcal{S} , causing one side to be “favored” for decay over the other. To understand this, let us cut \mathcal{S} into left (\mathcal{S}_L) and right (\mathcal{S}_R) sides. A fixed disorder realization $\{h_i\}$ will be stronger on one of the two sides; this can be roughly characterized by the standard deviation of $\{h_i, i \in \mathcal{S}_a\}$ for $a = L, R$. Let us assume, without loss of generality, that the disorder on \mathcal{S}_L is stronger than that on \mathcal{S}_R . Then, LIOM tails will have decay of the form $A_L \exp(-l/\xi_L)$ on the left but $A_R \exp(-l/\xi_R)$ on the right, with $\xi_R > \xi_L$. This implies that the operators P_S that achieve d_S on the right will have larger tails [$\propto \exp(-l/\xi_R)$] outside the boundary than operators $P_{S'}$ that achieve $d_{S'}$ on the left but *not* the right [with tail size $\propto \exp(-l/\xi_L)$].

Continuing our assumption of stronger decay on the left, in Fig. 4 we order the $2^5 - 1 = 31$ P_S operators by their distances d_L and d_R . We also include a schematic of the scenario described above. Unless $\xi_R \gg \xi_L$, we would expect $\exp(-l/\xi_L) > \exp[-(l+1)/\xi_R]$, and so the tail size should be determined first by $\min(d_L, d_R)$ and second by whether $\min(d_L, d_R) = d_R$. We see that this gives groups corresponding

j	2	3	4	5	6	d_i	j	2	3	4	5	6	d_i
1			τ			3	16					τ	1*
2		τ				2	17				τ	τ	1*
3		τ	τ			2	18			τ	τ	τ	1*
4				τ		2*	19			τ	τ	τ	1*
5			τ	τ		2*	20		τ		τ	τ	1*
6		τ		τ		2*	21		τ		τ	τ	1*
7		τ	τ	τ		2*	22		τ	τ		τ	1*
8	τ					1	23		τ	τ	τ	τ	1*
9	τ	τ				1	24	τ				τ	1*
10	τ	τ	τ			1	25	τ			τ	τ	1*
11	τ	τ	τ			1	26	τ		τ		τ	1*
12	τ			τ		1	27	τ		τ	τ	τ	1*
13	τ		τ	τ		1	28	τ	τ			τ	1*
14	τ	τ		τ		1	29	τ	τ		τ	τ	1*
15	τ	τ	τ	τ		1	30	τ	τ	τ		τ	1*
							31	τ	τ	τ	τ	τ	1*

FIG. 4. Top: Table of some combinations of local integrals of motion P_S in a subsystem of width 7 in a spin chain. On each row, the single τ operators in the product $\prod_{i \in S} \tau_i$ are indicated, and the distance to the boundary d_S is written on the right column. An asterisk denotes when the distance d_S is to the boundary of slower decay (see text). Bottom: Example of a P_S operator; in this case, we have $P_S = \tau_3 \tau_4$. If P_S were approximated by our method inside the boxed region, the closest distance to the boundary would be $d_S = 2$ on the left, and we would expect the average $\|[[P_S], \hat{H}]\| \propto \exp(-2/\xi)$, with ξ the localization length.

exactly to the gaps seen in the MBL phase in Fig. 3. This shows that our method can approximate not just one, but several LIOMs.

The argument above does not explain why the peak height drops in Fig. 3 (inset). We attribute this to interference from operators in the system other than the LIOMs, which may yet have a slow power-law decay. As the system size increases, we would expect these differences to grow until some saturation point, which would depend on the localization length ξ . We would also expect the symmetry-breaking effect to disappear. Unfortunately, we have not been able to study this due to our restrictions on subsystem size.

VII. ALIOM OPERATOR CHARACTERISTICS

If $\|\text{Tr}_{S,l} \tau_i\| / \|\tau_i\| \approx 1$, it is reasonable to expect the ALIOM $t_{i,l}$ to be a good approximation to τ_i . We now investigate whether $t_{i,l}$ has the characteristics of a local spin.

At large disorder strength, the tails of τ_i should sit well within a large subsystem, and so $t_{i,l}$ should itself be a sum of a small number of Pauli operators. We measure this by a means similar to the traditional inverse participation ratio for states: with, again, $t_{i,l} = \sum_I a_I \Sigma_I$, we can write

$$R(t_{i,l}) = \sum_I |a_I|^4. \quad (17)$$

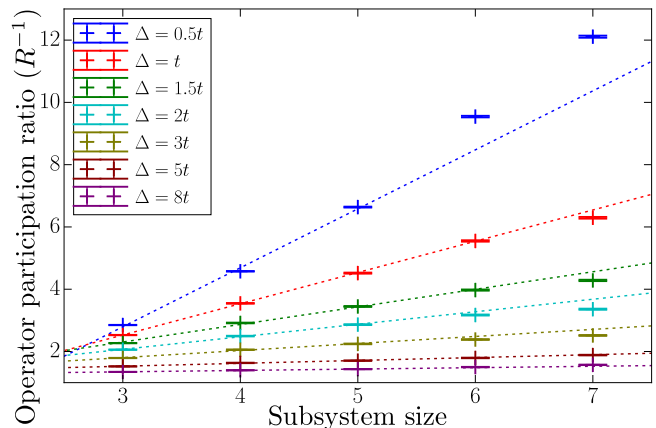


FIG. 5. Operator participation ratio, given by Eq. (17), for various disorder strengths as a function of subsystem size. All points are an average over 10000 disorder realizations (error bars in plot). Lines are a fit to the first few points to emphasize trends of curving above or below a linear increase.

This, or rather R^{-1} , can be thought of as a measure of how many of the Pauli operators contribute to $t_{i,l}$; if we have n equally contributing operators $t_{i,l} = \frac{1}{\sqrt{n}} \sum_{j=1}^n \sigma_j$, then $R^{-1}(t_{i,l}) = n$. In general, with uneven contributions, $R^{-1}(t_{i,l})$ is not an integer, but it is bounded below by 1. We dub $R^{-1}(t_{i,l})$ the *operator participation ratio* (OPR). Note that the value of R is dependent on the basis chosen for the commutant matrix.

In Fig. 5, we plot the disorder average of R^{-1} for the best-commuting operator as a function of subsystem size, for a range of disorders. We fit lines to the first few points of each curve to see the comparative trend. Deep within the localized phase, the OPR achieves a limiting value immediately and grows very little afterwards. At lower disorder, the OPR grows for a while before either bending sublinear ($\Delta = 1$) or to a higher power law ($\Delta = 0.5$).

Given a local integral of motion $\tau_i = \sum_l \partial \tau_{i,l}$, if $\|\partial \tau_{i,l}\| = A \exp(-l/\xi)$, then our normalization constant is

$$A^2 = \sum_{j=1}^{\infty} \exp(-2j/\xi). \quad (18)$$

The value of the limit of our OPR curves [$R^{-1}(\tau_i)$] depends on how many Pauli operators Σ_I from each $\partial \tau_{i,l}$ contribute to $\partial \tau_{i,l}$, making it impossible to explicitly determine. However, it is possible to make upper and lower bounds. The minimum value of the OPR is achieved if a single operator from each $\partial \tau_{i,l}$ contributes to $\partial \tau_{i,l}$, giving the following lower bound:

$$\begin{aligned} R_{\min}^{-1}(\tau_i) &= \frac{[\sum_{j=0}^{\infty} \exp(-2j/\xi)]^2}{\sum_{j=0}^{\infty} \exp(-4j/\xi)} \\ &= \frac{1 + \exp(-2/\xi)}{1 - \exp(-2/\xi)} = \coth(1/\xi), \end{aligned} \quad (19)$$

while a bound above for the OPR can be constructed by considering the case where only the largest operator (σ_i^z)

contributes,

$$R_{\max}^{-1}(\tau_i) = \left[\sum_{j=0}^{\infty} \exp(-2\xi^j) \right]^2 = \frac{1}{[1 - \exp(-2/\xi)]^2}. \quad (20)$$

For the case of $\Delta = 5t$, we have $R^{-1}(\tau_i) \approx 2$, which corresponds to a maximum possible decay length of 1.8 and a minimum of 1.6. This is over twice the decay length extracted from Fig. 2. This can be explained by the formation of singlets between pairs of LIOMs, as these would have a large effect on the OPR, but not the eigenvalue.

Whether $\lim_{l \rightarrow \infty} R^{-1}(\tau_{i,l})$ is finite is directly dependent on whether $\tau_i = \lim_{l \rightarrow \infty} \tau_{i,l}$ is normalizable, for we can write, for any l ,

$$\lim_{l \rightarrow \infty} R(\tau_{i,l}) \geq \max(a_l^l) \frac{\|\tau_{i,l}\|}{\|\tau_i\|} > 0, \quad (21)$$

with $\max(a_l^l)$ the largest coefficient of the expansion of $\tau_{i,l}$ in our Pauli basis. This is in turn related to whether the tails of τ_i decay faster than $1/L$ (as $1/x^n$ is a convergent series only when $n > 1$). This presents a conflict; the $\Delta = 0.5t$ curve is clearly diverging, while the corresponding eigenvalues in Fig. 2 decay at a rate $1/L^{2.3} > 1/L$. This is resolved by noting that the commutator $[\hat{H}, t_{i,l}]$, while bounded above by the size of the terms $\tau_{i,l}$ on the boundary, has no bound below and can possibly scale faster than the corresponding operator. In the Appendix A we demonstrate that this cannot cause a signature of exponential decay without the presence of LIOMs τ_i .

It is unfortunately not possible to distinguish whether the sublinear behavior of all other curves converges or is logarithmically divergent; as such we cannot accurately determine the decay behavior close to the phase transition. It should be noted though that our method does not fix a temperature on the system. Well within the localized phase, where we approximate well the LIOMs, the operators have nonzero eigenspectrum across the entire Hilbert space, and so are at infinite temperature. However, in the presence of a mobility edge, an operator projected into the localized energy band would have better scaling than one at infinite temperature, and so this may well be observed in our method. This would imply that we expect to see signatures of the MBL phase as soon as a non-negligible fraction of the Hilbert space becomes localized. However, as a sum of a large number of operators will be required to project out delocalized energy levels, we can expect the OPR to be large (but finite) in these situations, rather than immediately collapsing to near 1.

VIII. CORRELATIONS BETWEEN OPERATOR MEASURES

The model studied throughout this paper draws the disorder on any site from a uniform distribution, but this distribution is not necessarily represented well by the realization drawn for any small subsystem we study. In the MBL phase, the density of the actual LIOMs τ_i within such subsystems, and the size of the tails outside, will thus fluctuate between disorder realizations. If we attempt to approximate the real LIOMs τ_i in the MBL phase with our method, we should thus expect “better” performance for some realizations over others, in the

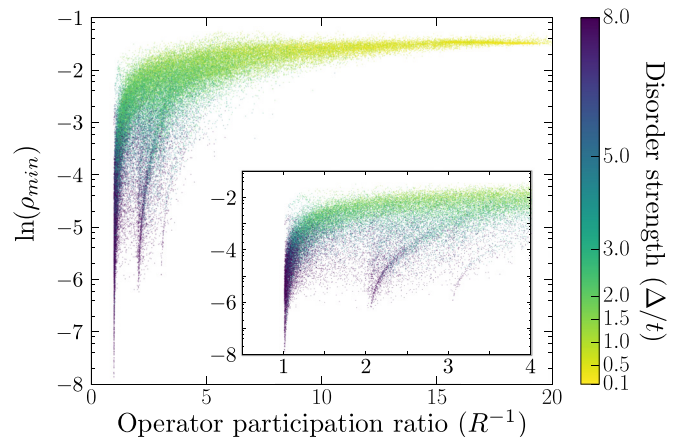


FIG. 6. Scatter plot of 10000 individual realizations of the lowest eigenvalue of the commutant matrix against the corresponding operator participation ratio for a range of disorders, colored by the disorder strength Δ . All points are calculated over a system size of length 7.

sense that the overlap between the actual LIOM τ_i and our ALIOM $t_{i,l}$ will be larger. This implies that in the MBL phase, all measures of the “goodness” of our approximation should roughly correlate with each other.

We have already used the commutant matrix eigenvalues and the OPR of the resultant operators to investigate the disordered XXZ spin chain. In the MBL phase, when the eigenvalue ρ corresponding to an ALIOM $t_{i,l}$ is small, we expect this to be a better approximation to the LIOM τ_i . Furthermore, we expect operators with smaller ρ to have smaller tails, and thus faster decay. This should then correspond with a smaller OPR. As $R^{-1}(\tau_i)$ is dependent on the local structure (and thus the local disorder as well as the tails), we expect a reasonable, but not perfect correlation.

In Fig. 6, we make a colored scatter plot of these two properties for a set of 10000 disorder realizations with $L = 7$, and a range of disorder strengths. We color the plots by their disorder strength. We find a very strong correlation between the OPR and the eigenvalue; our points tend to follow a hyperboloid shape between the two limits of large and small OPR/eigenvalue. The turning point in the hyperboloid occurs at around $\Delta = 2t$.

At strong Δ and small ρ_{\min} , we also notice a clear quantization of the OPR, with features that trend to $\text{OPR} = n$ as $\rho \rightarrow 0$ (this can be seen more clearly in the inset). These most likely correspond to the formation of singlets between two LIOMs. We note (as mentioned before) that these cause a large change to the OPR without immediately affecting the eigenvalue.

In Secs. V and VII, we made two independent measures of the decay length for the XXZ chain with parameters $\Delta = 5t$, $V = 0.5t$. These found $\xi = 0.75$ and $\xi = 1.6$ – 1.8 , respectively, a twofold deviation. We claim that the leading source of this deviation comes from the singlet presence, which should increase the IPR while not significantly affecting the eigenvalue. In order to test this claim, we remove the fraction of operators with $\text{OPR} > 2$, and recalculate both measures. The ξ calculated by the eigenvalue scaling (Sec. V) shifts to

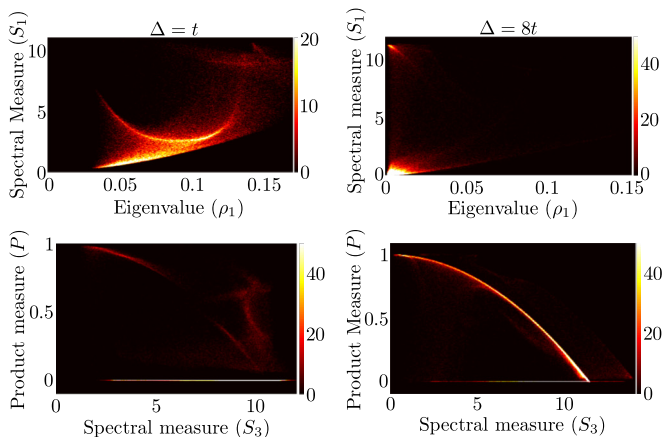


FIG. 7. 2D histogram plot of various correlations between eigenvalues and the spectral and product measures defined in the text. System parameters are $V = 0.5t$ and $\Delta = t$ (left) or $\Delta = 8t$ (right). All data taken at a resolution of 200×200 bins, from a total of 100000 samples.

$\xi = 0.72$, but the ξ calculated by the OPR (Sec. VII) shifts to 0.95–0.98, reducing the discrepancy to 25%. We believe the remaining difference comes from finite-size effects of the approximation.

The LIOMs τ_i are expected to be spin operators; they should divide the Hilbert space into two equal eigenspaces (corresponding to ± 1 eigenvalues). Equivalently, being traceless, they should square to the identity. Furthermore, the results from Sec. VI suggest that we should be able to isolate not only multiple τ_i operators, but also their products $\tau_i \tau_j$. If this is the case, and $\tau_i^2 = \mathbf{1}$, we can calculate $(\tau_i \tau_j) \times \tau_i \times \tau_j = \mathbf{1}$ also.

We propose two measures of the ability of the approximated $t_{i,l}$ operators to mimic these algebraic properties: the spectral measure $S(\hat{O})$ and the product measure $P(\hat{O}_1, \hat{O}_2, \hat{O}_3)$,

$$S(\hat{O}) = \sum_{\lambda_{\hat{O}}} ||\lambda_{\hat{O}}| - 1|, P(\hat{O}_1, \hat{O}_2, \hat{O}_3) = \text{Tr}[\hat{O}_1 \hat{O}_2 \hat{O}_3], \quad (22)$$

with $\lambda_{\hat{O}}$ the eigenvalues of \hat{O} . In Fig. 7, we plot a selection of heat maps (2D histograms) of correlations between these measures for a subsystem of size 4 with $\Delta = t$ (left) and $\Delta = 8t$ (right). The tails of operators centered at the border of the subsystem cannot exponentially decay before being cut by the approximation, so here we would expect to see a signature of two τ_i operators and their product. This should then produce in the $\Delta \rightarrow \infty$ limit a spectral measure $S_1 = S_2 = S_3 = 0$ for the first three eigenvalues, and $P_{123} = 1$.

In the MBL phase, we see that the sample is split into two sets by the spectral measure S_1 . The majority of the sample lies at very small S_1 and eigenvalue, and has a rough correlation between the two of them. A small fraction of the sample sits with spectral measure roughly 11.3. To explain this number, we note that the Hilbert space of a four-site subsystem is 16 dimensional, and so an operator of the form $\frac{1}{\sqrt{2}}(\sigma_i^z + \sigma_j^z)$ would have an eightfold degenerate 0 eigenvalue, and two

fourfold degenerate $\pm\sqrt{2}$ eigenvalues. Then, we can calculate

$$S \left[\frac{1}{\sqrt{2}}(\sigma_i^z + \sigma_j^z) \right] = 8 + 8(\sqrt{2} - 1) \approx 11.3. \quad (23)$$

This then gives support to our explanation of the quantization that appears in Fig. 6 (our Hilbert space is 16 dimensional for four sites, and so half of the 16 eigenvectors would have eigenvalue 0, and the other half $\sqrt{2}$).

Furthermore, we observe a very strong correlation in the MBL phase between the spectral measure of the operator with the third-lowest eigenvalue and the product measure, corresponding to around 90% of the sample. We argue that in these samples, the system is approximating two of the τ_i operators very well; the third is mixed somewhat with the delocalized operators, and as our approximation to the LIOM becomes better, it simultaneously matches the spectral and algebraic properties. We should note here that a large fraction of the sample does have a product measure fixed to 0 regardless of the spectral measure (about 10% of the entire sample). This complete lack of overlap possibly implies that some operator other than an approximation to the LIOMs had low enough eigenvalue to sit between them.

In the ergodic phase ($\Delta = t$), there is some trend towards a lower spectral measure with lower eigenvalue. Moreover, there is an intriguing, quadraticlike feature that appears in the center of the plot, which is currently not understood. Finally, we also note that though the product measure is almost always approximately zero, there is a faint but visible remnant of the correlation that was dominant in the MBL phase.

IX. CONCLUSIONS AND OUTLOOK

We have explicitly constructed a complete set of LIOMs, confirming the theoretical picture of the many-body localized phase of a spin system. Our method constructs the operator within a finite region that best commutes with the Hamiltonian, which at strong enough disorder is similar to cutting off the exponentially decaying tails of the corresponding LIOM. The size of the cut tails is reflected in the norm of the commutator of this operator with the Hamiltonian, which is the eigenvalue of the constructed commutant matrix. We have observed exponential decay of the lowest eigenvalue of the commutant matrix in the MBL system (which is not present in the ergodic system), and we have proved that this is an unambiguous sign of an exponentially decaying operator which commutes with the Hamiltonian.

We have also studied the properties of these operators, quantifying their locality via an operator participation ratio, and their algebraic properties via measures of how well they approximate a set of Pauli spins. In the MBL phase, the lowest eigenvector of the commutant matrix has been shown to have the algebraic properties of a dressed spin (with the exception of a small fraction of disorder realizations where it is clearly a linear combination of a few spins). Furthermore, we have shown that this method can reproduce the entire spectrum of LIOMs living in the studied subsystems. These could alternatively be reproduced by translating the subsystem to be centered on other sites.

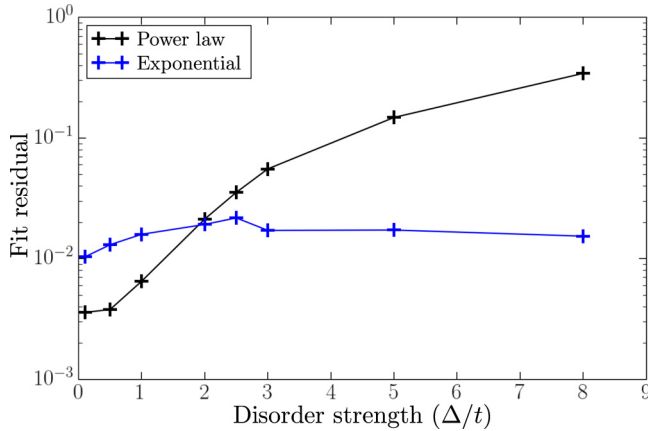


FIG. 8. Plot of the residual from the power-law and ergodic fits as in Fig. 2, for a range of disorder values. Each fit was taken using five data points from $L = 3$ to $L = 7$, with each data point being in turn averaged over 1000 disorder realizations.

The disadvantage of the method is that it scales even more prohibitively than full diagonalization of the Hamiltonian. The commutant matrix (9) has size $(4^L - 1) \times (4^L - 1)$, making the exact method unviable for anything beyond small L . One could employ various schemes of truncating the basis of the operator space, either by some physical insight dependent on the particular model and/or by using matrix product state (MPS) techniques [28]. For example, a natural way of truncating the basis could be to keep only the terms that explicitly appear in the Hamiltonian and those obtained from them via the action of the commutator. This should allow one to extend the method to much larger systems.

Finally, a natural future direction would be to use the method presented here to study the localization-delocalization transition. The residual of the exponential and power-law fits in Sec. V provide a means of distinguishing whether the system is in the ergodic or MBL phase. For $V = t$, we compare the two residuals in Fig. 8, and we extract a phase transition at around $\Delta = 2t$. The critical disorder strength predicted is thus lower than in other studies, which find a mobility edge between $\Delta = 2t$ and $\Delta = 4t$ [33,34]. This would suggest that our method “sees” a phase transition as soon as any of the eigenspaces is localized. For this to occur, LIOMs must appear with the onset of the MBL phase in any part of the eigenspectrum. This does not appear impossible, as there exist many operators that act only on a small part of the eigenspectrum (e.g., see Ref. [41]), and our method would return a positive result for the MBL phase if a single one was localized. Recent work [42] appears to suggest that some LIOMs could indeed be well defined in the presence of the mobility edge, but more detailed investigations are needed to resolve this interesting question.

ACKNOWLEDGMENTS

We acknowledge useful discussions with Tomaž Prosen. We also thank David Huse for several helpful comments on the paper. T.E.O. and G.V. acknowledge support by NSERC (discovery grant). T.E.O. also acknowledges support from the Foundation for Fundamental Research on Matter

(FOM), the Netherlands Organization for Scientific Research (NWO/OCW), and an ERC Synergy Grant. G.V. also acknowledges support by the Simons Foundation (Many Electron Collaboration). D.A. acknowledges support by SNF and by the Alfred Sloan Foundation. Research at the Perimeter Institute is supported by the Government of Canada through Industry Canada and by the Province of Ontario through the Ministry of Research and Innovation.

APPENDIX: EQUIVALENCE OF TWO TYPES OF EXPONENTIAL DECAY

We wish to show that exponential decay of the lowest eigenvalue of the commutant matrix is a definite signal of localized integrals of motion. In the presence of exponentially localized LIOMs τ_i , we can bound

$$\rho = \frac{\|[\hat{H}, t_{i,l}]\|^2}{\|t_{i,l}\|^2} \leq \frac{\|[\hat{H}, \tau_{i,l}]\|^2}{\|\tau_{i,l}\|^2} \leq C' \exp(-l/\xi). \quad (\text{A1})$$

This implies that our system does not have LIOMs unless ρ scales exponentially. However, we have not yet demonstrated that an exponentially scaling ρ must imply the presence of exponentially scaling modes. In this section, we argue why this must be the case.

There are two issues to address. Consider “amputating” an integral of motion τ_i to give successive $\tau_{i,l}$. We have that $[\tau_{i,l}, \hat{H}] = -[\tau_i - \tau_{i,l}, \hat{H}]$, and all terms in $\tau_{i,l}$ have support only within $S_{i,l}$, but all terms in the remainder $\tau_i - \tau_{i,l}$ have explicit support outside $S_{i,l}$. If \hat{H} is local, the commutator $[\hat{H}, \tau_i]$ cannot have support far from that of τ_i , but terms in $[\tau_i - \tau_{i,l}, \hat{H}]$ must retain support close to the outside of $S_{i,l}$, and so each term must live on the boundary. However, terms in $[\tau_{i,l}, \hat{H}]$ have no such restriction, as $t_{i,l}$ is not constructed as part of a larger operator. This gives $t_{i,l}$ additional degrees of freedom in its construction, and we must show (Proposition 1) that these cannot give exponential decay that would not otherwise be present.

Furthermore, there is no automatic lower bound for $\|[\tau_{i,l}, \hat{H}]\|/\|\partial\tau_{i,l}\|$. We must show that this fraction cannot show exponential decay in l , as otherwise the commutator will see exponential decay while the operator itself does not. Note that this is not obviously always the case, as in the MBL phase the fraction $\|[\tau_{i,l}, \hat{H}]\|/\|\tau_{i,l}\|$ does decay exponentially. Indeed, a general proof of this has proved elusive. In Proposition 2, we give a lower bound for $\|[\tau_{i,l}, \hat{H}]\|/\|\partial\tau_{i,l}\|$ in the case of the XXZ spin chain, and point out how this cannot be immediately extended to a general model (although we hypothesize that such a lower bound always exists in nontrivial cases).

Proposition 1. We note that the commutator between a local Hamiltonian and an operator increases linearly with the number of nontrivial tensor factors in the operator. This implies that the difference between $\tau_{i,l}$ and $t_{i,l}$ must be a sum of operators with small numbers of nontrivial tensor factors. The number \mathcal{N}_l of such operators increases polynomially in l , and so at large l , $\mathcal{N}_l/\mathcal{N}_{l+1} \approx 1$, and our spreading operation must be independent of l (i.e., it must converge in norm as a superoperator). This in turn implies that at large l , we should

have

$$\frac{||[t_{i,l}, \hat{H}]||}{||[\tau_{i,l}, \hat{H}]||} \approx \frac{||[t_{i,l-1}, \hat{H}]||}{||[\tau_{i,l-1}, \hat{H}]||}, \quad (\text{A2})$$

or, to be specific, the relative difference should go to zero in the $l \rightarrow \infty$ limit. Rearranging this gives

$$\lim_{l \rightarrow \infty} \frac{||[t_{i,l}, \hat{H}]||}{||[t_{i,l-1}, \hat{H}]||} = \lim_{l \rightarrow \infty} \frac{||[\tau_{i,l}, \hat{H}]||}{||[\tau_{i,l-1}, \hat{H}]||}. \quad (\text{A3})$$

This implies that power-law decay of $||[\tau_{i,l}, \hat{H}]||$ (right-hand side equal to 1) is equivalent to power-law decay of $||[t_{i,l}, \hat{H}]||$ (left-hand side equal to 1).

Proposition 2. We note that for τ_i to be an integral of motion, $[\tau_{i,l}, \hat{H}]$ must consist of terms from $[\partial \tau_{i,k}, \hat{H}]$ for k near l . So it remains to check whether $||[\partial \tau_{i,l}, \hat{H}]||/||\partial \tau_{i,l}||$ can scale exponentially. For the XXZ Hamiltonian, this is not the case: an operator with nontrivial tensor factors on site $i+l$ can be written in the form

$$\hat{\delta}_{i+l} = \sum_l A_l \sum_{a=x,y,z} c_{a,l} \sigma_{i+l}^a, \quad (\text{A4})$$

with A_l some operator acting on sites to the left of $i+l$. But then, the commutator of $\hat{\delta}_{i+l}$ with the $\sigma_{i+l}^x \sigma_{i+l+1}^x + \sigma_{i+l}^y \sigma_{i+l+1}^y$ term gives

$$\sum_l 2A_l (c_{x,l} \sigma_{i+l}^z \sigma_{i+l+1}^y - c_{y,l} \sigma_{i+l}^z \sigma_{i+l+1}^x + c_{z,l} \sigma_{i+l}^y \sigma_{i+l+1}^x - c_{z,l} \sigma_{i+l}^x \sigma_{i+l+1}^y), \quad (\text{A5})$$

giving a minimum bound,

$$||[\hat{\delta}_{i+l}, \hat{H}]||^2 / ||\hat{\delta}_{i+l}||^2 \geq 2. \quad (\text{A6})$$

The same argument applies for the operators acting on site $i-l$, and so $||[\partial \tau_{i,l}, \hat{H}]||/||\partial \tau_{i,l}||$ is bounded below, and exponential decay of the numerator implies exponential decay of the denominator.

Note that Proposition 2 relies explicitly on having two terms in the Hamiltonian that act on sites i and $i+1$. In general, this is not the case, and so this proof cannot be immediately extended to all systems with local Hamiltonians, although we believe the result to be true.

-
- [1] J. M. Deutsch, *Phys. Rev. A* **43**, 2046 (1991).
[2] M. Srednicki, *Phys. Rev. E* **50**, 888 (1994).
[3] M. Rigol, V. Dunjko, and M. Olshanii, *Nature (London)* **452**, 854 (2008).
[4] P. W. Anderson, *Phys. Rev.* **109**, 1492 (1958).
[5] I. V. Gornyi, A. D. Mirlin, and D. G. Polyakov, *Phys. Rev. Lett.* **95**, 206603 (2005).
[6] D. Basko, I. Aleiner, and B. Altshuler, *Ann. Phys.* **321**, 1126 (2006).
[7] V. Oganesyan and D. A. Huse, *Phys. Rev. B* **75**, 155111 (2007).
[8] M. Znidarić, T. Prosen, and P. Prelovsek, *Phys. Rev. B* **77**, 064426 (2008).
[9] J. H. Bardarson, F. Pollmann, and J. E. Moore, *Phys. Rev. Lett.* **109**, 017202 (2012).
[10] M. Serbyn, Z. Papić, and D. A. Abanin, *Phys. Rev. Lett.* **111**, 127201 (2013).
[11] R. Vosk and E. Altman, *Phys. Rev. Lett.* **110**, 067204 (2013).
[12] M. Serbyn, M. Knap, S. Gopalakrishnan, Z. Papić, N. Y. Yao, C. R. Laumann, D. A. Abanin, M. D. Lukin, and E. A. Demler, *Phys. Rev. Lett.* **113**, 147204 (2014).
[13] D. A. Huse, R. Nandkishore, V. Oganesyan, A. Pal, and S. L. Sondhi, *Phys. Rev. B* **88**, 014206 (2013).
[14] B. Bauer and C. Nayak, *J. Stat. Mech.: Theory Expt.* (2013) P09005.
[15] Y. Bahri, R. Vosk, E. Altman, and A. Vishwanath, [arXiv:1307.4092](https://arxiv.org/abs/1307.4092).
[16] J. A. Kjäll, J. H. Bardarson, and F. Pollmann, *Phys. Rev. Lett.* **113**, 107204 (2014).
[17] M. Serbyn, Z. Papić, and D. A. Abanin, *Phys. Rev. Lett.* **110**, 260601 (2013).
[18] D. A. Huse, R. Nandkishore, and V. Oganesyan, *Phys. Rev. B* **90**, 174202 (2014).
[19] B. Sutherland, *Beautiful Models: 70 Years of Exactly Solved Quantum Many-body Problems* (World Scientific, Singapore, 2004).
[20] A. Chandran, I. H. Kim, G. Vidal, and D. A. Abanin, *Phys. Rev. B* **91**, 085425 (2015).
[21] I. H. Kim, A. Chandran, and D. A. Abanin, [arXiv:1412.3073](https://arxiv.org/abs/1412.3073).
[22] J. Z. Imbrie, *J. Stat. Phys.* **163**, 998 (2016).
[23] V. Ros, M. Müller, and A. Scardicchio, *Nucl. Phys. B* **891**, 420 (2015).
[24] L. Rademaker and M. Ortuño, *Phys. Rev. Lett.* **116**, 010404 (2016).
[25] C. Monthus, *J. Phys. A* **49**, 305002 (2016).
[26] R.-Q. He and Z.-Y. Lu, [arXiv:1606.09509](https://arxiv.org/abs/1606.09509).
[27] S. Inglis and L. Pollet, *Phys. Rev. Lett.* **117**, 120402 (2016).
[28] H. Kim, M. C. Bañuls, J. I. Cirac, M. B. Hastings, and D. A. Huse, *Phys. Rev. E* **92**, 012128 (2015).
[29] A. Pal and D. A. Huse, *Phys. Rev. B* **82**, 174411 (2010).
[30] H. Kim, T. N. Ikeda, and D. A. Huse, *Phys. Rev. E* **90**, 052105 (2014).
[31] H. Kim and D. A. Huse, *Phys. Rev. Lett.* **111**, 127205 (2013).
[32] M. Žnidarić, A. Scardicchio, and V. K. Varma, *Phys. Rev. Lett.* **117**, 040601 (2016).
[33] D. J. Luitz, N. Laflorencie, and F. Alet, *Phys. Rev. B* **91**, 081103 (2015).
[34] M. Serbyn, Z. Papić, and D. A. Abanin, *Phys. Rev. X* **5**, 041047 (2015).
[35] W. De Roeck, F. Huveneers, M. Müller, and M. Schiulaz, *Phys. Rev. B* **93**, 014203 (2016).
[36] M. Friesdorf, A. H. Werner, W. Brown, V. B. Scholz, and J. Eisert, *Phys. Rev. Lett.* **114**, 170505 (2015).
[37] J. Eisert, M. Cramer, and M. B. Plenio, *Rev. Mod. Phys.* **82**, 277 (2010).
[38] T. Prosen, *Phys. Rev. Lett.* **106**, 217206 (2011).
[39] R. G. Pereira, V. Pasquier, J. Sirker, and I. Affleck, *J. Stat. Mech.: Theory Expt.* (2014) P09037.
[40] E. Ilievski, M. Medenjak, and T. Prosen, *Phys. Rev. Lett.* **115**, 120601 (2015).
[41] M. Mierzejewski, T. Prosen, and P. Prelovsek, *Phys. Rev. B* **92**, 195121 (2015).
[42] S. D. Geraedts, R. Bhatt, and R. Nandkishore, [arXiv:1608.01328](https://arxiv.org/abs/1608.01328).