Comparing many-body localization lengths via nonperturbative construction of local integrals of motion

Editors' Suggestion

Comparing many-body localization lengths via nonperturbative construction of local integrals of motion

Pai Peng (彭湃),[1,2] Zeyang Li,[3,2] Haoxiong Yan,[2] Ken Xuan Wei,[3,2,4] and Paola Cappellaro[5,2,*]

[1]*Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*
[2]*Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*
[3]*Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*
[4]*IBM T.J. Watson Research Center, Yorktown Heights, New York 10598, USA*
[5]*Department of Nuclear Science and Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*

(Received 2 January 2019; revised manuscript received 8 November 2019; published 13 December 2019)

Many-body localization (MBL), characterized by the absence of thermalization and the violation of conventional thermodynamics, has elicited much interest both as a fundamental physical phenomenon and for practical applications in quantum information. A phenomenological model which describes the system using a complete set of local integrals of motion (LIOMs) provides a powerful tool to understand MBL but can usually be computed only approximately. Here we explicitly compute a complete set of LIOMs with a nonperturbative approach by maximizing the overlap between LIOMs and physical spin operators in real space. The set of LIOMs satisfies the desired exponential decay of the weight of LIOMs in real space. This LIOM construction enables a direct mapping from the real-space Hamiltonian to the phenomenological model and thus enables studying the localized Hamiltonian and the system dynamics. We can thus study and compare the localization lengths extracted from the LIOM weights, their interactions, and dephasing dynamics, revealing interesting aspects of many-body localization. Our scheme is immune to accidental resonances and can be applied even at the phase transition point, providing a tool to study the microscopic features of the phenomenological model of MBL.

DOI: 10.1103/PhysRevB.100.214203

I. INTRODUCTION

How a many-body quantum system thermalizes—or fails to do so—under its own interaction is a fundamental yet elusive problem. Localization serves as a prototypical example for the absence of thermalization, first studied in the noninteracting single-particle regime known as Anderson localization [1,2] and then revived in the context of interacting systems [many-body localization (MBL)] [3]. The existence of MBL as a phase of matter was demonstrated theoretically [4–6] and numerically [7–11]. Recently, the MBL phase was observed in cold atoms [12–17], trapped ions [18,19], and natural crystals using nuclear magnetic resonances [20]. Most characteristics of MBL, such as area law entanglement [21,22], Poisson level statistics [8,9], logarithmic growth of entanglement [7,16,23–27], and power law dephasing [28–32], can be understood via a phenomenological model that expresses the Hamiltonian in terms of a complete set of conserved quantities with local support. However, the explicit computation of such local integrals of motion (LIOMs) [21,25] and their interactions is a challenging task, complicated by the fact that the set of LIOMs is not unique. LIOMs have been calculated by the infinite-time averaging of initially local operators [33,34]; however, the obtained LIOMs do not form a complete basis. A complete basis of LIOMs can be obtained using perturbative treatment of interactions [5,35–38], using Wegner-Wilson flow renormalization [39], minimizing the commutator with the Hamiltonian [40,41], ordering the eigenstates with the greedy method [42], and prompting the infinite-time averaged LIOMs [43]. The previous methods either require strong disorder field strength or assume a cutoff of LIOMs in real space, so a complete numerical study of localization lengths is missing.

Here we design and implement a method to compute a complete set of binary LIOMs (i.e., with eigenvalues ± 1) in a nonperturbative way by maximizing the overlap with physical spin operators. This criterion enables a recursive determination, similar to quicksort, of the LIOMs matrix elements in the energy eigenbasis, without the need to exhaust all the eigenstate permutations, which would be prohibitive for system size $L > 5$. We verify that in the MBL phase the LIOMs are exponentially localized in real space and their interaction strength decays exponentially as a function of interaction range. This typical behavior is usually investigated by defining two characteristics lengths, the LIOM localization lengths and interaction localization lengths, which can be extracted from the two exponential behaviors. Deep in the MBL phase, the two localization lengths are well characterized by the inequality derived in Ref. [3]. Near the transition point between localized and delocalized phases, which our construction enables exploring, the interaction localization length diverges, while the LIOM localization length remains finite: this should be expected given the constraints imposed by our construction, even if it contradicts the inequality in Ref. [3]. The explicit form of the LIOMs further enables

*pcappell@mit.edu

exploring the system evolution, which has been shown to display a dephasing behavior with a power law decay that is characterized by a third characteristic length, the dynamical localization length [28]. Here we show that the LIOMs we derive display a similar dynamics to the physical spin operators, and we are able to extract the dynamical localization length from the power law dephasing process. Interestingly, we find that the dynamical localization length is much shorter than what would be given by a conjectured relationship to the above two localization lengths [3,28], suggesting that the dynamics depends not only on the typical value of LIOMs and their interactions but also on higher-order correlations.

II. ALGORITHM

To understand the construction algorithm, we first review the properties of integrals of motion in the many-body localized phase. LIOMs $\{\tau_z^j\}$ are diagonal in the Hamiltonian H eigenbasis, $[H, \tau_z^j] = 0$. A complete set of LIOMs can be related to physical spin operators (described by Pauli matrices σ^j) by a local unitary transformation $\tau_z^j = U\sigma_z^j U^\dagger$, which implies that (i) half of the eigenvalues of τ_z^j are $+1$ and the other half are -1, (ii) LIOMs are mutually independent (orthonormal) $\mathrm{Tr}(\tau_z^j \tau_z^k)/2^L = \delta_{jk}$, and (iii) the weight of τ_z^j decays exponentially in real space for localized Hamiltonians. In particular, property (ii) requires that, for any j, in either the $+1$ or -1 sector of τ_z^j, half of the diagonal elements of τ_z^k are $+1$, and the other half are -1 for all $k \neq j$. In other words, the $+1$ and -1 sectors of τ_z^j are effectively two manifolds that represent two instances of a new system with $L-1$ spins containing all sites except j.

With only constraints (i) and (ii), there are $2^L!/L!$ different sets of integrals of motion (IOMs), among which we want to find the most local one. However, enumerating the $2^L!/L!$ different sets and quantifying the localization of the related τ_z^j are numerically prohibitive. Instead of explicitly demanding the exponential localization, the key idea of our construction is to maximize the overlap of LIOMs and physical spin operators $\mathrm{Tr}(\tau_z^j \sigma_z^j)$, which are themselves local. This strategy enables a systematic and efficient way to find a unique set of LIOMs, and we can then verify that these LIOMs are indeed exponentially localized in the MBL phase.

Expanding the IOMs τ_z^j in the energy eigenbasis $\{|n\rangle\}$, $n = 1, 2, \ldots, 2^L$, as $\tau_z^j = \sum_n a_n^j |n\rangle\langle n|$, our goal is to find $a_n^j \in \pm 1$ under constraints (i)–(iii). We note that the procedure assumes that we have diagonalized the Hamiltonian. The algorithm is reminiscent of quicksort (see Fig. 1):

(1) For all eigenstates $|n\rangle$ and spin j evaluate $s_n^j = \langle n|\sigma_z^j|n\rangle$.

(2) For each j, sort the eigenstates according to s_n^j, and define candidates $\tilde{\tau}_z^j = \sum_{n\in S_{\max}^j} |n\rangle\langle n| - \sum_{n\in S_{\min}^j} |n\rangle\langle n|$, where $S_{\max(\min)}^j$ is the set of eigenstates giving the $2^L/2$ largest (smallest) overlaps s_n^j.

(3) For each j, compute the overlaps $\langle \tilde{\tau}_z^j \sigma_z^j\rangle = \sum_{n\in S_{\max}^j} s_n^j - \sum_{n\in S_{\min}^j} s_n^j$ and find the site j_M that maximizes the overlap. For this site, set $\tau_z^{j_M} \equiv \tilde{\tau}_z^{j_M}$, thus assigning $a_n^{j_M}$.

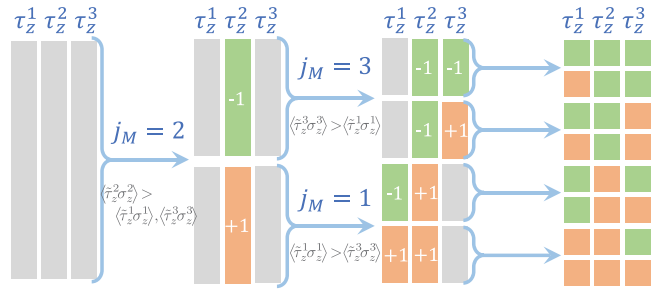

FIG. 1. The flow diagram shows an example of the construction of a complete set of LIOMs in a system with $L = 3$. Gray blocks represent undetermined matrix elements, and orange (green) blocks represent $+1$ (-1) matrix elements. The procedure works as follows: after diagonalizing the Hamiltonian, find the candidate LIOM τ_{j_M} that maximizes the overlap with the physical spin operator $\langle\tilde{\tau}_z^j\sigma_z^j\rangle$ ($j_M = 2$ here). Divide the eight eigenstates into two sectors, each containing four states according to $\langle n|\sigma_z^2|n\rangle$, and assign $\tau_z^2 = \tilde{\tau}_z^2$. For each sector, find j_M within the sector, divide into two sectors each containing two states, and assign $\tau_z^{j_M} = \tilde{\tau}_z^{j_M}$. Repeat the step one more time, and then all LIOMs are determined.

(4) Consider the two manifolds $\mathbb{S}_\pm^{j_M}$ corresponding to the ± 1 eigenstates of $\tau_z^{j_M}$. Each of these manifolds represents two instances of a new system with $L-1$ spins, containing all sites except j_M. In this new system, perform the same protocol in steps 1–3 to set another LIOM. This results in four sectors, each containing 2^{L-2} states.

(5) By repeating the previous steps $L-2$ times we finally reduce the dimension of each sector to just 1, and all a_n^j are assigned.

We note that our scheme does not necessarily find the most local set of τ_z^j since once the matrix elements of a LIOM are determined at a given step, the subsequent search for the rest of the LIOMs is restricted to its perpendicular complement to satisfy orthogonality (that is, we are not ensured of finding a global optimum). Therefore, we choose to divide sectors using the most local LIOM (largest $\langle\tilde{\tau}_z^j\sigma_z^j\rangle$), so that this division sets the least constraints on later divisions. Other choices are possible [43], but as we show in Fig. S4 of the Supplemental Material [44], this choice indeed gives the most local results among all alternate algorithms we tried. Because we utilize only the overlaps $s_n^j = \langle n|\sigma_z^j|n\rangle$ in the computation, the scheme is immune to accidental resonances in the spectrum.

III. RESULTS

A. Localization of operators and interactions

To test the proposed algorithm and characterize the LIOMs that it finds we consider a prototypical example of an MBL-supporting system, a Heisenberg spin-1/2 chain [45] with random fields,

$$H = \sum_{i=1}^{L-1} \vec{\sigma}^i \cdot \vec{\sigma}^{i+1} + \sum_{i=1}^{L} h_i \sigma_z^i, \tag{1}$$

where the first term is the isotropic exchange interaction (with unit strength) and the second term is a disorder longitudinal field with h_i uniformly distributed in $[-W, W]$ (we set

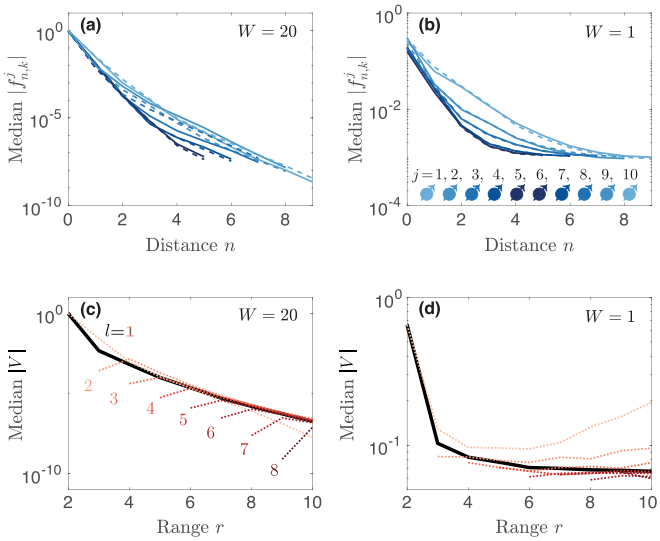

FIG. 2. (a) and (b) Median of the LIOM weights $|f^j_{n,k}|$ as a function of distance n for two disorder strengths: (a) $W = 20$, deep in the MBL phase, where the median decays exponentially to zero, and (b) $W = 1$, in the ergodic phase, where the median saturates to a nonzero value. For each j, the median is taken over the index k in $|f^j_{n,k}|$ as well as 20 different disorder realizations. The darker color represents LIOMs in the middle of the chain, $j = L/2$ [as shown in the bottom of (b)], and left (right) half of the LIOMs is represented by dashed blue curves. (c) and (d) Median of the interaction strength as a function of range r for two disorder strengths. Dotted curves represent l-body interaction terms $|V_{ij}|, |V_{ijk}|, \ldots$ $(l = 2, \ldots, 9)$, where the median is taken over all indices $i, j, \ldots$, as well as 100 disorder realizations. The solid curve represents the median of all interaction terms for a given range $V(r)$, regardless of how many LIOMs are involved. $L = 10$ in all subplots.

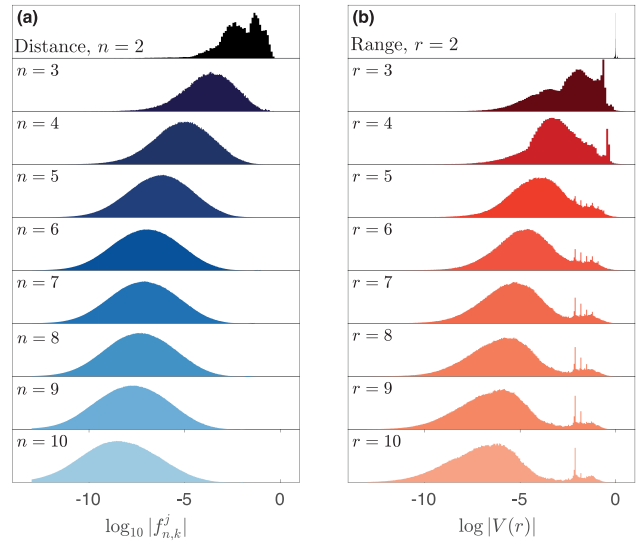

FIG. 3. (a) Probability distribution of LIOM weights $\log_{10}|f^j_{n,k}|$. For a given distance n, samples are taken from all possible j and k as well as 200 disorder realizations. The distribution shows one single Gaussian peak that shifts toward smaller weights with increasing distance n, signaling the localization of IOMs. (b) Probability distribution of the interaction strength $\log_{10}(|V|)$. For a given range r, samples are taken from all terms in Eq. (3) as well as 10 000 disorder realizations. Two peaks can be observed: the left peak is due to the localized cases as it shifts to smaller interaction strengths for longer range; the right peak shows the delocalized cases (rare events) as it is independent of interaction range. $L = 10$ and $W = 20$ for both (a) and (b).

$\hbar = 1$). It is known [8] that in the thermodynamic limit there is a MBL phase transition at $W_c \approx 7 \pm 2$. Although this model conserves the total magnetization along z, the validity of the algorithm does not depend on this symmetry. To quantitatively check the locality of LIOMs, we decompose them into tensor products of Pauli operators,

$$\tau^j_z = \sum_{n=0}^{L} \sum_k f^j_{n,k} \hat{O}^j_{n,k}, \tag{2}$$

where $\hat{O}^j_{n,k}$ is a tensor product of Pauli operators whose furthest nonidentity Pauli matrix from j is of *distance* n, e.g., $\sigma^1_x \otimes \sigma^2_x \otimes \sigma^3_y \otimes \mathbb{I}^4$ is of distance $n = 2$ to $j = 1$, because σ^3_y is the furthest nonidentity Pauli matrix. k labels operators with the same n. $f^j_{n,k} = \mathrm{Tr}(\tau^j_z \hat{O}^j_{n,k})$ is the weight of the jth LIOM on $\hat{O}^j_{n,k}$. Figures 2(a) and 2(b) show the median of $|f^j_{n,k}|$ as a function of distance n. In the MBL phase, the median weight decays exponentially with distance n, while in the ergodic phase it saturates at large n.

Because the LIOMs form an orthonormal basis, the Hamiltonian can be decomposed into this basis unambiguously and efficiently:

$$H = \sum_i \xi_i \tau^i_z + \sum_{ij} V_{ij} \tau^i_z \tau^j_z + \sum_{ijk} V_{ijk} \tau^i_z \tau^j_z \tau^k_z + \cdots. \tag{3}$$

For noninteracting models, only the ξ_i coefficients are nonzero. We can define the range r of each coupling term $V_{ij\ldots}$ as the largest difference among the indices. For example, the range for two-body interaction V_{ij} is simply $r = |i - j|$, while for three-body interactions it is $r = \max(|i - j|, |i - k|, |j - k|)$. Figures 2(c) and 2(d) show the median interaction strength as a function of interaction range. In the MBL phase, the interaction strength decays exponentially. The behaviors of two-body interactions $|V_{ij}|$ and three-body interactions $|V_{ijk}|, \ldots$ show no significant difference [36,39] and can essentially be captured by the median of all interaction terms for a given range $V(r)$. We considered the median instead of the mean in order to exclude rare events, i.e., instances where the disorder strength is small in a local region.

To gain more insight into the localization of IOMs and interactions and observe the occurrence of rare events, in Fig. 3 we further study the probability distribution of weight $f^j_{n,k}$ versus n and the probability distribution of interaction strength $V(r)$ versus r in the localized regime (strong disorder). The distribution of $\log_{10}(|f^j_{n,k}|)$ can be described by a single Gaussian peak, centered at smaller values of $|f^j_{n,k}|$ when the distance n increases, confirming the localization of IOMs. Instead, two peaks can be observed in the distribution of $\log_{10}(|V|)$. The left peak shifts to smaller $|V|$ with increasing r, while the right peak (larger $|V|$) shows no significant shift. Moreover, the area of the right peak decreases for larger W and smaller L. Therefore, we identify the left peak as describing localized cases and the right one as rare events. The

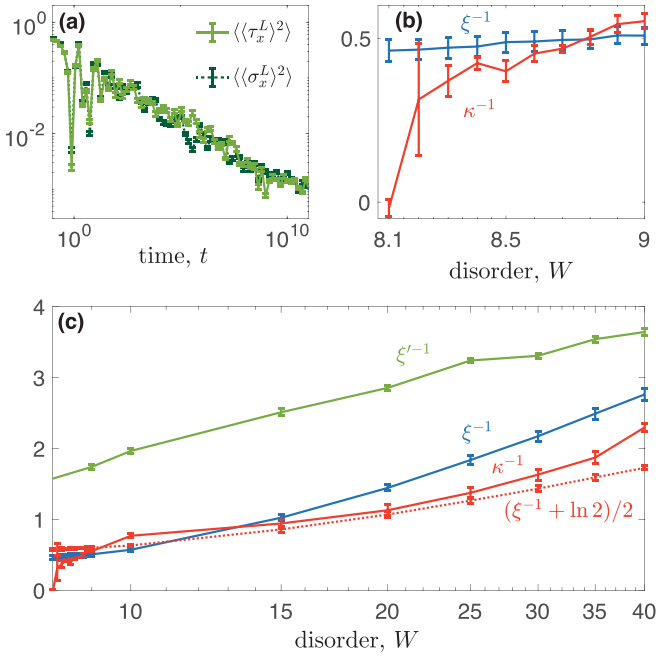

FIG. 4. (a) Dephasing of the physical spin operator σ_x^L (dark green dashed curve) and LIOM τ_x^L (green solid curve). The initial state is a product state with each spin pointing randomly in the xy plane, i.e., $|\psi(0)\rangle = \otimes_{j=1}^{L}(|+\rangle_j + e^{i\phi}|-\rangle_j)/\sqrt{2}$, with ϕ randomly sampled in $[0, 2\pi)$, $\sigma_z^j|+\rangle_j = |+\rangle_j$ $\sigma_z^j|-\rangle_j = -|-\rangle_j$ for the red curve and $\tau_z^j|+\rangle_j = |+\rangle_j$ $\tau_z^j|-\rangle_j = -|-\rangle_j$ for the blue curve. $L = 10$, $W = 20$. Averaging is performed over 20 different initial states and 20 disorder realizations. The error bar represents the standard deviation of all configurations. (b) and (c) Localization lengths as a function of disorder strength W for $L = 12$. The LIOM localization length ξ is extracted by fitting $\mathrm{Tr}(\tau_z^j\sigma_z^k) \sim \exp(-|k-j|/\xi)$ as a function of k (for $j = 1$). The interaction localization length is obtained from the fit of the interaction as a function of range, $V(r) \sim \exp(-r/\kappa)$. The dynamical localization length describes the LIOM dephasing shown in (a) and is obtained by the fit to $\langle\langle\tau_x^L\rangle^2\rangle \sim t^{-\xi'\ln 2}$. Here the error bars derive from the fitting error. The dephasing curve used to fit ξ' is extracted from the median of 50 disorder realizations and 50 initial states. ξ and κ are extracted from the median of 5000 disorder realizations. (b) is a zoom of (c) near the transition point.

exponential localization of the LIOMs and their interactions are usually the two criteria that define the LIOM. In the rare region of low disorder, however, the two requirements cannot be satisfied simultaneously, and there are no universal criteria on how to choose LIOMs in this case. Here we require the IOM τ_z to be local by construction, so the presence of a rare region shows up only in the interaction strengths (see Appendix B); choosing different criteria for the LIOM construction may lead to different results.

B. Localization lengths

From the explicit form of the LIOMs and their interactions, we can extract the *LIOM localization length* ξ via $|f_{n,k}^j| \sim \exp(-n/\xi)$ and *interaction localization length* κ via $|V(r)| \sim \exp(-r/\kappa)$ [3]. In Fig. 4 we show κ and ξ as a function of disorder strength W. The LIOM localization length ξ is extracted using the relation $\mathrm{Tr}(\tau_z^j\sigma_z^k) \sim \exp(-|k-j|/\xi)$ [33,36] because calculating $f_{n,k}^j$ is numerically demanding [44], and in Appendix A we verify the two, indeed, have the same localization length. The interaction localization length κ is extracted by fitting the distribution of $\log_{10}|V|$ (as in Fig. 3) to two Gaussian peaks and then fitting the localized peak center to a linear function of r. Because our method forces τ_z to be local, ξ is always finite, while κ diverges around $W = 8.1$ [Fig. 4(b)], which agrees with the critical point $W_c = 7 \pm 2$ reported in Ref. [8]. $\kappa = 0$ serves as an unambiguous metric to pinpoint the ergodic to MBL phase transition point. It has been shown in [3] that the two localization lengths satisfy the inequality $\kappa^{-1} \geqslant (\xi^{-1} - \ln 2)/2$. From the numerical results in Fig. 4(c), we find that this inequality is satisfied in the localized phase, except in the vicinity of the phase transition point.

C. Noninteracting model: Trade-off of localization

We can better understand why the interaction localization length κ diverges at the critical point while the LIOM localization length ξ remains finite by varying the ZZ coupling strength and study the noninteracting model $H = \sum_{i=1}^{L} h_i\sigma_z^i + \sum_{i=1}^{L-1}(\sigma_x^i\sigma_x^{i+1} + \sigma_y^i\sigma_y^{i+1} + J_z\sigma_z^i\sigma_z^{i+1})$. $J_z = 1$ is the Heisenberg model we investigated before, and $J_z = 0$ corresponds to a noninteracting model. For the noninteracting model, the system is effectively localized for arbitrarily small W. This Hamiltonian can be mapped to a free fermionic Hamiltonian via a Jordan-Wigner transformation [46]. The Hamiltonian can be diagonalized by single-particle IOMs $\{\Sigma_z^i\}$: $H = \sum_i \tilde{\xi}_i\Sigma_z^i$; that is, the interaction localization length in the $\{\Sigma_z^i\}$ basis is zero. However, note that the single-particle IOMs $\{\Sigma_z^i\}$ can be highly nonlocal for small W. We can instead apply our algorithm to find LIOMs $\{\tau_z^j\}$ for this model as done for the interacting Hamiltonian and compare $\{\Sigma_z^j\}$ and $\{\tau_z^j\}$ (see Fig. 5). For large disorder strength, $W = 20$, the Hamiltonian is practically interaction free even in the τ_z^j basis, and indeed, the LIOMs τ_z^j approach the IOMs, $\tau_z^j \approx \Sigma_z^j$. The trade-off between the two interaction strengths κ and ξ becomes evident for small disorder, $W = 0.5$, where $\tau_z^j \neq \Sigma_z^j$. In this regime, the single-particle IOMs Σ_z^j are delocalized, $\xi \gg 1$, but the Hamiltonian still has no interactions, $\kappa = 0$. Instead, the LIOMs obtained by our construction, $\{\tau_z^j\}$, are localized, but they give rise to long-range interactions in the Hamiltonian, $\kappa \gg 1$. Even though our method results in nonzero interaction among LIOMs, it is still able to distinguish interacting and noninteracting, as shown in Figs. 5(a) and 5(b), where a moderate J_z leads to a significant increase of the interaction among LIOMs for both weak and strong disorders. For interacting models, it is difficult to obtain IOMs that minimize the interactions in a nonperturbative way. Still, we expect that if one were indeed able to find such a set of IOMs, there would be a similar trade-off between how local they are (small ξ) versus how local the interactions are (small κ) outside the well-localized phase. Our choice of criterion for constructing LIOMs not only allows a simple and efficient algorithm; by keeping the operators local even when crossing the localization transition, τ_z^j are always well defined and can be used to explore properties of the system, such as its dynamics, around the localization-delocalization transition point.

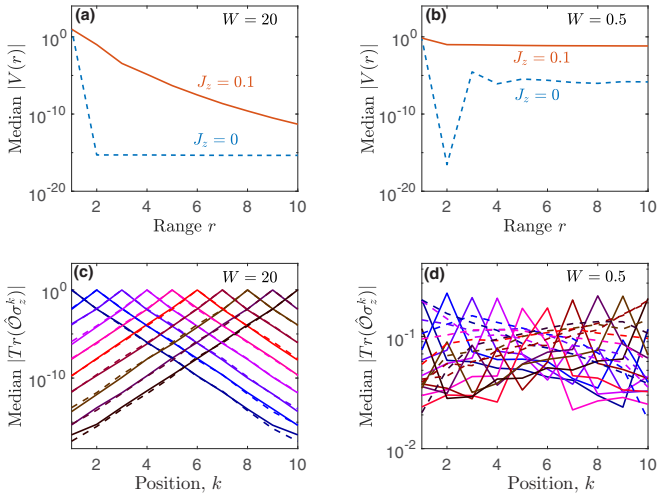

FIG. 5. LIOM in the noninteracting model. (a) and (b) Median interaction strength in the LIOM basis $\{\tau_z^j\}$ vs range r. $J_z = 0$ (blue dashed line) corresponds to a noninteracting model, and $J_z = 0.1$ (red line) corresponds to a model with a small interaction. We note that for large disorder the noninteracting model has only nonzero V for the range $r = 1$, which denotes the single-particle Hamiltonian $\xi_j \tau_z^j$, and even for small disorder $V(r = 1)$ is the dominant term. This is in contrast to the behavior of the interacting model. (c) and (d) Median overlap between LIOMs and physical spins $\mathrm{Tr}(\hat{\mathcal{O}}\sigma_z^k)$ for the noninteracting model $J_z = 1$, with $\hat{\mathcal{O}} = \Sigma_z^j$ (solid lines) for single-particle LIOM and $\hat{\mathcal{O}} = \tau_z^j$ (dashed lines) for LIOMs obtained using the scheme proposed in this paper. Blue stands for smaller j, and red stands for larger j except for the single-particle LIOMs in (d), where the color is randomly chosen because single-particle LIOMs are too delocalized to be ordered. In (a) and (c) $W = 20$. $r > 1$ interaction strength is below machine precision $\sim 10^{-15}$. Σ_z and τ_z show little difference. In (b) and (d) $W = 0.5$. τ_z^j is more localized at site j, but the interaction among LIOMs is not zero. $L = 10$ and 500 disorder realizations are used in all plots.

D. Dephasing dynamics

Since physical spin and LIOM operators are related by a local unitary transformation, they are expected to exhibit a similar dynamics [Fig. 4(a)]. In particular, the higher-order interaction terms in Eq. (3) induce dephasing of the transverse operators by creating an effective magnetic field H_{eff} at the location of spin j due to all the other spins. The dephasing of the expectation values $\langle\tau_x(t)\rangle$ and $\langle\sigma_x(t)\rangle$ is closely related to the logarithmic light cone in the MBL phase [28]. It was previously shown that $\langle\langle\sigma_x(t)\rangle^2\rangle \approx \langle\langle\tau_x(t)\rangle^2\rangle \propto t^{-\alpha}$, where we took the average of the expectation values over random initial states and disorder realizations. For an initial state given by a product state with each individual spin pointing randomly in the xy plane, $\alpha = 2\xi' \ln 2$ for bulk spins and $\alpha = \xi' \ln 2$ for boundary spins, where ξ' is a localization length different from ξ and κ [28]. This length ξ', which we name *dynamical localization length*, describes the strength of the contribution to the effective magnetic field felt by spin j due to spins at distance l: $H_{\mathrm{eff}}^l \sim \exp(-l/\xi')$ (see Appendix C). By assuming exponentially decaying interactions, $|V(r)| = \exp(-r/\kappa)$, it was conjectured that $\xi'^{-1} \leqslant \kappa^{-1} + (\ln 2)/2$ [3]. We find instead a much larger dephasing rate [Fig. 4(c)]. To investigate whether this is due solely to our LIOM construction, which does not explicitly enforce an exponentially decaying interaction strength, we artificially generate a Hamiltonian satisfying $|V(r)| \propto \exp(-r/\kappa)$ (see Appendix C). Still, although we indeed find a power law decay, this is even faster than what we observe in Fig. 4(a). We conjecture that the dephasing process cannot be simply described by a mean interaction strength (the model used to justify the relationship to κ), and higher-order correlations may play an important role.

IV. CONCLUSION AND OUTLOOK

We provided a method to efficiently compute the LIOMs for MBL systems by maximizing the overlap between LIOMs and physical spin operators. The method is nonperturbative and thus immune to resonances in the spectrum and can be applied at the phase transition point. The only quantity we used in computing the LIOMs and their interactions is the expectation value of physical spin operators on energy eigenstates $\langle n|\sigma_z^j|n\rangle$. Although we used exact diagonalization here, our scheme is compatible with renormalization group methods and matrix product state representations [38,47], which can potentially be applied to much larger systems and beyond one dimension. We showed the power of the constructed LIOMs by extracting the localization length of the LIOMs and the Hamiltonian interactions from their respective exponential decays. We also showed that in the MBL phase, the LIOMs and physical spin operators exhibit similar dephasing dynamics, even if it cannot be simply explained by the typical weights of LIOMs and typical interaction strengths.

ACKNOWLEDGMENT

This work was supported in part by the National Science Foundation under Grants No. PHY1734011 and No. PHY1915218.

APPENDIX A: COMPARISON OF LIOMS AND PHYSICAL SPIN OPERATORS

In the main text we defined the overlap $f_{n,k}^j$ as a quantifier of the locality of the LIOMs τ_z^j. Another metric that characterizes the LIOMs as a function of disorder strength is the distance of each τ_z^j from the corresponding physical spin-1/2 Pauli operator σ_z^j. Indeed, the larger the disorder is, the more local the LIOMs are, and therefore, the closer they are to the corresponding Pauli operators. We use the Frobenius norm of the matrix difference between the two operators at the same site [see Fig. 6(a)] to quantify the operator distance. At small disorder strength, the LIOM and physical spin operators are almost perpendicular,

$$\left\|\sigma_z^j - \tau_z^j\right\|_{W\to 0} \sim \sqrt{\left\|\sigma_z^j\right\|^2 + \left\|\tau_z^j\right\|^2} = 2^{(L+1)/2}. \tag{A1}$$

As the disorder strength increases, the distance decreases, as expected. At strong disorder strength $W > W_c \sim 7$, we find that the distance decreases as $1/W$, indicating that the system is in the MBL phase. This result shows that the Frobenius norm distance (or, equivalently, the trace norm) can be taken as a good proxy for the overlap $f_{n,k}^j$.

In the main text we state that the LIOM localization length can be extracted from $\mathrm{Tr}(\tau_z^j \sigma_z^k)$. To confirm this quantita-

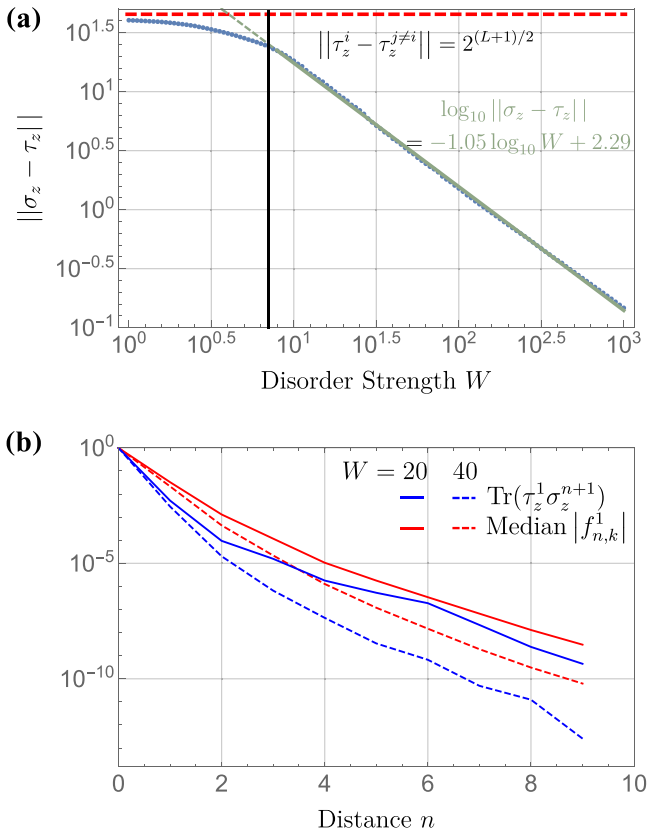

FIG. 6. Comparison of LIOMs and physical spin operators. (a) Blue dots show the Frobenius norm of the difference between same-site physical spin-1/2 Pauli matrices and local integral of motion $||\sigma_z^j - \tau_z^j||$ as a function of disorder strength in a size $L = 10$ system. For disorder $W > W_c \sim 7$ (vertical black line), the norm scales as $1/W$ (green line). The red dashed line shows the norm of the difference between two LIOMs at different sites for comparison. (b) Weight of the first LIOM $f_{n,k}^1$ (blue curves) and overlap of the first LIOM with physical spins $\sum_j \mathrm{Tr}(\tau_z^1 \sigma_z^{1+n})/L$ (red curves) as a function of distance n for $W = 20$ (solid curves) and $W = 40$ (dashed curves). $L = 10$, and the median is taken over k and 100 disorder realizations.

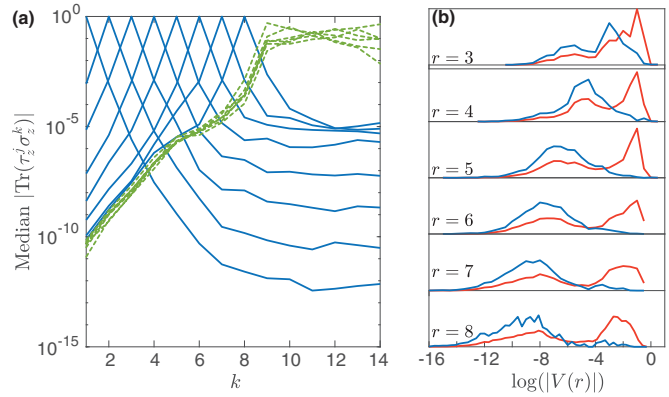

FIG. 7. (a) Median $|\mathrm{Tr}(\tau_z^j \sigma_z^k)|$ and (b) normalized probability distribution of $\log_{10}|V(r)|$ for a chain of size $L = 14$ with disordered field only on sites 1 to 8 ($i_e = 8$). In (a), green dashed (blue solid) curves represent the LIOMs in the disorder-free (disordered) region. In (b), blue curves show only the coupling terms within the disordered region, and red curves show the distribution of all coupling terms. $W = 50$ and 100 random realizations are used in both (a) and (b).

tively, in Fig. 6(b) we compare the weight of the first LIOM $f_{n,k}^1$ and $\mathrm{Tr}(\tau_z^1 \sigma_z^{1+n})$. Both of them show exponential decay with n, and the slopes (decay rates) are similar for $n \geqslant 2$. In numerics, calculating $f_{n,k}^1$ is demanding because it is defined in real space [44], while calculating $\mathrm{Tr}(\tau_z^1 \sigma_z^{1+n})$ can be done in the energy eigenbasis since the expectation value of σ_z^j on every eigenstate is already obtained during the construction process. Therefore, we use $\mathrm{Tr}(\tau_z^j \sigma_z^{1+n})$ with $n \geqslant 2$ to extract the LIOM localization length κ in the main text (Fig. 4).

APPENDIX B: DISTRIBUTION OF INTERACTION STRENGTHS AND RARE REGIONS

In the main text we linked the occurrence of rare events in the distribution of interaction strengths to rare regions of the disordered field. We can verify this conjecture by taking a closer look at one particular disorder realization that contains a low-disorder rare region (see Fig. S1 [44]). To further confirm the connection between a rare region and the rare event peak in the interaction strength distribution, we study a Heisenberg spin chain with the disorder field only on part of the chain, i.e., $h_i \in [-W, W]$ in the disordered region $i \leqslant i_e$ and $h_i = 0$ in the disorder-free region $i > i_e$. The LIOMs in the disordered region are localized, while the LIOMs in the disorder-free region are delocalized with an exponential tail extending into the disordered region [Fig. 7(a)]. Due to the existence of the disorder-free region, the probability distribution of $\log_{10}|V(r)|$ shows a large delocalized peak [blue curve in Fig. 7(b)], which is absent when considering only interactions inside the disordered region. We can further analyze how the occurrence of rare events changes with the system size (see Fig. S2 [44]). We find that for a given interaction range, the area of the delocalized peak gets larger for a longer chain because the frequency of having a local low-disorder region is higher for larger L.

APPENDIX C: DEPHASING WITH THE ARTIFICIAL HAMILTONIAN

It has been conjectured that the dephasing rate of $\langle\tau_x\rangle$ (and $\langle\sigma_x\rangle$) can be related to the interaction localization length via a simple, mean-field model. Using our LIOM construction, we found instead surprising results, as shown in Fig. 4. Here we want (i) to verify whether assuming an exponentially decaying interaction strength does indeed yield the relationship between localization lengths presented in Ref. [3] and (ii) to determine whether the Hamiltonian approximation with a simpler, exponentially decaying interaction strength is enough to capture the exact dephasing dynamics.

To answer these questions, we consider two artificially generated Hamiltonian: (1) $|V(r)| = \exp(-r/\kappa)$, with each interaction term randomly assigned a plus or minus sign, and (2) $|V(r)|$ randomly sampled from the simulated probability distribution [see Fig. 3(b) in the main text for an example] with a random sign.

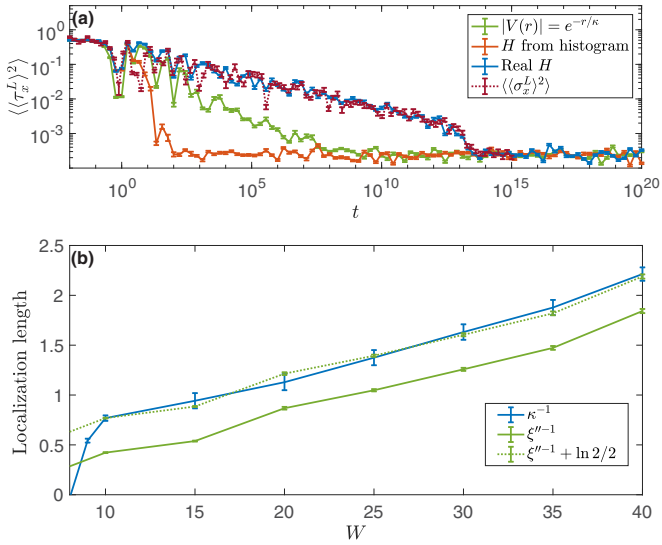

FIG. 8. (a) $\langle\langle\tau_x^L\rangle^2\rangle$ under real (green line) and two artificial Hamiltonians (blue and red lines). The dark green dashed curve shows $\langle\langle\sigma_x^L\rangle^2\rangle$ under the real Hamiltonian. $L = 12$, $W = 40$, so that the delocalized cases are negligible. All $\langle\langle\tau_x^L\rangle^2\rangle$ are averaged over 20 random initial states in the xy plane and 20 disorder realizations. $\langle\langle\sigma_x^L\rangle^2\rangle$ is averaged over ten random initial states in the xy plane and ten disorder realizations. (b) The red curve shows κ^{-1} as in Fig. 4 of the main text. The green solid curve represents the dephasing localization length ξ'' extracted from the artificial Hamiltonian with $|V(r)| = e^{-(r/\kappa)}$. The green dashed curve shows $\xi'' + \ln 2/2$, which overlaps with κ^{-1} within the error bars.

The first Hamiltonian exactly satisfies the hypothesis under which the relation between the interaction localization length and dynamical localization length was derived in Ref. [3]. Therefore, we fit the power law dephasing obtained under this Hamiltonian [see Fig. 8(a)] and extract the dynamical localization length ξ'' as done in Fig. 4 of the main text. We find that $\kappa^{-1} \approx \xi''^{-1} + \ln 2/2$ [Fig. 8(b)], which gives a more stringent relation than the bound $\kappa^{-1} \geqslant \xi''^{-1} - \ln 2/2$ given in Ref. [3]. We can provide a heuristic argument for the relation between ξ'' and κ, under the assumption $|V(r)| = \exp(-r/\kappa)$. As described in Ref. [28], the dephasing can be understood as arising from an effective magnetic field at site j due to all other spins τ_z. Starting from the phenomenological model in Eq. (3), the effective magnetic field at site j is

$$H_j = \mathrm{Tr}\left(\tau_z^j H\right) = \xi_j + H_j^1 + H_j^2 + \cdots, \tag{C1}$$

where H_j^l denotes the magnetic field created by spins within the distance l from spin j. For example, the first term is given by

$$H_j^1 = V_{j,j+1}\tau_z^{j+1} + V_{j-1,j}\tau_z^{j-1} + V_{j-1,j,j+1}\tau_z^{j-1}\tau_z^{j+1}. \tag{C2}$$

Similarly, H_j^l contains interactions of the range $l+1, l+2, \ldots, 2l+1$. As the interaction strength decays as $|V(r)| = \exp(-r/\kappa)$ and the number of terms grows as $\sim 2^r$, the Frobenius norm of H_j^l is estimated as

$$\left\|H_j^l\right\| = \left[\sum_{r=l+1}^{2l+1} 2^r e^{-2r/\kappa}\right]^{1/2} \approx \frac{(2e^{-2/\kappa})^{(l+1)/2}}{1 - 2e^{-2/\kappa}}. \tag{C3}$$

In the last term we assumed that $l \gg 1$ and the system is deep in the MBL phase so that $2\exp(-2/\kappa) < 1$. We thus find that H_j^l also exhibits an exponential decay $H_j^l \propto \exp(-l/\xi'')$, with $\xi''^{-1} = \kappa^{-1} + \ln 2/2$, yielding the dephasing [28] $\langle\langle\tau_x^L(t)\rangle^2\rangle \sim t^{-\xi''\ln 2}$, as shown in Fig. 8.

While we confirm that the dephasing under the approximated Hamiltonian satisfying $|V(r)| = \exp(-r/\kappa)$ follows the predicted relation to κ, we still find that dephasing under the "real" Hamiltonian is different. The physical spin and LIOMs under the real Hamiltonian in Eq. (1) show similar dephasing, as expected. Under either artificial Hamiltonian, however, $\langle\langle\tau_x^L\rangle^2\rangle$ dephases much faster than under the real Hamiltonian, suggesting that the dephasing dynamics cannot be fully captured by the interaction localization length κ or even the probability distribution of $|V(r)|$ [Fig. 8(a)]. For instance, in the real system for a given disorder realization the interaction terms may have some correlation, which gives rise to a slower dephasing, but this is not captured by the probability distribution.

[1] P. W. Anderson, Phys. Rev. **109**, 1492 (1958).
[2] *50 Years of Anderson Localization*, edited by E. Abrahams (World Scientific, Singapore, 2010).
[3] D. A. Abanin, E. Altman, I. Bloch, and M. Serbyn, Rev. Mod. Phys. **91**, 021001 (2019).
[4] D. Basko, I. Aleiner, and B. Altshuler, Ann. Phys. (NY) **321**, 1126 (2006).
[5] J. Z. Imbrie, J. Stat. Phys. **163**, 998 (2016).
[6] J. Z. Imbrie, Phys. Rev. Lett. **117**, 027201 (2016).
[7] M. Žnidarič, T. Prosen, and P. Prelovšek, Phys. Rev. B **77**, 064426 (2008).
[8] A. Pal and D. A. Huse, Phys. Rev. B **82**, 174411 (2010).
[9] V. Oganesyan and D. A. Huse, Phys. Rev. B **75**, 155111 (2007).
[10] T. C. Berkelbach and D. R. Reichman, Phys. Rev. B **81**, 224429 (2010).
[11] I. V. Gornyi, A. D. Mirlin, and D. G. Polyakov, Phys. Rev. Lett. **95**, 206603 (2005).
[12] 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).
[13] 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).
[14] P. Bordia, H. P. Lüschen, S. S. Hodgman, M. Schreiber, I. Bloch, and U. Schneider, Phys. Rev. Lett. **116**, 140401 (2016).
[15] S. S. Kondov, W. R. McGehee, W. Xu, and B. DeMarco, Phys. Rev. Lett. **114**, 083002 (2015).
[16] A. Lukin, M. Rispoli, R. Schittko, M. E. Tai, A. M. Kaufman, S. Choi, V. Khemani, J. Léonard, and M. Greiner, Science **364**, 256 (2019).
[17] F. A. An, E. J. Meier, and B. Gadway, Phys. Rev. X **8**, 031045 (2018).
[18] J. Smith, A. Lee, P. Richerme, B. Neyenhuis, P. W. Hess, P. Hauke, M. Heyl, D. A. Huse, and C. Monroe, Nat. Phys. **12**, 907 (2016).

[19] P. Roushan, C. Neill, J. Tangpanitanon, V. M. Bastidas, A. Megrant, R. Barends, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, A. Fowler, B. Foxen, M. Giustina, E. Jeffrey, J. Kelly, E. Lucero, J. Mutus, M. Neeley, C. Quintana, D. Sank, A. Vainsencher, J. Wenner, T. White, H. Neven, D. G. Angelakis, and J. Martinis, Science **358**, 1175 (2017).
[20] K. X. Wei, C. Ramanathan, and P. Cappellaro, Phys. Rev. Lett. **120**, 070501 (2018).
[21] M. Serbyn, Z. Papić, and D. A. Abanin, Phys. Rev. Lett. **111**, 127201 (2013).
[22] B. Bauer and C. Nayak, J. Stat. Mech. (2013) P09005.
[23] J. H. Bardarson, F. Pollmann, and J. E. Moore, Phys. Rev. Lett. **109**, 017202 (2012).
[24] M. Serbyn, Z. Papić, and D. A. Abanin, Phys. Rev. Lett. **110**, 260601 (2013).
[25] D. A. Huse, R. Nandkishore, and V. Oganesyan, Phys. Rev. B **90**, 174202 (2014).
[26] R. Vosk and E. Altman, Phys. Rev. Lett. **110**, 067204 (2013).
[27] I. H. Kim, A. Chandran, and D. A. Abanin, arXiv:1412.3073.
[28] M. Serbyn, Z. Papić, and D. A. Abanin, Phys. Rev. B **90**, 174302 (2014).
[29] 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).
[30] G. De Tomasi, S. Bera, J. H. Bardarson, and F. Pollmann, Phys. Rev. Lett. **118**, 016804 (2017).
[31] X. Chen, T. Zhou, D. A. Huse, and E. Fradkin, Ann. Phys. (Berlin, Ger.) **529**, 1600332 (2017).
[32] M. Serbyn and D. A. Abanin, Phys. Rev. B **96**, 014202 (2017).
[33] A. Chandran, I. H. Kim, G. Vidal, and D. A. Abanin, Phys. Rev. B **91**, 085425 (2015).
[34] S. D. Geraedts, R. N. Bhatt, and R. Nandkishore, Phys. Rev. B **95**, 064204 (2017).
[35] V. Ros, M. Müller, and A. Scardicchio, Nucl. Phys. B **891**, 420 (2015).
[36] L. Rademaker and M. Ortuño, Phys. Rev. Lett. **116**, 010404 (2016).
[37] L. Rademaker, M. Ortuño, and A. M. Somoza, Ann. Phys. (Berlin, Ger.) **529**, 1600322 (2017).
[38] Y. Z. You, X. L. Qi, and C. Xu, Phys. Rev. B **93**, 104205 (2016).
[39] D. Pekker, B. K. Clark, V. Oganesyan, and G. Refael, Phys. Rev. Lett. **119**, 075701 (2017).
[40] T. E. O'Brien, D. A. Abanin, G. Vidal, and Z. Papić, Phys. Rev. B **94**, 144208 (2016).
[41] T. B. Wahl, A. Pal, and S. H. Simon, Phys. Rev. X **7**, 021018 (2017).
[42] A. K. Kulshreshtha, A. Pal, T. B. Wahl, and S. H. Simon, Phys. Rev. B **99**, 104201 (2019).
[43] M. Goihl, M. Gluza, C. Krumnow, and J. Eisert, Phys. Rev. B **97**, 134202 (2018).
[44] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevB.100.214203 for detailed analysis of the rare regions, algorithm stability, and computational complexity, which includes Refs. [43, 47, 38].
[45] H. Bethe, Z. Phys. **71**, 205 (1931).
[46] P. Jordan and E. Wigner, Z. Phys. **47**, 631 (1928).
[47] V. Khemani, F. Pollmann, and S. L. Sondhi, Phys. Rev. Lett. **116**, 247204 (2016).