



## Diagonalization and Many-Body Localization for a Disordered Quantum Spin Chain

John Z. Imbrie

*Department of Mathematics, University of Virginia, Charlottesville, Virginia 22904-4137, USA*

(Received 13 April 2016; published 5 July 2016)

We consider a weakly interacting quantum spin chain with random local interactions. We prove that many-body localization follows from a physically reasonable assumption that limits the extent of level attraction in the statistics of eigenvalues. In a Kolmogorov-Arnold-Moser-style construction, a sequence of local unitary transformations is used to diagonalize the Hamiltonian by deforming the initial tensor-product basis into a complete set of exact many-body eigenfunctions.

DOI: 10.1103/PhysRevLett.117.027201

In the past few years, there has been a surge of interest in the phenomenon of many-body localization (MBL). In the well-studied Anderson tight-binding model [1] a particle moving in a sufficiently strong random potential is localized; eigenstates decay exponentially away from localization centers and transport is absent. A number of authors have argued that localization persists in the presence of weak interparticle interactions [2–8]. In particular, the detailed perturbative analysis of [7] provided strong evidence for MBL. More recently, numerical studies on one-dimensional spin systems and particle systems [9–11] gave evidence for a transition from a thermalized phase to a many-body localized phase, as the strength of the disorder increases. See [12] for a review of recent work on MBL.

On a theoretical level, it is important to get past perturbative analysis, as rare regions with weak disorder (Griffiths regions [13]) have the potential to spoil localization. Rigorous results on localization in many-body systems include a proof of dynamical localization for an isotropic random spin chain, using the Jordan-Wigner transformation to reduce the problem to an equivalent one-body Hamiltonian [14]. Localization in the ground state of the interacting Aubry-André model was established in [15].

In this Letter we establish rigorously that for a one-dimensional disordered spin chain, MBL follows from a physically reasonable assumption on level statistics. We consider a random field, random transverse field, random exchange Ising model on an interval  $\Lambda = [-K, K'] \cap \mathbb{Z}$ ,

$$H = \sum_{i=-K}^{K'} h_i S_i^z + \sum_{i=-K}^{K'} \gamma_i S_i^x + \sum_{i=-K-1}^{K'} J_i S_i^z S_{i+1}^z. \quad (1)$$

Here,  $S_i^{x,z}$  are Pauli matrices, with  $S_i^z \equiv 1$  for  $i \notin \Lambda$ . We take  $\gamma_i = \gamma \Gamma_i$  with  $\gamma$  small. Thus, the Hamiltonian is close to one that is diagonal in the basis given by tensor products of  $S_i^z$  eigenstates. We take the random variables  $h_i, \Gamma_i, J_i$  to be independent and bounded, with bounded probability densities. This is a variant of the model considered in [10]. We need to make an assumption of limited level attraction

(LLA) for the spectrum of  $H$ , for some values of  $\nu > 0$  and  $C < \infty$ .

*Assumption LLA*( $\nu, C$ ). Consider the Hamiltonian  $H$  in boxes of size  $n$ . Its eigenvalues satisfy

$$P\left(\min_{\alpha \neq \beta} |E_\alpha - E_\beta| < \delta\right) \leq \delta^\nu C^n, \quad (2)$$

for all  $\delta > 0$  and all  $n$ .

Physically, this is a mild assumption specifying that with high probability the minimum level spacing should be no smaller than some exponential in the volume. Note that random matrices normally have either neutral statistics ( $\nu = 1$ , e.g., Poisson) or repulsive ones ( $\nu > 1$ , e.g., Gaussian orthogonal ensemble). Our analysis works even for attractive statistics, i.e.,  $0 < \nu < 1$ . Mathematically, techniques to prove estimates such as (2) are not yet available for many-body systems, but a promising approach is available for single-body Hamiltonians [16].

We give an explicit construction of a sequence of unitary rotations that diagonalizes the Hamiltonian. Each rotation is generated by quasilocal operators. This means that a rotation generator that involves  $\ell$  spins is exponentially small in  $\ell$ , with high probability. Resonant regions where the required rotations are far from the identity are dilute; the probability that two sites a distance  $D$  apart are in the same resonant region decays faster than any power of  $D$ . These rotations define a way to deform the original basis states [tensor products of (1,0) or (0,1) at each site] into the exact eigenstates. Away from resonant regions, each eigenstate resembles the basis state it came from, and classical spin configurations  $\sigma = \{\sigma_i\} \in \{-1, 1\}^{|\Lambda|}$  can be used as eigenstate labels. This is made evident by the following result:

*Theorem.*—Let  $\nu, C$  be fixed. There exists a  $\kappa > 0$  such that for  $\gamma$  sufficiently small, LLA( $\nu, C$ ) implies the following estimates:

$$\mathbb{E} \text{av}_\alpha |\langle S_0^z \rangle_\alpha| = 1 - O(\gamma^\kappa), \quad (3)$$

where  $\mathbb{E}$  denotes the disorder average,  $\text{av}_\alpha$  denotes an average over  $\alpha$ , and  $\langle \cdot \rangle_\alpha$  denotes the expectation in the eigenstate  $\alpha$ . For any  $i \neq j$ ,

$$\max_{\alpha} |\langle \mathcal{O}_i; \mathcal{O}_j \rangle_{\alpha}| \leq \gamma^{|i-j|/3}, \quad (4)$$

with probability  $1 - (\gamma^{\kappa})^{1+(\log|i-j|)^2}$ . Here,  $\langle \mathcal{O}_i; \mathcal{O}_j \rangle_{\alpha} \equiv \langle \mathcal{O}_i \mathcal{O}_j \rangle_{\alpha} - \langle \mathcal{O}_i \rangle_{\alpha} \langle \mathcal{O}_j \rangle_{\alpha}$ , with  $\mathcal{O}_i$  being any operator formed from products of  $S_{i'}^x$  or  $S_{i'}^z$ , for  $i'$  near  $i$ . All bounds are uniform in  $\Lambda$ .

We may take  $av_{\alpha}$  to be any normalized average over the  $2^n$  values of  $\alpha$  (for a box of size  $n$ ), e.g., uniform weights (infinite temperature) or  $(\text{const}) \exp(-\beta E_{\alpha})$ . We see that with high probability, most states have the property that the expectation of  $S_0^z$  is close to  $+1$  or  $-1$ , just like the basis vectors. The spins are effectively frozen in place in each eigenstate. This would contrast with a thermalized phase, wherein states resemble thermal ensembles (a consequence of the eigenstate thermalization hypothesis [17–19]). At infinite temperature, thermalization implies that eigenstate expectations of  $S_0^z$  go to 0 in the infinite volume limit. Thus, the bound (3) implies a failure of thermalization, a key feature of the MBL phase.

In the course of the proof, we construct a sequence of rotations (or changes of basis) that gives an explicit quasiloca deformation between basis vectors and the exact eigenstates. This is an important feature of the fully many-body localized phase [11,20]; the property is essentially equivalent to the existence of a complete set of quasilocal integrals of motion [21–23]. It is the many-body analog of the fact that one can deform (or label) single-particle eigenstates by their localization centers [24].

*Proof of theorem.*—We present the key elements of the proof; more technical aspects are published separately [20]. We obtain a complete diagonalization of  $H$  by successively eliminating low-order off-diagonal terms as in Newton’s method. The process runs on a sequence of length scales  $L_k = (\frac{15}{8})^k$ , with off-diagonal elements of order  $\gamma^m$ ,  $m \in [L_k, L_{k+1})$  eliminated in the  $k$ th step. The orthogonal rotations that accomplish this can be written as a convergent graphical expansion, provided nonresonant conditions are satisfied. Resonant regions are diagonalized as blocks in quasidegenerate perturbation theory. As in the Kolmogorov-Arnold-Moser (KAM) theorem constructing integrals of motion in Hamiltonian dynamical systems, one needs to control the measure of resonant sets where perturbation arguments break down. See [24] for an application of these ideas to the Anderson model. KAM ideas have been useful in quantum models with quasiperiodic potentials [15,25–27].

We begin by outlining the first step of the procedure ( $k = 1$ ).

*Resonances.*—Perturbation theory works if there are gaps between eigenvalues. Initially, the only off-diagonal term is  $\gamma_i S_i^x$ , which is local. We only need to worry about single-flip resonances (for the moment). Let the spin configuration  $\sigma^{(i)}$  be equal to  $\sigma$  with the spin at  $i$  flipped. The associated change in energy is

$$\Delta E_i \equiv E(\sigma) - E(\sigma^{(i)}) = 2\sigma_i(h_i + J_i\sigma_{i+1} + J_{i-1}\sigma_{i-1}). \quad (5)$$

We say that the site  $i$  is resonant if  $|\Delta E_i| < \varepsilon \equiv \gamma^{1/20}$  for at least one choice of  $\sigma_{i-1}, \sigma_{i+1}$ . Then for nonresonant sites the ratio  $\gamma_i/\Delta E_i$  is  $\leq \gamma^{19/20}$ . A site is resonant with probability  $\sim 4\varepsilon$ . Hence resonant sites form a dilute set where perturbation theory breaks down.

*Perturbation theory.*—Let  $H = H_0 + \mathcal{J}$  with  $H_0$  being diagonal and  $\mathcal{J}$  off diagonal. Put  $\mathcal{J} = J^{\text{res}} + J^{\text{per}}$ , where  $J^{\text{res}}$  contains terms  $J(i) \equiv \gamma_i S_i^x$  with  $i$  being resonant, and  $J^{\text{per}}$  contains the rest. Then define an antisymmetric matrix

$$A \equiv \sum_{\text{nonresonant } i} A(i) \quad \text{with} \quad A(i)_{\sigma\sigma^{(i)}} = \frac{J(i)_{\sigma\sigma^{(i)}}}{E_{\sigma} - E_{\sigma^{(i)}}}. \quad (6)$$

First-order perturbation theory can be implemented by using  $\Omega = e^{-A}$  for a basis change. This leads to a renormalized Hamiltonian,

$$H^{(1)} = e^{-A} H e^A. \quad (7)$$

By construction (first-order perturbation theory),

$$[A, H_0] = -J^{\text{per}}, \quad (8)$$

which cancels all but the resonant terms to leading order,

$$\begin{aligned} H^{(1)} &= e^A H e^{-A} = H + [A, H] + \frac{[A, [A, H]]}{2!} + \dots \\ &= H_0 + J^{\text{res}} + \sum_{n=1}^{\infty} \frac{(\text{ad } A)^n}{n!} \left( \frac{n}{n+1} J^{\text{per}} + J^{\text{res}} \right) \\ &\equiv H_0 + J^{\text{res}} + J^{(1)}, \end{aligned} \quad (9)$$

where  $(\text{ad } A)B \equiv [A, B]$ . A similar transformation was used in [28].

*Properties of the new Hamiltonian.*—After the change of basis,  $J^{\text{per}}$  is gone, while  $J^{\text{res}}$  remains. The new interaction  $J^{(1)}$  is quadratic and higher order in  $\gamma$ . Note that  $A(i)$  commutes with  $A(j)$  or  $J(j)$  if  $|i-j| > 1$ . Thus, we preserve quasilocality of  $J^{(1)}$ ; it can be written as  $\sum_g J^{(1)}(g)$ , where  $g$  is a sum of connected graphs involving spin flips  $J(i)$  and associated energy denominators as in (6). Specifically, a graph is determined by a sequence of sites  $i_0, \dots, i_n$  such that  $\text{dist}(i_p, \{i_0, \dots, i_{p-1}\}) \leq 1$  for  $1 \leq p \leq n$ ; this specifies a nonvanishing term in  $(\text{ad } A)^n J$  that operates on the spins at those sites. A graph involving  $\ell$  spin flips has  $\ell - 1$  energy denominators and is bounded by  $\gamma(\gamma/\varepsilon)^{\ell-1}$ .

We define resonant blocks by taking connected components of the set of sites belonging to resonant graphs. We perform exact rotations  $O$  in small, isolated resonant blocks to diagonalize the Hamiltonian there. This paves the way

for reintegrating such regions into the perturbative framework in subsequent steps.

*Expectations in approximate eigenstates.*— $S_0^z$  is diagonal in the  $\sigma$  basis, so  $(S_0^z)_{\sigma\tilde{\sigma}} = \sigma_0\delta_{\sigma_0\tilde{\sigma}_0}$ . We prove a version of (4) for the eigenfunctions as approximated so far (that is, for the columns of  $\Omega O$ ),

$$\mathbb{E} \text{av}_\alpha \left| \left| \sum_\sigma (O^{\text{tr}} \Omega^{\text{tr}})_{\alpha\sigma} \sigma_0 (\Omega O)_{\sigma\alpha} \right| - 1 \right| \leq O(\varepsilon). \quad (10)$$

To handle resonances, one may throw out a set of small probability  $O(\varepsilon)$  where the origin is in a resonant block and the rotation  $O$  acts nontrivially. Then the rotation  $\Omega$  is close to the identity. Hence the two terms cancel except for interactions involving  $A(i)$ , which is bounded by  $\gamma/\varepsilon$ . To obtain the theorem, we need to prove this for the complete diagonalization of  $H$ . Similar methods should lead to a uniform bound on the entanglement entropy of a subsystem, because the only contributions come from rotations straddling the boundary.

*Multiscale iteration.*—We continue the process on a sequence of length scales  $L_k = (\frac{15}{8})^k$ , so that off-diagonal elements of  $H^{(k)}$  are of the order of  $\gamma^{L_k}$ . In each step, the diagonal elements of  $H^{(k)}$  are renormalized by interactions up to the  $k$ th scale; they are denoted  $E_\sigma^{(k)}$ . We say that  $g$  is resonant if

$$A_{\sigma\tilde{\sigma}}^{(k+1)}(g) \equiv \frac{J_{\sigma\tilde{\sigma}}^{(k)}(g)}{E_\sigma^{(k)} - E_{\tilde{\sigma}}^{(k)}} \quad (11)$$

is greater than  $(\gamma/\varepsilon)^{|g|}$  in magnitude, where  $|g|$  denotes the number of spin flips in  $g$ ; we have  $|g| \geq L_k$  in the  $k$ th step (as in Newton's method, the procedure has a convergence rate that is close to quadratic). For nonresonant graphs, we may use (11) to generate the next rotation. Then the next interaction  $J^{(k+1)}$  is again given by a sum of graphs  $\sum_g J^{(k+1)}(g)$ . This is a recursive construction; repeated application of the ad expansion (9) leads to graphs  $g$  that are given by a sequence of graphs from the previous scale. The precise definition of  $g$  is somewhat involved (see Appendix 1 of [20]), but as in the first step, we may use the condition of nonvanishing commutators to enforce connectivity. Then if one unwraps the expansions, one obtains a sequence of spin flips at a set of sites that is nearest-neighbor connected. Each subgraph from an earlier scale comes with an energy denominator representing the result of flipping the spins of the subgraph; there are a total of  $|g| - 1$  energy denominators.

Let us discuss a key estimate giving control over the probability of resonances. For the moment we make some simplifying assumptions: (a) The graph  $g$  involves spin flips in regions not previously found to be resonant; (b)  $g$  does not flip any spin more than once; and (c) the ad expansion (9) is developed only up to order  $n_0$  on each scale (this avoids complications with factorials in  $n$ ).

We need to maintain uniform exponential decay on the probability that  $g$  is resonant. One cannot simply bound denominators from below, as was done in the first step. There is a mixture of energy denominators from different scales, and combining lower bounds  $O(\varepsilon^{L_j})$  for denominators on scales  $j \leq k$  causes the rate of decay for graphs to degenerate as  $k$  increases. Instead, we prove that each graph  $g$  obeys a fractional moment bound with  $s = \frac{2}{7}$  when averaged over the disorder,

$$\mathbb{E} |A_{\sigma\tilde{\sigma}}^{(k)}(g)|^s \leq \gamma^{s|g|} \mathbb{E} \prod_{\tilde{\tau} \in g} |E_{\tilde{\tau}}^{(j)} - E_{\tilde{\tau}}^{(j)}|^{-s} \leq (c\gamma)^{s|g|}. \quad (12)$$

Here, the constant  $c$  does not depend on  $k$ . By (a), energy denominators  $E_{\tilde{\tau}}^{(j)} - E_{\tilde{\tau}}^{(j)}$  are given by a sum of  $\pm 2h_i$  over the sites  $i$  flipped in the transition  $\tau \rightarrow \tilde{\tau}$  (up to corrections of order  $\gamma$ ). By (b), each energy denominator contains an independent integration variable  $h_i$ . Boundedness of the density for  $h_i$  implies that  $\mathbb{E} |\Delta E|^{-s}$  is bounded, and (12) follows. Then Markov's inequality implies that

$$P(|A_{\sigma\tilde{\sigma}}^{(k)}(g)| > (\gamma/\varepsilon)^{|g|}) \leq (c\varepsilon)^{s|g|}. \quad (13)$$

This procedure estimates all of the denominators of  $g$  together; it yields the desired exponential bound on the rotation generator with probability  $1 - (c\varepsilon)^{s|g|}$ .

The set of sites that belong to resonant graphs  $g$  is decomposed into connected components. The result is the set of resonant blocks in the  $k$ th step. The number of graphs containing a given site is exponential in  $|g|$ , so (13) allows us to sum over collections of graphs connecting one site to another. We obtain an exponentially decaying connectivity function, demonstrating that resonant blocks do not percolate.

In order to relax simplifying assumption (b), we consider graphs with a significant fraction of flips appearing at previously flipped sites. Such graphs no longer have the independent denominator property that was needed to obtain (13). We resum collections of graphs for which the interval spanned by  $g$  is smaller than  $\frac{7}{8}|g|$ . The shortened span implies an increased rate of decay with distance, so inductive bounds from previous steps are sufficient, and probabilistic estimates like (13) are not needed. For the remaining "straight" graphs, most of the denominators are independent, and we obtain a bound similar to (13) on the probability of resonance. To relax assumption (c), we need to incorporate the available  $1/n!$  from the ad expansion into both sides of (13) so as to allow summation over graphs of arbitrarily high order at a given scale  $k$ ; see [20] for details.

In order to completely diagonalize  $H$ , it is important to relax assumption (a) by considering graphs that involve resonant blocks. We know that resonant blocks are dilute, but their interactions could lead to long-range effects. A resonant block  $B$  with diameter  $d$  may have level spacings of order  $2^{-d}$ . In step  $k$  we work with

interactions of order  $\gamma^{L_k}$ , so we are able to resolve energy differences of order  $\varepsilon^{L_k}$ . Therefore, we need to wait until  $L_k$  becomes comparable to  $d$  before considering interactions involving  $B$ . But larger interaction terms of order  $\gamma^{L_j}$  from steps  $j < k$  still connect  $B$  to its immediate neighborhood. To get rid of these, we define a fattened block  $\bar{B}$  by adding an  $L_k$  neighborhood to  $B$ . Any graphs  $g$  connecting  $B$  to  $\bar{B}^c$  (the complement of  $\bar{B}$ ) must have  $|g| \geq L_k$ . We separate the interaction terms of  $H^{(k)} = H_0 + J^{(k)}$  into two parts:  $J^{(k)\text{int}}$  (involving terms internal to  $\bar{B}$ ) and  $J^{(k)\text{ext}}$  (the rest). Let  $O^{(k)}$  be the matrix that diagonalizes  $H_0 + J^{(k)\text{int}}$ , and use it to rotate  $H^{(k)}$ . In the new basis, we have an equivalent system with  $2^{|\bar{B}|}$  eigenstates of  $\bar{B}$  interacting with  $\bar{B}^c$  through terms of order  $\gamma^{L_k}$  or smaller. In effect,  $\bar{B}$  represents a “fat” site with a spin variable labeling the  $2^{|\bar{B}|}$  basis states. We call this variable the metaspin for  $\bar{B}$ . Henceforth, eigenstates are labeled by spin variables  $\sigma_i = \pm 1$  away from resonant blocks, and metaspin variables in blocks  $\bar{B}$ . This is necessary because when a rotation is far from the identity, there is no natural way to assign eigenstates to the original basis states.

In the new basis, we may rotate away interactions of order  $\gamma^{L_k}$  and higher as discussed above. But to control the probability that a graph involving  $\bar{B}$  is resonant, we need estimates like (13), which depend on the finiteness of  $\mathbb{E}|\Delta E|^{-s}$ . This is where we need the assumption  $\text{LLA}(\nu, C)$ . By construction, the dependence of  $\Delta E$  for any level change in  $\bar{B}$  depends on the state in  $\bar{B}^c$  only out to a distance comparable to  $L_k$ . Therefore, we can approximate  $\Delta E$  within an error  $\gamma^{O(L_k)}$  by examining the equivalent transition for the Hamiltonian in an  $L_k$  neighborhood of  $\bar{B}$ . Using  $\text{LLA}(\nu, C)$  for a box of size  $n$  comparable to  $L_k$ , we have that with high probability the minimum level spacing is at least  $\tilde{C}^{-L_k}$  for some  $\tilde{C} < \infty$ . Furthermore, since all three terms in the Hamiltonian (1) are random, the radial degree of freedom in the variables  $\{h_i, \Gamma_i, J_i\}$  scales all energy differences equally. Hence as long as there is a minimum level spacing, we can control the probability that a transition energy  $\Delta E$  is resonant with a given nearby transition—even in the case of a block-block interaction as in Fig. 1. An energy difference  $\Delta E$  of  $\bar{B}$  moves linearly with the radial variable, so as before the fractional moment  $\mathbb{E}|\Delta E|^{-s}$  is bounded (in this case the bound is  $\tilde{C}^{L_k}$ ).

We see that a resonant region of size  $d$  requires a buffer zone of width comparable to  $d$  to enable it to become disentangled from neighboring degrees of freedom. Consequently, one should unite two such neighboring blocks into one if they are within a distance  $d$ . This leads



FIG. 1. A block-block resonance.

to hierarchically organized, loosely connected resonant regions whose connectivity function decays as a stretched exponential. This idea is explored in a simple theory of the MBL transition in [29]. In dimensions greater than 1, the volume of  $\bar{B}$  may be substantially larger than  $d$ . Consequently, the level spacing in  $\bar{B}$  may become smaller than the size of the interaction terms, and our procedure breaks down.

Resonant blocks create gaps in decay for graphs passing through them. We need to ensure that there is minimal loss in decay from collections of resonant blocks on all scales. We want the usual metric in  $\mathbb{Z}$  to be comparable to the one where blocks are contracted to points. Then graphs decay exponentially in the usual metric, and those that traverse the buffer zone of  $B$  are small enough to control the exponential number of states in  $\bar{B}$ . To this end, we use a more generous notion of connectivity wherein a block of volume  $V$  (potentially much smaller than the diameter) is united with similar or larger blocks up to a distance  $\exp[O(V^{1/2})]$  away. Then the separation between blocks grows much faster than their diameter, and the fraction of distance lost to blocks is summable and small. A similar construction was used in [30]. This rule implies that block volumes scale up with their diameters at least as  $O(|\log d|^2)$ , which leads to a connectivity function that decays as  $(\gamma^\kappa)^{1+(\log|i-j|)^2}$ , as in the theorem.

To complete the proof of the theorem, let  $k$  tend to infinity. This eliminates all off-diagonal entries of the Hamiltonian. All rotations are close to the identity, except in resonant blocks. As in the discussion of (10),  $\langle S_0^z \rangle_\alpha$  is close to  $\pm 1$  except for a set of measure  $O(\varepsilon) = O(\gamma^{1/20})$ , and (3) follows. An eigenstate correlation as in (4) involves graphs extending from  $i$  to  $j$ . If no more than half the distance from  $i$  to  $j$  is covered by resonant blocks, then we obtain exponential decay as in (4). Otherwise, we have an event whose probability is dominated by the block connectivity function, which decays as indicated above (faster than any power of  $1/|i-j|$ ).

In summary, we have described a multiscale approach to the problem of deforming the tensor-product basis into the exact eigenfunctions. An assumption on level spacing allows probabilistic control over cases where perturbation theory is not well behaved; as a result we show that resonant regions fail to percolate, and MBL follows. We give convergent expansions for eigenvalues and eigenfunctions demonstrating closeness to the states from whence they came.

- 
- [1] P. W. Anderson, *Phys. Rev.* **109**, 1492 (1958).
  - [2] L. Fleishman and P. W. Anderson, *Phys. Rev. B* **21**, 2366 (1980).
  - [3] T. Giamarchi and H. Schulz, *Europhys. Lett.* **3**, 1287 (1987).
  - [4] B. L. Altshuler, Y. Gefen, A. Kamenev, and L. S. Levitov, *Phys. Rev. Lett.* **78**, 2803 (1997).

- [5] L. F. Santos, G. Rigolin, and C. O. Escobar, *Phys. Rev. A* **69**, 042304 (2004).
- [6] I. V. Gornyi, A. D. Mirlin, and D. G. Polyakov, *Phys. Rev. Lett.* **95**, 206603 (2005).
- [7] D. M. Basko, I. L. Aleiner, and B. L. Altshuler, *Ann. Phys. (Amsterdam)* **321**, 1126 (2006).
- [8] M. Žnidarič, T. Prosen, and P. Prelovšek, *Phys. Rev. B* **77**, 064426 (2008).
- [9] V. Oganesyan and D. A. Huse, *Phys. Rev. B* **75**, 155111 (2007).
- [10] A. Pal and D. A. Huse, *Phys. Rev. B* **82**, 174411 (2010).
- [11] B. Bauer and C. Nayak, *J. Stat. Mech.* (2013) P09005.
- [12] R. Nandkishore and D. A. Huse, *Annu. Rev. Condens. Matter Phys.* **6**, 15 (2015).
- [13] R. B. Griffiths, *Phys. Rev. Lett.* **23**, 17 (1969).
- [14] E. Hamza, R. Sims, and G. Stolz, *Commun. Math. Phys.* **315**, 215 (2012).
- [15] V. Mastropietro, *Phys. Rev. Lett.* **115**, 180401 (2015).
- [16] J. Z. Imbrie and R. Mavi, *J. Stat. Phys.* **162**, 1451 (2016).
- [17] J. M. Deutsch, *Phys. Rev. A* **43**, 2046 (1991).
- [18] M. Srednicki, *Phys. Rev. E* **50**, 888 (1994).
- [19] M. Rigol, V. Dunjko, and M. Olshanii, *Nature (London)* **452**, 854 (2008).
- [20] J. Z. Imbrie, *J. Stat. Phys.* **163**, 998 (2016).
- [21] D. A. Huse, R. Nandkishore, and V. Oganesyan, *Phys. Rev. B* **90**, 174202 (2014).
- [22] M. Serbyn, Z. Papić, and D. A. Abanin, *Phys. Rev. Lett.* **111**, 127201 (2013).
- [23] V. Ros, M. Müller, and A. Scardicchio, *Nucl. Phys.* **B891**, 420 (2015).
- [24] J. Z. Imbrie, *Commun. Math. Phys.* **341**, 491 (2016).
- [25] J. Bellissard, R. Lima, and D. Testard, *Commun. Math. Phys.* **88**, 207 (1983).
- [26] V. Chulaevsky and Y. Sinai, *Rev. Math. Phys.* **03**, 241 (1991).
- [27] L. Eliasson, *Acta Math.* **179**, 153 (1997).
- [28] N. Datta, R. Fernández, and J. Fröhlich, *J. Stat. Phys.* **84**, 455 (1996).
- [29] L. Zhang, B. Zhao, T. Devakul, and D. A. Huse, [arXiv:1603.02296](https://arxiv.org/abs/1603.02296).
- [30] J. Fröhlich and T. Spencer, *Commun. Math. Phys.* **88**, 151 (1983).