


# Local integrals of motion in many-body localized systems

John Z. Imbrie<sup>1</sup>, Valentina Ros<sup>2,4,\*</sup>, and Antonello Scardicchio<sup>3,4</sup> 

Received 27 September 2016, accepted 10 April 2017

Published online 12 May 2017

We review the current (as of Fall 2016) status of the studies on the emergent integrability in many-body localized models. We start by explaining how the phenomenology of fully many-body localized systems can be recovered if one assumes the existence of a complete set of (quasi)local operators which commute with the Hamiltonian (local integrals of motions, or LIOMs). We describe the evolution of this idea from the initial conjecture, to the perturbative constructions, to the mathematical proof given for a disordered spin chain. We discuss the proposed numerical algorithms for the construction of LIOMs and the status of the debate on the existence and nature of such operators in systems with a many-body mobility edge, and in dimensions larger than one.

## 1 Introduction

With the aim of describing the energy transport in spin systems, Anderson formulated in [1] a model of a quantum random walker in a stochastic potential landscape, now known as the “Anderson model”. He argued that when the randomness is sufficiently strong, the quantum random walker is localized by the quenched disorder, meaning that localized initial conditions do not decay, and diffusive transport is suppressed. His theoretical work laid the foundations for the theory of the quantum dynamics in a strongly-disordered environment, with implications that go far beyond the realm of solid state physics; the occurrence of localization challenges indeed the basic assumptions underlying the theory of equilibration and thermalization in isolated quantum many-body systems [2, 3].

Whether localization occurs even in presence of interactions between the constituent degrees of freedom is a question of theoretical interest and practical relevance, which motivated the search for the so called Many-Body Localized (MBL) phase [4]. The stability of the localized phase to the addition of weak scattering

has been addressed theoretically in [5, 6], by means of a perturbative treatment applied to Hamiltonians of interacting fermions in a disordered potential (similar arguments have been extended to bosonic Hamiltonians [7]). Subsequently, a large body of numerical works has revealed the occurrence of localization in one dimensional systems of interacting fermions on a lattice and spin chains in random fields [8–18]. Signatures of this phenomenon have been found in the structure of the highly-energetic many-body eigenstates, for values of the interactions that lie outside the perturbative regime. More recently, there are claims of experimental observations of MBL in artificial quantum systems of cold atoms [19–21] or trapped ions [22]. Since these experiments focus on the suppression of transport, it is still debated how well they can distinguish MBL from single-particle Anderson localization (a review of the experimental situation can be found in another article in this issue).

From the theoretical analysis it emerges that MBL systems exhibit a strongly non-ergodic dynamics, characterized by the suppression of diffusive transport of the global conserved quantities (such as energy, spin or particle number) and by the slow, power-law relaxation of local observables towards stationary values that are highly dependent on the initial condition. Remarkably, even in presence of interactions and at finite energy density, disordered quantum systems fail to thermally equilibrate following their own dynamics and remain permanently out-of-equilibrium. As such, they open interesting possibilities for the storage of quantum information, that can be locally manipulated and retrieved [23], for the protection of topological order at finite temperature, or for the realization of long-range

\* Corresponding author E-mail: valentina.ros@cea.fr

<sup>1</sup> Department of Mathematics, University of Virginia, Charlottesville, VA 22904-4137, USA

<sup>2</sup> SISSA- International School for Advanced Studies, via Bonomea 265, 34136 Trieste, Italy

<sup>3</sup> The Abdus Salam ICTP, Strada Costiera 11, 34151 Trieste, Italy

<sup>4</sup> INFN Sezione di Trieste, Via Valerio 2, 34127 Trieste, Italy

order and finite temperature phase transitions in  $d = 1$ , that would be forbidden by the equilibrium statistical mechanics [24–26].

As it is well known, the breaking of quantum ergodicity is not an exclusive feature of disordered systems, as it is realized in “conventional” integrable systems satisfying the Yang-Baxter relations [27]. For such systems, ergodicity breaking is understood in terms of an extensive set of non-trivial conservation laws. It is natural to expect that a similar “integrable” structure emerges also in the disordered case; such a structure needs however to account for the suppression of diffusive transport, which is a distinguishing property of disordered systems. Many-Body Localized systems may thus be considered as a peculiar class of integrable systems, characterized by extensively many conserved operators (Local Integrals of Motion, or “LIOMs” in this review) whose structure in space prevents not only ergodicity and thermalization, but also transport over macroscopic scales [28–30]. In the following, we review the main efforts made to substantiate this perspective.

This Review is structured as follows: In Sec. 2, we recall the main features of Many-Body Localized systems. In Sec.3 we discuss how such features can be justified in terms of the “emergent integrability”. In Secs. 4 and 5 we review the main analytical and numerical constructions of conserved quantities that have been proposed in the literature. We devote Sec. 6 to the discussion of debated issues concerning the fate of the conserved quantities in the case in which a mobility edge is present, and in dimensions higher than one.

## 2 MBL systems: a multifaceted phenomenology

In its most direct formulation, localization corresponds to the fact that local excitations do not decay. In the single particle setting, this was first shown to occur [1] for the Anderson model on a lattice  $\Lambda$ , with Hilbert space  $\ell^2(\Lambda)$  and Hamiltonian

$$H_A = \sum_{x \in \Lambda} \epsilon_x c_x^\dagger c_x - \gamma \sum_{\langle x, y \rangle} (c_x^\dagger c_y + c_y^\dagger c_x), \quad (1)$$

where  $\langle x, y \rangle$  are edges in  $\Lambda$ ,  $\gamma$  is the kinetic (“hopping”) amplitude, and  $\epsilon_x$  are independent random variables uniformly distributed in  $[-W/2, W/2]$ , defining a stochastic process indexed by the sites  $x \in \Lambda$ . It was argued in [1] that the level width of a local excitation at a site  $x \in \Lambda$ , given by  $\Gamma_x(z) = -\Im S_x(z)$  with  $S_x(z)$  the local self-energy defined by  $G_x(z) = \langle x | (z - H_A)^{-1} | x \rangle =$

$(z - \epsilon_x - S_x(z))^{-1}$ , goes to zero as the energy variable  $z = E + i\eta$  approaches the real line. This holds in probability, i.e. for almost all realizations of the random landscape. It implies that the spectrum of local excitations remains discrete in the thermodynamic limit. An analogous statement can be formulated [5, 6] for many-body fermionic Hamiltonians of the generic form

$$H_{\text{MB}} = \sum_{\alpha} E_{\alpha} n_{\alpha} + \frac{1}{2} \sum_{\alpha\beta, \gamma\delta} U_{\alpha\beta, \gamma\delta} c_{\alpha}^{\dagger} c_{\beta}^{\dagger} c_{\gamma} c_{\delta} = H_0 + U, \quad (2)$$

with  $\alpha$  labeling the eigenstates of a quadratic part (1). In this case,  $\Gamma_{\alpha}(z)$  is replaced with the imaginary part  $\Gamma_{\alpha}(\epsilon, t)$  of the (Wigner transform of the) many-body self energy associated to the retarded Green function  $G_{\alpha}^R(t_1 = t - \frac{\tau}{2}, t_2 = t + \frac{\tau}{2}) = -i\theta(t_2 - t_1) \langle \{c_{\alpha}(t_1), c_{\alpha}^{\dagger}(t_2)\} \rangle$ , and the limit  $\eta \rightarrow 0$  is replaced by the limit of vanishing coupling with an external thermal reservoir. A vanishing typical value of  $\Gamma_{\alpha}(\epsilon, t)$  implies that the irreversible evolution towards thermal equilibrium is hindered, as it follows from the fact that  $\Gamma_{\alpha}(\epsilon, t)$  enters in the collision integral of the quantum Boltzmann equation: in absence of an external reservoir, the system fails to act as a heat bath for itself.

In the theoretical works [1, 5, 6], the above criteria given in terms of the level width of local excitations are recast as a problem of convergence (in probability) of the perturbative expansion for the decay rates around the trivially localized limits  $\gamma = 0$  in (1) and  $U = 0$  in (2). The mechanism for localization extends nonetheless also beyond the perturbative regime, as it has been shown by means of numerical analysis of one-dimensional fermionic or spin Hamiltonians on the lattice. Prototypical models for the numerics are XXZ-chains in random longitudinal field [11, 13, 16, 31]:

$$H_{\text{XXZ}} = \sum_{i=-K}^{K'-1} [J (S_i^x S_{i+1}^x + S_i^y S_{i+1}^y) + J_z S_i^z S_{i+1}^z] + \sum_{i=-K}^{K'} h_i S_i^z, \quad (3)$$

or Ising chains in random transverse [14] and longitudinal fields, among which the model considered in [32]:

$$H_I = \sum_{i=-K}^{K'} h_i S_i^z + \sum_{i=-K}^{K'} \gamma_i S_i^x + \sum_{i=-K}^{K'-1} J_i S_i^z S_{i+1}^z. \quad (4)$$

In both cases,  $S_i^{\alpha}$ ,  $\alpha = x, y, z$  are Pauli matrices on sites of a lattice  $\Lambda = [-K, K'] \cap \mathbb{Z}$ , with  $S_i^z \equiv 1$  for  $i \notin \Lambda$ ,  $\gamma_i = \gamma \Gamma_i$  with  $\gamma$  small and the random variables  $h_i, \Gamma_i, J_i$  are independent and bounded, with bounded probability densities.

As a result of the effort to characterize MBL for these models, several diagnostics have been developed and exploited. In the following, we briefly recap the main features that are commonly considered as a signature of this phase; as we shall argue in Sec. 3, all of these features can be justified in terms of the “emergent integrability” of the localized system, which thus furnishes a comprising definition of MBL.

**(i) Absence of DC transport** In the single particle case, Anderson’s criterion on the level width of local excitations implies that local states  $|x\rangle \in \mathcal{L}(\Lambda)$  are bound states. Stronger statements on the dynamics of localized initial conditions  $|x\rangle$  can be rigorously proved, referred to as “dynamical localization”. They correspond to the following bound holding almost surely:

$$\sup_{t \geq 0} \left( \sum_y |y|^{2q} |\langle y | e^{-iHt} |x\rangle|^2 \right) < \infty, \quad (5)$$

implying that localized initial conditions  $|x\rangle$  have, with probability one, uniformly in time bounded moments of all orders  $q > 0$ . Such a condition rules out the possibility of transport, in particular of diffusive transport (which is instead expected in the weak disorder regime for  $d \geq 3$ ). The proof of (5) makes use of exponential bounds on either the local matrix elements of the resolvent [33], or on their fractional moments [34], or on the eigenstates correlators [35]. Such bounds encode the exponentially-decaying structure of the eigenstates  $\phi_\alpha$  of (1), whose envelope satisfies

$$|\phi_\alpha(x)| \sim A_\alpha \exp\left(-\frac{|x - r_\alpha|}{\xi_\alpha}\right) \quad (6)$$

where  $r_\alpha$  is the localization center of  $\phi_\alpha$  and  $\xi_\alpha$  its localization length.

Arguments for the vanishing of the diffusion constant are given also for the many-body case [5, 6]; similarly to the single particle case, they rely on the exponential decay of the correlations of the local density operators  $\rho_r = c_r^\dagger c_r$  on the exact many-body eigenstates  $|E_n\rangle$ :

$$\mathcal{L}_{nm}^\rho(r) = \sum_r \langle E_n | \rho_r | E_m \rangle \langle E_m | \rho_{r+r} | E_n \rangle \lesssim \exp\left(-\frac{|r|}{\xi(E)}\right), \quad (7)$$

where  $E_n \approx E_m \approx E$  and  $\xi(E)$  is an energy dependent localization length. The absence of diffusion has been investigated numerically by analyzing the low-frequency behavior of the dynamical conductivity and its dc limit [36, 37] and the spin-spin or density-density correlation functions in the infinite-time limit [9, 13, 38]. The van-

ishing of the transport coefficients is a peculiar feature of MBL systems, which contrasts with the efficient transport properties of clean integrable systems [39–41]. Nevertheless, it is not an exclusive feature of this phase, as in one dimension the delocalized phase also displays subdiffusive transport [42–46].

**(ii) Anderson localization in configuration space.** The perturbative treatment in [5, 6] is consistent with the picture, originally proposed in the influential work [47] (see also the construction in Sec. 4.2 in this review), of MBL as Anderson localization on an abstract graph whose sites correspond to the Fock states diagonalizing the non-interacting part of (2), and whose geometry is determined by the interactions. A similar picture can be formulated for the spin Hamiltonians (3) and (4), with the Fock states replaced by the tensor products of the simultaneous eigenstates of the operators  $S_i^z$ . This scenario entails that the many-body eigenstates are “weak deformations” of the non-interacting states, a point of view emphasized in [10]. This picture is corroborated by several numerical diagnostics, such as the scaling of the IPR of MBL eigenstates in the basis of non interacting states [11, 16], or the expectation values of the non-interacting occupation numbers or  $S_i^z$  operators, which are shown to be close to their non-interacting value also on MBL eigenstates [13]. The latter statement is rigorously proven in [32] for the chain (4), see Theorem 1 in Sec. 4.3.

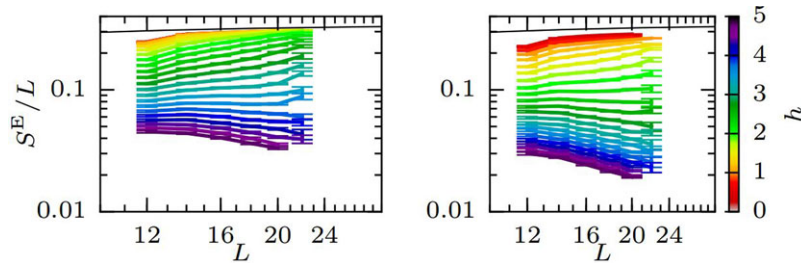
**(iii) Area-law entanglement in highly excited states.** A key implication of the above scenario is the low entanglement entropy in the excited eigenstates  $|E_n\rangle$  of MBL Hamiltonians. For one-dimensional systems, this is captured by the bipartite eigenstate entanglement entropy, obtained by splitting the chain into a right half R and a left half L, and by tracing the degrees of freedom corresponding to one of the halves,

$$S = -\text{Tr}(\rho_R \log_2 \rho_R) \quad (8)$$

with

$$\rho_R = \text{Tr}_L(|E_n\rangle\langle E_n|). \quad (9)$$

It is shown in [10, 14, 16] that  $S$  does not scale with system size but obeys an area-law, see Fig. 1. This is a typical ground-state property for generic gapped Hamiltonian, which in the MBL case extends to the whole spectrum. It implies that MBL eigenstates of extensive energy can be efficiently represented via Density-Matrix-RG or Matrix Product States [48–51] and Tensor Networks [52]. Moreover, it allows to access to the dynamical properties



**Figure 1** Bipartite entanglement entropy  $S^E$  defined in Eq. (8) for the Hamiltonian (3) with  $J = J_z = 1$  and  $h_a \in [-h, h]$ , as a function of system size  $L$  for different disorder strengths in the middle of the spectrum (left) and in the upper part (right). For strong disorder,  $S^E/L$  decreases signaling area-law. The figure is taken from Ref. [16].

in the MBL phase adapting computational approaches originally designed for the ground-state physics, such as the strong disorder renormalization group exploited in [24] to construct approximate many-body eigenstates, and in [15, 53] to describe the dynamical evolution of initial product states. Further signatures of MBL are encoded in the structure of the full entanglement spectrum [51, 54, 55]: in particular, a distinctive feature of MBL is the power-law decrease [51] of the size of the eigenvalues of the reduced density matrix (9).

**(iv) Violation of the Eigenstate Thermalization Hypothesis.** The structure of the MBL eigenstates is inherently incompatible with the Eigenstate Thermalization Hypothesis (ETH), formulated in [56–59] as a microscopic justification for quantum thermalization. The ETH conjectures that the individual eigenstates  $|E_n\rangle$  of a thermalizing Hamiltonian locally reproduce the thermodynamic ensembles, meaning that the eigenstates expectation values  $\langle E_n|O|E_n\rangle$  of local observables  $O$  depend smoothly on the energy of the state, and coincide with the microcanonical expectation values at energy  $E \approx E_n$ . This hypothesis, together with the assumption of exponential decay (with the system size) of the off-diagonal matrix elements  $\langle E_n|O|E_m\rangle$ , guarantees that out-of-equilibrium initial state relax to states that are “locally thermal”.

Contrary to the ETH requirements, MBL eigenstates in the same energy shell are locally distinguishable and non-thermal: the expectation values of the local observables are far from their equilibrium value, and they strongly fluctuate between states that are close in energy [60], see Fig. 2(a). ETH is also incompatible with the area-law scaling of the bipartite entanglement, as it requires the entanglement to be equal to the thermal equilibrium entropy of the subsystem, that scales with the number of its degrees of freedom. This dichotomy has been exploited extensively to pinpoint the phase diagram of disordered systems by probing the violation of ETH in individual eigenstates obtained from the exact diagonalization; it is however useful to keep in mind that while localization is a statement on the long-time limit

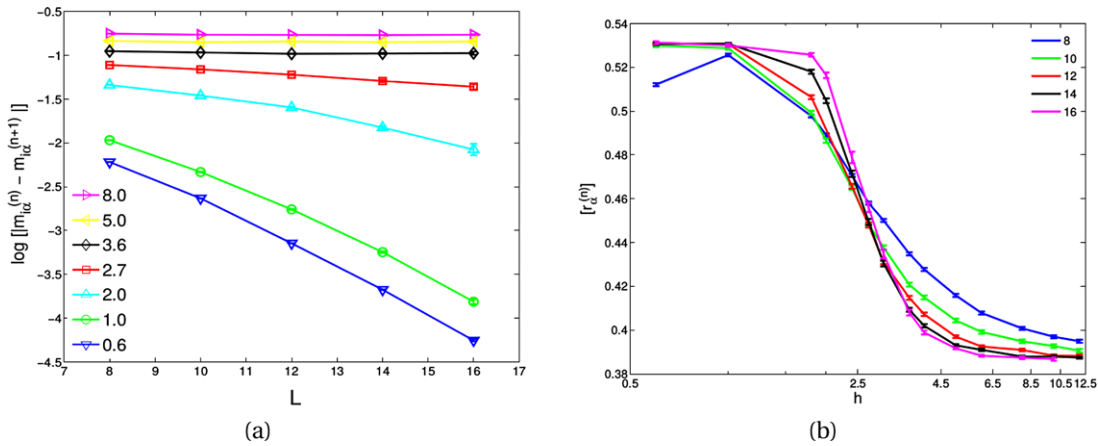
of thermodynamically large systems, the ETH ansatz determines the stationary behavior of finite-size samples: these two approaches are not necessarily equivalent [61].

**(v) Absence of level repulsion.** Starting from the earlier works [62, 63], the absence of level repulsion, generically regarded as a signature of integrability [64–66], has been extensively exploited as a characterization of the MBL phase, see also [13, 31, 63, 67–69]. An extensively used indicator of the absence of repulsion [8] is the average value  $\langle r \rangle$  of the dimensionless ratio

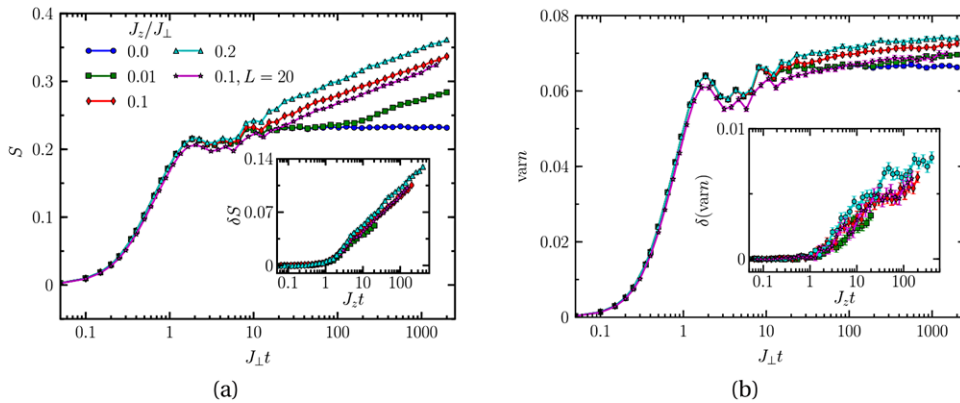
$$r_n = \frac{\min\{E_{n+1} - E_n, E_{n+2} - E_{n+1}\}}{\max\{E_{n+1} - E_n, E_{n+2} - E_{n+1}\}} \quad (10)$$

obtained from the gaps between consecutive eigenvalues. In the localized phase, for increasing system size the latter approaches the Poisson theoretical value  $\langle r \rangle \approx 0.39$ , see Fig. 2(b).

**(vi) Slow growth of entanglement and slow relaxation.** Despite the structure of MBL eigenstates is not significantly altered by the interactions, the latter strongly affect the quantum dynamics. Their signatures are traced in the real-time evolution of the bipartite entanglement entropy  $S(t)$ , obtained as in (8), with the substitution of  $|E_n\rangle$  in (9) with the time-evolved pure state  $|\psi(t)\rangle$  of the entire system. Numerical simulations performed on disordered spin chains initialized in a low-entangled/product state [9, 70, 71] display a transient fast growth of  $S(t)$  (dominated by the direct nearest-neighbor interactions across the cut and extending to times of the order of the inverse interaction coupling), followed by a slow logarithmic growth, see Fig. 3(a). The increase in entropy is expected to continue indefinitely for an infinite system, while for finite systems  $S(t)$  saturates to a value which depends on the initial state only, it scales with the system size but it is nevertheless smaller than the one to be expected in the thermal regime [72, 73].



**Figure 2** (a) Logarithm of the averaged difference between the local magnetizations  $m_{i\alpha}^{(n)} = \langle n_i S_i^z | n \rangle_\alpha$  in adjacent eigenstates of the Hamiltonian (3) with  $J_z = J = 1$  and  $h$  indicated in the legend. The average is over the disorder realization  $\alpha$  and the pairs of eigenstates. For large  $h$ , the differences remain large as the length of the chain  $L$  is increased. (b) Ratio of adjacent energy gaps defined in Eq. (10) for different system sizes  $L$  indicated in the legend. For large  $h$ , the level statistics are Poisson. Figures taken from Ref. [13].



**Figure 3** (a) Unbounded growth of the bipartite entanglement after a quench starting from a site-factorized  $S^z$  eigenstate of the Hamiltonian (3) with  $J = J_\perp$ ,  $h_\alpha \in [-5, 5]$ ,  $L = 10$  and different interaction strengths  $J_z$ . The inset shows the same data with a rescaled time axis and subtracted  $J_z = 0$  values. (b) Growth of the particle number fluctuations of a half chain after the quench. The behavior is qualitatively different than the entanglement entropy: the interactions do enhance the particle number fluctuations, but while there are signs of a logarithmic growth as for the entanglement, this growth slows down with time. Figures taken from Ref. [71].

The logarithmic scaling is a peculiar feature of the MBL phase<sup>1</sup>, that allows to distinguish it from the non-interacting localized phase, where  $S(t)$  saturates to a finite value independent of the system size. The unbounded growth of entanglement is perfectly compatible with the absence of transport, see Fig. 3(b). It is ascribed to the interaction-induced dephasing between the eigenstates involved in the decomposition of the initial product state. The same mechanism is at the root of

other distinguishing dynamical features of MBL systems, such as the power-law decay in time of the response to a (properly designed) spin-echo protocol [23], the power-law relaxation (towards non-thermal values) of the expectation value of local observables [79] or of the concurrence (a measure of the non-classicality of the system) [80] after a quench, and the suppression, with respect to the non-interacting case, of the revival rates of single-site observables [81].

We discuss in the following section how these different aspects of MBL can be understood by assuming the existence of extensively-many, non-trivial conservation laws. The discussion is done assuming

<sup>1</sup> As opposed to the linear growth in clean integrable or non-integrable systems [74, 75] and to the sub-ballistic growth in disordered, delocalized systems close to the MBL phase [76–78].

the *full* many-body spectrum is localized, irrespectively of the energy density. This assumption applies to lattice Hamiltonians having a bounded local spectrum, as pointed out in [8]; the corresponding systems are often called ‘fully MBL’ in the literature. Comments on the occurrence of mobility edges are postponed to Sec. 6.

### 3 Quasilocals integrals of motions: a unifying framework

It is a common expectation that the failure of ergodicity in closed, interacting systems is related to some form of integrability. Indeed, the phenomenology of MBL systems summarized in Sec.2 suggests that conservation laws exist for strongly disordered systems. This expectation has been made concrete in [28–30], where it was suggested that MBL Hamiltonians are non-linear functionals of a complete set of LIOMs  $I_\alpha$ , of the form

$$H_{diag} = h_0 + \sum_{\alpha} h_{\alpha} I_{\alpha} + \sum_{\alpha, \beta} h_{\alpha\beta} I_{\alpha} I_{\beta} + \sum_{\alpha, \beta, \gamma} h_{\alpha\beta\gamma} I_{\alpha} I_{\beta} I_{\gamma} + \dots, \quad (11)$$

where the dots stand for higher order products. The  $I_\alpha$  are expected to be functionally independent and mutually commuting. The set is complete in the sense that every many-body eigenstate can be labeled in a unique way with the eigenvalues of the  $I_\alpha$ .

While the expansion (11) is to some extent generic (in particular, to specify  $H_{diag}$  it is necessary to determine a number of coefficients which scales with the size of the Hilbert space [82]), the fingerprint of localization is the “*quasilocality*” of the operators  $I_\alpha$ . The notion of quasilocality extends, at the operator level, the structure of the single-particle eigenstates  $\psi_\alpha$  of (1), cfr. Eq. (6): the operator norm of  $I_\alpha$  is expected to decay exponentially away from a compact region of space of a typical size  $\xi_{op}$  around a center  $R_\alpha$ . More precisely, let

$$I_\alpha = \sum_{\mathcal{I}} \mathcal{A}_{\mathcal{I}}^{(\alpha)} O_{\mathcal{I}} \quad (12)$$

be the expansion of  $I_\alpha$  on a basis of local operators  $O_{\mathcal{I}}$  labeled by  $\mathcal{I}$ . For the spin chains (3) and (4), a suitable basis is made by the tensor products of local spin operators  $S_{i_1}^{\alpha_1} \otimes \dots \otimes S_{i_n}^{\alpha_n}$  with  $\alpha_i \in \{x, y, z\}$  and  $i_k \in \Lambda$ , while in the fermionic case (2) the normal-ordered tensor products of creation and annihilation operators of single-particle states can be considered, see Eq. (26). Let  $S(\mathcal{I})$  denote the “support” of  $O_{\mathcal{I}}$ , i.e. the set of sites/local degrees of

freedom on which the operators acts non-trivially. Quasilocality entails that

$$|\mathcal{A}_{\mathcal{I}}^{(\alpha)}| \lesssim C_\alpha \exp\left(-\frac{d[R_\alpha, S(\mathcal{I})]}{\xi_{op}}\right), \quad (13)$$

where  $d[R_\alpha, S(\mathcal{I})]$  is the distance between  $R_\alpha$  and the furthest degree of freedom in  $S(\mathcal{I})$ . This encodes the property of the  $I_\alpha$  of being weak deformations of the local, physical degrees of freedom (such as the local density or spin operators).

This spatial structure implies that the expansion (11) is not structureless, as the typical value of the coefficients  $h_{\alpha\beta\dots}$  decays exponentially in the distance between the localization centers of the corresponding operators, on a typical scale that we denote with  $\xi_{int}$ . This structure is at the root of the genuinely many-body dynamical features of MBL systems, as it entails that the dephasing between the  $I_\alpha$  induced by the interactions in (11) occurs over a broad range of time scales.

The phenomenological models in [28–30] additionally assume that (i) the  $I_\alpha$  have binary spectrum, and thus can be considered as effective spins or occupation number operators (also termed “logical bits” or “1-bits”), and (ii) the full spin (or fermionic) algebra can be constructed, with ladder operators  $I_\alpha^\pm$ . This follows straightforwardly from the construction in [32], where the conserved quantities of the disordered chain (4) are obtained from the Pauli operators  $S_a^z$  by means of a quasilocal rotation  $\Omega$ ,  $I_\alpha = \Omega S_a^z \Omega^*$  and  $I_\alpha^\pm = \Omega S_a^\pm \Omega^*$ , that obviously preserves the spectrum and the commutation relations.

This set of assumptions is sufficient to derive the full phenomenology of MBL systems, as we shortly recap in the following.

- (i) *Absence of DC transport*: we report the argument given in [83]. The Kubo formula for the dc conductivity  $\sigma$  associated to a local current-density  $J_r$  (supported on a finite set of sites in the vicinity of  $r$ ) reads:

$$\begin{aligned} \text{Re}[\sigma(\omega \rightarrow 0)] &= \frac{\pi\beta}{|\Lambda|} \sum_{r,r'} \sum_{m \neq m'} \frac{e^{-\beta E_m}}{\mathcal{Z}} \times \\ &\times \langle E_{m'} | J_{r+r} | E_m \rangle \langle E_m | J_{r'} | E_{m'} \rangle \delta_\eta(E_{m'} - E_m), \end{aligned} \quad (14)$$

where  $|\Lambda|$  is the system’s volume,  $\mathcal{Z}$  is the partition function at inverse temperature  $\beta$ ,  $|E_m\rangle$  the system’s eigenstates and  $\delta_\eta(x) = \pi^{-1} \eta / (x^2 + \eta^2)$  a regularized  $\delta$ -function. Since the set of LIOMs is complete, for any pair  $m, m'$  there is a  $\tilde{I}$  such that  $\tilde{I}|E_m\rangle = \tilde{I}|E_{m'}\rangle$  and  $\tilde{I}|E_{m'}\rangle = \tilde{I}|E_m\rangle$  and  $\tilde{I}_{m'} \neq \tilde{I}_m$ . If the LIOMs are *strictly local* (in the sense that their support is finite and compact, of size  $\xi_{op}$ ), then for  $r > \xi_{op}$ , one of the

two matrix elements

$$\begin{aligned}\langle E_{m'} | J_{r'} | E_m \rangle &= \frac{\langle E_{m'} | [J_{r'}, \tilde{I}] | E_m \rangle}{(\tilde{I}_m - \tilde{I}_{m'})}, \\ \langle E_{m'} | J_{r'+r} | E_m \rangle &= \frac{\langle E_{m'} | [J_{r'+r}, \tilde{I}] | E_m \rangle}{(\tilde{I}_m - \tilde{I}_{m'})}\end{aligned}\quad (15)$$

is exactly zero: in Eq. (14) the sum over  $r$  is restricted to  $r \lesssim \xi_{op}$ . Furthermore, for any  $m$  the sum over  $m'$  is restricted to a finite set, since  $J_{r'} | E_m \rangle$  can differ from  $| E_m \rangle$  only in a finite number of LIOMs. As a consequence, in the thermodynamic limit, when  $\eta \rightarrow 0$ , the contribution to the  $\delta$ -function vanishes with probability one, and  $\text{Re}[\sigma(\omega = 0)] = 0$ . For *quasilocal* conserved quantities, the matrix elements  $\langle m' | J_{r'} | m \rangle$  are not exactly zero also for those eigenstates for which  $\tilde{I}$  is supported at distance  $x\xi_{op}$  from  $r'$ : they are instead exponentially small in  $x$ . Since there are also exponentially many states  $m$ ,  $m'$  which satisfy these criteria, some energy differences  $E_m - E_{m'}$  in (14) become exponentially small. The competition between the matrix elements and the energy denominators is however dominated by the exponential decay of the matrix elements with probability one: this is the key statement that guarantees the stability of the MBL phase as well as the existence of quasilocal  $I_\alpha$ . It follows that the conductivity remains zero also when quasilocality is properly taken into account.

- (ii) *Anderson localization in configuration space*: the exponential decay of the amplitudes (13) can be interpreted as exponential localization in a space with sites labeled by  $\mathcal{I}$ , which can be put in one to one correspondence with Fock states, see Sec. 4.2, or non-entangled product states in the spin case. Notice however that contrary to the Anderson model, in the many-body case an exponential decay of the amplitude in Fock space starting from a reference configuration does not imply  $O(1)$  participation ratios for the wave function, due to the fast growth of the number of configurations with the distance from the reference configuration.
- (iii) *Area-law entanglement in highly excited states*: all the eigenstates of (11) are product states of the  $I_\alpha$ ; the bipartite entanglement entropy is thus area-law, as it is effectively contributed only by those  $I_\alpha$  that are localized in the vicinity (within  $\xi_{op}$ ) of the cut.
- (iv) *Violation of the ETH*: as in conventional integrable systems, the locality of the conserved quantities implies that local memory of the initial condition is preserved at any time, thus preventing thermalization. Moreover, since the many-body eigen-

states are simultaneous eigenstates of the  $I_\alpha$  that are weak deformations of the non-interacting occupation numbers  $n_\alpha$  (or of the spin operators  $S_\alpha^z$ ), it follows that the expectation value of the latter on exact eigenstates does not depart significantly from  $\pm 1$ . It is thus non-thermal and it fluctuates significantly between states that are close in energy, as such states differ by the eigenvalues of integrals of motion having a large overlap with the local operator under consideration.

- (v) *Absence of level repulsion*: the absence of level repulsion arises because adjacent states in the spectrum typically differ by an extensive number of eigenvalues of the  $I_\alpha$ , they are unable to hybridize and thus do not repel at the scale of the mean level spacing.
- (vi) *Slow growth of entanglement and slow relaxation*: the interaction terms in (11) induce dephasing between the conserved operators; given a pair of operators  $I_\alpha, I_\beta$ , their dephasing time  $\tau$  scales with the inverse of the interaction between them, and thus it depends on the distance  $l$  between the corresponding localization centers,  $\tau \sim H_0^{-1} e^{J/\xi_{int}}$  where  $H_0$  is the typical interaction scale. It follows that at a given time  $t$ , only the degrees of freedom that are within the distance  $l(t) \sim \xi_{int} \log(H_0 t)$  are dephased. The logarithmic growth of the bipartite entanglement follows from the fact that the entanglement entropy produced at time  $t$  (starting from a product state) is proportional to the volume of degrees of freedom that have dephased up to that time [72, 73]. A similar argument implies that the expectation value of operators  $O$  with finite support decays as a power-law in time [79]. Indeed, when expanded in the basis of the  $I_\alpha, I_\alpha^\pm$ , the operator reads

$$O = \bar{O} + O_{osc}, \quad (16)$$

where  $O_{osc}$  contains  $I_\alpha^\pm$  terms, while  $\bar{O}$  commutes with (11), and can be written as

$$\bar{O} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T O(t) dt = \sum_{\vec{\alpha}} C_{\vec{\alpha}} I_{\alpha_1} I_{\alpha_2} \cdots I_{\alpha_n} \quad (17)$$

where  $\vec{\alpha} = (\alpha_1, \dots, \alpha_n)$  (due to the quasilocality of the  $I_\alpha$  and the locality of  $O$ , the coefficients  $C_{\vec{\alpha}}$  have themselves an exponentially decaying structure). Let  $|\psi\rangle = \sum_I A_I |I\rangle$  be an initial state expanded in the basis of simultaneous eigenstates  $|I\rangle$  of the  $I_\alpha$ . It

is argued in [79] that the second term in

$$\langle \psi(t) | O | \psi(t) \rangle = \sum_I |A_I|^2 \langle I | \bar{O} | I \rangle + \sum_{I, J} A_I A_J^* e^{i(E_I - E_J)t} \langle J | O_{osc} | I \rangle \quad (18)$$

decays as a power law in time, due to the randomization of the relative phases in the eigenstates decomposition generated by the interactions of the  $I_\alpha, I_\beta$ . For example, let  $O_{osc} = I_\gamma^+$ , and consider for simplicity the case in which  $|A_I|^2 \sim 2^{-N}$  with  $N$  the number of spins. For any state  $|I\rangle$  with  $i_\gamma = -1$ , let  $|J\rangle = |\bar{I}\rangle$  be the state with  $\tilde{i}_\gamma = +1$  and  $\tilde{i}_\alpha = i_\alpha$  for  $\alpha \neq \gamma$ . The second term in the right-hand side of (18) using (11),

$$\frac{1}{2^N} \sum_{I: i_\gamma = -1} \exp \left[ 2i \left( h_\gamma + \sum_\beta h_{\gamma\beta} i_\beta + \sum_{\beta, \delta} h_{\gamma\beta\delta} i_\beta i_\delta + \dots \right) t \right]. \quad (19)$$

For any time  $t$ , (19) can be splitted into the sum over quantum numbers  $i_\alpha$  whose operators are localized within distance  $l(t) \sim \xi_{int} \log(H_0 t)$  from  $I_\gamma$ , and the sum over quantum numbers of operators localized at larger distance. Any of the terms at the exponent involving quantum numbers in the first set is dephased at time  $t$ : (19) is therefore proportional to a sum over  $N(t) \sim 2^{2l(t)}$  random phases with zero average, which scales as  $\sim 1/\sqrt{N(t)} = (H_0 t)^{-\xi_{int} \log 2}$ . The expectation value (18) thus exhibits a power law relaxation to the constant (non-thermal) value  $\sum_I |A_I|^2 \langle I | \bar{O} | I \rangle$ .

#### 4 Construction of conservation laws: analytic schemes

It was pointed out in [29] that, for an arbitrary many-body system with  $N$  degrees of freedom, any bijection between the eigenstates and the  $2^N$  binary strings  $|i_{\alpha_1} i_{\alpha_2} \dots i_{\alpha_N}\rangle$  with  $i_\alpha = \pm 1$  (or  $i_\alpha \in \{0, 1\}$ ) defines a complete set of conserved quantities  $I_\alpha$  with binary spectrum  $\{i_\alpha\}$ . The latter are simply the operators whose quantum numbers  $i_\alpha$  label the eigenstates. They can be obtained as  $I_\alpha = \sum_{I'} |I'\rangle \langle I' | - |\bar{I}'\rangle \langle \bar{I}'|$ , assuming  $i_\alpha = \pm 1$  and denoting with  $|I'\rangle$  a many-body eigenstate with  $i_\alpha = +1$ , and with  $|\bar{I}'\rangle$  the state with all quantum numbers equal except  $i_\alpha = -1$  (see Eq. (54)). In [29] it was conjectured that,

for MBL systems, an optimal assignment exists that results in (quasi-)local operators  $I_\alpha$ . Although this has not led to a practical way to find LIOMs, that early work was instrumental for further development of the idea.

In this section, we analyze the constructive analytic recipes that have been proposed to construct the conserved operators starting from microscopic models. The discussion on numerical schemes is postponed to Sec. 5.

##### 4.1 Infinite-time averages of local densities

It was pointed out in [84] that the infinite-time average of any local operator  $O$  is a conserved quantity. If an underlying algebra of quasilocal operators  $I_\alpha, I_\alpha^\pm$  exists, the average (17) of operators with finite, compact support is itself quasilocal. For a spin Hamiltonian such as (3), an extensive set of LIOMs can be obtained as the infinite-time average of the local energy- or spin-densities operators, i.e.  $O_i \rightarrow S_i^z$  or  $O_i \rightarrow H_i \equiv J(S_i^x S_{i+1}^x + S_i^y S_{i+1}^y) + J_z S_i^z S_{i+1}^z + h_i S_i^z$  [84]. The conserved operators which result from this procedure are not pseudospins, as the time-averaging does not preserve the spectrum. However, they have the advantage of being measurable in the following sense: for a chain with  $N$  spins with  $O_i = S_i^z$ , the coefficients in the expansion

$$\bar{S}_i^z = \sum_{L, \kappa} M_i^\kappa S_i^{\kappa_1} \dots S_n^{\kappa_n}, \quad (20)$$

with  $\kappa_i \in \{z, x, y\}$ , can be obtained measuring multi-spin correlations on a time-averaged density matrix  $\bar{\rho}$ ,

$$M_i^\kappa = 2^{-N} \text{Tr} (S_i^{\kappa_1} \dots S_n^{\kappa_n} \bar{\rho}), \quad (21)$$

where  $\rho$  at  $t = 0$  describes the state with magnetization one at site  $i$  and zero everywhere else,  $\rho = 2^{-N} (1 + S_i^z) \otimes \prod_{k \neq i} \mathbf{1}_k$ . The  $\bar{S}_i^z$  thus provide information on the spreading of the spin through the infinite temperature ensemble.

The alternative choice  $\bar{O}_i \rightarrow \bar{H}_i$  is used in [85] to derive a Lieb-Robinson bound for the information propagation in the MBL phase, with a logarithmic lightcone. The bound implies that the growth of the bipartite entanglement entropy is at most logarithmic. Its derivation relies on a stronger condition of quasilocality for the  $\bar{H}_i$  (formulated in terms of averages rather than of typical values), requiring that there exists a constant  $\xi$  such that

$$\mathbb{E} (\| [\bar{H}_i, O] \|) \leq e^{-x/\xi} \|O\| \quad (22)$$

for any operator  $O$  whose support is at distance  $x$  from the site  $i$ , where  $\mathbb{E}[\cdot]$  denotes the disorder average.



## 4.2 Perturbative dressing of the non-interacting occupation numbers

Consider the fermionic Hamiltonian (2): in the absence of interactions  $U = 0$ , the occupation numbers  $n_\alpha$  associated to the single particle eigenstates  $\phi_\alpha$  are mutually commuting, quasilocal conserved quantities. Their quasilocal nature follows directly from the spatial localization of the single particle wave-functions. Indeed, the operator  $c_i^\dagger c_j$  contributes to the operator expansion:

$$n_\alpha = \sum_{i,j} \phi_\alpha^*(i) \phi_\alpha(j) c_i^\dagger c_j \quad (23)$$

with a weight  $\phi_\alpha^*(i) \phi_\alpha(j)$  which decays exponentially in the distance between its support (the sites  $i, j$ ) and the localization center  $r_\alpha$  of  $\phi_\alpha$ . By truncating the sum (23) to terms with support only within a neighborhood of  $m\xi$  of  $r_\alpha$  one obtains an operator whose commutator with the Hamiltonian vanishes up to exponentially small terms. As  $m \rightarrow \infty$  the operator rapidly converges (in the operator norm) to the conserved density  $n_\alpha$ . It is natural to expect that for weak  $U$ , quasilocal operators that are conserved by the *full* interacting Hamiltonian can be built as weakly dressed versions of the (23). In order to confirm this expectation, it is necessary to show that the perturbation theory in operator space converges in the regime of parameters corresponding to weak scattering processes; this is argued in [83] for a fermionic Hamiltonian on a  $d$ -dimensional lattice.

Before summarizing the main steps of the argument, we introduce the relevant energy scales in the problem. The convergence of the perturbative expansion is argued for a simplified version of the model (2), that is coarse-grained at the scale of the localization length  $\xi$  of the single particle states  $\phi_\alpha$ . As in [6], the lattice  $\Lambda$  is partitioned into volumina of size  $\xi$  (henceforth called “localization volumina”), and each single particle state is assigned to a volume according to the position of its localization center. An energy scale is associated to the localization volumina: it is the average energy-gap between the single-particle states belonging to it,  $\delta_\xi = 1/\nu \xi^d$  with  $\nu$  the density of states. We assume  $\delta_\xi \ll \mathcal{W}$  where  $\mathcal{W}$  denotes the total width of the single particle spectrum. The interaction matrix elements  $U_{\alpha\beta,\gamma\delta}$  are taken to be non-zero only if the corresponding single-particle states are in the same volume or in adjacent ones, and provided that

$$|E_\alpha - E_\delta|, |E_\beta - E_\gamma| \lesssim \delta_\xi \quad \text{or} \quad |E_\alpha - E_\gamma|, |E_\beta - E_\delta| \lesssim \delta_\xi. \quad (24)$$

In this case, they are set equal to  $U_{\alpha\beta,\gamma\delta} = \lambda \delta_\xi \eta_{\alpha\beta,\gamma\delta}$  where  $\eta_{\alpha\beta,\gamma\delta}$  is a uniform variable in  $[-1, 1]$  and  $\lambda$  is a dimensionless constant measuring the interaction strength. Correlations between the single-particle energies  $E_\alpha$  are neglected.

The setup for the construction is the following: an expansion is assumed for the conserved operators  $I_\alpha$ , of the form:

$$I_\alpha = n_\alpha + \sum_{N \geq 1} \sum_{\substack{\mathcal{I} \neq \mathcal{J} \\ |\mathcal{I}|=N=|\mathcal{J}|}} \mathcal{A}_{\mathcal{I},\mathcal{J}}^{(\alpha)} \left( O_{\mathcal{I},\mathcal{J}} + O_{\mathcal{I},\mathcal{J}}^\dagger \right) \quad (25)$$

where  $\mathcal{I} = (\beta_1, \dots, \beta_N)$  and  $\mathcal{J} = (\gamma_1, \dots, \gamma_N)$  are sets of indices labeling the single particle states, and

$$O_{\mathcal{I},\mathcal{J}} = \prod_{\beta \in \mathcal{I}} c_\beta^\dagger \prod_{\gamma \in \mathcal{J}} c_\gamma \quad (26)$$

is a normal ordered operator (an ordering between the single particle indices is also assumed). The operator expansion (25) corresponds to a number operator dressed with strings of excitations. It is non-generic due to the constraint  $\mathcal{I} \neq \mathcal{J}$ , which is imposed to ensure that the coefficients  $\mathcal{A}_{\mathcal{I},\mathcal{J}}^{(\alpha)}$  are *uniquely* fixed by the conservation condition  $[I_\alpha, H] = 0$  (see the following discussion). The resulting coefficients depend on the interaction strength  $\lambda$ , and approach zero as  $\lambda \rightarrow 0$ . Whenever the operator expansion converges in norm, the Ansatz (25) defines a complete set of conserved quantities that are expected to be functionally independent at finite  $\lambda$  (as they are for  $\lambda = 0$ ) and mutually commuting (the commutation relations  $[I_\alpha, I_\beta] = 0$  for  $\alpha \neq \beta$  are not explicitly enforced, but assumed to hold due to the fact that the spectrum is almost-surely non degenerate.).

The goal is to argue that for sufficiently small  $\lambda$  the expansion (25) converges in probability, that is, that there is a  $\lambda_c$  such that, for  $\lambda < \lambda_c$  and for any  $\epsilon > 0$ , the following condition is satisfied:

$$\lim_{R \rightarrow \infty} \mathbb{P} \left( \sum_{\substack{\mathcal{I} \neq \mathcal{J} \\ r(\mathcal{I},\mathcal{J}) > R}} |\mathcal{A}_{\mathcal{I},\mathcal{J}}^{(\alpha)}| < \epsilon \right) = 1, \quad (27)$$

where  $r(\mathcal{I}, \mathcal{J})$  is the maximal distance between the localization center of  $\phi_\alpha$  and any of the states  $\phi_\beta$  that are acted upon by  $O_{\mathcal{I},\mathcal{J}}$ . This ensures that the series defining the operator  $I_\alpha$  converges almost surely, since  $\|O_{\mathcal{I},\mathcal{J}}\| = 1$  for all  $\mathcal{I}, \mathcal{J}$ . The arguments for the convergence rely on a mapping to an equivalent single-particle problem, obtained imposing  $[I_\alpha, H] = 0$  and interpreting the resulting linear constraints for the amplitudes  $\mathcal{A}_{\mathcal{I},\mathcal{J}}^{(\alpha)}$  as the equations for a single particle hopping in

“operator space”. The latter is a disordered lattice with sites labeled by the Fock indices ( $\mathcal{I}, \mathcal{J}$ ) and links determined by the conditions (24) on the interactions. Within this formalism, the exponential decay of the coefficients  $\mathcal{A}_{\mathcal{I}, \mathcal{J}}^{(\alpha)}$  corresponds to the localization of the effective single particle in operator space. At the operator level, it implies that the resulting operator  $I_\alpha$  is quasilocal in the sense of (13).

We now discuss in slightly more detail the main steps leading to (27). We anticipate that the exponential convergence is proved within a “forward approximation” [1, 86–88] (henceforth FA), which in this setting boils down to replacing each amplitude in (25) with its lowest order expansion in the strength of the interaction  $\lambda$ . This approach, although approximate, has the advantage of being explicit, and to allow one to obtain analytic expressions for relevant quantities characterizing the MBL phase [89]. For an exact treatment that goes beyond this approximation, we refer the reader to Sec. 4.3.

#### 4.2.1 Formal perturbation theory and the problem of local resonances

When attempting to construct conserved quantities perturbatively, it is natural to consider the formal expansion

$$I_\alpha = n_\alpha + \sum_{n \geq 1} \lambda^n \Delta I_\alpha^n, \quad (28)$$

where  $\Delta I_\alpha^n$  is determined recursively from  $\Delta I_\alpha^{n-1}$ , as the solution of the conservation equation

$$[U, \Delta I_\alpha^{n-1}] + [H_0, \Delta I_\alpha^n] = 0. \quad (29)$$

This equation can be solved for  $\Delta I_\alpha^n$  provided that the commutator with  $H_0$  can be inverted; this requires that at any order  $n$ , the operator  $[U, \Delta I_\alpha^{n-1}]$  does not belong to the kernel  $K$  of the map  $(\text{ad } H_0)X = [H_0, X]$ , which can be easily shown to be true for time-reversal symmetric Hamiltonians. The condition (29) does not determine  $\Delta I_\alpha^n$  uniquely: indeed, arbitrary terms  $\Delta K_\alpha^n$  belonging to  $K$  can be added to  $\Delta I_\alpha^n$ , and the resulting operator  $\Delta I_\alpha^n + \Delta K_\alpha^n$  would still satisfy (29). We may thus write

$$\begin{aligned} \Delta I_\alpha^n &= i \lim_{\eta \rightarrow 0} \int_0^\infty d\tau e^{-\eta\tau} e^{iH_0\tau} [U, \Delta I_\alpha^{n-1}] e^{-iH_0\tau} + \Delta K_\alpha^n \\ &= \Delta J_\alpha^n + \Delta K_\alpha^n, \end{aligned} \quad (30)$$

where the first term in (30) belongs to the space spanned by the normal ordered operators (26) with  $\mathcal{I} \neq \mathcal{J}$  (which is the orthogonal complement of  $K$ ), while  $\Delta K_\alpha^n \in K$ . It is shown in [83] that there exists a unique choice of  $\Delta K_\alpha^n \in K$  that ensures that the spectrum of (28) is binary, i.e.  $I_\alpha^2 = I_\alpha$ , at the given perturbative order. This term can be written as

$$\Delta K_\alpha^{(n)} = (1 - 2n_\alpha) \left[ \sum_{m=1}^{n-1} \Delta I_\alpha^m \Delta I_\alpha^{n-m} + \left\{ n_\alpha - \frac{1}{2}, \Delta J_\alpha^n \right\} \right]. \quad (31)$$

Thus, the formal perturbative recipe uniquely defines a set of conserved number operators obtained summing (30). Let us now address the problem of convergence and quasilocality.

Despite the perturbative equations are solvable at any finite order, the resulting series (28) diverges almost surely, due to occurrence of terms with large norm that, even if rare, appear *repeatedly* in (28), at any order in  $\lambda$ , leading to a divergence. These are the same “trivial” divergences affecting the perturbation theory for the self energies discussed in P. W. Anderson’s work [1]; they signal the presence of *rare resonances* between almost degenerate states of the unperturbed Hamiltonian, that are strongly hybridized by the hopping/interactions. The presence of isolated resonances does not necessarily entail delocalization, as long as they do not proliferate in space; in [1], the corresponding divergences are addressed by means of the “multiple scattering technique”, which consists in a re-summing the subsequences containing repeating terms. As a result, a modified perturbative expansion is obtained as a sum over non-repeating terms, with self energy corrections arising from the re-summed subsequences. The latter are subsequently neglected in the so called forward or “upper limit” approximation, which leads to an overestimate of the critical value of the disorder for the onset of delocalization. The arguments given in [83] for the convergence of (25) are exactly at this level of approximation; however, to recover an analogous expansion in terms of non-repeating terms, it turns out to be technically convenient to drop the terms  $\Delta K_\alpha^n$  in (30). The resulting operator  $I_\alpha - n_\alpha$  belongs then to the orthogonal complement of  $K$ , and so it admits the expansion (25). This choice guarantees that the coefficients are uniquely fixed *just* by imposing  $[H, I_\alpha] = 0$ , without any additional requirement. The drawback is that the ansatz (25) does not define an operator with binary spectrum; however, it is expected that the statement on the finiteness of the radius of

convergence is not spoiled once the neglected terms are reinserted and the “normalization” of (25) is imposed.

#### 4.2.2 Mapping to a single-particle problem and forward approximation

The condition  $[I_\alpha, H] = 0$  translates into a set of linear constraints for the amplitudes  $\mathcal{A}_{\mathcal{I}, \mathcal{J}}^{(\alpha)}$ , cfr. Eq.(40) in [83]. They define an effective non-hermitian hopping problem on a lattice with sites labeled by  $(\mathcal{I}, \mathcal{J})$ , on-site disorder  $\mathcal{E}_{\mathcal{I}, \mathcal{J}} = \sum_{n=1}^N (E_{\beta_n} - E_{\gamma_n})$  and hopping given by the interactions: precisely, two sites  $(\mathcal{I}, \mathcal{J})$  and  $(\mathcal{I}', \mathcal{J}')$  are connected if their index sets differ by indices belonging to a non-zero matrix element  $U_{\alpha\beta, \gamma\delta}$ . A pictorial representation of the resulting lattice is given in Fig. 4. The main feature is its hierarchical structure: sites are organized into generations  $N$  according to the length of the index sets, and the hopping is only within the same generation or from one generation to the next; in the latter case, it is unidirectional (hence the non-hermiticity of the hopping-problem). The typical connectivity is estimated accounting for the space (particles need to be in the same or in an adjacent localization volume) and phase-space restrictions (24) imposed to the interaction matrix elements. For hoppings  $U_{\alpha\beta, \gamma\delta}$  between consecutive generations, it scales as

$$\mathcal{K} = 4 \frac{\mathcal{W}}{\delta_\xi}. \quad (32)$$

Such hoppings indeed require a particle (or hole)  $\alpha$  in  $(\mathcal{I}, \mathcal{J})$  to scatter to the state  $\gamma$  within the localization volume that is closer in energy (above or below  $\alpha$ ), while another particle-hole pair of levels  $(\beta, \delta)$  with adjacent energies is created. The particle  $\beta$  can be chosen in  $\mathcal{W}/\delta_\xi$  ways, and there are two choices for  $\gamma$  and  $\delta$ , respectively, hence (32). In contrast, hoppings to sites of the same generation correspond to processes where each member of a pair of particles (or holes) scatter to one of the two closest energy levels: there are order  $O(1)$  possible final states to which a fixed given pair can decay.

The forward approximation consists in neglecting the links connecting sites in the same generation of the lattice, which is justified provided (32) is large. This significantly simplifies the lattice topology, as some sites are eliminated (the corresponding amplitudes in (25) approximated to zero), see Fig. 4. The equations for the amplitudes on the retained sites become recursive equations for increasing generations, with initial condition  $\mathcal{A}_{\alpha, \alpha}^{(\alpha)} = 1$ ; this allows to derive a closed expression for the

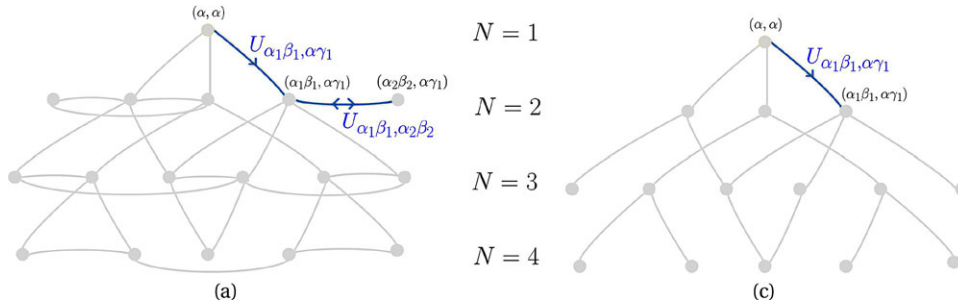
amplitude at the sites  $(\mathcal{I}, \mathcal{J})$  as a sum over all directed, non-repeating paths [88] from the root  $(\alpha, \alpha)$  to the given site:

$$\begin{aligned} \mathcal{A}_{\mathcal{I}, \mathcal{J}}^{(\alpha)} &= \sum_{\substack{\text{directed paths} \\ (\alpha, \alpha) \rightarrow (\mathcal{I}, \mathcal{J})}} (-1)^{\sigma_{\text{path}}} \prod_{i=1}^{N-1} \frac{\lambda \eta_{\alpha_i \beta_i, \gamma_i \delta_i} \delta_\xi}{\sum_{k=1}^i \mathcal{E}_{\alpha_k \beta_k, \gamma_k \delta_k}} \\ &\equiv \sum_{\substack{\text{directed paths} \\ (\alpha, \alpha) \rightarrow (\mathcal{I}, \mathcal{J})}} \omega_{\text{path}}^{(N)}, \end{aligned} \quad (33)$$

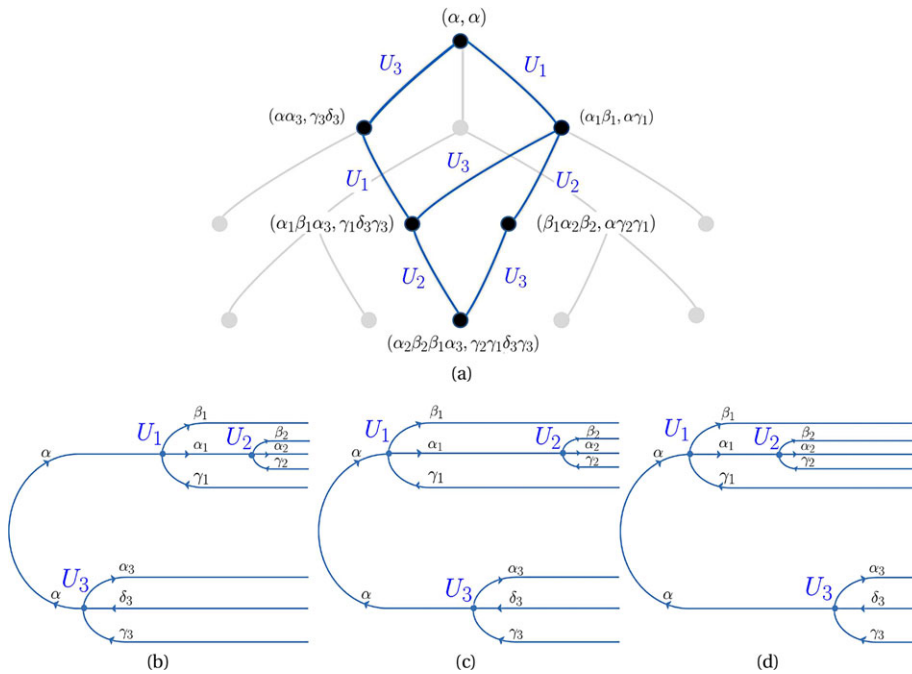
where the factor  $(-1)^{\sigma_{\text{path}}}$  is a global sign. The expression (33) is of order  $\lambda^{N-1}$ , which is the lowest possible order for amplitudes of operators involving  $2N$  particle-hole indices. Thus, for any  $N$ , any term in (25) that is of order  $O(\lambda^N)$  or smaller is set to zero. The remaining terms satisfy  $r(\mathcal{I}, \mathcal{J}) \leq N\xi$  due to the locality of the interaction; it follows that within this approximation  $r(\mathcal{I}, \mathcal{J})$  in (27) can be replaced by  $N$ , so that the convergence is controlled by the generation  $N$ . By eliminating intra-generation hoppings, the FA addresses the problem of arbitrary repetitions of small denominators; further insights on its meaning is derived from the diagrammatic representation of paths illustrated in Fig. 5. The expansion in directed paths corresponds in fact to selecting only the processes where at each vertex an additional particle-hole pair is created; this is similar to the imaginary self consistent Born approximation exploited in [6].

#### 4.2.3 Issue of factorials and re-summation of correlated graphs

Within the FA, the number of sites  $(\mathcal{I}, \mathcal{J})$  at distance  $N$  from the root scales exponentially, as  $\sim \mathcal{K}^N$ . Moreover, the *typical* path weights in (33) decay exponentially with  $N$  for small enough  $\lambda$ ; thus, one might expect that a transition occurs, arising from the competition between these two exponentials. Although this will turn out to be precisely the case, an intermediate step has to be performed. The sum (33) is indeed over a number of terms which grows faster-than-exponential with  $N$ , since to any given path of length  $N$ , one can associate a class of  $O(\sim N!)$  other paths corresponding to graphs sharing the same interaction vertexes, but occurring in different order, see Fig. 5. Due to the structure of the energy denominators, each such graph has a different weight contributing to (33). A crucial step in arguing the exponential decay of (33) consists in showing that this factorial is compensated, as cancellations occur between the different terms that are due to their statistical correlations. For



**Figure 4** Structure of the operator lattice before (a) and after (c) making the forward approximation. Vertices correspond to Fock indices  $(\mathcal{I}, \mathcal{J})$ ; links are drawn between index pairs, which are connected by the interaction  $U$ .



**Figure 5** Directed paths in the operator lattice and associated scattering graphs: the sites along the path are intermediate states of the graph, the hopping terms correspond to vertices  $U_{\alpha_1 \alpha_2, \beta_1 \beta_2}$  in the graph (denoted with the simplified notation  $U_1, U_2, U_3$ ), and the energy  $\mathcal{E}_{\mathcal{I}, \mathcal{J}}$  is the sum of the energy differences  $\mathcal{E}_{\alpha_1 \alpha_2, \beta_1 \beta_2} = E_{\alpha_1} + E_{\alpha_2} - E_{\beta_1} - E_{\beta_2}$  associated with all preceding scatterings. The three highlighted paths form loops in the lattice, associated to graphs differing only in the order in which the interactions  $U_1, U_2, U_3$  act.

instance, in the case of Fig. 5 it holds

$$\left[ \frac{1}{\mathcal{E}_3(\mathcal{E}_3 + \mathcal{E}_1)} + \frac{1}{\mathcal{E}_1(\mathcal{E}_1 + \mathcal{E}_3)} + \frac{1}{\mathcal{E}_1(\mathcal{E}_1 + \mathcal{E}_2)} \right] \frac{1}{\mathcal{E}_1 + \mathcal{E}_2 + \mathcal{E}_3} = \frac{1}{\mathcal{E}_3 \mathcal{E}_1 (\mathcal{E}_1 + \mathcal{E}_2)}, \quad (34)$$

where  $\mathcal{E}_i$  is the linear combination of single particle energies associated to the interaction vertex  $U_i$ . This problem of factorials can be dealt systematically, giving an

integral representation of the sum over path-weights differing by permutations of the interactions, and optimizing on the sequence of pole integrations that can be performed on the representation. As a result, the sum over the factorially-many paths is explicitly rewritten as a sum over only exponentially-many terms  $\tilde{\omega}_{\text{path}}^{(N)} = O(\lambda^N)$  having a similar structure as (33), termed “effective paths”:

$$\mathcal{A}_{\mathcal{I}, \mathcal{J}}^{(\alpha)} = \sum_{\text{effective paths } (\alpha, \alpha) \rightarrow (\mathcal{I}, \mathcal{J})} \tilde{\omega}_{\text{path}}^{(N)}. \quad (35)$$

#### 4.2.4 The estimate of the convergence radius

For large  $N$ , the generating function of the random variables  $\tilde{\omega}_{\text{path}}^{(N)}$  (analogous to the quantity in (47)) is computed by means of a transfer matrix technique, accounting for the correlations of energies in the denominators. The tail of the distribution of the (effective)-path weights is obtained by inverse Fourier transform, within a saddle point approximation. It is shown that the (effective)-path weights are heavy-tailed distributed, so that the sum (35) is dominated by the maximal summand. The typical value of (35) is then estimated with a calculation that is close to the one performed in [47] for the Hamiltonian (1) on a Bethe lattice: neglecting the residual correlations between the effective paths, one finds that for  $z < 1$

$$\begin{aligned} \log \mathbb{P} \left( \sum_{\substack{\mathcal{I} \neq \mathcal{J} \\ |\mathcal{I}|=N=|\mathcal{J}|}} |\mathcal{A}_{\mathcal{I},\mathcal{J}}^{(\omega)}| < z^N \right) &\approx -\mathcal{N}_N \mathbb{P} \left( |\tilde{\omega}_{\text{path}}^{(N)}| > z^N \right) \\ &\approx -\exp \left( N \log \mathcal{G}(\lambda, z, \mathcal{K}) + o(N) \right), \end{aligned} \quad (36)$$

where the function  $\mathcal{G}(\lambda, z, \mathcal{K})$  combines the exponential growth (with  $N$ ) of the number  $\mathcal{N}_N$  of effective paths of length  $N$  with the exponential decay of the large-deviation probability  $\mathbb{P}(|\tilde{\omega}_{\text{path}}^{(N)}| > z^N)$ . The number  $\mathcal{N}_N$  is determined by means of a combinatorial estimate of the number of diagrams representing the scattering processes with fixed final state, which accounts for the locality of the interactions by enforcing that only particle-hole pairs in nearby localization volumina can be involved in the same interaction vertex. Imposing  $\mathcal{G}(\lambda, z, \mathcal{K}) < 1$  for  $z \rightarrow 1$ , the following condition for the convergence radius is obtained:

$$\lambda_c = \frac{C}{\nu_F(1 - \nu_F)} \frac{1}{\mathcal{K} \log \mathcal{K}}, \quad (37)$$

with  $C$  a numerical constant estimated to be  $18 < C < 37$ ,  $\nu_F$  the filling fraction and  $\mathcal{K}$  given in (32). The many-body nature of the processes involved is evident in the fact that  $\lambda_c \rightarrow \infty$  when  $\nu_F \rightarrow 0$ , while the divergence at  $\nu_F \rightarrow 1$  is a Fermi blockade effect. Thus, we conclude that for  $\lambda < \lambda_c$  and within the set of approximations made, quasilocal conserved quantities for the interacting Hamiltonian (2) can be constructed as weak perturbations of the non-interacting occupation numbers  $n_\alpha$ .

The estimate (37) is in agreement with the perturbative results in [5, 6], where a finite temperature transition is predicted. Indeed, so far the convergence of the con-

struction was analyzed at the operator level, with no assumption made on the occupation of the single-particle energy levels. When acting on some typical many-body states at temperature  $T$ , some terms in (25) get annihilated when attempting to create particles on occupied states or holes on already empty states. This translates into a reduction of the phase space associated to the decay processes, which in turns implies that  $\mathcal{K} \rightarrow \mathcal{K}_{\text{eff}} \sim T/\delta_\xi$ : substitution into (37) gives the condition for the critical temperature in [5, 6]. The agreement is not surprising, as the class of diagrams that are statistically analyzed is the same as within the forward approximation. It is however argued in [90] (see also the discussion in [83]) that this scenario of a finite-temperature transition becomes unstable beyond this set of approximations, due to rare fluctuations within typical, putatively localized states, that are argued to restore ergodicity in the long time limit. We comment further on this point in Sec.6.

#### 4.3 Nonperturbative construction of local integrals of motion

We now come to the discussion of the nonperturbative construction of local integrals of motion given in [32]. The work contains a rigorously proof of many-body localization for the spin chain (4) (which we denote simply with  $H$  in the following) in its strongly disordered limit, following from a physically reasonable assumption on level statistics. The proof proceeds by constructing a sequence of quasilocal unitary rotations that diagonalize the Hamiltonian; the conserved quantities are obtained as a by-product of the diagonalization procedure. We begin by reviewing the construction in [32, 91], and subsequently discuss its implications for local integrals of motion.

As the spin flip interactions in (4) are of order  $\gamma \ll 1$ , the Hamiltonian is effectively in the strong-disorder regime. In the present discussion the focus is on the disorder in the random magnetic field  $h_i$ : the randomness is added to the other two terms in the Hamiltonian as it helps with level separation arguments, but should not be essential to the physics of this model. For small  $\gamma$ , the Hamiltonian is close to one that is diagonal in the basis given by the tensor products of  $S_i^z$  eigenstates. The MBL transition in this model can be thought of as a many-body version of what happens for a single spin: when  $\Lambda$  is a single site, the Hamiltonian reduces to

$$\begin{pmatrix} h & \gamma \\ \gamma & -h \end{pmatrix}, \quad (38)$$

having eigenvectors close to (1, 0) and (0, 1) for small  $\gamma$ , whereas for large  $\gamma$ , they are close to the fully hybridized vectors (1, 1) and (1, -1). The result in [32] extends this picture to the many-body Hamiltonian, demonstrating (under the assumption on level statistics) the existence of an MBL phase where the eigenstates of  $H$  resemble the basis vectors, *i.e.* they are largely unhybridized. Precisely, the following statement is proven:

**Theorem 1.** *Assume that for some  $\nu, C$ , the eigenvalues of  $H$  in boxes of size  $n$  satisfy*

$$P\left(\min_{\alpha \neq \beta} |E_\alpha - E_\beta| < \delta\right) \leq \delta^\nu C^n, \quad (39)$$

for all  $\delta > 0$  and all  $n$ . Then there exists a  $\kappa > 0$  such that for  $\gamma$  sufficiently small,

$$\mathbb{E} \text{Av}_\alpha |\langle S_0^z \rangle_\alpha| = 1 - O(\gamma^\kappa), \quad (40)$$

for all  $\Lambda$ . Here  $\mathbb{E}$  denotes the disorder average,  $\langle \cdot \rangle_\alpha$  denotes the expectation in an eigenstate  $\alpha$  of  $H$ , and  $\text{Av}_\alpha$  denotes an average over  $\alpha$ .

The normalized average  $\text{Av}_\alpha$  over the eigenstates of  $H$  can be taken with thermal weights  $(\text{const}) \exp(-\beta E_\alpha)$ , which at infinite temperature becomes a uniform weighting. Thus 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. This would contrast with a thermalized phase, with states approximating thermal ensembles, as expected from the Eigenstate Thermalization Hypothesis recalled in Sec. 2. Indeed, at infinite temperature thermalization would imply that eigenstate expectations of  $S_0^z$  would go to zero in the infinite volume limit; the bound (40) implies a failure of thermalization, a key feature of the MBL phase.

The level-statistics assumption (39) specifies that with high probability the minimum level spacing should be no smaller than an exponential in the volume. Physically, one expects to see (39) satisfied with  $\nu = 1$  in a localized phase (Poisson statistics) or with  $\nu > 1$  in a thermalized phase (repulsive statistics). Although these bounds are expected to hold, the tools for proving them are not yet available in the many-body context (see [92] for a potentially useful approach in the one-body context). We discuss in the following the main steps in the diagonalization procedure.

#### 4.3.1 A multiscale procedure to diagonalize an MBL Hamiltonian

The proof proceeds by diagonalizing  $H$  through successive elimination of low-order off-diagonal terms as in

Newton's method. Elements of the tensor product basis can be labeled by classical spin configurations  $\sigma = \{\sigma_i\}_{i \in \Lambda}$ , with  $\sigma_i = \pm 1$  indicating the eigenvalue of  $S_i^z$ . Initially, the only off-diagonal term is  $\gamma_i S_i^x$ , which is local. This operator flips  $\sigma_i \rightarrow -\sigma_i$ . Let  $\sigma^{(i)}$  denote the result of flipping the spin at  $i$  in the spin configuration  $\sigma$ . Then the spin flip produces a change in energy

$$\Delta E_i \equiv E(\sigma) - E(\sigma^{(i)}) = 2\sigma_i(h_i + J_i\sigma_{i+1} + J_{i-1}\sigma_{i-1}). \quad (41)$$

The site  $i$  is said to be resonant if  $|\Delta E_i| < \varepsilon \equiv \gamma^{1/20}$  for at least one choice of  $\sigma_{i-1}, \sigma_{i+1}$ . Due to the small energy denominator, resonant sites may require a rotation that is far from the identity, and perturbation theory is not useful. A site is resonant with probability  $\sim 4\varepsilon$ , so resonant sites form a dilute set.

Ignore the resonant sites, for the moment. A rotation can be designed that eliminates the off-diagonal terms  $J(i) \equiv \gamma_i S_i^x$  in the Hamiltonian. Let  $H = H_0 + \mathcal{J}$ , where  $H_0$  contains the diagonal part of  $H$  (first and third terms of (4)) and  $\mathcal{J}$  contains the off-diagonal part (second term of (4)). Then define an antisymmetric matrix

$$A \equiv \sum_i A(i), \text{ where } A(i)_{\sigma\sigma^{(i)}} = \frac{J(i)_{\sigma\sigma^{(i)}}}{E_\sigma - E_{\sigma^{(i)}}}. \quad (42)$$

This is a local operator, which can be used to generate a basis change  $e^{-A}$ . The result is a rotated (or renormalized) Hamiltonian:

$$H^{(1)} = e^A H e^{-A}. \quad (43)$$

This rotation is correct to first order in perturbation theory, because

$$[A, H_0] = -\mathcal{J}; \quad (44)$$

this enables the cancellation of the off-diagonal terms to leading order:

$$\begin{aligned} H^{(1)} &= e^A H e^{-A} = H + [A, H] + \frac{[A, [A, H]]}{2!} + \dots \\ &= H_0 + \sum_{n=1}^{\infty} \left( \frac{1}{n!} - \frac{1}{(n+1)!} \right) (\text{ad } A)^n \mathcal{J} \\ &\equiv H_0 + J^{(1)}. \end{aligned} \quad (45)$$

Here  $(\text{ad } A)B \equiv [A, B]$ . See [93] for a similar construction. See also [35], where this method is used to diagonalize the Anderson model Hamiltonian with weak hopping. The new interaction  $J^{(1)}$  contains terms that are at least second order in  $\gamma$ . Note that  $A(i)$  commutes with  $A(j)$  or  $J(j)$  if  $|i - j| > 1$ . One can write  $J^{(1)} = \sum_g J^{(1)}(g)$ , where  $g$  is a connected graph specifying a nonvanishing term

of  $(\text{ad } A)^n J$ . A graph involving  $\ell$  spin flips has  $\ell - 1$  energy denominators and is bounded by  $\gamma(\gamma/\varepsilon)^{\ell-1}$ . Thus the size of terms in  $J^{(1)}$  decay exponentially with their range: the interaction is quasilocal in the sense of (13).

If there are resonant sites present, then one may restrict the sum over  $i$  in (42) to nonresonant sites, thereby avoiding any small denominator  $E_\sigma - E_{\sigma(i)}$ . Then the rotation (43) eliminates terms of order  $\gamma$  in the nonresonant region. To handle the resonant region, define resonant blocks by taking connected components of the set of resonant sites. As in quasidegenerate perturbation theory, one performs exact rotations in resonant blocks to diagonalize the Hamiltonian there. These rotations can be far from the identity, as resonances can lead to significant hybridization of spins. Thus, the notion of quasilocality has to be broadened to include the possibility of a dilute set of sites where a nontrivial basis change is required.

Introducing a sequence of length scales  $L_k = (15/8)^k$ , the procedure continues by defining the  $k^{\text{th}}$  rotated Hamiltonian  $H^{(k)}$ , which has off-diagonal terms eliminated up through order  $\gamma^{L_k}$ . (As in Newton's method, the iteration converges at a rate that is close to quadratic.) In each step, the diagonal elements of  $H^{(k)}$  are renormalized by interactions up to the  $k^{\text{th}}$  scale; they are denoted  $E_\sigma^{(k)}$ . In step  $k \geq 2$ , one needs to work with a graph-based notion of resonance. By definition, a graph  $g$  is resonant if

$$A_{\sigma\bar{\sigma}}^{(k+1)}(g) \equiv \frac{J_{\sigma\bar{\sigma}}^{(k)}(g)}{E_\sigma^{(k)} - E_{\bar{\sigma}}^{(k)}} \quad (46)$$

is greater than  $(\gamma/\varepsilon)^{|g|}$  in magnitude. Here  $|g|$  denotes the number of spin flips in  $g$ . For nonresonant graphs, the generator (46) is used to generate the next rotation. Then the ad expansion (45) produces the next interaction  $J^{(k+1)}$ , which is again given by a sum of graphs  $\sum_g J^{(k+1)}(g)$ . This is a recursive construction. As in the first step, we may use the condition of nonvanishing commutators to enforce connectivity. Unwrapping the expansions, one finds that each  $g$  corresponds to a sequence of spin flips at a set of sites that is nearest-neighbor connected. At each stage of the procedure, quasilocality of the interaction is preserved.

The set of resonant blocks in the  $k^{\text{th}}$  step is obtained by taking the connected components of the set of sites that belong to resonant graphs  $g$ . In order to preserve the diluteness of the resonant region (an essential part of the notion of quasilocality), one needs to maintain uniform exponential decay on the probability that  $g$  is resonant.

To this end, one proves that for  $s = \frac{2}{7}$ , there is a bound

$$\mathbb{E} |A_{\sigma\bar{\sigma}}^{(k)}(g)|^s \leq \gamma^{s|g|} \mathbb{E} \prod_{\tau \bar{\tau} \in g} |E_\tau^{(j)} - E_{\bar{\tau}}^{(j)}|^{-s} \leq (c\gamma)^{s|g|} \quad (47)$$

on the fractional moment of the rotation generator. Then Markov's inequality implies that

$$P\left(|A_{\sigma\bar{\sigma}}^{(k)}(g)| > (\gamma/\varepsilon)^{|g|}\right) \leq (c\varepsilon)^{s|g|}. \quad (48)$$

The number of graphs containing a given site is exponential in  $|g|$ , so (48) controls the sum over collections of graphs connecting one site to another. Then it is clear that resonant blocks are dilute and do not percolate. As in the first step, rotations are performed in resonant blocks to diagonalize the Hamiltonian there.

#### 4.3.2 Going beyond the forward approximation

However, there are complications in obtaining (47) if the graph  $g$  involves a significant number of repeated spin flips. With repeated spin flips, energy denominators can appear to a high power, or there can be a large number of relations between them. If this is the case, then the fractional moment will no longer be finite, because of the lack of integrability of  $|h|^{-sp}$  for  $p \geq 1/s$ . In effect, (47) is valid only in the forward approximation, wherein graphs looping back to previously visited sites are not allowed. In order to handle the troublesome graphs, one relies on previously obtained bounds  $|A_{\sigma\bar{\sigma}}^{(j)}(\bar{g})| \leq (\gamma/\varepsilon)^{|\bar{g}|}$  for  $j < k$ . By induction, these can be used in the nonresonant region, and they lead to exponentially decaying estimates on  $A_{\sigma\bar{\sigma}}^{(k)}(g)$ . Such estimates would tend to degenerate with  $k$ , were it not for the fact that graphs with many repeated spin flips do not span as great a distance as "self-avoiding" ones. One can, however, obtain uniform decay in the span of the graph, and this is sufficient for quasilocality and convergence in subsequent steps.

Similar arguments allow control over the factorial number of terms in the ad expansion (45). Even in the first step, there are  $\sim n!$  choices for indices  $\{i_1, \dots, i_n\}$  such that  $\prod_{p=1}^n (\text{ad } A(i_p)) J(i_0) \neq 0$ . This is compensated in perturbation theory by the inverse factorials in (45). However, in the Markov inequality, one needs some improvement, because the available  $1/n!$  has to control two sums: (1) the sum over graphs in the ad expansion and (2) the sum over events  $\{A_{\sigma\bar{\sigma}}^{(k)}(g) \text{ is resonant}\}$ . Improvement is possible for graphs with few repeated spin flips, because in that case the number of graphs grows only as an exponential in  $n$ . As explained above, inductive bounds



Figure 6 A block-block resonance.

are sufficient when working with graphs with many repeated spin flips.

#### 4.3.3 The role of dimensionality

In later steps, graphs may connect resonant blocks with nearby sites or with other blocks. For graphs connecting different blocks, the fractional-moment bound depends on having some control over the probability that an energy difference in a block is close to that of a given nearby transition. See Fig. 6. One can obtain the necessary bounds using the level-spacing assumption (39). See [32] for details.

For a block  $B$  of diameter  $L$ , connections to sites at a distance  $d$  have size  $\sim \gamma^d$ . This needs to be smaller than  $2^{-L} = 1/(\text{number of states in } B)$  if they are to be treated perturbatively. So  $B$  needs to be expanded to  $\tilde{B}$ , which includes all sites within a distance  $d = O(L)$  of  $B$ . Then exact diagonalization in  $\tilde{B}$  will eliminate the problematic nearby interactions. Expanding the blocks leads to enhanced connectivity, as nearby blocks have to be combined. The extended blocks may be thought of as connected clusters of a percolation problem. One can show that the probability that sites  $i$  and  $j$  lie in the same extended block decays as  $(\gamma^k)^{1+(\log|i-j|)^2}$ . This provides quantitative control over the size of regions where rotations far from the identity may be required, *i.e.* a quantitative statement of quasilocality.

In dimension 2 or more, the proof of MBL breaks down, because for  $L$  large, there is no  $d$  such that  $\gamma^d$  is smaller than  $2^{-\text{volume}(\tilde{B})}$ . It is argued in [94] that this issue destabilizes the MBL phase at very long times in dimension 2 or more. We postpone further comments about this problem to Sec. 6.

#### 4.3.4 Eigenstate labels and local integrals of motion

Taking the limit as  $k$  tends to infinity, all off-diagonal entries of the Hamiltonian are eliminated. Thus the cumulative rotation from the procedure diagonalizes  $H$ . The columns of this rotation matrix are the many-body eigenfunctions.

Let us discuss how diagonalization of the Hamiltonian with quasilocal rotations leads to a complete set of local integrals of motion. As explained above, diagonalizing  $H$  can be accomplished with small, quasilocal rotation generators throughout most of  $\Lambda$ . One may call this region the perturbative region. There is a dilute, nonpercolating set of resonant blocks where rotations far from the identity may be required. The resonant blocks constitute the nonperturbative region. In the perturbative region, the cumulative rotation is close to the identity. Thus each eigenstate arises as a small rotation performed on one of the tensor product basis vectors; hence one can unambiguously label it by the spin configuration  $\sigma$  that defines that basis vector. Adding in the nonperturbative region, it is evident that the eigenstate labels have to be supplemented with additional information in each resonant block. The need for state labels arises when rotations are performed in blocks to diagonalize the Hamiltonian there. The choice of labels is rather arbitrary, as we saw earlier in the  $2 \times 2$  example (38). In [32], the labels are called “metaspins.” For a block of size  $n$ , the metaspin takes  $2^n$  values. The following discussion assumes a choice of 1:1 correspondence between the metaspins of a block and the spin configuration  $\sigma$  in that block.

Having assigned spin labels to the eigenstates of  $H$ , one may proceed to define the local integrals of motion. Let  $\Omega$  be the rotation that diagonalizes  $H$ , so that  $\tilde{H} = \Omega^* H \Omega$  is diagonal. Note that the columns of  $\Omega$  are the eigenvectors. As the eigenvectors are labeled by spin configurations  $\sigma$ , one may use  $\sigma$  as the column index for  $\tilde{H}$ . Similarly, the rows of  $\tilde{H}$  are indexed by spin configurations. Thus  $\tilde{H}$  is a diagonal matrix that can act on vectors in the original tensor product basis. Although there is some arbitrariness in how this action is defined, the ambiguity is limited to a dilute set of resonant blocks. Note that the spin operators  $S_i^z$  are also diagonal in this basis. Therefore,  $[\tilde{H}, S_i^z] = 0$  for each  $i \in \Lambda$ . Transferring the rotations in this equation to the spin operators, one obtains that  $[H, I_i] = 0$ , where

$$I_i \equiv \Omega S_i^z \Omega^*. \quad (49)$$

The rotations that diagonalize  $H$  have been used to rotate the spin operators  $S_i^z$  into local integrals of motion  $I_i$ .

Note that  $I_i$  is defined by acting on  $S_i^z$  with the rotation opposite to the one used to diagonalize  $H$ . The forward rotation  $\Omega^* S_i^z \Omega$  produces the matrix elements of  $S_i^z$  between eigenstates. In particular,  $[\Omega^* S_i^z \Omega]_{\sigma\sigma}$  gives the expectation of  $S_i^z$  in the eigenstate with label  $\sigma$ . Away from resonant blocks, this rotation is close to the identity.



Thus the eigenstate with label  $\sigma$  resembles the basis vector with the same label: we have that  $\langle S_i^z \rangle_\sigma = \sigma_i + O(\gamma^\kappa)$  for some  $\kappa > 0$ . The probability that  $i$  lies in a resonant block is  $O(\gamma^\kappa)$  as well. This explains the conclusion (40) of Theorem 1.

One may get another perspective on this construction of local integrals of motion by considering the analogous problem in the context of the Anderson model (1) on a rectangle  $\Lambda \subset \mathbb{Z}^d$ , with weak hopping  $0 < \gamma \ll 1$ . In [35], a sequence of quasilocal rotations were constructed to diagonalize  $H_\Lambda$ . This work served as a model case on which the many-body construction of [32] was based. Paralleling the constructions in (41)–(48), one may specify that a site  $x$  is resonant if  $|\epsilon_x - \epsilon_y| < \varepsilon = \gamma^{1/20}$  for any  $y$  with  $|x - y| = 1$ . Then, away from the set of resonant sites, one may define an antisymmetric matrix

$$A_{xy} = \frac{J_{xy}}{\epsilon_x - \epsilon_y}, \text{ for } x \neq y, \quad (50)$$

where

$$J_{xy} = \begin{cases} \gamma, & \text{if } |x - y| = 1; \\ 0, & \text{otherwise.} \end{cases} \quad (51)$$

contains the hopping terms of  $H_\Lambda$ . Then  $A$  may be used to generate the first-step rotation. On the dilute set of resonant blocks, large rotations may be needed to diagonalize the Hamiltonian locally. Continuing with the multiscale procedure outlined above in the many-body case, one obtains again a rotation matrix  $\Omega$  generated via quasilocal operators. Away from resonant blocks, the rotation is close to the identity, so as before one can label eigenstates by the original basis vectors (allowing for some arbitrariness in the choice of 1:1 correspondence in resonant blocks). Writing  $\psi_x$  or  $|\psi_x\rangle$  for the eigenfunction with label  $x$ , one finds that  $|\psi_x(y)| \leq \gamma^{|x-y|/2}$  for  $|x - y| > R$ . Here  $R$  depends on the disorder, but  $\text{Prob}(R > L) \leq \varepsilon^L$ .

As above, one finds that  $\tilde{H}_\Lambda = \Omega^* H_\Lambda \Omega$  is diagonal, and  $\Omega_{yx} = \psi_x(y)$ . Let  $|x\rangle$  denote the basis vector which is 1 at  $x$  and 0 elsewhere. Then  $|x\rangle\langle x|$  is the projection onto functions supported at  $x$ . As it is a diagonal operator, it commutes with  $\tilde{H}_\Lambda$ . As in (49), one can define

$$\tilde{I}_x \equiv \Omega |x\rangle\langle x| \Omega^* = |\psi_x\rangle\langle\psi_x|, \quad (52)$$

and then  $[H_\Lambda, \tilde{I}_x] = 0$ . So in this case the local integral of motion  $\tilde{I}_x$  is simply the projection onto the subspace spanned by the eigenstate with label  $x$ .

In an analogous fashion, one may rewrite (49) in the many-body case. Write  $|\sigma\rangle$  for the basis vector associated with the spin configuration  $\sigma$ , and  $|\psi_\sigma\rangle$  for the associated

interacting eigenfunction. Then

$$S_i^z = \sum_\sigma \text{sgn}(\sigma_i) |\sigma\rangle\langle\sigma| \quad (53)$$

and

$$I_i = \sum_\sigma \text{sgn}(\sigma_i) |\psi_\sigma\rangle\langle\psi_\sigma|. \quad (54)$$

## 5 Construction of conservation laws: numerical schemes

Beside being the suitable framework in which to prove rigorously the occurrence of MBL, the schemes to diagonalize quantum Hamiltonians iteratively are amenable to be implemented numerically [95]. They give prescriptions to construct sequences of rotated Hamiltonian  $H^{(n+1)} = U_n^\dagger H^{(n)} U_n$  converging to a diagonal form; each such prescription defines a set of conserved quantities (as in (49)), that are expected to be quasilocal in the MBL phase. Similar iterative schemes are also at the basis of the RG procedures used to characterize the MBL phase: conservation laws emerge naturally from this setting as well. Approximately conserved operators can also be obtained by means of variational procedures minimizing the (Frobenius) norm of their commutator with the Hamiltonian. We shortly review these numerical approaches in the following.

### 5.1 Conserved quantities obtained from diagonalizing flows

Unitary rotations can be repeatedly applied to diagonalize the Hamiltonian by means of the systematic elimination of its strongest off-diagonal term with respect to some chosen basis, in analogy with the Jacobi algorithm for the diagonalization of matrices [96]. The extension of the Jacobi algorithm to fermionic Hamiltonians has been proposed in [97] and discussed, with specific reference to the MBL problem, in [98]. The off-diagonal terms of the Hamiltonian are in this case strings of fermionic operators (in the local basis of single-particle states) that cannot be rewritten as products of only number operators. Each partial rotation eliminating an off-diagonal string generates new strings of equal or longer length, allowing to eliminate the off-diagonal terms order by order in the length of the corresponding operator. The transformation preserves the many-body Hilbert space as no degree of freedom is integrated out after each partial

rotation: as a consequence, terms eliminated at a given step are typically regenerated at subsequent steps, with renormalized coefficients whose distribution flows to zero if the method converges.

Alternative, continuous diagonalizing flows have been discussed in [95, 99, 100]. The Wegner-Wilson flow has been exploited in [99] to diagonalize the Heisenberg chain (3) with  $J = J_z = 1/4$  and random fields  $h_i \in [-w/2, w/2]$ . The distribution  $\mathcal{F}_{r,w}(h)$  of the coefficients  $h_{ij\dots}$  in (11) is analyzed as a function of the maximal distance  $r$  between the sites  $i, j, \dots$ . Evidences of a “flow” of the distribution  $\mathcal{F}_{r,w}(h)$  with  $r$  towards a  $1/f$  law are given for the MBL phase, while the distribution appears to be scale-free (independent on  $r$ ) in the critical regime. An adaptation of the Toda flow to random spin chains, having the advantage of preserving the sparsity of the Hamiltonian, has been proposed in [95].

## 5.2 Conservation laws emerging from the RG schemes

The RG approaches to MBL are a direct extension of the Strong Disorder Renormalization Group (SDRG) scheme developed to capture the low-temperature thermodynamical properties of random magnets [101–103]. In its conventional form, the renormalization is performed on the system’s Hamiltonian by means of an iterative elimination of its largest coupling constant, whose magnitude  $\Omega$  defines the decreasing energy-cutoff. The corresponding spin subsystem is diagonalized, projected onto the lowest local energy state (or manifold), and decimated from the chain by introducing an effective coupling between its neighboring spins (that accounts perturbatively for the transitions induced by quantum fluctuations involving the virtual occupation of the excited state of the subsystem). The outcome of the RG procedure is an effective ground state, built as an approximate tensor product of few-spins lowest-energy states.

This procedure has been adapted in [24, 53] to access the properties of the excited states of a one dimensional interacting transverse field Ising model

$$H = \sum_i (J_i^z S_i^z S_{i+1}^z + h_i S_i^x + J_i^x S_i^x S_{i+1}^x), \quad (55)$$

with uniformly distributed couplings  $J_i^z, h_i$  with comparable size, and perturbatively small interactions  $J_i^x$ . The generalization consists in keeping track of the projection of the decimated subsystem on both its lowest and higher energy manifold. A full set of approximate

eigenstates is thus constructed by progressively resolving energy gaps of the order of the running cutoff  $\Omega^2$ .

Conserved pseudospin emerge naturally from the RG scheme, as the operators whose eigenvalues  $\pm 1$  label the energy subspaces selected at each RG step. For example [104], assume that at a given RG step the strongest coupling in (55) is  $|J_j^z| = \Omega$ , and let  $H_\Omega = J_j^z S_j^z S_{j+1}^z$ ,  $V = H - H_\Omega$ , and  $v$  be the magnitude of the largest coupling in  $V$ , ( $v < \Omega$ ). The local two-sites Hilbert space at  $j, j+1$  is partitioned into two almost degenerate energy manifolds (spanned by  $|\uparrow\uparrow\rangle, |\downarrow\downarrow\rangle$  and  $|\uparrow\downarrow\rangle \pm |\downarrow\uparrow\rangle$ , respectively) separated by a gap of order  $\Omega$ . The decimation is performed by means of a unitary transformation  $H' \equiv e^{iA} H e^{-iA}$  chosen in such a way that  $[H', H_\Omega] = 0$ ; following the standard SDRG, this program is realized perturbatively by expanding the generator  $A$  to lowest order in  $v/\Omega$ . Two different renormalized Hamiltonians are obtained projecting the second order truncation of  $H'$  into the subspaces corresponding to a fixed eigenvalue of  $S_j^z S_{j+1}^z$ ,

$$\begin{aligned} H_{eff}^\pm &= V' \pm \frac{1}{J_j^z} (h_{j+1} J_{j+1}^x S_{j+2}^x + h_j J_{j-1}^x S_{j-1}^x) \\ &\pm \frac{1}{J_j^z} (h_j J_{j+1}^x \tilde{S}_j^x S_{j+2}^x + h_{j+1} J_{j-1}^x S_{j-2}^x \tilde{S}_j^x) \\ &+ \left( J_j^x \pm \frac{h_j h_{j+1}}{J_j^z} \right) \tilde{S}_j^x \pm \frac{J_{j-1}^x J_{j+1}^x}{J_j^z} S_{j-2}^x \tilde{S}_j^x S_{j+2}^x, \end{aligned} \quad (56)$$

where  $V'$  contains all terms in  $V$  not involving the spins at the sites  $j, j+1$  and  $\tilde{S}_j^x$  is an effective spin operator breaking the internal degeneracy in the subspaces. Analogous steps are performed when the largest coupling is a local field. As a consequence of the projection, the effective dynamics within each energy manifold decouples. The operator  $\tilde{I}_j = e^{-iA} (S_j^z S_{j+1}^z) e^{iA}$  is an emergent conserved quantity for  $H$  (approximate since its dynamics is suppressed with the decimation), and the full set of approximate many-body eigenstates is clustered into two subgroups labeled by its eigenvalues, and separated by a gap of order  $\Omega$ . At the end of the procedure, the full spectrum is constructed with the resulting eigenstates given as product states in the basis of the approximate conserved quantities. Since resonances between distant

<sup>2</sup> This approach admits a natural dynamical interpretation [15, 53], as the renormalized Hamiltonian can be interpreted as the operator giving the effective dynamics at the largest time-scales, when the fastest degrees of freedom oscillating at the scale  $t \sim \Omega^{-1}$  have been integrated out. The phenomenological Hamiltonian (11) is recovered as the fixed point of the renormalization.

degrees of freedom are not accounted for, the scheme is expected to provide accurate results in the strong disorder limit.

An alternative real space RG scheme has been proposed in [82], based on a block renormalization running on an increasing sequence of length scales  $L_k = 2^k$ . At each step  $k$ , a conserved quantity  $\tau_i^{(k)}$  is produced in parallel for each of the partitioning block, as a pseudospin operator whose eigenvalues label the degenerate energy manifolds obtained diagonalizing an intra-block Hamiltonian. The interaction between the blocks is renormalized assuming that the pseudospin operators are constant of motion, producing an effective Hamiltonian at a larger length scale.

### 5.3 Local operators that are approximately conserved

If quasilocal conserved operators exist, a good approximation for them can be obtained by cutting off the exponentially decaying tails, truncating the operators to a finite region of size  $L \gtrsim \xi_{op}$ . The resulting operators are *approximately* conserved local operators, or “ALIOMs”, as they commute with the Hamiltonian up to terms that are exponentially small in  $L$ . In [105], a set of ALIOMs supported on a finite interval of length  $L$  is constructed by expanding them in a complete basis of local operators, and by fixing the coefficients by direct minimization of the Frobenius norm of the commutator with the Hamiltonian. The problem is equivalent to an eigenvalue problem for a  $(4^L \times 4^L - 1)$  matrix, whose low-eigenvalue manifold maps to the set of approximately conserved operators and their products. Similar quantities have been constructed heuristically in [106], where a quantum Monte Carlo approach to the highly excited MBL eigenstates is proposed.

An alternative approach has been adopted in [107], by diagonalizing the one-body density matrix defined on individual many-body eigenstates. In the MBL phase, the resulting eigenvalues  $\{n_\alpha\}_{\alpha=1,\dots,L}$  have been found to be close to either 0 or 1, with eigenvectors that are spatially localized. A step-like discontinuity in the occupation spectrum is present in the MBL phase, analogous to the one of Fermi liquids. This suggests that the operators  $\tilde{n}_\alpha = c_\alpha^\dagger c_\alpha$ , have large overlap with the exactly conserved LIOMs.

## 6 Discussion and conclusive remarks

In this section we discuss recent works on the extension of the idea of LIOMs beyond their original inception.

Although some of these works are not universally accepted in the community, we discuss them anyway for their possible important implications.

### 6.1 Many-body mobility edges and LIOMs

Firstly, we should mention that it is not completely clear how to reconcile the interplay of an extensive mobility edge (separating the lower-energy, localized many-body eigenstates from the higher energy, extended ones) with the existence of LIOMs. In the approximate perturbative constructions of [5, 6, 83, 108, 109], the temperature dependence of the critical interaction is obtained by considering, in the perturbative analysis, typical non-interacting states at the given temperature, which are treated within the forward approximation (or its analogues). This treatment leads to a temperature or energy-density dependent critical value of the interaction  $\lambda_c(W, T)$  (where  $W$  is the disorder), which, by parallel with the critical disorder in Anderson localization, defines a mobility edge  $\epsilon(\lambda, W)$ .

Signatures of mobility edges have been found in numerics of infinite-dimensional and 1d systems [14, 16, 108]. Recent works [90] maintain however that, for any  $\lambda > \min_T \lambda_c(T)$ , rare spontaneous local fluctuations in the energy-density within putative localized states (termed “bubbles”) are sufficient to restore the conductivity of the system at any temperature, as they can move resonantly through the system and act as a mobile bath. This argument rules out the scenario of a transition driven by temperature; it is reconciled with the perturbative treatment discussed in Sec.4.2 or analogous ones (which instead entail a finite-temperature transition) by noticing that bubbles are not captured within the forward approximation exploited in those contexts<sup>3</sup>. The authors of [90] provided evidence for the absence of many-body mobility edges following from the “bubbles” scenario, which is however still subject to debate.

If on the other hand a mobility edge is present, it is natural to expect that certain conservation laws are recovered once a projection onto the localized portion of the Hilbert space is performed. One step in this direction has been made in [110] for a Heisenberg chain (although

<sup>3</sup> Within the operator formalism of Sec. 4.2, transport driven by mobile bubbles would correspond to a divergence of the expansion (25) due to a subsequence of operators having support that is compact (with bounded index level  $N$ ), but localized at increasingly far distance from the localization center of  $I_\alpha$ . These processes are not accounted for in the forward approximation.

for values of the parameters corresponding to a fully MBL phase), by projecting the local spin operators onto a subspace  $\mathcal{H}_{1-f}$  of the total Hilbert space  $\mathcal{H} = \mathcal{H}_{1-f} \oplus \mathcal{H}_f$ , spanned by a finite fraction  $(1 - f)$  of eigenstates. Conserved quantities are obtained from the resulting projected operators by performing an infinite-time average of the operators evolved with the (non-local) Hamiltonian governing the dynamics in  $\mathcal{H}_{1-f}$ , in analogy with the recipe discussed in Sec.4.1 of this review. The resulting operators are argued to be local in the sense that their “weight” in compact regions of the chain remains finite in the thermodynamic limit, although there is a global dressing of the operators, whose total weight grows with  $f^2$ .

## 6.2 Higher dimensions

Effects of the dimensionality on (even a fully) MBL region have been discussed in [6]. However, since in that work (as well as in [83]) only typical diagrams are considered, not much difference is seen in the MBL phase in different dimensions, nor in the critical properties (which in [6] are paralleled to those of a directed polymer). On the other hand, some non-typical effects can spoil the phenomenology of LIOMs: this was considered in a particular setup in [61], and in a more general setup in [94]. The starting point of both works is that of identifying (either by construction, or by estimating its probability of appearance) a rare region in which the disorder is anomalously low so that, if the region were typical, it would be ergodic.

In [94], rare “ergodic grains” within regions of localized spins, generated by atypical disorder fluctuations, were considered. Exploiting a random-matrix description of such “bath-like” regions, it was argued that full MBL is destabilized by the ergodic grains in any dimension  $d \geq 2$ . This implies that transport is restored at large times, and that the non-perturbative construction of LIOMs discussed in Sec. 4.3 would break down in dimension higher than one.

<sup>4</sup> More precisely, the locality of the projected operators  $O$  is understood in the following way: given the decomposition  $O = O^A + O^\perp$ , where  $O^A$  is supported on some compact interval of the spin chain of size  $N_A$ , the ratio of the Frobenius norms  $\lambda = \|O^A\|^2 / \|O\|^2$  remains finite (of the order of  $1 - f$ ) as the thermodynamic limit is taken with  $N_A$  kept fixed. However,  $\|O^\perp\|$  does not decay to zero exponentially as  $N_A$  is increased, as it would be required for the operator to be quasilocal.

In the setup of [61], instead, such an ergodic region was the  $d - 1$  dimensional face of a cube of side  $L$  of fully MBL spins (characterized by the existence of  $L^d$  LIOMs). This region of size  $L^{d-1}$  has typical level spacing  $e^{-aL^{d-1}}$ , which in  $d > 2$  is much smaller than the typical matrix element  $e^{-bL}$  of an operator coupling a spin in the bulk with a spin on the boundary (irrespective of  $a, b$ ). This implies that any spin in the bulk can thermalize using the boundary, and therefore the eigenstates should look thermal. However, eigenstates are the infinite-time limit of a finite-size system: it was argued in [61] that a timescale diverging with  $L$  arises, which hinders transport on macroscopic distances. This phenomenology is described, in the words of [61], by means of  $I^*$ -bits  $I_\alpha^*$ , namely approximately conserved, quasilocal operators, whose commutator with the Hamiltonian is not exactly zero but vanishing in the thermodynamic limit:

$$[I_\alpha^*, H] = O(e^{-bL}). \quad (57)$$

This discussion can be framed in that on coupling MBL systems to an external bath. This is only tangential to the topic of this review, and we refer the interested reader to the existing literature [111–113].

## 6.3 Conclusive remarks

The aim of this work was to illustrate how the notion of emergent integrability has clarified the phenomenology of MBL systems, and to review the main recipes proposed in the literature to construct the conserved operators. Besides the topics already mentioned in this section, there are other possibly relevant implications that have not been touched in the review. For instance, it may be interesting to re-consider the implications of MBL on well known protocols in quantum computation (such as the quantum adiabatic algorithm [114, 115]) in view of this characterization in terms of the emergent integrability. Moreover, although several constructive schemes are available, few attempts have been made so far to exploit the LIOMs to explicitly compute physically relevant quantities (see [89] for an example). The existence of many non-trivial conserved quantities in quantum interacting, disordered systems has potential implications that are not only of theoretical interest but also of practical relevance, in view of the potential use of MBL systems in quantum devices; it is thus certainly worth to keep investigating the problem discussed in this review.

**Key words.** many-body localization, local integrals of motion.

## References

- [1] P. W. Anderson, *Phys. Rev.* **109**, 1492 (1958).
- [2] R. Nandkishore and D. A. Huse, *Annual Review of Condensed Matter Physics* **6**(1), 15–38 (2015).
- [3] C. Gogolin and J. Eisert, *Reports on Progress in Physics* **79**(5), 056001 (2016).
- [4] L. Fleishman and P. W. Anderson, *Phys. Rev. B* **21**, 2366–2377 (1980).
- [5] I. Gornyi, A. Mirlin, and D. Polyakov, *Phys. Rev. Lett.* **95**, 206603 (2005).
- [6] D. M. Basko, I. L. Aleiner, and B. L. Altshuler, *Annals of Physics* **321**(May), 1126 (2006).
- [7] I. L. Aleiner, B. L. Altshuler, and G. V. Shlyapnikov, *Nat Phys* **6**(11), 900–904 (2010).
- [8] V. Oganesyan and D. A. Huse, *Phys. Rev. B* **75**, 155111 (2007).
- [9] M. Žnidarič, T. Prosen, and P. Prelovšek, *Phys. Rev. B* **77**, 064426 (2008).
- [10] B. Bauer and C. Nayak, *J. Stat. Mech* **2013**, P09005 (2013).
- [11] A. De Luca and A. Scardicchio, *Europhys. Lett.* **101**, 37003 (2013).
- [12] I. Mondragon-Shem, A. Pal, T. L. Hughes, and C. R. Laumann, *Phys. Rev. B* **92**(Aug), 064203 (2015).
- [13] A. Pal and D. A. Huse, *Phys. Rev. B* **82**(Nov), 174411 (2010).
- [14] J. A. Kjäll, J. H. Bardarson, and F. Pollmann, *Phys. Rev. Lett.* **113**(Sep), 107204 (2014).
- [15] R. Vosk and E. Altman, *Phys. Rev. Lett.* **110**(Feb), 067204 (2013).
- [16] D. J. Luitz, N. Laflorencie, and F. Alet, *Phys. Rev. B* **91**(Feb), 081103 (2015).
- [17] Y. Bar Lev, G. Cohen, and D. R. Reichman, *Phys. Rev. Lett.* **114**(Mar), 100601 (2015).
- [18] J. Goold, C. Gogolin, S. R. Clark, J. Eisert, A. Scardicchio, and A. Silva, *Phys. Rev. B* **92**(Nov), 180202 (2015).
- [19] M. Schreiber, S. S. Hodgman, P. Bordia, H. P. Lüschen, M. H. Fischer, R. Vosk, E. Altman, U. Schneider, and I. Bloch, *Science* **349**(6250), 842–845 (2015).
- [20] 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**(6293), 1547–1552 (2016).
- [21] P. Bordia, H. P. Lüschen, S. S. Hodgman, M. Schreiber, I. Bloch, and U. Schneider, *Phys. Rev. Lett.* **116**(Apr), 140401 (2016).
- [22] J. Smith, A. Lee, P. Richerme, B. Neyenhuis, P. W. Hess, P. Hauke, M. Heyl, D. A. Huse, and C. Monroe, *Nature Physics* **12**, 907–911 (2016).
- [23] 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**(Oct), 147204 (2014).
- [24] D. Pekker, G. Refael, E. Altman, E. Demler, and V. Oganesyan, *Phys. Rev. X* **4**, 011052 (2014).
- [25] D. A. Huse, R. Nandkishore, V. Oganesyan, A. Pal, and S. Sondhi, *Phys. Rev. B* **88**, 014206 (2013).
- [26] A. Chandran, V. Khemani, C. R. Laumann, and S. L. Sondhi, *Phys. Rev. B* **89**(Apr), 144201 (2014).
- [27] M. Jimbo, *Yang-Baxter Equation in Integrable Systems* (World Scientific, 1990).
- [28] M. Serbyn, Z. Papić, and D. A. Abanin, *Physical review letters* **111**, 127201 (2013).
- [29] D. A. Huse, R. Nandkishore, and V. Oganesyan, *Phys. Rev. B* **90**(Nov), 174202 (2014).
- [30] B. Swingle, arXiv:1307.0507 [cond-mat.dis-nn].
- [31] E. Canovi, D. Rossini, R. Fazio, G. E. Santoro, and A. Silva, *Physical Review B* **83**(9), 094431 (2011).
- [32] J. Z. Imbrie, *Journal of Statistical Physics* **163**(5), 998–1048 (2016).
- [33] J. Fröhlich and T. Spencer, *Commun. Math. Phys.* **88**, 151–184 (1983).
- [34] M. Aizenman and S. Molchanov, *Comm. Math. Phys.* **157**(2), 245–278 (1993).
- [35] J. Z. Imbrie, *Communications in Mathematical Physics* **341**(2), 491–521 (2016).
- [36] O. S. Barišić, J. Kokalj, I. Balog, and P. Prelovšek, *Phys. Rev. B* **94**(Jul), 045126 (2016).
- [37] T. C. Berkelbach and D. R. Reichman, *Physical Review B* **81**(22), 224429 (2010).
- [38] Y. B. Lev and D. R. Reichman, *Physical Review B* **89**(22), 220201 (2014).
- [39] P. Mazur, *Physica* **43**(4), 533–545 (1969).
- [40] M. Suzuki, *Physica* **51**(2), 277–291 (1971).
- [41] X. Zotos, F. Naef, and P. Prelovsek, *Phys. Rev. B* **55**(May), 11029–11032 (1997).
- [42] Y. Bar Lev and D. R. Reichman, *Physical Review B* **89**(22), 220201 (2014).
- [43] K. Agarwal, S. Gopalakrishnan, M. Knap, M. Müller, and E. Demler, *Phys. Rev. Lett.* **114**(Apr), 160401 (2015).
- [44] M. Žnidarič, A. Scardicchio, and V. K. Varma, *Phys. Rev. Lett.* **117**, 040601 (2016).
- [45] V. Kerala Varma, A. Leroose, F. Pietracaprina, J. Goold, and A. Scardicchio, ArXiv:1511.09144[cond-mat.dis-nn].
- [46] I. Khait, S. Gazit, N. Y. Yao, and A. Auerbach, *Phys. Rev. B* **93**, 224205 (2016).
- [47] B. L. Altshuler, Y. Gefen, A. Kamenev, and L. S. Levitov, *Phys. Rev. Lett.* **78**, 2803 (1997).
- [48] X. Yu, D. Pekker, and B. K. Clark, <https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.118.017201>
- [49] M. Friesdorf, A. H. Werner, W. Brown, V. B. Scholz, and J. Eisert, *Phys. Rev. Lett.* **114**(May), 170505 (2015).
- [50] V. Khemani, F. Pollmann, and S. L. Sondhi, *Phys. Rev. Lett.* **116**(Jun), 247204 (2016).
- [51] M. Serbyn, A. A. Michailidis, D. A. Abanin, and Z. Papić, *Phys. Rev. Lett.* **117**(Oct), 160601 (2016).
- [52] A. Chandran, J. Carrasquilla, I. H. Kim, D. A. Abanin, and G. Vidal, *Phys. Rev. B* **92**(Jul), 024201 (2015).
- [53] R. Vosk and E. Altman, *Phys. Rev. Lett.* **112**(May), 217204 (2014).
- [54] Z. C. Yang, C. Chamon, A. Hamma, and E. R. Mucciolo, *Phys. Rev. Lett.* **115**(Dec), 267206 (2015).

- [55] S. D. Geraedts, R. Nandkishore, and N. Regnault, *Phys. Rev. B* **93**(May), 174202 (2016).
- [56] R. V. Jensen and R. Shankar, *Phys. Rev. Lett.* **54**(Apr), 1879–1882 (1985).
- [57] J. M. Deutsch, *Phys. Rev. A* **43**(Feb), 2046 (1991).
- [58] M. Srednicki, *Phys. Rev. E* **50**, 888 (1994).
- [59] M. Rigol, V. Dunjko, and M. Olshanii, *Nature* **452**, 854–858 (2008).
- [60] P. Ponte, Z. Papić, F. Huveneers, and D. A. Abanin, *Phys. Rev. Lett.* **114**, 140401 (2015).
- [61] A. Chandran, A. Pal, C. R. Laumann, and A. Scardicchio, *Phys. Rev. B* **94**(Oct), 144203 (2016).
- [62] R. Berkovits and B. I. Shklovskii, *Journal of Physics: Condensed Matter* **11**(3), 779 (1999).
- [63] Y. Avishai, J. Richert, and R. Berkovits, *Phys. Rev. B* **66**(Aug), 052416 (2002).
- [64] M. V. Berry and M. Tabor, *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* **356**(1686), 375–394 (1977).
- [65] O. Bohigas, M. J. Giannoni, and C. Schmit, *Phys. Rev. Lett.* **52**(Jan), 1–4 (1984).
- [66] D. Poilblanc, T. Ziman, J. Bellissard, F. Mila, and G. Montambaux, *EPL (Europhysics Letters)* **22**(7), 537 (1993).
- [67] E. Cuevas, M. Feigel'Man, L. Ioffe, and M. Mezard, *Nature communications* **3**, 1128 (2012).
- [68] R. Modak and S. Mukerjee, *New J. Phys.* **16**(9), 093016 (2014).
- [69] M. Serbyn and J. E. Moore, *Phys. Rev. B* **93**, 041424 (2016).
- [70] G. D. Chiara, S. Montangero, P. Calabrese, and R. Fazio, *Journal of Statistical Mechanics: Theory and Experiment* **2006**(03), P03001 (2006).
- [71] J. H. Bardarson, F. Pollmann, and J. E. Moore, *Phys. Rev. Lett.* **109**(Jul), 017202 (2012).
- [72] A. Nanduri, H. Kim, and D. A. Huse, *Phys. Rev. B* **90**, 064201 (2014).
- [73] M. Serbyn, Z. Papić, and D. A. Abanin, *Phys. Rev. Lett.* **110**(June), 260601 (2013).
- [74] P. Calabrese and J. Cardy, *Journal of Statistical Mechanics: Theory and Experiment* **2005**(04), P04010 (2005).
- [75] H. Kim and D. A. Huse, *Phys. Rev. Lett.* **111**, 127205 (2013).
- [76] D. J. Luitz, N. Laflorencie, and F. Alet, *Phys. Rev. B* **93**(Feb), 060201 (2016).
- [77] R. Vosk, D. A. Huse, and E. Altman, *Phys. Rev. X* **5**(Sep), 031032 (2015).
- [78] A. C. Potter, R. Vasseur, and S. A. Parameswaran, *Phys. Rev. X* **5**(Sep), 031033 (2015).
- [79] M. Serbyn, Z. Papić, and D. A. Abanin, *Phys. Rev. B* **90**, 174302 (2014).
- [80] F. Iemini, A. Russomanno, D. Rossini, A. Scardicchio, and R. Fazio, <https://journals.aps.org/prb/abstract/10.1103/PhysRevB.94.214206>
- [81] R. Vasseur, S. A. Parameswaran, and J. E. Moore, *Phys. Rev. B* **91**(Apr), 140202 (2015).
- [82] C. Monthus, *Journal of Statistical Mechanics: Theory and Experiment* **2016**(3), 033101 (2016).
- [83] V. Ros, M. Müller, and A. Scardicchio, *Nuclear Physics B* **891**, 420–465 (2015), [Corrigendum. *Nuclear Physics B*, 900:446–448, 2015].
- [84] A. Chandran, I. H. Kim, G. Vidal, and D. A. Abanin, *Physical Review B* **91**(8), 085425 (2015).
- [85] I. H. Kim, A. Chandran, and D. A. Abanin, [arXiv:1412.3073](https://arxiv.org/abs/1412.3073) [cond-mat.dis-n].
- [86] E. Medina and M. Kardar, *Phys. Rev. B* **46**, 9984 (1992).
- [87] R. Abou-Chacra, D. Thouless, and P. Anderson, *J. Phys. C* **6**, 1734 (1973).
- [88] F. Pietracaprina, V. Ros, and A. Scardicchio, *Phys. Rev. B* **93**(Feb), 054201 (2016).
- [89] V. Ros and M. Mueller, [arXiv:1608.06225](https://arxiv.org/abs/1608.06225) [cond-mat.dis-nn].
- [90] W. De Roeck, F. Huveneers, M. Müller, and M. Schiulaz, *Physical Review B* **93**(1), 014203 (2016).
- [91] J. Z. Imbrie, *Phys. Rev. Lett.* **117**, 027201 (2016).
- [92] J. Z. Imbrie and R. Mavi, *J. Stat. Phys.* **162**(6), 1451–1484 (2016).
- [93] N. Datta, R. Fernández, and J. Fröhlich, *J. Stat. Phys.* **84**, 455–534 (1996).
- [94] W. De Roeck and F. Huveneers, *Phys. Rev. B* **95**, 155129 (2017).
- [95] C. Monthus, *Journal of Physics A: Mathematical and Theoretical* **49**(30), 305002 (2016).
- [96] C. Jacobi, *Crelle's Journal* **30**, 51–94 (1846).
- [97] S. R. White, *The Journal of Chemical Physics* **117**(16), 7472–7482 (2002).
- [98] L. Rademaker and M. Ortuño, *Phys. Rev. Lett.* **116**, 010404 (2016).
- [99] D. Pekker, B. K. Clark, V. Oganesyan, and G. Refael, [arXiv:1607.07884](https://arxiv.org/abs/1607.07884) [cond-mat.dis-n].
- [100] V. L. Quito, P. Titum, D. Pekker, and G. Refael, <https://journals.aps.org/prb/abstract/10.1103/PhysRevB.94.104202>
- [101] C. Dasgupta and S. k. Ma, *Phys. Rev. B* **22**, 1305–1319 (1980).
- [102] D. S. Fisher, *Phys. Rev. Lett.* **69**(Jul), 534–537 (1992).
- [103] F. Iglói and C. Monthus, *Physics Reports* **412**, 277–431 (2005).
- [104] E. Altman and R. Vosk, *Annual Review of Condensed Matter Physics* **6**(1), 383–409 (2015).
- [105] T. E. O'Brien, D. A. Abanin, G. Vidal, and Z. Papić, <https://journals.aps.org/prb/abstract/10.1103/PhysRevB.94.144208>
- [106] S. Inglis and L. Pollet, *Phys. Rev. Lett.* **117**(Sep), 120402 (2016).
- [107] S. Bera, H. Schomerus, F. Heidrich-Meisner, and J. H. Bardarson, *Phys. Rev. Lett.* **115**(Jul), 046603 (2015).
- [108] C. R. Laumann, A. Pal, and A. Scardicchio, *Phys. Rev. Lett.* **113**(Nov), 200405 (2014).
- [109] C. L. Baldwin, C. R. Laumann, A. Pal, and A. Scardicchio, *Phys. Rev. B* **93**(Jan), 024202 (2016).
- [110] S. D. Geraedts, R. N. Bhatt, and R. Nandkishore, <https://journals.aps.org/prb/abstract/10.1103/PhysRevB.95.064204>
- [111] R. Nandkishore, S. Gopalakrishnan, and D. A. Huse, *Physical Review B* **90**(6), 064203 (2014).

- [112] R. Nandkishore, and S. Gopalakrishnan, *Ann. Phys. (Berlin)* **529**, 1600181 (2017).
- [113] D. A. Huse, R. Nandkishore, F. Pietracaprina, V. Ros, and A. Scardicchio, *Phys. Rev. B* **92**(Jul), 014203 (2015).
- [114] B. Altshuler, H. Krovi, and J. Roland, *Proc. Natl. Acad. Sci. USA* **107**(July), 12446 (2010).
- [115] C. Laumann, R. Moessner, A. Scardicchio, and S. Sondhi, *The European Physical Journal Special Topics* **224**(1), 75–88 (2015).