Local conservation laws and the structure of the many-body localized states

Maksym Serbyn,1 Z. Papić,2 and Dmitry A. Abanin3,4

1Department of Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
2Department of Electrical Engineering, Princeton University, Princeton, NJ 08544, USA
3Perimeter Institute for Theoretical Physics, Waterloo, ON N2L 2Y5, Canada
4Institute for Quantum Computing, Waterloo, ON N2L 3G1, Canada

(Dated: May 27, 2013)

We construct a complete set of local integrals of motion that characterize the many-body localized phase. Our approach relies on the assumption that local perturbations only act locally on the eigenstates in the interacting localized phase, which we support with numerical simulations of a random-field XXZ spin chain. Our study provides a description of the structure of the many-body localized states, and shows that the many-body localized phase is “integrable” in a local sense. We discuss the implications of the local conservation laws for quantum dynamics in the interacting localized phase, arguing that the many-body localization can be used to protect coherence in the system by preventing relaxation between eigenstates with different local integrals of motion.

Introduction. Localization of eigenstates of a single particle moving in a disorder potential is among the most remarkable consequences of quantum mechanics. Although the single-particle localization and localization-delocalization transition are well understood [1, 2], much less is known about the nature of the eigenstates in interacting many-body disordered systems. The interest in the problem of the many-body localization (MBL) was rekindled when recent works [3, 4] suggested that the localized phase is stable with respect to weak interactions. This conjecture was corroborated by subsequent numerical studies [5–16].

The dynamics in the non-interacting localized phase is simple because any initial wave function can be decomposed into a superposition of localized single-particle eigenstates. However, when interactions are introduced, the dynamics becomes notably richer [6, 16–18]. Although particle transport is still expected to be blocked, the time evolution of initial product states in the interacting localized phase generates a universal slow growth of entanglement entropy [17]. The saturated entropy was established to be proportional to the system size [6, 16–18], and such growth of the entanglement was argued to reflect “partial thermalization” of the system. However, the type of ensemble describing the MBL phase is not known.

On the experimental side, studying the dynamics of interacting disordered systems has become feasible due to the advances in the field of ultracold atomic gases [19, 20]. In particular, one can engineer nearly isolated quantum systems of cold atoms, prepare them in a variety of initial states, including product states, and probe their subsequent time evolution. These opportunities call for developing a better understanding of the laws that govern the dynamics in the MBL phase.

Here we analyze the structure of the eigenstates in the MBL phase, showing that it is characterized by a large number of emergent local integrals of motion which correspond to multiple local conservation laws. These integrals of motion form a complete set, in the sense that their values completely determine the eigenstates. Our work supports the hypothesis that the many-body localized phase is “integrable” in the above sense. However, unlike in conventional translationally invariant integrable 1D systems, the conservation laws in this case are local.

Local conservation laws strongly constrain quantum dynamics in the MBL phase, preventing a complete thermalization of any given subsystem. Any initial state can be decomposed in terms of the eigenstates, each of which possesses given values of integrals of motion. During time evolution, the weights of different states cannot change. However, because of the exponentially weak interaction between distant degrees of freedom, the relative phases between the states with different values of local integrals of motion become randomized. Any local observable in the long-time limit is therefore determined only by the set of probabilities of local integrals of motion that affect the degrees of freedom in the region where the observable is measured. We refer to this as the local diagonal ensemble. The dephasing due to interactions between distant subsystems is a distinct feature of the MBL phase compared to the non-interacting one, and underlies the slow growth of entanglement [16, 17].

Integrals of motion. First, we note that for the case of non-interacting particles in a disorder potential, the local integrals of motion are simply given by $I_i = c_i^\dagger c_i$, where $c_i^\dagger$ creates a localized single-particle state. For fermions, the possible eigenvalues for this integral of motion are $I_i \in \{0, 1\}$. In a system with $K$ orbitals, there are $2^K$ eigenstates, which are uniquely labeled by the eigenvalues of $K$ integrals of motion.

In order to explicitly construct local integrals of motion for the case of interacting systems, we assume the following property of the localized phase: local perturbations lead only to local modifications of the eigenstates in the MBL phase. That is, if we act on a MBL eigenstate with some local perturbation, introduced either adiabatically or instantaneously, the degrees of freedom sit-
uated at a distance $L \gg \xi$ (here $\xi$ is the localization length) away from the support of the perturbation operator, are generally affected exponentially weakly. Although this statement appears plausible, we will support it below by numerically studying a specific model of the many-body localization – the XXZ chain with a random magnetic field along the $z$ direction also considered in Refs. [6, 9, 11, 15, 16].

Let us consider a MBL system described by a local Hamiltonian, and let us divide it into subsystems of size $l \gg \xi$. For simplicity, we consider a 1d system, and number the subsystems by $i = 1, ..., N$ from left to right. From our discussion it will become clear that the dimensionality is not essential, and the conclusions apply to localized phases in any number of dimensions.

We will construct $N \times M$ integrals of motion, where $M$ is the number of degrees of freedom in each subsystem (e.g., for $K$ spins, $M = 2^K$; also for simplicity we assume that all subsystems are of equal size). Let us fix an index $i$, and denote parts of the system to the left and to the right of $i$ by $\mathcal{L}_i$ and $\mathcal{R}_i$, respectively. The Hamiltonian can be written in the following form:

$$ H = H_\mathcal{L} + H_i + H_\mathcal{R} + H_{\mathcal{L}i} + H_{\mathcal{R}i}, \quad (1) $$

where $H_\mathcal{L}, H_\mathcal{R}, H_i$ act only on the degrees of freedom in $\mathcal{L}, \mathcal{R}, i$, while $H_{\mathcal{L}i}, H_{\mathcal{R}i}$ couple $\mathcal{L}, i$ and $\mathcal{R}, i$.

For the Hamiltonian in which the subsystems $\mathcal{L}, i, \mathcal{R}$ are disconnected from each other (that is, $H_\mathcal{L}i$, $H_{\mathcal{R}i}$ are set to zero), the eigenstates are given by products of eigenstates in the three subsystems:

$$ |\alpha \beta \gamma \rangle_0 = |\alpha \rangle_\mathcal{L} \otimes |\beta \rangle_i \otimes |\gamma \rangle_\mathcal{R}, \quad (2) $$

where $\alpha \in \{1, ..., M^{i-1}\}$, $\beta = \{1, ..., M\}$, $\gamma = \{1, ..., M^{N-i}\}$. Once the three systems are connected, the eigenstates of the full Hamiltonian (1) can be obtained from the product states $|\alpha \beta \gamma \rangle_0$ (2) by nearly local rotations. We will label the resulting eigenstates by their “ancestors”, omitting subscript 0:

$$ |\alpha \beta \gamma \rangle = \hat{O}_{\mathcal{L}i}\hat{O}_{\mathcal{R}i}|\alpha \rangle_\mathcal{L} \otimes |\beta \rangle_i \otimes |\gamma \rangle_\mathcal{R}. \quad (3) $$

Operators $\hat{O}_{\mathcal{L}i}$ and $\hat{O}_{\mathcal{R}i}$ are unitary rotations in the many-body Hilbert space which strongly transform only the degrees of freedom within a distance $\sim \xi$ away from the boundary between $\mathcal{L}$ and $i$, and the boundary between $\mathcal{R}$ and $i$. Their commutator, as well as the action on the degrees of freedom far away, decays exponentially. We note that the assignment (3), which links the eigenstates of the system to the eigenstates of subsystems, is not unique. We will assume that a certain one-to-one correspondence is chosen.

We define the integral of motion for the subsystem $i$:

$$ \hat{I}_i = \sum_{\beta} \sum_{\alpha} \sum_{\gamma} \hat{P}_{\alpha \beta \gamma}, \quad \hat{P}_{\alpha \beta \gamma} = |\alpha \beta \gamma \rangle \langle \alpha \beta \gamma|. \quad (4) $$

Being a linear combination of projectors onto the exact eigenstates, this operator necessarily commutes with the Hamiltonian. The integral of motion has eigenvalues $1, ..., M$. Intuitively, different many-body states with the same eigenvalue of the operator $\hat{I}_i$ look nearly identical within subsystem $i$ at distances larger than $\xi$ away from the boundaries with subsystems $\mathcal{L}, \mathcal{R}$.

By construction, the operator (4) is a trivial integral of motion. Generally, a linear combination of projectors onto the eigenstates is a non-local operator affecting all degrees of freedom in the system. However, the operator in Eq. (4) is special in that it is local, i.e. it weakly affects the degrees of freedom living in $\mathcal{L}$ or $\mathcal{R}$ at a distance $x \gg \xi$ away from the boundaries with the $i$th subsystem. The locality of $\hat{I}_i$ follows directly from the locality of operators $\hat{O}_{\mathcal{L}i}, \hat{O}_{\mathcal{R}i}$, which implies that the sum of projectors becomes very close to the identity operator far away from the boundaries. Below, we will test the locality of the projector operator entering Eq. (4) in a specific model.

Having defined the integral of motion localized near the subsystem $i$, we can similarly define $N - 1$ integrals of motion for the remaining $N - 1$ subsystems, such that in total we have $N$ integrals, $\hat{I}_i, i = 1, ..., N$. To be more specific, $\hat{I}_i$ can be viewed as the $z$-component $\hat{S}_z$ of a “spin” of size $S = (M - 1)/2$. Raising and lowering operators can then be used to construct the entire set of eigenstates, starting from any given eigenstate $|I_1 I_2 ... I_N \rangle$ characterized by the integrals of motion $I_1, I_2, ..., I_N$.

Note that different $\hat{I}_i$ commute with each other by construction, $[\hat{I}_i, \hat{I}_j] = 0$, and their number is $N \times M$, which coincides with the dimensionality of the Hilbert space. Therefore, specifying the eigenvalues of all integrals of motion defined above completely determines the eigenstates of the system, showing that the MBL phase is in a certain sense integrable.

**Hamiltonian and its relation to integrals of motion.** The Hamiltonian takes an especially simple form when written in terms of the integrals of motion:

$$ H = \sum_i \sum_{I_1}^{M} E_{I_1} \hat{P}_{I_1}^i + \sum_{i \neq j} \sum_{I_1, I_j}^{M} E_{I_1 I_j} \hat{P}_{I_1}^i \hat{P}_{I_j}^i + \sum_{i,j,k}^{N} \sum_{I_k}^{M} E_{I_1 I_2 I_k} \hat{P}_{I_1}^i \hat{P}_{I_2}^j \hat{P}_{I_k}^k + ..., \quad (5) $$

where $\hat{P}_{I_i}^i$ is the projector onto the subspace for which the eigenvalue of $i$th integral of motion is equal to $I_i$. In the above equation, $E_{I_i}$ can be roughly viewed as the energy of the $i$th subsystem for the sector $I_i$. $E_{I_1 I_2}$ is the interaction energy between $i$ and $j$ subsystems, etc. There are interactions between any given $n$ subsystems, however, they are exponentially small. Generally, we expect that energies $E_{I_i}$ are proportional to the size of subsystems, e.g., $E_{I_i}$ are proportional to $\xi$ when $i = j \pm 1$, and...
are suppressed as $\xi e^{-i(l(|l-j|^{-1})/\xi}$ otherwise (the interactions between the neighboring subsystems are limited to the boundary and are therefore proportional to $\xi$). The above representation of the Hamiltonian gives us a way to describe the dynamics in the MBL phase for various kinds of initial states [6, 9–11, 16, 17].

**Dynamics.** As a first step, we consider the dynamics of the eigenstate which is perturbed locally. We assume a sudden action of local unitary operator $U$ on the eigenstate $|\Psi_0\rangle = |I_1I_2...I_N\rangle$. Operator $U$ acts only on the degrees of freedom in the subsystem 1, and its support is situated far from the boundary between subsystems 1 and 2. The initial wave function is given by

$$|\Psi(t = 0)\rangle = U|\Psi_0\rangle = \sum_{I_1'} U_{I_1'I_1}|I_1'I_2..I_N\rangle + \ldots, \quad (6)$$

where on the right-hand side we decomposed the wave function $|\Psi(t = 0)\rangle$ in terms of eigenstates. The particular form of the decomposition is dictated by the fact that the values of the integrals of motion $I_2,...,I_N$ can be changed only with an exponentially small probability [terms with other values of $I_2, I_3,...$ in Eq. (6) are represented by ellipses]. Neglecting these terms, the subsequent dynamics becomes trivial:

$$|\Psi(t)\rangle = \sum_{I_1'} U_{I_1'I_1} e^{-iE(I_1'I_2...I_N)t}|I_1'I_2..I_N\rangle, \quad (7)$$

where $E(I_1'I_2...I_N)$ is the energy of the state $|I_1'I_2...I_N\rangle$.

Generally, we expect that a number of different values $I_1'$ which have significant matrix elements $U_{I_1'I_1}$ is finite, typically comparable to the dimensionality of a subsystem of size $\sim \xi$. Therefore, time evolution described by Eq. (7) corresponds to coherent oscillations that involve a finite number of states. Any local observable in region 1 would therefore oscillate at a number of frequencies, showing revivals but no dephasing. This situation changes if the state $|\Psi_0\rangle$ is not an eigenstate, but a superposition of several eigenstates which involve different values of $I_2, I_3,...,I_k$. In this case, exponentially slow dephasing will arise, suppressing the revivals and oscillations of local observables in the long-time limit. The long-time values of observables will be determined by the probabilities $|U_{I_1'I_1}|^2$.

Second, we describe the global evolution of states which differ from the eigenstates everywhere, not just locally. Generally, it is easy to envision experiments in which the system is prepared in an initial product state. For definiteness, consider an initial product state of subsystems $1, 2, ..., N$:

$$|\Psi\rangle = \otimes_{i=1}^N \left( \sum_{\alpha_i=1}^M A_{\alpha_i}|\alpha_i\rangle \right), \quad (8)$$

where $|\alpha_i\rangle$ is an eigenstate of the Hamiltonian $H_i$.

Each product state $\otimes_{i=1}^N |\alpha_i\rangle$ can be related to an eigenstate of the whole system, $|I_1I_2...I_N\rangle$, by a set of local rotations acting near the boundaries between different subsystems. The dynamics corresponding to this effect will be limited to the boundaries between pairs of subsystems. However, for each wave function, degrees of freedom at a distance $x \gg \xi$ away from the boundary will remain undisturbed. Such dynamics therefore does not generate long-range entanglement.

More importantly, since we are dealing with a superposition of different product states $\otimes_{i=1}^N |\alpha_i\rangle$, the degrees of freedom in the subsystem $i$ will be in a superposition of states with different values of the integral of motion $I_i$. Different states entering this superposition are eigenstates, so their relative weights cannot change under time evolution. However, _their phases would become random_, due to interactions with distant subsystems, as is evident from the Hamiltonian (5). Such dephasing, although exponentially slow, will produce long-range entanglement, and will give rise to the entanglement entropy that is extensive in the system size, determined by the participation ratios of different eigenstates. The growth of entanglement entropy for a particular model was studied in detail in our previous paper [16].

**Numerical simulations.** Finally, we present results of the numerical simulations which justify our assumption that local perturbations act locally in the MBL phase. We perform exact diagonalization for the XXZ spin chain of $L$ spins with exchange $J_z = 1$ (corresponding to hopping in the language of fermions), and coupling strength controlled by $J_z = V$. Random on-site magnetic field is uniformly distributed in the interval $\pm W$ (see Ref. [16] for details). First, we study fidelity, defined as the squared overlap of a given initial state with itself.
after time $t$, $F(t) = |\langle \psi(t) | \psi(0) \rangle|^2 = |\langle \psi | e^{-iHt} | \psi \rangle|^2$. Initial state is chosen to be an eigenstate $| \lambda_i \rangle$, perturbed at the right boundary $| \psi \rangle = (1/2 + 2S_L \cdot S_{L-1}) | \lambda_i \rangle$, where $S_{L,L-1}$ are spin operators at the two rightmost sites. Fig. 1(a) shows the fidelity as a function of time averaged over different initial states and realizations of disorder. For strong disorder, the saturated fidelity, Fig. 1(b), weakly depends on the system size, demonstrating the local character of the perturbation introduced at the right boundary.

To provide further support for our construction of the integrals of motion, we numerically implemented the projector operator similar to the one defined in Eq. (4). Exact diagonalization limits the accessible system sizes, thus we divide our system into two subsystems. Every eigenstate of the full system is labeled by its “ancestor” in the left subsystem as in Eqs. (2)-(3). An ancestor is defined as an eigenstate of the Hamiltonian restricted to the left subsystem which has the largest probability in the given eigenstate of the full system. (The probability is calculated from the density matrix of the left subsystem and the exact eigenstates of the left subsystem, see Ref. [16] for details.) Although we do not assign labels for the right subsystem, such labelling is sufficient to implement the operator $P_\alpha = \sum_\beta P_{\alpha \beta}$ as a projector onto the subspace of all eigenstates with the same label $\alpha$ for the left subsystem. As a simple test, we study the locality of the projector $P_\alpha$: by construction it must have trivial action in the right subsystem. To test this property, we perturb the eigenstate with label $\alpha$, $| \lambda^\alpha \rangle$, at the right boundary by the same operator as above to get $| \psi^\alpha \rangle$. Because we are interested in the weight of $| \psi^\alpha \rangle$ in the subspace with the same label $\alpha$, we plot the averaged $\langle \psi^\alpha | P_\alpha | \psi^\alpha \rangle$ as a function of disorder in Fig. 2. For strong disorder, even when the interaction strength is $V = 1$, the perturbed state $| \psi^\alpha \rangle$ has almost all of its weight in the subspace with index $\alpha$, indicating the locality of the constructed $P_\alpha$ in the MBL phase. We note that the specific construction described above allows for more explicit tests to be done, which will be presented in the future work [21].

**Discussion.** We established that the MBL phase is characterized by many local integrals of motion, supporting the hypothesis put forward in Ref. [18]. This implies that the MBL phase does not thermalize, and only partial thermalization of the initial product states, constrained by the local conservation laws, is possible.

It should be noted that there are many ways to define local integrals of motion. For example, in certain problems, it might be helpful to define integrals of motion by a set of $1/2$-pseudo spins – an approach taken in Ref. [22]. Let us assume that $M = 2^K$. Then, $M$ possible values of a given integral of motion $\tilde{I}_\eta$ can be labelled by a state of $K$ pseudospins $\sigma^\eta_i$, $\eta = 1, \ldots, K$. The $z$-projections of these pseudospins form a compete set of integrals of motion, and the Hamiltonian only involves $\sigma^\eta_z$ operators and their products. Operators $\sigma^\eta_i$ can be viewed as effective degrees of freedom, in terms of which dynamics becomes trivial: up-down states of spins are eigenstates, so time evolution can only lead to dephasing between them.

Another implication of our work concerns the structure of the eigenstates in the MBL phase. The eigenstates are short-range entangled, and the area law is obeyed. The eigenstates can generally be represented as product of the eigenstates of subsystems of size $\gg \xi$ which have been locally “corrected” near the boundaries between different subsystems. This suggests a route to numerically finding the eigenstates of large MBL systems: one would break the system into blocks of size much larger that $\xi$, and find the spectrum of each block, first neglecting the terms in the Hamiltonian that couple different blocks. In order to “glue” different blocks together, one would then repeatedly act with the terms in the Hamiltonian that act on degrees of freedom in neighbouring blocks (boundary terms). The boundary terms, repeatedly acting on a given eigenstate of a block, will generate only a finite-dimensional space; diagonalizing the boundary Hamiltonian for each finite-dimensional subspace, it should be possible to find the eigenstates of two coupled blocks, etc. Generally, because of their short-range entanglement, it should be possible to efficiently describe MBL states in 1D using matrix product states.

Finally, our picture suggests a realistic route to extending coherence times in nearly isolated quantum systems, where decoherence is induced by interactions. Examples of such systems, in addition to systems of ultracold atoms, include nuclear spins and NV centers in diamond [23]. Assuming that one could induce strong static disorder leading to MBL, the coherence times of a subsystem can be made very long. To achieve this, one needs to prepare a subsystem of size $\gg \xi$ (e.g., subsystem 1 in
the above example), as well as its immediate neighborhood (e.g., subsystem 2) in some eigenstate. Then, local operations on the subsystem’s degrees of freedom would couple states with different integrals of motion $I_1$, but with fixed values of $I_2$. Therefore, even though the rest of the system is in some complicated superposition state, it will only give rise to an exponentially weak dephasing, with the rate proportional to $e^{-t/\xi}$.

Acknowledgements. We thank J. Moore for a useful discussion. Research at Perimeter Institute is supported by the Government of Canada through Industry Canada and by the Province of Ontario through the Ministry of Economic Development & Innovation. Z.P. was supported by DOE Grant No. DE-SC0002140. The simulations presented in this article were performed on computational resources supported by the High Performance Computing Center (PICSciE) at Princeton University.

Note added. During the completion of this manuscript, we became aware of a related work [22] discussing the existence of local integrals of motion in the MBL phase.

[1] P. Anderson, Phys. Rev. 109, 1492 (1958).
[2] P. Lee and T. V. Ramakrishnan, Rev. Mod. Phys. 57, 287 (1985).
[3] D. Basko, I. Aleiner, and B. Altshuler, Annals of Physics 321, 1126 (2006), ISSN 0003-4916.
[4] I. V. Gornyi, A. D. Mirlin, and D. G. Polyakov, Phys. Rev. Lett. 95, 206603 (2005).
[5] V. Oganesyan and D. A. Huse, Phys. Rev. B 75, 155111 (2007).
[6] M. Žnidarič, T. Prosen, and P. Prelovšek, Phys. Rev. B 77, 064426 (2008).
[7] C. Monthus and T. Garel, Phys. Rev. B 81, 134202 (2010).
[8] T. C. Berkelbach and D. R. Reichman, Phys. Rev. B 81, 224429 (2010).
[9] A. Pal and D. A. Huse, Phys. Rev. B 82, 174411 (2010).
[10] C. Gogolin, M. P. Müller, and J. Eisert, Phys. Rev. Lett. 106, 040401 (2011).
[11] E. Canovi, D. Rossini, R. Fazio, G. E. Santoro, and A. Silva, Phys. Rev. B 83, 094431 (2011).
[12] M. Rigol, V. Dunjko, V. Yurovsky, and M. Olshanii, Phys. Rev. Lett. 98, 050405 (2007).
[13] F. Buccheri, A. De Luca, and A. Scardicchio, Phys. Rev. B 84, 094203 (2011).
[14] E. Cuevas, M. Feigel’man, L. Ioffe, and M. Mezard, Nat Commun 3, 1128 (2012).
[15] A. De Luca and A. Scardicchio, Europhys. Lett. 101, 37003 (2013).
[16] Maksym Serbyn, Z. Papić and Dmitry A. Abanin, arXiv:1305.4915 (2013).
[17] J. H. Bardarson, F. Pollmann, and J. E. Moore, Phys. Rev. Lett. 109, 017202 (2012).
[18] R. Vosk and E. Altman, Phys. Rev. Lett. 110, 067204 (2013).
[19] A. Polkovnikov, K. Sengupta, A. Silva, and M. Vengalattore, Rev. Mod. Phys. 83, 863 (2011).
[20] I. Bloch, J. Dalibard, and W. Zwerger, Rev. Mod. Phys. 80, 885 (2008).
[21] Maksym Serbyn, Z. Papić and Dmitry A. Abanin, (to appear).
[22] D. A. Huse and V. Oganesyan, arXiv:1305.4915 (2013).
[23] L. Childress, M. V. Gurudev Dutt, J. M. Taylor, A. S. Zibrov, F. Jelezko, J. Wrachtrup, P. R. Hemmer, and M. D. Lukin, Science 314, 281 (2006).