Local constants of motion imply information propagation



2015 New J. Phys. 17 113054

(http://iopscience.iop.org/1367-2630/17/11/113054)

New Journal of Physics
The open access journal at the forefront of physics

Deutsche Physikalische Gesellschaft DPG
IOP Institute of Physics
Published in partnership with: Deutsche Physikalische Gesellschaft and the Institute of Physics

CrossMark

OPEN ACCESS

RECEIVED
14 June 2015

REVISED
23 September 2015

ACCEPTED FOR PUBLICATION
1 October 2015

PUBLISHED
27 November 2015



PAPER

# Local constants of motion imply information propagation

M Friesdorf[1], A H Werner[1], M Goihl[1], J Eisert[1] and W Brown[1,2]

[1] Dahlem Center for Complex Quantum Systems, Freie Universität Berlin, D-14195 Berlin, Germany
[2] Computer Science Department, University College London, London WC1E 6BT, London

E-mail: jense@physik.fu-berlin.de

Keywords: quantum information, many-body localisation, information propagation, entanglement, disordered models

## Abstract

Interacting quantum many-body systems are expected to thermalize, in the sense that the evolution of local expectation values approaches a stationary value resembling a thermal ensemble. This intuition is notably contradicted in systems exhibiting many-body localisation (MBL). In stark contrast to the non-interacting case of Anderson localisation, the entanglement of states grows without limit over time, albeit slowly. In this work, we establish a novel link between quantum information theory and notions of condensed matter physics, capturing this phenomenon in the Heisenberg picture. We show that the mere existence of local constants of motion, often taken as the defining property of MBL, together with a generic spectrum of the Hamiltonian, is already sufficient to rigorously prove information propagation: these systems can be used to send a classical bit over arbitrary distances, in that the impact of a local perturbation can be detected arbitrarily far away. This counterintuitive result is compatible with and further corroborates the intuition of a slow entanglement growth following global quenches in MBL systems. We perform a detailed perturbation analysis of quasi-local constants of motion and also show that they indeed can be used to construct efficient spectral tensor networks, as recently suggested. Our results provide a detailed and at the same time model-independent picture of information propagation in MBL systems.

## 1. Introduction

When driven out of equilibrium, interacting quantum many-body systems are usually expected to thermalize [1–3], in the sense that local expectation values can be described by thermal ensembles. For this to be at all possible, local expectation values need to equilibrate to an apparent stationary state and energy has to be transported through the entire system. Such an expected generic behaviour is prominently violated by many-body localized (MBL) systems [4, 5] that show a strong suppression of transport [6–9] and fail to serve as their own heat bath [10, 11]. Thus, these systems do not thermalize and energy remains largely confined within certain regions.

On the level of static properties of the Hamiltonian, equilibration in expectation is guaranteed by non-degenerate energy values and gaps [3, 12–14]; a condition that is expected to hold with unit probability if small random interactions are added to the system. While these equilibration results are rather well understood, the question to what extend the local equilibrium values can be captured by thermal ensembles is still open to debate. A direct way to ensure thermal behaviour is given by the eigenstate thermalisation hypothesis [15–17], one reading of which assumes that most individual eigenstates are already highly entangled and locally indistinguishable from Gibbs states.

For MBL systems, the static properties are markedly different. While the randomness typically occurring in these models will almost surely guarantee non-degenerate energy values and gaps, the individual eigenstates generically have low entanglement [10, 11, 18] and are expected to be efficiently described in terms of tensor networks [18–20]. Moreover, one typically finds that the system has local constants of motion [21–24] that are invariant in time. In fact, it has been shown that such local constants of motion can be used to infer the structure of the eigenstates and obtain an efficient tensor network description of the eigenprojectors [20].

The investigation of information propagation in interacting many-body systems has a long tradition, with upper bounds, giving an effective speed of sound, being provided early on by Lieb and Robinson [25]. For localized systems, the non-interacting case notably leads to a full suppression of propagation, at least in the limit of infinite systems. It came as some surprise that this is no longer the case in the presence of interactions and that entanglement entropies very slowly grow without limit over time [7, 8, 26]. These numerical findings indicate that information is allowed to propagate in these models, at least for the infinite energy states usually considered, in a sense made more precise subsequently.

In this work, we present a rigorous proof for information propagation in MBL systems, using remarkably few and innocent assumptions: only the existence of local constants of motions and a generic spectrum. While local constants of motion slow down the spreading of information in comparison to the ballistic behaviour expected from thermalizing systems, in our proof, we are able to use those constants of motion to show information propagation counter physical intuition. Our approach is entirely model independent and assumes no specific structure of the Hamiltonian. This is achieved by basing the proof on a recently established link between spreading in the Heisenberg picture and equilibration behaviour [3, 27]. Our results are a considerable step forward in the quest to prove that information is allowed to propagate in generic quantum many-body systems, which so far has only been achieved in highly specific systems.

## 2. Many-body localisation

In this work, we will focus on systems exhibiting MBL, which can be seen as a generalisation of Anderson localisation to interacting quantum many-body systems. While a comprehensive definition of this phenomenon is still lacking, it is generally expected that it is closely connected to the existence of approximately local constants of motion [10, 20]. These are operators that commute with the Hamiltonian

$$[H,\,\mathcal{Z}]=0, \tag{1}$$

but are nevertheless to some extent local. In order to access the locality of operators, we consider systems on a cubic lattice $\Lambda$ of fixed dimension, with a spin or fermionic degree of freedom at each site with local dimension $d$. Hence, the system's Hilbert space is given by $\otimes_{x\in\Lambda}\ \mathcal{H}_{\mathrm{loc}}$, where $\mathcal{H}_{\mathrm{loc}}$ is the Hilbert space of the local degree of freedom. We will denote the total number of sites by $L$ and local regions will be denoted by $S$ or $X$. The support of an operator is the region where it acts different from the identity. Of particular importance for our work are operators that are not strictly local, but only approximately local. To this end, it is convenient to introduce a map $\Gamma_S$, which restricts an operator $A$ to a region $S$

$$\Gamma_S(A) := \mathbb{I}_{S^c}\otimes d^{-|S^c|}\mathrm{tr}_{S^c}(A), \tag{2}$$

where $S^c=\Lambda\backslash S$ denotes the complement of $S$ and $\mathrm{tr}_{S^c}(A)$ denotes the partial trace of $A$ over $S^C$. The normalisation $d^{-|S^c|}$ is chosen such that the norm of operators that are already local on $S$ remains unchanged. In order to analyze the locality of an operator, we assume a central support region $X$, enlarged regions $X_l$ that also contain all $l$-nearest neighbours and investigate how the approximation error scales with the size of the enlarged regions $X_l$. We choose the following description

$$\|A-\Gamma_{X_l}(A)\|\leqslant\begin{cases}0 & :\text{ strictly local},\\ g(l) & :\text{ approximately local},\end{cases} \tag{3}$$

where $g$ maps positive integers to positive real numbers is some rapidly decaying function and $\|\cdot\|$ denotes the operator norm, amounting for Hermitian operators to the largest eigenvalue. Naturally, for strictly local observables, the initial region $X$ needs to be taken large enough to include the full support.

In order to investigate the structure of the local constants of motion, we employ two simple models of MBL. In one setting, it is assumed that the Hamiltonian is a sum of these commuting approximately local terms [9]

$$H=\sum_j\tilde{h}_j. \tag{4}$$

In the second setting, the Hamiltonian is a higher order polynomial in terms of the approximately local constants of motion [22, 23]

$$H=\sum_j h_i\mathcal{Z}_i+\sum_{i,j}J_{i,j}\mathcal{Z}_i\mathcal{Z}_j+\cdots, \tag{5}$$

where $J_{i,j}\in\mathbb{R}$ for all $i,\ j$ and the interactions decay both with their order and with the distance of the involved spins. There is a very important difference between these two models. MBL is typically associated to a randomly chosen local potential as in the original work of Anderson [28]. Due to this disorder, it is strongly expected that the spectrum of the corresponding Hamiltonians is generic in the sense that it has non-degenerate energies and gaps.

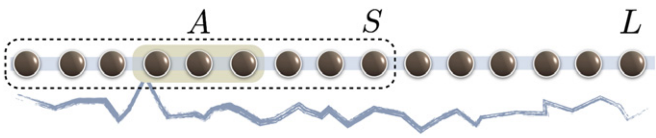


**Figure 1.** The geometry underlying definition 2 on information propagation on average. An observable $A$ evolving in the Heisenberg picture under a many-body Hamiltonian of a many-body system of size $L$, initially locally supported, will necessarily leak out of any region $S$ containing the originally local support of $A$, leading to a measurable signal under measurements in the complement of $S$, for suitably long times.

These assumptions already give us some information about the spectrum of the constants of motion. Since the Hamiltonian in equation (4) is defined as a sum of commuting operators, it can only have a generic spectrum if each local constant of motion already has a generic spectrum. In contrast, the Hamiltonian in equation (5) allows for higher order interactions in the constants of motion. Thus, each constant of motion might only have a small number of distinct eigenvalues similar to a simple local simple Pauli-$Z$-matrix, even though the total Hamiltonian has generic spectrum.

This is the situation expected to occur in MBL systems and we will follow the intuition provided by the Hamiltonian with these higher order interaction between constants of motion, leading us to the following definition [9].

**Definition 1 (Local constant of motion).** Let $\mathcal{Z}$ be an operator that commutes with the Hamiltonian and has $M$ disjoint eigenvalues, all separated by a spectral gap lower bounded by $\gamma > 0$, independent of the system size. $\mathcal{Z}$ is an exactly local constant of motion, if it is strictly local and an approximately local constant of motion, if it is approximately local.

The precise value of $M$ is not crucial for our purposes, since our results apply as long as the number $M$ is independent of the system size. One direct consequence of the simple spectrum of the constants of motion is that the dimension of their eigenspaces has to grow exponentially in the system size. This can be seen from a perturbation theoretical point of view, where the exponentially small tails are not sufficient to create transitions between distinct eigenvalues. Thus the spectrum is approximately given by that of a strictly local operator, which has the feature of exponentially growing eigenspaces since it is extended by tensoring with identity. Moreover, using such constants of motion also allows to prove that the spectral tensor networks construction using exactly local constants of motion [20] can still be carried out in the approximately local case (section A.2 of the appendix).

This leads to the interesting situation that for systems exhibiting MBL, eigenstates will typically efficiently be captured in terms of matrix-product states with low bond dimension. Whereas, product states will build up arbitrary large entanglement over time [7, 8, 26]. We now turn to our main result, namely a rigorous proof of information propagation based only on the existence of a single constant of motion and a non-degenerate spectrum.

## 3. Main result: proof of information propagation

In order to capture how classical information can be send through these models, we imagine that there are two parties, for brevity referred to as Alice and Bob, who have control over different parts of a spin system [29]. For simplicity, let this model be 1D and let Alice control some part at the right end (figure 1). We further assume a fixed separation between the parties and finally that Bob controls the rest of the chain. Alice encodes a classical bit by either doing nothing or acting on her part left the spin chain with a local unitary $V$. At later time, Bob measures some local operator $A$. How well these two parties can communicate with such a protocol is captured by the difference in expectation value for Bob, depending on Alice's action

$$\langle\psi|\ VA_tV^\dagger\ |\psi\rangle - \langle\psi|\ A_t\ |\psi\rangle. \tag{6}$$

Such a procedure amounts to a positive channel capacity in the language of information theory. In less information theoretic terms: a local modification will necessarily lead to a measurable state modification far away in the chain for later times. At time zero, the support of $V$ and $A$ are spatially separated and the above quantity is zero. Over time, the support of $A_t$ might grow and thus eventually lead to a signal. Thus, whether the two parties can communicate crucially depends on the growth of an operator in the Heisenberg picture. In this way, the following quantity is a meaningful way to capture the capability of a Hamiltonian to propagate information.

**Definition 2 (Information propagation on average).** A Hamiltonian allows for information propagation on average, if for any $\epsilon > 0$, there exists a strictly local observable $A$ with unit operator norm, such that, for each finite region $S$, truncation of the Heisenberg evolution to that finite region necessarily leads to a fixed error $\varepsilon$, as long as the system size $L$ is large enough

$$\forall \epsilon > 0 \ \exists A \ \text{local} \ \forall S \ \exists L_0 \ \forall L > L_0: \\ \overline{\|A_t - \Gamma_S(A_t)\|} \geqslant 1 - \epsilon. \tag{7}$$

Here $\overline{A_t} = \lim_{T\to\infty} \frac{1}{T}\int_0^T \mathrm{d}tA_t$ denotes the infinite time average of $A$.

This definition is very restrictive, as it demands that the lower bound can be chosen arbitrarily close to 1. On the other hand, it does not require any information on the corresponding time scale and allows for the support of the observable to take exponentially long to grow. In order to connect this definition to the information propagation protocol introduced above, we use that

$$\Gamma_S(A) = \int_{S^c} \mathrm{d}UUAU^\dagger, \tag{8}$$

which gives

$$\int_{S^c} \mathrm{d}U\, \overline{\|[U, A]U^\dagger\|} \geqslant \overline{\left\|\int_{S^c} \mathrm{d}U\,[U, A]U^\dagger\right\|} = \overline{\|A - \Gamma_S(A)\|} \geqslant 1 - \epsilon. \tag{9}$$

Thus, for systems showing information propagation according to definition 2, we can find a state vector $|\psi\rangle$, such that an excitation created by some local unitary $V$ will be, on average, detectable at distances arbitrary far away

$$\overline{\langle\psi|\ VA_tV^\dagger\ |\psi\rangle - \langle\psi|\ A_t\ |\psi\rangle} = \overline{\langle\psi|\left[V, A_t\right]V^\dagger\ |\psi\rangle} > 1 - \epsilon. \tag{10}$$

Hence, we can interpret $1 - \epsilon$ as a signal strength. Thus, if a Hamiltonian allows for information propagation in the above sense, Alice and Bob can indeed communicate by the described protocol, no matter how large their separation, as long as Bob is allowed to measure on a large enough subsystem. Our main result states that this information propagation can be rigorously deduced from only the existence of a local constant of motion and a suitably non-degenerate spectrum of the Hamiltonian.

**Theorem 1 (Information propagation).** *Let $H$ be a Hamiltonian with non-degenerate energies and gaps and $\mathcal{Z}$ be an approximately local constant of motion with decay function $g$, localisation region $X$, spectral gap $\gamma > 0$ and eigenspaces with dimension larger than $\tilde{d}_{\min}$. Then $H$ necessarily has information propagation on average in the sense that there exists a local operator $A$ initially supported on $X_l \supset X$ with $\|A\| = 1$ such that $A_t$, on average, has support outside any finite region $S$*

$$\overline{\|A_t - \Gamma_S(A_t)\|} \geqslant 1 - 13\frac{g(l)}{\gamma} - \frac{d^{|S|}}{2\tilde{d}_{\min}^{1/2}}. \tag{11}$$

The first non-constant term in equation (11) can be chosen arbitrarily small by picking the initial support $X_l$ large enough and the second term decays exponentially with system size $L$, due to the growth of the degeneracy $\tilde{d}_{\min}$. This result shows that for MBL Hamiltonians, a zero velocity Lieb–Robinson bound does not occur and it is always possible to use the system to send classical information as long as a specific initial state consisting of a superposition of arbitrary energy states is assumed. Hence, it is perfectly compatible with the existence of a dynamical mobility edge, in the sense of a zero-velocity Lieb–Robinson bound for low energy states [19, 30]. Let it be noted that while MBL systems provably allow for information propagation, we expect that they do not exhibit particle or energy transport.

## 4. Proof idea: equilibration implies information propagation

Our results only rely on the existence of an approximately local constant of motion and assume no specific structure of the Hamiltonian. In order to first present the argument in its simplest form, however, we will use the following MBL model

$$H = \sum_j h_j\sigma_j^z + \sum_{i,j} J_{i,j}\sigma_i^z\sigma_j^z, \tag{12}$$

where $\sigma_j^z$ are the local Pauli-$Z$-matrices and $J_{i,j} \in \mathbb{R}$ decays exponentially with the distance between the spins.

To carry out the proof in this simplified setting, we construct two objects. Firstly, a state that is the equal superposition of all eigenstates of the Hamiltonian and secondly, a local operator $A$ that initially has expectation value one with respect to this state, but at the same time has a zero diagonal in the energy eigenbasis and thus zero expectation value for the equilibrated infinite time average. Since equilibration guarantees that local expectation values are described by the infinite time average, this will allow us to conclude that the Heisenberg evolution of the operator $A$ has to be non-local.

**Lemma 1 (Diagonal Hamiltonians).** *Let $H$ be a diagonal Hamiltonian on a spin-1/2 lattice with non-degenerate eigenvalues and gaps. Let $A = \sigma_j^x$ be the Pauli-X-matrix supported on spin j. Then H necessarily has information propagation on average in the sense that the operator $A_t$ has, on average, support outside any finite region S*

$$\overline{\|A_t - \Gamma_S(A_t)\|} \geqslant 1 - d^{|S|-N/2}. \tag{13}$$

**Proof.** From the concrete form of the Hamiltonian in equation (12), we could calculate the time evolution of the $X$-operator and see that it acquires strings of $Z$-operators that sooner or later extend over the whole chain. In the following, we will show how this spreading behaviour can be derived when no specific Hamiltonian structure is used. For the argument, we use a state vector $|\psi\rangle$ that is initially a product with $|+\rangle$ on all sites, which also is the equal superposition of all eigenstates of the system. Since we assume that the Hamiltonian has non-degenerate energies, we know that the infinite time average of $\rho = |\psi\rangle\langle\psi|$ is diagonal, since all off-diagonal elements correspond to non-zero energy gaps and are dephased away. Moreover, as the diagonal is invariant under the time evolution, the time-averaged state $\omega$ is the normalized identity matrix. Considering a subsystem $S$, we can use the non-degenerate energy gaps to employ known equilibration results [12] for the expected deviation from the time average

$$\overline{\|\mathrm{tr}_{S^c}\left(\rho_t - d^{-N}\mathbb{I}\right)\|_1} \leqslant \frac{d_{\mathrm{sys}}}{2d_{\mathrm{eff}}^{1/2}} \leqslant d^{|S|-N/2}. \tag{14}$$

Here $N$ is the total number of spins and the effective dimension counts how many eigenvectors $|k\rangle$ of the Hamiltonian are part of the state

$$d_{\mathrm{eff}} = \left(\sum_k |\langle k|\psi\rangle|^4\right)^{-1}. \tag{15}$$

The above result states that, for most times, the reduced state of $\rho_t$ looks like the identity. Due to the way equilibration is proven, the results also applies to the inverse evolution $\rho_{-t}$ [12].

To investigate how information propagates under the Hamiltonian, we look at the time evolution of an observable $A$ consisting of a single Pauli-$X$-operator somewhere in the region $S$. The key trick is to use the initial expectation value and to insert time evolution operators

$$1 = \mathrm{tr}\left(A_0\rho_0\right) = \mathrm{tr}\left(A_t\rho_{-t}\right). \tag{16}$$

Since we know that the equilibrated state is the normalized identity, the expectation value of any local traceless operator $B$ has to vanish on average

$$\overline{\mathrm{tr}\left(B\rho_{-t}\right)} = 0. \tag{17}$$

Since $A_0$ is traceless and the time-evolution leaves the trace invariant, the operator $A_t$ on average cannot be local anymore. More precisely, we have that

$$\begin{aligned}\overline{\|A_t - \Gamma_S\left(A_t\right)\|} &= \overline{\|A_t - d^{|S|-N}\mathbb{I}_{S^c}\otimes\mathrm{tr}_{S^c}\left(A_t\right)\|}\\ &\geqslant \overline{\left|\mathrm{tr}\left(A_t\rho_{-t}\right) - d^{|S|-N}\,\mathrm{tr}_S\left(\rho_{-t}^S\,\mathrm{tr}_{S^c}\left(A_t\right)\right)\right|},\end{aligned} \tag{18}$$

where we have used that $|\mathrm{tr}(A\rho)| \leqslant \|A\|\|\rho\|_1$ and have defined $\rho^S = \mathrm{tr}_{S^c}(\rho)$. Next we use the inverse triangle inequality, equation (16) and insert $0 = \mathrm{tr}(d^{-|S|}\mathbb{I}_S\,\mathrm{tr}_{S^c}(A))$ which is using the fact that the reduced observable has zero expectation value with the infinite time average

$$\begin{aligned}\overline{\left|\mathrm{tr}\left(A_t\rho_{-t}\right) - \mathrm{tr}_S\left(\rho_{-t}^S d^{|S|-N}\mathrm{tr}_{S^c}\left(A_t\right)\right)\right|} &\geqslant 1 - \overline{\left|\mathrm{tr}_S\left(\rho_{-t}^S d^{|S|-N}\mathrm{tr}_{S^c}\left(A_t\right)\right)\right|}\\ &\geqslant 1 - \overline{\left|\mathrm{tr}_S\left(\left(\rho_{-t}^S - d^{-|S|}\mathbb{I}_S\right)d^{|S|-N}\mathrm{tr}_{S^c}\left(A_t\right)\right)\right|}.\end{aligned} \tag{19}$$

Another application of $|\mathrm{tr}(A\rho)| \leqslant \|A\|\|\rho\|_1$ allows to use the equilibration results discussed previously. Using $\|d^{|S|-N}\mathrm{tr}_{S^c}(A_t)\| \leqslant 1$ concludes the estimate

$$\overline{\|A_t - \Gamma_S(A_t)\|} \geqslant 1 - \|\rho_{-t}^S - d^{-|S|}\mathbb{I}_S\|_1 \|d^{|S|-N}\mathrm{tr}_{S^c}(A_t)\| \geqslant 1 - d^{|S|-N/2}. \tag{20}$$

□

The above proof explicitly establishes a recently proposed connection between information propagation and equilibration [3, 27]. Thus, the assumption of non-degenerate energy gaps is only needed to guarantee equilibration, which means that the condition can be substantially relaxed [13]. The main idea for the proof still can be carried out in the setting where the Hamiltonian is no longer assumed to be diagonal, but where only the existence of an approximately local constant of motion is guaranteed.

## 5. Constants of motion imply information propagation

In order to generalize the proof to rely only on the existence of a single approximately local constant of motion, it is first assumed that the constant of motion is exactly local. This implies that it is possible to distinguish different sets of eigenvectors locally and thus allows to construct local observables that have zero diagonal in the eigenbasis of the Hamiltonian. Moreover, a state with large expectation value with respect to this observable can be constructed, again allowing to use equilibration results, together with the off-diagonality of the observable in order to prove information propagation. Finally, a perturbation analysis extends the argument to approximately local constants of motion. We will now present this argument in detail.

In case the Hamiltonian has exactly local constants of motion [20] $\mathcal{Z}_X$ supported on some region $X$, they can be employed to obtain the operators $A$ in the above construction. For this, let us assume that

$$\mathcal{Z}_X = \sum_{k=1}^{M} \lambda_k P_k \tag{21}$$

with exactly local projectors $P_k$ supported on $X$ and $M$ distinct eigenvalues. The goal is then to construct an operator that is block-off-diagonal with respect to the projectors $P_k$. For this, let $d_{\min}$ be the smallest dimension of the eigenspaces of $\mathcal{Z}_X$, when viewed as a local operator. For the construction, we fix two eigenspaces of $\mathcal{Z}_X$. The larger of the two is then truncated down to the dimension $d_{\mathrm{trunc}}$ of the smaller one. Note that the resulting dimension of both spaces is lower bounded by $d_{\min}$. In these subspaces, we further fix some basis labelled by two indices $|k, r\rangle$ where $k$ labels the eigenspaces of $\mathcal{Z}_X$ and $r$ the basis vectors in each of these subspaces. We will denote the eigenspaces by $k = 0$ and $k = 1$. The operator $A$ is constructed to be supported on the small region $X$ and taken to be the flip operator between the subspaces

$$A = \sum_{r}^{d_{\mathrm{trunc}}} |k = 0, r\rangle\langle k = 1, r| + |k = 1, r\rangle\langle k = 0, r|. \tag{22}$$

The operator norm of this observable is one and we will proceed by constructing an initial state that is an eigenstate of $A$ to eigenvalue 1, but still has large effective dimension. For this, we pick the subspace with smaller dimension and take the equal superposition of all eigenvectors in this subspace denoted by $|v\rangle$. For this, it is crucial to choose the subspace with smaller dimension, as the truncation in general, is not aligned with the eigenstates of the global Hamiltonian. The number of eigenvectors in the untruncated subspace will be lower bounded by $\tilde{d}_{\min} = d_{\min}\, d^{N-|X|}$, which is simply the smallest eigenspace dimension of $\mathcal{Z}_X$ when viewed as an operator on the full lattice. The initial state vector is then taken to be

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|v\rangle + A\,|v\rangle). \tag{23}$$

It is straightforward to check that this is indeed an eigenstate of $A$, since $A^2\,|v\rangle = |v\rangle$. What is more, the state vector $|\psi\rangle$ has an effective dimension lower bounded by $\tilde{d}_{\min}$. This gives the following result.

**Corollary 3.** (*Information propagation: strictly local constants*). *Let $H$ be a Hamiltonian with non-degenerate energies and gaps and $\mathcal{Z}_X$ be a strictly local constant of motion supported on $X$, with eigenspaces with dimension larger than $d_{\min}$. Then $H$ necessarily has information propagation in the sense that for any finite region $S$ containing $X$ there exists a local operator $A$ initially supported on $X$ with $\|A\| = 1$ such that $A_t$, on average, has support outside $S$*

$$\overline{\|A_t - \Gamma_S(A_t)\|} \geqslant 1 - \frac{d^{|S|+|X|/2}}{d_{\min}^{1/2}} d^{-N/2}. \tag{24}$$

**Proof.** To prove this statement, we can directly follow the proof of lemma 1. Using the construction of the initial state and the observable $A$ described above, we immediately obtain

$$\overline{\|A_t - \Gamma_S(A_t)\|} \geqslant 1 - \frac{d_{\text{sys}}}{d_{\text{eff}}^{1/2}}, \tag{25}$$

from the equilibration results in equation (14). Inserting the effective dimension described above $d_{\text{eff}} \geqslant \tilde{d}_{\min} = d_{\min} d^{N-|X|}$ and $d_{\text{sys}} = d^{|S|}$ concludes the proof. □

In many localized systems, one does not expect the constants of motion to be strictly local, but only approximately local [20]. Using perturbation theory, it follows that this bound is sufficient to obtain local approximations for the eigenprojectors and makes it possible to once again construct an observable $A$ that propagates through the system. We will now present this argument in detail and prove theorem 1 .

**Proof.** The first step of the proof is to show that the approximate locality of the constant of motion also implies quasi-local eigenprojectors. Let

$$\mathcal{Z} = \sum_{k=1}^{M} \lambda_k P_k \tag{26}$$

and let $\gamma$ denote the smallest spectral gap. Due to locality, we can express $\mathcal{Z}$ for each fixed $l$, as

$$\mathcal{Z} = \Gamma_{X_l}(\mathcal{Z}) + V_l, \tag{27}$$

with a bounded perturbation $V_l$ satisfying

$$V_l = \mathcal{Z} - \Gamma_{X_l}(\mathcal{Z}), \tag{28}$$

$$\|V_l\| < g(l). \tag{29}$$

Let $P_k^l$ be the eigenprojectors for the truncated observable. Perturbation theory assures us that the perturbed eigenspaces stay approximately orthogonal (theorem VII.3.1 in [31])

$$\|P_k(\mathbb{I} - P_k^l)\| \leqslant \frac{\|V_l\|}{\gamma} = \frac{g(l)}{\gamma}, \tag{30}$$

which also implies

$$\|P_k - P_k^l\| \leqslant \frac{2\|V_l\|}{\gamma} = \frac{2\,g(l)}{\gamma}. \tag{31}$$

Choosing the distance $l$ large enough such that the function $g$ becomes smaller than $\gamma/2$, we know that the perturbed and unperturbed eigenspaces have the same dimension [31]. This local approximation of the eigenprojectors of the constant of motion will be the basis for the construction of the observable $A$ as well as the initial state $\rho$.

In order to construct the observable, we will work with the truncated constant of motion $\Gamma_{X_l}(\mathcal{Z})$, fix two subspaces and construct the same flip operator as in the case of exactly local constants of motion

$$A = \sum_{r}^{d_{\text{trunc}}} |k = 0, r\rangle\langle k = 1, r| + |k = 1, r\rangle\langle k = 0, r|. \tag{32}$$

Without loss of generality, let $k = 0$ be the space with smaller dimension and $k = 1$ the one truncated to $d_{\text{trunc}}$. Let $P_1^l\,|_I$ be the projector on the truncated subspace of $P_1^l$ corresponding to the image of $A$.

For the initial state, we will use the corresponding subspaces, again labelled by $k = 0{,}1$ of the full constant of motion $\mathcal{Z}$. Again we pick the smaller of the two subspaces and define $|v\rangle$ to be the equal superposition of all eigenstates within this space. The initial state vector is then

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|v\rangle + A\,|v\rangle). \tag{33}$$

By construction, the effective dimension and the equilibration results will be as in the case of a strictly local constant of motion.

Crucial in the above construction is that we use the truncated constant of motion $\Gamma_{X_l}(\mathcal{Z})$ for the observable $A$ in order to ensure locality, while we use the full object $\mathcal{Z}$ for the initial state in order to achieve a large effective dimension. What remains to be shown is that despite this locality difference in the construction, we still achieve a large expectation value of $A$ with $|\psi\rangle$, but an almost vanishing expectation value with the infinite time average.

As a first step, we will show that $A$ is almost block-off-diagonal with respect to the eigenprojectors of the full constant of motion $\mathcal{Z}$. Introducing the identity $\mathbb{I} = P_k^l + Q_k^l$, this takes the following form

$$\|P_k A P_k\| \leqslant \|P_k P_k^l A Q_k^l P_k\| + \|P_k Q_k^l A P_k^l P_k\| \leqslant 2\frac{g(l)}{\gamma}. \tag{34}$$

Here we used that $A$ is block-off-diagonal with respect to the truncated constant of motion $\Gamma_{X_l}(\mathcal{Z})$. The same estimate holds for the projectors $Q_k$. Using this, bounding the expectation value with the infinite time average is straightforward

$$\mathrm{tr}(A\omega) = \mathrm{tr}\left(AP_0\omega P_0\right) + \mathrm{tr}\left(AQ_0\omega Q_0\right) \leqslant \|P_0AP_0\| + \|Q_0AQ_0\| \leqslant 4\frac{g(l)}{\gamma}. \tag{35}$$

We now have to show that the expectation value of $A$ with $\rho$ is large initially

$$\langle\psi|\ A\ |\psi\rangle \geqslant \langle v|\ AA\ |v\rangle - \frac{1}{2}\,|\langle v|\ A\ |v\rangle + \langle v|\ AAA\ |v\rangle|. \tag{36}$$

In the following, we will show that the first term is almost one, while the other two almost vanish due to the block-off-diagonality. For the first term, we will use that $A^2 = P_0^l + P_1^l\,|_I$, where $P_1^l\,|_I$ is the projector onto the image of $A$ in $P_1^l$. Using that $\langle v|\ Q\ |v\rangle$ can only increase if we enlarge the subspace of the projector $Q$, we obtain

$$\langle v|\ AA\ |v\rangle \geqslant \langle v|\ P_0^l\ |v\rangle - \left|\langle v|\ Q_0^l\ |v\rangle\right| \geqslant \langle v|\ P_0\ |v\rangle - \left|\langle v|\ Q_0\ |v\rangle\right| - 4\frac{g(l)}{\gamma} = 1 - 4\frac{g(l)}{\gamma}, \tag{37}$$

where we have used (31). The second term can be bounded directly using block-off-diagonality

$$|\langle v|\ A\ |v\rangle| \leqslant \|P_0AP_0\| \leqslant 2\frac{g(l)}{\gamma}. \tag{38}$$

The last term, finally can be bounded as follows

$$\begin{aligned}|\langle v|\ AAA\ |v\rangle| &= \left|\langle v|\ P_0A\left(P_0^l + P_1^l\,|_I\right)P_0\ |v\rangle\right| \\ &\leqslant \|P_0AP_0^l\| + \|P_1^lP_0\| \\ &\leqslant \|P_0^lAP_0^l\| + 2\frac{g(l)}{\gamma} + \frac{g(l)}{\gamma} \\ &\leqslant 3\frac{g(l)}{\gamma}.\end{aligned} \tag{39}$$

To summarize, we have

$$\langle\psi|\ A\ |\psi\rangle \geqslant 1 - 9\frac{g(l)}{\gamma}. \tag{40}$$

Putting together the estimates for the expectation value of $A$ with the initial state, the equilibration result and the expectation value of $A$ with the infinite time average $w$, we obtain the desired bound. More precisely, we choose $\rho = |\psi\rangle\langle\psi|$ and proceed as follows

$$\begin{aligned}\overline{\|A_t - \Gamma_S\left(A_t\right)\|} &\geqslant \mathrm{tr}(A\rho) - \overline{\left|\mathrm{tr}\left(\Gamma_S\left(A_{t_0}\right)\rho_{-t}\right)\right|} \\ &\geqslant \mathrm{tr}(A\rho) - \overline{\|w - \rho_{-t}\|_1} - \left|\mathrm{tr}\left(\Gamma_S\left(A_{t_0}\right)w\right)\right| \\ &\geqslant \mathrm{tr}(A\rho) - \frac{d_{\mathrm{sys}}}{2d_{\mathrm{eff}}^{1/2}} - \left|\mathrm{tr}\left(\Gamma_S\left(A_{t_0}\right)w\right)\right|.\end{aligned} \tag{41}$$

Inserting the effective dimension $d_{\mathrm{eff}} = \tilde{d}_{\min}$ and using equations (35) and (40)

$$\begin{aligned}\overline{\|A_t - \Gamma_S\left(A_t\right)\|} &\geqslant 1 - 9\frac{g(l)}{\gamma} - \frac{d_s}{2\tilde{d}_{\min}^{1/2}} - 4\frac{g(l)}{\gamma} \\ &\geqslant 1 - 13\frac{g(l)}{\gamma} - \frac{d_s}{2\tilde{d}_{\min}^{1/2}}\end{aligned} \tag{42}$$

concludes the proof. □

## 6. Discussion and outlook

In this work, we have shown that for systems with suitably non-degenerate spectrum a single approximately local constant of motion is sufficient to rigorously derive information propagation. We explicitly construct local excitation operators whose effect spreads over arbitrary distances, thus giving rise to a protocol for using MBL systems for signalling. This implies that the recently derived logarithmic light cone [9] can never be tightened to a zero-velocity Lieb–Robinson bound, at least if one allows for infinite energy in the system. These results strengthen and are concomitant with the observation of a logarithmic growth of entanglement entropies in MBL models. It is notable how little has to be assumed to arrive at the conclusion of having information propagation. Not only do the approximately local constants of motion not prohibit information propagation in this sense: they give rise to a protocol for the transmission of classical information for all Hamiltonians with a generic spectrum, which is quite a counterintuitive result.

As future work, it would be interesting to clear up the precise role of local constants of motion in MBL systems. As described above, current analytical descriptions of MBL are usually connected to an extensive set of approximately local constants of motion and are capable of accurately capturing the phenomenology [10, 23, 26]. However from a numerical point of view, it seems still unclear whether important models such as the *XXZ* chain allow such a description [21, 32]. In fact, the occurrence of rare states with large entanglement [18, 30] seems to contradict this, as it is not compatible with the eigenstates described in terms of a spectral tensor network [20]. Assuming that those rare eigenstates are indeed present in the *XXZ* chain, they seem to have little influence on the physical behaviour such as the entanglement growth or the breakdown of thermalisation [7, 8, 10, 30]. However, a more careful analysis, especially of the role of the disorder strength is surely needed. We would therefore speculate that systems could still show MBL properties even if no complete set of constants of motion can be constructed. To what extent models with only a single local constant of motion can be devised is currently unclear, even though MBL systems coupled to generic thermalizing models are certainly candidates in this direction.

Aside from a clarifying discussion of constants of motion, it would be interesting to explore the speed of the information propagation, which naturally is linked to the open problem of deriving physically meaningful time scales of equilibration in local models. Another important question is how the information propagation derived above is linked to the available energy scale in the system. In particular, it would be interesting how our results relate to the possibility of having a mobility edge and how they are connected to the presumed suppression of energy and particle transport in MBL systems.

## Acknowledgments

We thank Christian Gogolin for many interesting discussions on possible links between equilibration and transport and Henrik Wilming and Tobias Osborne for numerous discussions on many-body localisation. We are grateful for support by the EU (SIQS, RAQUEL, AQuS, COST), the ERC (TAQ), the BMBF, the Studienstiftung des Deutschen Volkes, and the EPSRC.

## Appendix

### A.1. Simple MBL model implies information propagation

In this appendix, we prove information propagation in two simple MBL models, one where the Hamiltonian is diagonal in the computational basis and one where the eigenstates are deformed by an *f*-local unitary (definition 4). We start with the following simple model

$$H = \sum_j h_i \sigma_i^z + \sum_{i,j} J_{i,j} \sigma_i^z \sigma_j^z, \tag{43}$$

which consists of interacting Pauli-*Z*-matrices with suitably random coefficients $J_{i,j}$ that decay with the distance between sites *i* and *j*. In this system, the eigenvectors are the computational basis and the energies and gaps will be non-degenerate due to the randomness in the model. For local dimension $d > 2$, one can easily extend the model, by allowing for coupling with arbitrary local and diagonal matrices. In this case, we further require a generalized Pauli-*X*-matrix on site *j* defined via matrix elements

$$\left(\tilde{\sigma}_j^x\right)_{r,s} = \frac{1 - \delta_{r,s}}{d - 1}. \tag{44}$$

As discussed in the main text, this model implies information propagation in the following way.

**Lemma 1 (Diagonal Hamiltonians).** *Let H be a diagonal Hamiltonian with non-degenerate eigenvalues and gaps. Let* $A = \tilde{\sigma}_j^x$ *be the generalized Pauli-X-matrix supported on spin j. Then H necessarily has information propagation on average in the sense that the operator* $A_t$ *has, on average, support outside any finite region S*

$$\overline{\|A_t - \Gamma_S(A_t)\|} \geqslant 1 - d^{|S|-N/2}. \tag{45}$$

The proof is contained in the main text and directly carries over to the case with local dimension $d > 2$. Here, we will extent this result and show that the same construction can still be carried out in the case of approximately local eigenvectors. For this, we introduce approximately local operators and unitaries.

**Definition 3 (Local observable).** An operator $A$ will be called $(g, X)$-local, if there exists a finite localisation region $X$ such that

$$\|A - \Gamma_{X_l}(A)\| \leqslant \|A\| g(l) \tag{46}$$

for some function $g: \mathbb{N} \to \mathbb{R}_0^+$ with suitable decay in $l$.

**Definition 4 (Local unitary).** A unitary operator $U$ will be called $f$-local, if the conjugation of a local observable $A$ with localisation region $X$ remains local in the sense that

$$\|UAU^\dagger - \Gamma_{X_l}\left(UAU^\dagger\right)\| \leqslant \|A\| f(l) \tag{47}$$

for some function $f: \mathbb{N} \to \mathbb{R}_0^+$ with suitable decay in $l$.

Correspondingly, we will say that a Hamiltonian has $f$-local eigenvectors, if the unitary diagonalizing it is $f$-local. With this, we are able to present a generalized version of lemma 1 to the case of approximately local eigenvectors.

**Corollary 2 ($f$-local eigenvectors imply information propagation).** *Let $H$ be a Hamiltonian with $f$-local eigenvectors and non-degenerate energies and gaps. Then $H$ necessarily has information propagation on average in the sense that for any fixed finite region $X_l$ of diameter $l$, there exists a local operator $A$ initially supported on $X_l$ with $\|A\| = 1$ such that $A_t$ has, on average, support outside any finite region $S$*

$$\overline{\|A_t - \Gamma_S\left(A_t\right)\|} \geqslant 1 - d^{|S|-N/2} - 2f(l). \tag{48}$$

**Proof.** We will now use lemma 1 in order to provide a proof for corollary 2 . For this, we use that the Hamiltonian can be diagonalised by a $f$-local unitary $V$ and work with the observable

$$A = V\tilde{\sigma}_j^x V^\dagger, \tag{49}$$

where $\tilde{\sigma}_j^x$ is the generalized Pauli-$X$-matrix on some spin $j$ within the set $S$. This operator will no longer be strictly local, but due to the $f$-locality of the unitary $V$, the operator can be truncated

$$\|A - \Gamma_{X_l}(A)\| \leqslant f(l), \tag{50}$$

where $X_l$ denotes the set that contains the inital support, namely site $j$ and all $l$-nearest neighbours. Here we used that the operator norm of $A$ is one. From this, we can use the local reduction $A^l = \Gamma_{X_l}(A)$ as the local operator that will display information propagation. We will further need the time evolution of this truncated operator $A_t^l = \mathrm{e}^{\mathrm{i}tH} A^l \mathrm{e}^{-\mathrm{i}tH}$, where we first truncate and then evolve it in time. Naturally the unitary time evolution does not change the norm difference. The proof relies on a series of triangle inequalities. First we use that for any state

$$\|A_t^l - \Gamma_S\left(A_t^l\right)\| \geqslant |\mathrm{tr}\left(A_t^l \rho_{-t}\right)| - |\ \mathrm{tr}\left(\Gamma_S\left(A_t^l\right)\rho_{-t}\right)|. \tag{51}$$

Next, we look at the two terms separately, which gives

$$\begin{aligned} |\mathrm{tr}\left(A_t^l \rho_{-t}\right)| &= |\mathrm{tr}\left(\mathrm{e}^{\mathrm{i}tH}\Gamma_{X_l}(A)\mathrm{e}^{-\mathrm{i}tH}\rho_{-t}\right)| \\ &\geqslant |\mathrm{tr}\left(A_t \rho_{-t}\right)| - \|A - \Gamma_{X_l}(A)\| \\ &\geqslant |\mathrm{tr}\left(A_0 \rho_0\right)| - f(l) = 1 - f(l), \end{aligned} \tag{52}$$

where we have inserted a zero term $\pm\mathrm{tr}(A_t \rho_{-t})$ and have used the above truncation estimate in equation (50) . The other term can be estimated as follows

$$\begin{aligned} |\mathrm{tr}\left(\Gamma_S\left(A_t^l\right)\rho_{-t}\right)| &= |\mathrm{tr}\left(\Gamma_S\left(\mathrm{e}^{\mathrm{i}tH}\Gamma_X(A)\mathrm{e}^{-\mathrm{i}tH}\right)\rho_{-t}\right)| \\ &\leqslant \|\Gamma_S\left(\mathrm{e}^{\mathrm{i}tH}\left(\Gamma_X(A) - A\right)\mathrm{e}^{-\mathrm{i}tH}\right)\| \\ &\quad + |\mathrm{tr}_S\left(\Gamma_S\left(A_t\right)\mathrm{tr}_{S^c}(\rho_{-t})\right)|. \end{aligned} \tag{53}$$

Here we again inserted a zero term

$$\pm\mathrm{tr}_S\left(\Gamma_S\left(A_t\right)\right)\mathrm{tr}_{S^c}(\rho_{-t}) \tag{54}$$

and used a norm estimate. To proceed, we insert one more zero term $\pm\mathrm{tr}_S\left(\Gamma_S(A_t)d^{-|S|}\mathbb{I}_S\right)$ and use the triangle inequality

$$|\mathrm{tr}\,\Gamma_S\left(A_t^l\right)\rho_{-t}| \leqslant \|\Gamma_X(A) - A\| + \|\mathrm{tr}_{S^c}(\rho_{-t}) - d^{-|S|}\mathbb{I}_S\|_1 + |\mathrm{tr}_S\left(\Gamma_S\left(A_t\right)d^{-|S|}\mathbb{I}_S\right)|, \tag{55}$$

and use that $\Gamma$ is a norm contractive map. These three terms can now easily be bounded. The first is small due to the $f$-locality of the unitary $V$ involved in constructing $A$, see equation (50) . The second term becomes small, once the time average is taken, which allows us to use known equilibration results [12]

$$\overline{\left|\mathrm{tr}_S\left(\Gamma_S\left(A_t\right)d^{-|S|}\mathbb{I}\right)\right|} \leqslant d^{|S|-N/2}. \tag{56}$$

The third term vanishes completely, since the observable $A$ has zeros on its diagonal. This completes the estimate of the second term in equation (51)

$$\overline{\left|\mathrm{tr}\left(\Gamma_S\left(A_t^l\right)\rho_{-t}\right)\right|} \leqslant f(l) + d^{|S|-N/2}. \tag{57}$$

Patching the estimates in equations (52) and (57) together concludes the proof

$$\|A_t^l - \Gamma_S\left(A_t^l\right)\| \geqslant 1 - 2f(l) - d^{|S|-N/2}. \tag{58}$$

□

The above Hamiltonians are special instances of systems having local constants of motion. In the diagonal case, the constants of motion are simply the local Pauli-$Z$-matrices. Once they are deformed by a quasi-local unitary, exact locality is lost, but one still obtains a full set of approximately local constants of motion.

### A.2. Approximate spectral tensor networks

In [20], it is shown that if a Hamiltonian has suitable local constants of motion, then each eigenprojector can be efficiently represented as a matrix product operator. Moreover, it is rigorously derived that there exists an efficient spectral tensor network for all eigenprojectors at the same time. Reference [20] then proceeds to sketch the case of approximately local constant of motion, for which similar conclusions are reached. In this appendix, we show that indeed, even for approximately local constants of motion with robust spectrum (definition 1 in the main text), one can rigorously obtain a spectral tensor network.

**Result.** (Efficient spectral tensor networks). Let $H$ be a Hamiltonian with an extensive number of approximately local constants of motion (definition 1 in the main text) with $|X| \leqslant L$ and $g(l) \leqslant c_1 \exp(-c_2 l)$, for suitable constants $c_1, c_2 > 0$. We assume that the approximately local constants of motion are algebraically independent, commute with each other and have suitable distributed support on the lattice. Then there exists an efficient spectral tensor network representation for all eigenprojectors of $H$.

The proof of this statement directly follows from [20], together with our corollary 3 in the main text. We start from the observation that the approximate locality of the constant of motion also implies quasi-local eigenprojectors, in the sense that

$$\|P_k - P_k^l\| \leqslant \frac{2\,g(l)}{\gamma}, \tag{59}$$

and that the perturbed and unperturbed eigenspaces have the same dimension. Using this approximation, one finds that projectors onto an eigenspace of an approximately local constant of motion can be efficiently approximated by matrix-product operators. For a given site $j$, call $A$ the subset of sites for which the MPO approximations have a support that includes $j$. With the same argument as in [20], choosing a path in the supports of the strictly local constants of motion and performing singular value decompositions, as outlined in [20], one finds that the collection of all approximately local constants of motion in $A$ can again be written as a matrix-product operator. The stability lemma 2 below concludes the proof.

**Lemma 2 (Stability).** *Let $\{\mathcal{Z}_j\}$ be a set of N approximately local constants of motion with a lower uniform bound $\gamma > 0$ on their minimal spectral gaps and uniform upper bounds L and g(l) on size and decay of their localisation regions $X_j$, such that*

$$\|\mathcal{Z}_j - \Gamma_{X_j^l \supset X_j}\left(\mathcal{Z}_j\right)\| \leqslant g(l) \tag{60}$$

*for any $X_j^l$ containing $X_j$ together with a buffer region of size l. Then if $P_{j,m}$ denotes the eigenprojector of $\mathcal{Z}_j$ for eigenvalue m we have*

$$\|P_{j_1,m_1}\cdots P_{j_N,m_n} - P^l_{j_1,m_1}\cdots P^l_{j_N,N_1}\| \leqslant 2\,N\frac{g(l)}{\gamma} \tag{61}$$

*with* $P^l_{j_i,m_i} = \Gamma_{X^l_j}(P_{j_i,m_i})$ *being strictly local.*

**Proof.** The proof utilizes perturbation theory on the level of the single eigenprojectors $P^l_{j_i,m_i}$ similar to the proof of corollary 3 in the main text. Using the triangle inequality we can upper bound the norm difference as

$$\|P_{j_1,m_1}\cdots P_{j_N,m_n} - P^l_{j_1,m_1}\cdots P^l_{j_N,N_1}\| \leqslant \sum_{k=1}^{N}\|P_{j_k,m_k} - P^l_{j_k,m_k}\|. \tag{62}$$

The result now follows from equation (31) and our uniformity assumptions. □

## References


[1] Eisert J, Friesdorf M and Gogolin C 2015 Quantum many-body systems out of equilibrium *Nat. Phys.* **11** 124–30
[2] Polkovnikov A, Sengupta K, Silva A and Vengalattore A 2011 Non-equilibrium dynamics of closed interacting quantum systems *Rev. Mod. Phys.* **83** 863
[3] Gogolin C and Eisert J 2015 Equilibration, thermalization, and the emergence of statistical mechanics in closed quantum systems arxiv:1503.07538
[4] Basko D M, Aleiner I L and Altshuler B L 2006 Metal-insulator transition in a weakly interacting many-electron system with localized single-particle states *Ann. phys.* **321** 1126–205
[5] Gornyi I V, Mirlin A D and Polyakov D G 2005 Interacting electrons in disordered wires: Anderson localization and low-*t* transport *Phys. Rev. Lett.* **95** 206603
[6] Burrell C K and Osborne T J 2007 Bounds on the speed of information propagation in disordered quantum spin chains *Phys. Rev. Lett.* **99** 167201
[7] Bardarson J H, Pollmann F and Moore J E 2012 Unbounded growth of entanglement in models of many-body localization *Phys. Rev. Lett.* **109** 017202
[8] Znidaric M, Prosen T and Prelovsek P 2008 Many-body localization in the Heisenberg XXZ magnet in a random field *Phys. Rev.* B **77** 064426
[9] Kim I H, Chandran A and Abanin D A 2014 Local integrals of motion and the logarithmic lightcone in many-body localized systems (arXiv:1412.3073)
[10] Nandkishore R and Huse D A 2015 Many body localization and thermalization in quantum statistical mechanics *Ann. Rev. Condens. Matter Phys.* **6** 15–38
[11] Gogolin C, Mueller M P and Eisert J 2011 Absence of thermalization in non-integrable systems *Phys. Rev. Lett.* **106** 040401
[12] Linden N, Popescu S, Short A J and Winter A 2009 Quantum mechanical evolution towards thermal equilibrium *Phys. Rev.* E **79** 061103
[13] Malabarba A S L, Garcia-Pintos L P, Linden N, Farrelly T C and Short A J 2014 Quantum systems equilibrate rapidly for most observables *Phys. Rev.* E **90** 012121
[14] Reimann P and Kastner M 2012 Equilibration of isolated macroscopic quantum systems *New. J. Phys.* **14** 043020
[15] Srednicki M 1994 Chaos and quantum thermalization *Phys. Rev.* E **50** 888–901
[16] Deutsch J M 1991 Quantum statistical mechanics in a closed system *Phys. Rev.* A **43** 2046–9
[17] Rigol M, Dunjko V and Olshanii M 2008 Thermalization and its mechanism for generic isolated quantum systems *Nature* **452** 854–8
[18] Bauer B and Nayak C 2013 Area laws in a many-body localized state and its implications for topological order *J. Stat. Mech.* **P09005**
[19] Friesdorf M, Werner A H, Brown W, Scholz V B and Eisert J 2015 Many-body localization implies that eigenvectors are matrix-product states *Phys. Rev. Lett.* **114**
[20] Chandran A, Carrasquilla J, Kim I H, Abanin D A and Vidal G 2015 Spectral tensor networks for many-body localization *Phys. Rev.* B **92** 024201
[21] Chandran A, Kim I H, Vidal G and Abanin D A 2015 Constructing local integrals of motion in the many-body localized phase *Phys. Rev.* B **91** 085425
[22] Nanduri A, Kim H and Huse D 2014 Entanglement spreading in a many-body localized system *Phys. Rev.* B **90** 064201
[23] Huse D A, Nandkishore R and Oganesyan V 2014 Phenomenology of fully many-body-localized systems *Phys. Rev.* B **90** 174202
[24] Serbyn M, Papić Z and Abanin D A 2013 Local conservation laws and the structure of the many-body localized states *Phys. Rev. Lett.* **111** 127201
[25] Lieb E H and Robinson D W 1972 The finite group velocity of quantum spin systems *Commun. Math. Phys.* **28** 251–7
[26] Serbyn M, Papić Z and Abanin D A 2013 Universal slow growth of entanglement in interacting strongly disordered systems *Phys. Rev. Lett.* **110** 260601
[27] Gogolin C 2014 Equilibration and thermalization in quantum systems *PhD Thesis* FU Berlin
[28] Anderson P W 1958 Absence of diffusion in certain random lattices *Phys. Rev.* **109** 1492
[29] Osborne T J and Linden N 2004 Propagation of quantum information through a spin system *Phys. Rev.* A **69** 052315
[30] Luitz D J, Laflorencie N and Alet F 2015 Many-body localization edge in the random-field Heisenberg chain *Phys. Rev.* B **91** 081103
[31] Bhatia R 1996 *Matrix Analysis* (Berlin: Springer)
[32] Pekker D and Clark B K 2014 Encoding the structure of many-body localization with matrix product operators (arXiv:1410.2224)