Many-body localisation implies that eigenvectors are matrix-product states

M. Friesdorf,[1] A. H. Werner,[1] W. Brown,[2] V. B. Scholz,[3] and J. Eisert[1]

[1]*Dahlem Center for Complex Quantum Systems, Freie Universität Berlin, 14195 Berlin, Germany*
[2]*Computer Science Department, University College London, London WC1E 6BT*
[3]*Institute for Theoretical Physics, ETH Zurich, Wolfgang-Pauli-Str. 27, 8093 Zurich, Switzerland*
(Dated: November 19, 2014)

The phenomenon of many-body localisation received a lot of attention recently, both for its implications in condensed-matter physics of allowing systems to be an insulator even at non-zero temperature as well as in the context of the foundations of quantum statistical mechanics, providing examples of systems showing the absence of thermalisation following out-of-equilibrium dynamics. In this work, we establish a novel link between dynamical properties – a vanishing group velocity and the absence of transport – with entanglement properties of individual eigenvectors. Using Lieb-Robinson bounds and filter functions, we prove rigorously under simple assumptions on the spectrum that if a system shows strong dynamical localisation, all of its many-body eigenvectors have clustering correlations. In one dimension this implies directly an entanglement area law, hence the eigenvectors can be approximated by matrix-product states. We also show this statement for parts of the spectrum, allowing for the existence of a mobility edge above which transport is possible.

The concept of disorder induced localisation has been introduced in the seminal work by Anderson [1] who captured the mechanism responsible for the absence of diffusion of waves in disordered media. This mechanism is specifically well understood in the single-particle case, where one can show that in the presence of a suitable random potential, all eigenfunctions are exponentially localised [1]. In addition to this spectral characterisation of localisation there is a notion of dynamical localisation, which requires that the transition amplitudes between lattice sites decay exponentially [2, 3].

Naturally, there is a great interest in extending these results to the many-body setting [4, 5]. In the case of integrable systems that can be mapped to free fermions, such as the XY chain, results on single particle localisation can be applied directly [6, 7]. A far more intricate situation arises in interacting systems. Such *many-body localisation* [8, 9] has received an enormous attention recently. In terms of condensed-matter physics, this phenomenon allows for systems to remain an insulator even at non-zero temperature [10], in principle even at infinite temperature [11]. In the *foundations of statistical mechanics*, such many-body localised systems provide examples of systems that fail to thermalise. When pushed out of equilibrium, signatures of the initial condition will locally be measurable even after long times, in contradiction to what one might expect from quantum statistical mechanics [4, 12, 13].

Despite great efforts to approach the phenomenon of many-body localisation, many aspects are not fully understood and a comprehensive definition is still lacking. Similar to the case of single-particle Anderson localisation, there are two complementary approaches to capture the phenomenon. On the one hand probes involving real-time dynamics [14, 15] have been discussed, showing excitations "getting stuck", or seeing suitable signatures in density-auto-correlation functions [16], leading to a dynamic reading of the phenomena. On the other hand, statistics of energy levels has been considered as an indicator [11] as well as a lack of entanglement in the eigenbasis [5] and a resulting violation of the eigenstate thermalisation hypothesis (ETH) [17–19]. In such a static description of many-body localisation, it has for example been suggested to define many-body localised states as those states that can be prepared from product states or Slater determinants of single-particle localised states with a finite-depth local unitary transformation. Accordingly, a many-body localised phase is one where most eigenvectors have this property [5].

It seems fair to say that the connection between these static and dynamic readings is still unclear. In this work, we make a substantial first step in relating these approaches: If the evolution of an interacting many-body system is dynamically localising (and its spectrum is generic), then all eigenvectors have exponentially clustering correlations in general and, in one dimension, satisfy an area law and can be efficiently written as a matrix-product state (MPS) [20, 21] with a low bond dimension. Our proof uses advanced mathematical tools developed for the analysis of interacting many-body systems, such as energy filtering and does not rely on the precise details of the considered model. We show that short ranged correlations in individual energy eigenstates necessarily follow from absence of transport in localising models.

Setting. For simplicity, we will state all our results in the language of spin systems of local dimension d, but they can equivalently be derived for fermionic systems. We consider local Hamiltonians $H = \sum_j h_j$, where each interaction term h_j is supported on finitely many sites of a lattice with N vertices, and the physical system associated with each site is finite-dimensional. We make for most of the argument no assumptions concerning the dimensionality or the kind of the lattice. In what follows, we will consider observables A and B that are supported on finitely many sites on the lattice. There is a natural distance $d(A, B)$ in the lattice related to the length of the shortest path of interactions connecting the supports of A and B. For convenience of notation, we will choose $\|A\| = 1 = \|B\|$ from here on, which does not restrict generality [22–26]. For any local Hamiltonian of this sort there is an upper bound $v > 0$ to the largest group velocity, referred to as Lieb-Robinson velocity. There are several possible formulations of such a bound (see Appendix A for a detailed discussion). One convenient and common way is to compare the time-evolution of a local observable $A(t) = e^{itH} A e^{-itH}$ under the full Hamiltonian H with its time-evolution under a truncated Hamiltonian H_A^l that only includes interactions h_j

contained in a region of distance no more than l from the support of A. The Lieb-Robinson theorem then states that there exists a constant $c > 0$ and a velocity $v > 0$ such that

$$\|A(t) - e^{itH_A^l} A e^{-itH_A^l}\| \le c e^{-\mu(l - vt)}$$

is true for all times $t \ge 0$. Such a bound limits the *speed of information propagation*: In any such lattice with short-ranged interactions, all interactions are causal in this sense. A related, but in the case of $v = 0$ weaker estimate is given by the following commutator bound (see also A)

$$\|[A, B(t)]\| \le C e^{-\mu(\mathrm{d}(A,B) - vt)}. \qquad (1)$$

Our results will rely on two generic and natural assumptions on the spectrum of the Hamiltonian. To state those, we express the many-body Hamiltonian H in its eigenbasis as

$$H = \sum_k E_k \, |k\rangle\langle k| \, .$$

- *Non-degenerate energies (AI):* The energies of the full Hamiltonian are assumed to be non-degenerate. The smallest gap between these energies will be called γ.

- *Locally independent gaps (AII):* The energies of reduced Hamiltonians H_A and H_B which include all interactions inside rectangular regions A and B respectively, are assumed to be non-degenerate when viewing them as operators on their respective truncated Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$. The smallest gap will be called $\tilde{\gamma}$. Moreover, the gaps of these Hamiltonians need to be locally independent with respect to each other, in the sense that

$$E_a - E_{a'} = E_b - E_{b'} \Rightarrow a = a', b = b',$$

 where a, a' and b, b' label the eigenvalues of H_A and H_B respectively. The smallest difference of these gaps will be called η.

- *Non-symmetric gaps (AIII):* For all eigenvalues E_k of the full Hamiltonian the spectrum is no-where symmetric in the sense that

$$\min_{r \ne s}\big||E_r - E_k| - |E_s - E_k|\big| \ge \zeta > 0 \; \forall k.$$

Similar reasonable conditions have already been considered in the context of equilibration of closed systems and are assumed to hold when a small amount of random noise is added to a local Hamiltonian [27]. Especially in the context of localising systems which are typically based on large random terms in the Hamiltonian, these assumptions are very natural indeed. Our assumptions are, however, considerably weaker than assuming fully non-degenerate energy gaps [27]. Another quantity that will be important later on is the number of states up to energy E given by $\Theta(E) := \sum_{l, E_l \le E} 1$.

Notions of dynamical localisation. One defining feature of localisation is that the system shows no transport of information. There are many possible readings of this. The first and strongest is that the time evolution can be truncated to a finite region independent of time and is directly connected to an absence of thermalisation in the model (see Appendix G).

Definition 1 (Strong dynamical localisation)**.** *A Hamiltonian H exhibits strong dynamical localisation iff its time evolution satisfies*

$$\|A(t) - e^{itH_A^l} A e^{-itH_A^l}\| \le c_{\mathrm{loc}} e^{-\mu l} \, ,$$

for a suitable $\mu > 0$, where A is an arbitrary local observable, $c_{\mathrm{loc}} > 0$ a constant independent of the system size and H_A^l denotes a Hamiltonian which includes all interactions contained in a region of distance no more than l from the support of A.

This corresponds to setting the Lieb-Robinson velocity v to zero in Eq. (1). A potential candidate for such a model is given by free fermionic systems with strong local disorder. There, it is known that transport is strongly suppressed and it is still being debated to what extent they satisfy zero velocity Lieb-Robinson bounds [6, 28]. This definition also implies that information cannot be transported by encoding it locally in the chosen Hamiltonian (see Appendix A for details). As such, it does not allow for a growth of entanglement entropies logarithmically in time, a phenomenon that has been numerically observed in spin chains with random onsite noise [15].

In a second step, we relax our dynamical assumption to a weaker variant of zero velocity Lieb-Robinson bounds and restricting it to a suitable subspace. It is natural to assume that the transport properties of an electronic system are closely connected to the energy available in the system. In fact, often the notion of a *mobility edge* is introduced [28, 29]. The intuition is that for all energies below this mobility edge, transport is fully blocked, while it might be possible for higher energies. The following definition is making this precise and allows for transport in highly excited sectors.

Definition 2 (Mobility edge)**.** *A Hamiltonian H is said to have a mobility edge at energy E_{mob} iff its time evolution satisfies for all times t, $\forall \rho \in \mathrm{span}\{|l\rangle\langle k| : \; E_l, E_k \le E_{\mathrm{mob}}\}$*

$$|\mathrm{tr}\left(\rho[A(t), B]\right)| \le \min(t, 1) c_{\mathrm{mob}} e^{-\mu \mathrm{d}(A,B)} \, , \qquad (2)$$

with constants $c_{\mathrm{mob}}, \mu > 0$ independent of t. That is, all transport is suppressed for states supported only on the low-energy sector below energy E_{mob}.

This definition relies on a rather weak version of zero velocity Lieb-Robinson bounds (see Appendix A). While our results will be applicable to all eigenvectors below the mobility edge, they can naturally also be used to capture the ground state, yielding an extension of the previous results [28].

In experiments, a natural setting for dynamical localisation is one where a system is excited locally. For example, in case of a ground state vector $|0\rangle$, this corresponds to the action of a unitary operator U with local support $|\psi\rangle = U|0\rangle = e^{-iG}|0\rangle$ where G denotes its generator. In this setting, suppression of transport means that the excitation cannot be measured at far distances, even for long times, in the sense that

$$|\langle 0| A(t) - e^{iG} A(t) e^{-iG} |0\rangle| \le \min(t, 1)\|G\| C e^{-\mu d(A,U)},$$

for arbitrary observables A. Here $C, \mu > 0$ are constants independent of the time t and the system size. If excitations are

stuck in the above sense in an entire certain energy subspace, then this implies our zero velocity Lieb-Robinson estimate in Eq. (2), thus relating Definition 2 to a quantity that can in principle be experimentally measured (see Appendix A).

Main result. We are now in the position to state our main result. It provides a clear link between the dynamical properties of the system and the static correlation behaviour of the corresponding eigenvectors and can be applied to quantum systems of arbitrary dimension. In one dimension it directly implies an area law and the existence of an approximating MPS (see Corollary 2).

Theorem 1 (Clustering of correlations of eigenvectors). *The dynamical properties of a local Hamiltonian imply clustering correlations of its eigenvectors in the following way.*

> **a** *If the Hamiltonian shows strong dynamical localisation and its spectrum fulfils assumptions **AI** and **AII** then all its eigenvectors $|k\rangle$ have exponentially clustering correlations, i.e.*

$$|\langle k| AB |k\rangle - \langle k| A |k\rangle\langle k| B |k\rangle| \le 4c_{\mathrm{loc}}e^{-\mu \mathrm{d}(A,B)/2}.$$

> **b** *If the Hamiltonian has a mobility edge at energy E_{mob} and its spectrum fulfils assumptions **AI** and **AIII**, then all eigenvectors $|k\rangle$ up to that energy E_{mob} cluster exponentially, in the sense that for all $\kappa > 0$*

$$|\langle k| AB |k\rangle - \langle k| A |k\rangle\langle k| B |k\rangle|$$
$$\le \left(12\pi\Theta(E_k + \kappa)c_{\mathrm{mob}} + \ln\frac{\pi\mu\mathrm{d}(A,B)e^{4+2\pi}}{\kappa^2}\right)\frac{e^{-\mu\mathrm{d}(A,B)}}{2\pi},$$

> *where $\Theta(E)$ is the number of eigenstates up to energy E and κ can be chosen arbitrarily to optimise the bound.*

Part **b** of the theorem only requires Lieb-Robinson bounds on a subspace and is thus perfectly compatible with a growth of entanglement entropies following quenches. In fact, for the proof it is sufficient to assume Eq. (2) on the level of the individual eigenstates (see Appendix E). It further allows to freely choose an energy cut-off κ that serves as an artificial energy gap and can be used to optimise the bound. For typical local models, we expect the density of states to behave like a Gaussian, an intuition that can be numerically tested for small systems (see Fig. 1), and rigorously proved in a weak sense [30]. In this case the number of states will behave like a low-order polynomial in the system size at energies close to the ground state energy $E_k + \kappa$. Thus fixing a $\kappa > 0$ independent of the system size will lead to a pre-factor that scales like a low-order polynomial in the system size where the order of the polynomial will increase as one moves to higher energies. In the bulk of the spectrum, the number of states will grow exponentially with the system size, thus rendering our bounds useless as one moves to high energies. Interestingly, this feature of stronger correlations, associated with a larger entanglement in the state, at higher energies seems to be shared by the Heisenberg chain with random on-site magnetic field (see Fig. 1). Naturally our theorem can also be applied to the ground state where it extends previous results [28] and

for example for the case of an almost degenerate ground state will be a substantial improvement. Naturally, our results can also be applied to the highly excited states at the other end of the spectrum.

The proofs for both parts of the theorem rely on energy filtering techniques [29, 31]. This is a versatile tool especially in the study of perturbation bounds for Hamiltonian systems and allows to partly diagonalise a local observable in the energy eigenbasis, while still keeping some locality structure. Energy filtering of an observable is defined as

$$I_f^H(A) = \int_{-\infty}^{\infty} dt f(t)A(t).$$

Here $f : \mathbb{R} \to \mathbb{R}^+$ is a *filter function* that can be used to interpolate between locality of the resulting observable and the strength of the off-diagonal elements in the Hamiltonian basis (see Appendix B).

In order to show part **a** of the theorem, we will make use of energy filterings of the local observables A and B with respect to the full Hamiltonian H as well as with respect to local restrictions H_A and H_B. Those restrictions contain the support of the corresponding observable and are chosen as large as possible, while still satisfying $[H_A, H_B] = 0$. The main idea is to choose a suitable (Gaussian) filter function I_α^H and use the gap assumptions **AI** and **AII** to show that the joint filter of the observables AB decouples into energy filters for A and B separately. Then strong dynamical localisation can be used to make these filters local. Finally choosing the width of the energy filter α small enough, and in particular exponentially small in the system size, allows to conclude the proof (for details see Appendix D). The proof of part **b** of the theorem is more involved, relies on a high-pass filter (see Appendix B) and is contained in Appendix E.

Implications on area laws and matrix-product states. In one dimension, the conclusions of our main theorem about the correlation behaviour of eigenvectors can be turned into a statement about their entanglement structure. It has been noted before that many-body localisation should be connected to eigenvectors fulfilling an area law (see Conjecture 1 in Ref. [5]), and eigenvectors being well approximated by matrix-product state vectors of the form

$$|\mathrm{MPS}\rangle = \sum_{i_1,\dots,i_N=1}^{D} \mathrm{tr}(M_{i_1}M_{i_2}\dots M_{i_N})|i_1,\dots,i_N\rangle,$$

where D is the bond dimension and $M_j \in \mathbb{C}^{D \times D}$ for all j. Our main theorem allows to rigorously prove this connection.

Corollary 2 (Area laws and matrix-product states). *An eigenvector $|k\rangle$ of a localising Hamiltonian can be approximated by an MPS with fidelity $|\langle k|\mathrm{MPS}\rangle| \ge 1 - \epsilon$, where the bond dimension for a sufficiently large system is given as follows.*

> **a** *If the Hamiltonian shows strong dynamical localisation and its spectrum fulfils assumptions **AI** and **AII**, then the statement holds for all eigenvectors $|k\rangle$ and, for some constant $C > 0$, the approximation has a bond dimension*

$$D = C\left(N/\epsilon\right)^{\frac{16}{\mu \log_2 e}}.$$
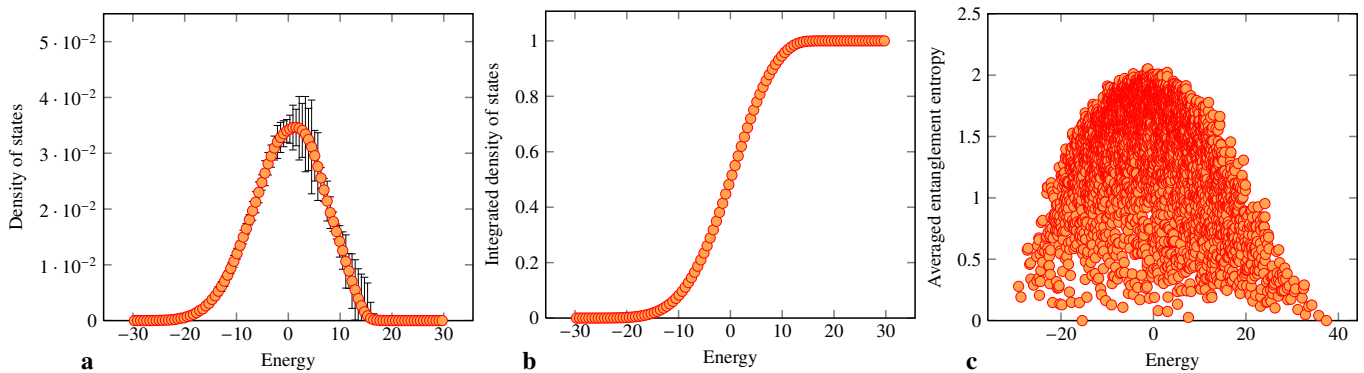
FIG. 1.   We look at a one-dimensional Heisenberg model with random on-site magnetic field of the form $H = \sum_i (X_i X_{i+1} + Y_i Y_{i+1} + Z_i Z_{i+1} + \mu_i Z_i)$, where μ_i is drawn uniformly from the interval $[-h, h]$. **a** Shown is the number of eigenstates for 14 sites as a function of the discretised energy for $h = 1$ averaged over 100 realisations and the variance of this averaging is indicated by vertical error bars. The resulting function is close to a Gaussian. **b** Here, the integrated density of states of this model with the same parameters is shown. The corresponding error bars are too small to appear. This quantity is related to the number of states Θ by an exponential factor 2^N. This plot indicates that indeed the number of states will have small tails and thus few states at low energies. **c** Depicted is the (von Neumann) entanglement entropy of the eigenvectors as a function of the energy. The results were obtained for the disorder strength $h = 4$ on a system of 12 sites and for each individual state an average over the different cuts through the 1D chain was taken. This plot corroborates the intuition that the entanglement entropy and thus the associated bond dimension of the corresponding MPS increase with the energy.

b *If the Hamiltonian has a mobility edge at energy E_{mob}, and its spectrum fulfils assumptions **AI** and **AIII**, then the statement holds for all eigenvectors below this energy E_{mob} and the bond dimension is given by*

$$D = \mathrm{poly}\left(\Theta(E_k + \kappa), N\right),$$

for any fixed κ which enters in the precise form of the polynomial.

The proof is a direct consequence of our main theorem together with the fact that exponential clustering in one dimension implies strong bounds on entanglement entropies between any bipartite cut of the chain [32]. Using techniques from [33] this leads to an efficient MPS approximation with the above bounds on the bond dimension (see Appendix F for details).

Summary & outlook. Despite considerable progress in understanding the effects of random potentials on quantum many-body systems, a precise definition of the phenomenon of many-body localisation continues to be elusive. Attempts to capture the phenomenology can largely be classified into two complementary approaches: One of them puts characteristic properties of the eigenfunctions into the focus of attention and asks for a lack of entanglement and a violation of the ETH. The other one takes the suppression of transport as the basis, which seems closer to being experimentally testable.

In this work, we have established a clear link between these two approaches, by showing that dynamical localisation implies that eigenvectors cluster exponentially. This result, together with the existence of an approximating MPS description in one dimension, reminds of the definition of Ref. [5], that defines many-body localisation in terms of matrix-product state approximations of eigenstates. In contrast, in our work, this feature is shown to follow from an absence of

transport. For future research, it would be interesting to further explore this connection and to address the converse direction, namely to establish that an area-law for all eigenvectors implies that excitations cannot travel through the system.

A different approach towards approximating the individual eigenstates with a matrix-product state could potentially be provided by constructing meaningful local constants of motion that give a set of local quantum numbers [34]. It seems likely that tools using energy filtering will again prove useful in this context, a prospect that we briefly discuss in Appendix C. On the practical side, further numerical effort will be needed to understand the behaviour of individual models and to fully understand the transport properties for different energy scales. Our work could well provide a first stepping stone for further endeavours in this direction: Showing that eigenstates are well approximated by matrix-product states implies that not only ground states, but in fact also excited states can efficiently be described in terms of tensor networks. Our result as such does not yet provide an efficient algorithm to find the respective matrix-product states: this reminds of the situation of the existence of lattice models for which the ground states are exact matrix-product states, but it amounts to a computationally difficult problem to find them [35]. Still, this appears to be a major step in the direction of formulating such numerical prescriptions of describing the low-energy sector of many-body localizing systems. Eventually, the leading vision in any of these endeavours appears to be a rigorous proof of many-body localisation in the spirit of the original results by Anderson. For this, creating a unifying framework and linking the possible definitions seems a key first step.

[1] Anderson, P. W. Absence of diffusion in certain random lattices. *Phys. Rev.* **109**, 1492 (1958).

[2] Klein, A. Multiscale analysis and localization of random operators (2007). ArXiv:0708.2292.

[3] Stolz, G. An introduction to the mathematics of Anderson localization. In *Entropy and the Quantum II*, vol. 552 of *Contemp. Math.* (American Mathematical Society, 2011).

[4] Nandkishore, R. & Huse, D. A. Many body localization and thermalization in quantum statistical mechanics. ArXiv:1404.0686.

[5] Bauer, B. & Nayak, C. Area laws in a many-body localised state and its implications for topological order. *J. Stat. Mech.* P09005 (2013).

[6] Burrell, C. K. & Osborne, T. J. Bounds on the speed of information propagation in disordered quantum spin chains. *Phys. Rev. Lett.* **99**, 167201 (2007).

[7] Hamza, E., Sims, R. & Stolz, G. Dynamical localization in disordered quantum spin systems. *Commun. Math. Phys.* **315**, 215 (2012).

[8] Basko, D. M., Aleiner, I. L. & Altshuler, B. L. Metal-insulator transition in a weakly interacting many-electron system with localised single-particle states. *Ann. Phys.* **321**, 1126 (2006).

[9] Basko, D. M., Aleiner, I. L. & Altshuler, B. L. On the problem of many-body localization. ArXiv:cond-mat/0602510.

[10] Aleiner, I. L., Altshuler, B. L. & Shlyapnikov, G. V. A finite-temperature phase transition for disordered weakly interacting bosons in one dimension. *Ann. Phys.* **321**, 1126 (2006).

[11] Oganesyan, V. & Huse, D. A. Localization of interacting fermions at high temperature. *Phys. Rev. B* **75**, 155111 (2007).

[12] Pal, A. & Huse, D. A. The many-body localization transition. *Phys. Rev. B* **82**, 174411 (2010).

[13] Gogolin, C., Mueller, M. P. & Eisert, J. Absence of thermalization in nonintegrable systems. *Phys. Rev. Lett.* **106**, 040401 (2011).

[14] Nanduri, A., Kim, H. & Huse, D. A. Entanglement spreading in a many-body localised system. ArXiv:1404.5216.

[15] Bardarson, J. H., Pollmann, F. & Moore, J. E. Unbounded growth of entanglement in models of many-body localization. *Phys. Rev. Lett.* **109**, 017202 (2012).

[16] Znidaric, M., Prosen, T. & Prelovsek, P. Many-body localization in the Heisenberg XXZ magnet in a random field. *Phys. Rev. B* **77**, 064426 (2008).

[17] Srednicki, M. Chaos and quantum thermalization. *Phys. Rev. E* **50**, 888 (1994).

[18] Deutsch, J. M. Quantum statistical mechanics in a closed system. *Phys. Rev. A* **43**, 2046 (1991).

[19] Rigol, M., Dunjko, V. & Olshanii, M. Thermalization and its mechanism for generic isolated quantum systems. *Nature* **452**, 854 (2008).

[20] Fannes, M., Nachtergaele, B. & Werner, R. F. Finitely correlated states on quantum spin chains. *Commun. Math. Phys.* **144**, 443 (1992).

[21] Perez-Garcia, D., Verstraete, F., Wolf, M. M. & Cirac, J. I. Matrix product state representations. *Quant. Inf. Comp.* **5**&**6**, 401 (2006).

[22] Hastings, M. B. & Koma, T. Spectral gap and exponential decay of correlations. *Commun. Math. Phys.* **265**, 781 (2006).

[23] Nachtergaele, B. & Sims, R. Lieb-Robinson bounds and the exponential clustering theorem. *Commun. Math. Phys.* **265**, 119 (2006).

[24] Eisert, J. & Osborne, T. J. General entanglement scaling laws from time evolution. *Phys. Rev. Lett* **97**, 150404 (2006).

[25] Bravyi, S., Hastings, M. B. & Verstraete, F. Lieb-Robinson bounds and the generation of correlations and topological quantum order. *Phys. Rev. Lett.* **97**, 050401 (2006).

[26] Nachtergaele, B., Ogata, Y. & Sims, R. Propagation of correlations in quantum lattice systems. *J. Stat. Phys.* **124**, 1 (2006).

[27] Linden, N., Popescu, S., Short, A. J. & Winter, A. Quantum mechanical evolution towards thermal equilibrium. *Phys. Rev. E* **79**, 061103 (2009).

[28] Hamza, E., Sims, R. & Stolz, G. Dynamical localization in disordered quantum spin systems. *Commun. Math. Phys.* **315**, 215 (2012).

[29] Hastings, M. B. Quasi-adiabatic continuation for disordered systems: Applications to correlations, Lieb-Schultz-Mattis, and Hall conductance. ArXiv:1001.5280.

[30] Keating, J. P., Linden, N. & Wells, H. J. Spectra and eigenstates of spin chain Hamiltonians. ArXiv:1403.1121.

[31] Bachmann, S., Michalakis, S., Nachtergaele, B. & Sims, R. Automorphic equivalence within gapped phases of quantum lattice systems. *Commun. Math. Phys.* **309**, 835 (2012).

[32] Brandao, F. G. S. L. & Horodecki, M. An area law for entanglement from exponential decay of correlations. *Nat. Phys.* **9**, 721 (2013).

[33] Verstraete, F. & Cirac, J. I. Matrix product states represent ground states faithfully. *Phys. Rev. B* **73**, 094423 (2006).

[34] Chandran, A., Kim, I. H., Vidal, G. & Abanin, D. A. Constructing local integrals of motion in the many-body localised phase. ArXiv:1407.8480.

[35] Schuch, N., Cirac, J. I. & Verstraete, F. The computational difficulty of finding MPS ground states. *Phys. Rev. Lett.* **100**, 250501 (2008).

[36] Them, K. Towards experimental tests and applications of Lieb-Robinson bounds. *Phys. Rev. A* **89**, 022126 (2014).

[37] Barthel, T. & Kliesch, M. Quasilocality and efficient simulation of Markovian quantum dynamics. *Phys. Rev. Lett.* **108**, 230504 (2012).

[38] Kliesch, M., Gogolin, C. & Eisert, J. Lieb-Robinson bounds and the simulation of time evolution of local observables in lattice systems. ArXiv:1306.0716.

[39] Nachtergaele, B. & Sims, R. Lieb-Robinson bounds in quantum many-body physics. In Sims, R. & Ueltschi, D. (eds.) *Entropy and the Quantum*, vol. 529 of *Contemporary Mathematics* (AMS, 2010).

[40] Schollwoeck, U. The density-matrix renormalization group in the age of matrix product states. *Ann. Phys.* **326**, 96 (2011).

APPENDIX

A. Lieb-Robinson bounds

Zero velocity Lieb-Robinson bounds are a way to rigorously capture the suppression of transport. Since there is no comprehensive discussion of such bounds available in the literature, we precisely discuss the relationships between different possible definitions here. Depending on the precise notion of transport adapted, the formulation of the bound slightly varies. The strongest variant is that information cannot be send through a system by locally changing the Hamiltonian and is rigorously captured by our definition of strong dynamical localisation (Def. 1).

Definition 3. *(Zero velocity Lieb-Robinson bounds: Truncated Hamiltonian) A Hamiltonian is said to satisfy zero velocity Lieb-Robinson bounds in the truncated Hamiltonian formulation, iff*

$$\|A(t) - e^{itH_l}Ae^{-itH_l}\| \leq \min(t,1)Ce^{-\mu l}. \qquad (4)$$

Here $A(t)$ is an observable in the Heisenberg picture, H^l the truncated Hamiltonian including all Hamiltonian terms within a distance l of the support of A and $C, \mu > 0$ are constants.

As in this whole work, $\|A\| = 1 = \|B\|$ was chosen for convenience. In constrast to the definition in the main text (Def. 1), we here include a factor $\min(t,1)$, which is stronger than what we need for our proof of Theorem 1 part **a**. However, since it seems reasonable to assume that the left hand side of Eq. (4) initially grows continuously with t and this way of stating the bound allows for an easier connection between the different Lieb-Robinson versions, we decided to include the factor here. Indeed the strongest standard Lieb-Robinson bounds in nearest-neighbour systems also include such a factor [36].

Another possible and very physical way of stating that a system has no transport, is to say that excitations cannot spread through the system. Mathematically, this can be captured by demanding that the action of a local unitary cannot be detected far away, even for long times.

Definition 4. *(Zero velocity Lieb-Robinson bounds: Excitations) A Hamiltonian is said to satisfy zero velocity Lieb-Robinson bounds in the excitation formulation, iff for any locally supported Hermitian operator G with $\|G\| = 1$*

$$| \langle\psi| A(t) |\psi\rangle - \langle\psi| e^{isG}A(t)e^{-isG} |\psi\rangle |$$
$$\leq s \min(t,1)C'e^{-\mu d(A,G)},$$

where $|\psi\rangle$ is an arbitrary state vector and $C', \mu > 0$ are t independent constants.

Most useful for mathematical manipulations is often the formulation in terms of a commutator [28, 29].

Definition 5. *(Zero velocity Lieb-Robinson bounds: Commutator) A Hamiltonian is said to satisfy zero velocity Lieb-Robinson bounds in the commutator formulation, iff*

$$\|[A(t), B]\| \leq \min(t,1)C''e^{-\mu d(A,B)},$$

with suitable t independent constants $C'', \mu > 0$.

Interestingly, these definitions can be closely related, similar to the case of standard Lieb-Robinson bounds [37–39].

Lemma 6. *(Connecting zero velocity Lieb-Robinson bounds) The different formulations of zero velocity Lieb-Robinson bounds are related in the following way*

Truncation (Def. 3) $\Rightarrow$ Commutator (Def. 5)
Excitations (Def. 4) $\Leftrightarrow$ Commutator (Def. 5),

with suitable constants C, C', C''.

Proof. Going from the truncation formulation (Def. 3) to the commutator formulation (Def. 5) is straightforward

$$\|[A(t), B]\| \leq \|[A(t) - e^{itH_l}Ae^{-itH_l}, B]\|$$
$$+ \|[e^{itH_l}Ae^{-itH_l}, B]\|$$
$$\leq \min(t,1)C''e^{-\mu d(A,B)},$$

where $l = d(A, B)$ has been chosen to make the second commutator zero.

To show the one implication of the second part of the lemma, going from Def. 4 to Def. 5, we rewrite the excitations formulation as

$$\left| \langle\psi| \sum_{k>0} \frac{s^k i^k}{k!} \tau_G^k(A(t)) |\psi\rangle \right| \leq s \min(t,1)C'e^{-\mu d(A,G)},$$

where $\tau_G(A) = [G, A]$ is the commutator and τ_G^k denotes multiple applications of it. Dividing by s and taking the limit $s \to 0$ yields

$$| \langle\psi| [A(t), G] |\psi\rangle | \leq \min(t,1)C'e^{-\mu d(A,G)} .$$

Since the bound holds for arbitrary Hermitian observables G and arbitrary state vectors $|\psi\rangle$, this concludes the proof.

To show the converse direction, going from Def. 5 to Def. 4, we express the left hand side of Def. 4 as an integral

$$| \langle\psi| A(t) |\psi\rangle - \langle\psi| e^{isG}A(t)e^{-isG} |\psi\rangle |$$
$$= | \langle\psi| \int_0^s dr e^{irG}[G, A(t)]e^{-irG} |\psi\rangle |$$
$$\leq \int_0^s dr \|[G, A(t)]\|$$
$$\leq s \min(t,1)C''e^{-\mu d(A,B)},$$

where we used Def. 5 in the final step. $\square$

Using common tools [39], one can also use the commutator formulation to arrive at estimates similar to the truncation formulation, but with a slightly weaker pre-factor, depending on the time t.

B. Energy filtering

A tool that plays an important role in this work is energy-filtering with respect to a suitable filter function. Energy-filtering of a local observable is defined as follows

$$I_f^H(A) = \int_{-\infty}^{\infty} dt f(t) A(t) .$$

Here $f : \mathbb{R} \to \mathbb{R}^+$ is a so-called filter function, usually taken as a C^∞-function. Here, $A(t) = e^{itH} A e^{-itH}$ refers to time evolution under the full Hamiltonian H, but we will later also consider filters with respect to truncated Hamiltonians, such as $I_f^{H_A}$. Energy-filtering allows to alter the matrix elements of a local observable in the eigenbasis of a Hamiltonian, while still keeping some form of locality [29]. We will work with two types of filter functions. Gaussian filter functions in particular provide a good compromise between locality in Fourier space and decay behaviour in real time and hence allow us to pick out narrow energy windows, while still preserving the approximate locality of the observable.

Definition & Lemma 7 (Gaussian filters). *A Gaussian filter is defined as*

$$I_\alpha^H(A) := \sqrt{\frac{\alpha}{\pi}} \int_{-\infty}^{\infty} dt e^{-\alpha t^2} A(t),$$

where $\alpha > 0$ defines the sharpness of the filter. The matrix elements in the eigenbasis of the Hamiltonian fulfil

$$\langle r| I_f^H(A) |s\rangle = \langle r| A |s\rangle e^{-(E_s - E_r)^2/(4\alpha)}.$$

For strongly localizing systems (Def. 1), local observables filtered with a Gaussian filter still remain approximately local in the sense that

$$\|I_\alpha^H(A) - I_\alpha^{H_l}(A)\| \le c_{\text{loc}} e^{-l}.$$

For systems with a mobility edge (Def. 2), we have

$$|\langle k| [I_f(A), B] |k\rangle| \le c_{\text{mob}} e^{-d(A,B)}.$$

Proof. The Gaussian suppression of off-diagonal elements readily follows from the definition

$$\langle r| I_\alpha(A) |s\rangle := \sqrt{\frac{\alpha}{\pi}} \int dt e^{-\alpha t^2} \langle r| A(t) |s\rangle$$

$$= \langle r| A |s\rangle \sqrt{\frac{\alpha}{\pi}} \int dt e^{-\alpha t^2} e^{it(E_r - E_s)},$$

and the fact that the Fourier transform of a Gaussian is again a Gaussian. Assuming strong dynamical localisation, deriving locality is straightforward,

$$\|I_\alpha^H(A) - I_\alpha^{H_A}(A)\|$$
$$= \left\| \sqrt{\frac{\alpha}{\pi}} \int_{-\infty}^{\infty} dt e^{-\alpha t^2} \left(A(t) - e^{itH_A} A e^{-itH_A} \right) \right\|$$
$$\le c_{\text{loc}} e^{-\mu l},$$

where have used that the Gaussian filter is normalised. The case with a mobility edge can be shown in the same way. □
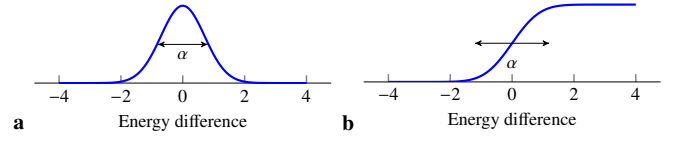

FIG. 2. Presented is a schematic sketch of the energy filtering under a Gaussian (**a**) and a high-pass (**b**) filter and the role of the sharpness α.

Another important filter is the high-pass filter, which blocks large negative energy differences.

Definition & Lemma 8 (High-pass filters). *A high-pass filter is defined by*

$$\Gamma_\alpha(A) := \lim_{\varepsilon \to 0} \frac{i}{2\pi} \int_{-\infty}^{\infty} dt \frac{e^{-\alpha t^2}}{t + i\varepsilon} A(t).$$

Here $\alpha > 0$ describes the sharpness of the filter. The matrix elements can be bounded for any $|E_s - E_r| \ge \sigma \ge 0$ by

$$\frac{\langle r| \Gamma_\alpha(A) |s\rangle}{\langle r| A |s\rangle} = \begin{cases} \frac{1}{2} e^{-\sigma^2/(4\alpha)} & \text{for } E_r \ge E_s \\ 1 - \frac{1}{2} e^{-\sigma^2/(4\alpha)} & \text{for } E_r \le E_s \end{cases} .$$

Local observables remain approximately local under a high-pass filter, in the sense that

$$| \langle k| [\Gamma_\alpha(A), B] |k\rangle | \le \frac{e^{-\mu d(A,B)}}{2\pi} \left(4 + \ln \frac{\pi}{4\alpha} \right).$$

Proof. Calculating the off-diagonal elements of such a high-pass filter relies on the subsequent bound on the error function of a Gaussian random variable proven in Ref. [22], stated here as Lemma 9. Making use of the identity proven in Ref. [22]

$$\lim_{\varepsilon \to 0} \frac{i}{2\pi} \int_{-\infty}^{\infty} \frac{e^{-\alpha t^2} e^{-i\Delta E t}}{t + i\varepsilon} dt$$
$$= \frac{1}{2\pi} \sqrt{\frac{\pi}{\alpha}} \int_{-\infty}^{0} e^{-(\omega + \Delta E)^2/(4\alpha)} d\omega,$$

the bounds for the matrix elements of $\Gamma_\alpha(A)$ follow from Lemma 9. The locality statement can be shown by splitting the integral into three parts [28] and using our assumption of dynamical localisation

$$\langle k| [\Gamma_\alpha(A), B] |k\rangle \le \frac{e^{-\mu d(A,B)}}{2\pi} \times$$
$$\left(\int_{|t| \le 1} e^{-\alpha t^2} dt + \int_{1 \le |t| \le \lambda} \frac{e^{-\alpha t^2}}{t} dt + \int_{|t| \ge \lambda} \frac{e^{-\alpha t^2}}{t} dt \right)$$
$$\le \frac{e^{-\mu d(A,B)}}{\pi} \left(1 + \ln \lambda + \frac{1}{2\lambda} \sqrt{\frac{\pi}{\alpha}} \right).$$

Here estimating the first term used the $\min(1, t)$ factor included in the definition of a mobility edge (Def. 2). Choosing

$$\lambda = \frac{\sqrt{\pi}}{2\sqrt{\alpha}}$$

concludes the proof. □

Lemma 9 (Hastings-Koma). *Let $E \in \mathbb{R}$, $\alpha > 0$ then for all $\gamma > 0$ with $E \leq -\gamma$*

$$\frac{1}{2\pi}\sqrt{\frac{\pi}{\alpha}}\int_{-\infty}^{0} e^{-\frac{(\omega+E)^2}{4\alpha}}\,\mathrm{d}\omega \leq \frac{1}{2}e^{-\gamma^2/(4\alpha)}$$

and for all $\gamma > 0$ with $E \geq \gamma$

$$\left|\frac{1}{2\pi}\sqrt{\frac{\pi}{\alpha}}\int_{-\infty}^{0} e^{-\frac{(\omega+E)^2}{4\alpha}}\,\mathrm{d}\omega - 1\right| \leq \frac{1}{2}e^{-\gamma^2/(4\alpha)}.$$

For later use, we give the Gaussian and high pass filters of sharpness α the symbols I_α^H and Γ_α^H, respectively. In case the local observable is filtered with the full system Hamiltonian, we will often omit the H. It is an interesting insight that the locality structure of a Gaussian filter is independent of its sharpness α for dynamically localising systems, while it still depends on α for a high-pass filter. It is precisely this property of Gaussian filters that will allow us to prove part **b** of Theorem 1.

C. Local constants of motion

In this appendix, we briefly comment on how local constants of motion could be constructed using energy filtering. A natural candidate for local constants of motion are the energy-filtered versions of the local Hamiltonian terms

$$M_j := I_\alpha^H(h_j).$$

Since they act non-trivially on most eigenstates, it is clear that they are retained by the energy-filtering without being averaged to zero and thus indeed provide useful quantum numbers. What is more, strong dynamical localisation implies that they will stay approximately local. This construction can be optimised by choosing suitable filter functions, allowing for an interpolation between the locality of the observables and making them completely constant in time. In particular, taking the time-average is equivalent to choosing a constant filter function [34] leading to quantities that are completely preserved in time, while they will usually not be strictly local anymore. Naturally, this approach is not limited to taking the local Hamiltonian terms and in principle any local observable that does not vanish after applying the energy filter will provide useful quantum numbers. Specifically, polynomials of the local Hamiltonian terms can equally well be used. In systems with a mobility edge, where not all transport is suppressed, this construction is still possible, but the resulting constants of motion might no longer be approximately local.

Since local constants of motion by definition commute with the Hamiltonian, their local eigenbasis is compatible with the global energy eigenbasis of the Hamiltonian. Thus, they could be used to directly construct local states in a matrix-product state form, as long as enough constants of motion exist in order to use their quantum numbers to divide the full Hilbert space into suitable small fractions.
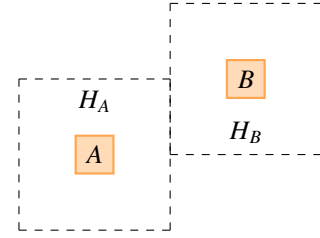

FIG. 3. This figure shows the Hamiltonian decomposition used in the proof for part **a** of our main theorem.

D. Proof of Theorem 1 part a

In this appendix, we will formulate the details of the proof of Theorem 1 part **a**, which shows that strong dynamical localisation implies exponential clustering of all eigenvectors. For this, we will rely on the following lemma.

Lemma 10 (Decoupled energy filtering). *Under the assumption of locally independent gaps (**AII**), energy-filtering of two observables can be factorized into local energy filters*

$$\|I_\alpha^{H_A+H_B}(AB) - I_\alpha^{H_A}(A)I_\alpha^{H_B}(B)\| \leq 2^{4N+1}e^{-\xi^2/(4\alpha)},$$

with $\xi = \min\{\eta, \sqrt{2}\tilde{\gamma}\}$. Here H_A and H_B are chosen to include all Hamiltonian terms within distance $d(A,B)/2$ of the support of A and B respectively (See Fig. 3).

Proof. To show that instead of applying an energy filter to AB, we can also apply it to the observables individually, we need to use that the eigenvalues of $H_A + H_B$ are disconnected on the two regions and we can thus label them by two different quantum numbers a, b. We hence get

$$\left\|I_\alpha^{H_A+H_B}(AB) - \sum_{a,b}|a,b\rangle\langle a,b|\,AB\,|a,b\rangle\langle a,b|\right.$$
$$\left. + \sum_{a,b}|a,b\rangle\langle a,b|\,AB\,|a,b\rangle\langle a,b| - I_\alpha^{H_A}(A)I_\alpha^{H_B}(B)\right\|.$$

We will use the triangle inequality and proceed to show that both energy filters give only the diagonal entires up to a small error. Starting with the first term gives the following estimate,

$$\|I_\alpha^{H_A}(A)I_\alpha^{H_B}(B) - \sum_{a,b}|a,b\rangle\langle a,b|\,AB\,|a,b\rangle\langle a,b|\,\|$$
$$=\left\|\sum_{a,b}\sum_{a'\neq a,b'\neq b}|a,b\rangle\langle a,b|\,AB\,|a',b'\rangle\langle a',b'|\right.$$
$$\left.\exp\left(-\frac{(E_a - E_{a'} + E_b - E_{b'})^2}{4\alpha}\right)\right\|$$
$$=\sum_{a,b}\sum_{a'\neq a,b'\neq b}\left\||a,b\rangle\langle a,b|\,AB\,|a',b'\rangle\langle a',b'|\right.$$
$$\left.\exp\left(-\frac{(E_a - E_{a'} + E_b - E_{b'})^2}{4\alpha}\right)\right\|$$
$$=2^{4N}e^{-\eta^2/(4\alpha)},$$

where we assumed locally independent gaps (**AII**). The second term yields the following estimate,

$$\| I_\alpha^{H_A}(A) I_\alpha^{H_B}(B) - \sum_{a,b} |a,b\rangle\langle a,b| \, AB \, |a,b\rangle\langle a,b| \, \|$$

$$= \Big\| \sum_{a,b} \sum_{a'\neq a, b'\neq b} |a,b\rangle\langle a,b| \, AB \, |a',b'\rangle\langle a',b'|$$

$$\exp\left(-\frac{(E_a - E_{a'})^2}{4\alpha}\right) e^{-\frac{(E_b - E_{b'})^2}{4\alpha}} \Big\|$$

$$= \Big\| \sum_{a,b} |a,b\rangle\langle a,b| \, AB \widetilde{D}_{\tilde\gamma}^a \widetilde{D}_{\tilde\gamma}^b \Big\| \leq 2^{2N} e^{-2\tilde\gamma^2/(4\alpha)}.$$

Here, $\widetilde{D}_{\tilde\gamma}^a$ is a diagonal matrix with entries $e^{-(E_a - E_a')^2/(4\alpha)}$ for $a \neq a'$ and 0 otherwise and we assumed locally independent gaps (**AII**). $\qquad\square$

Finally let us briefly comment that the local independence of gaps can alternatively be ensured by demanding that the joint Hamiltonian $H = H_A + H_B$ supported on two disjoint rectangular regions has non-degenerate energies, since these energies are all possible sums of the local energies $E_a + E_b$.

Proof of Theorem 1 We are now in the position to prove part **a** of our main theorem.

Theorem 1 (Clustering of correlations of eigenvectors). *The dynamical properties of a local Hamiltonian imply clustering correlations of its eigenvectors in the following way.*

    *a If the Hamiltonian shows strong dynamical localisation and its spectrum fulfils assumptions **AI** and **AII** then all its eigenvectors have exponentially clustering correlations*

$$|\langle k| AB |k\rangle - \langle k| A |k\rangle\langle k| B |k\rangle| \leq 4 c_{\mathrm{loc}} e^{-\mu \mathrm{d}(A,B)/2}.$$

Proof. The proof uses the local Hamiltonians H_A and H_B (See Fig. 3) and the following estimates

$$\langle k| AB |k\rangle = \langle k| I_\alpha^H(AB) |k\rangle$$
$$\overset{=}{I} \langle k| I_\alpha^{H_A + H_B}(AB) |k\rangle + \Omega(e^{-\mu l})$$
$$\overset{=}{II} \langle k| I_\alpha^{H_A}(A) I_\alpha^{H_B}(B) |k\rangle + \Omega(e^{-\mu l})$$
$$\overset{=}{III} \langle k| I_\alpha^{H}(A) I_\alpha^{H}(B) |k\rangle + \Omega(e^{-\mu l})$$
$$\overset{=}{IV} \langle k| A |k\rangle\langle k| B |k\rangle + \Omega(e^{-\mu l}).$$

In each step, an error term of the form $C_I e^{-\mu l}$ with C_I to C_{IV} will be introduced. The constant C appearing in the main theorem will simply be the sum of them. Constants C_I, C_{III}, C_{IV} directly follow from the properties of the applied Gaussian filter (Def. 7) and constant C_{II} is derived in Lemma 10

$$C_I = 2 c_{\mathrm{loc}},$$
$$C_{II} = 2^{4N+1} e^{-\xi^2/(4\alpha)} e^{\mu l},$$
$$C_{III} = c_{\mathrm{loc}},$$
$$C_{IV} = e^{-\gamma^2/(4\alpha)} e^{\mu l},$$
$$C := C_I + C_{II} + C_{III} + C_{IV}.$$

The constants C_{II} and C_{IV} can be chosen arbitrarily small, by picking a sharp enough filter function, meaning a sufficiently small α. In particular, we can choose $C_{II}, C_{IV} \leq c_{\mathrm{loc}}/2$, which yields

$$C = 4 c_{\mathrm{loc}}.$$

Choosing $l = \mathrm{d}(A,B)/2$ concludes the proof. $\qquad\square$

E. Proof of Theorem 1 part b

In this appendix, we will show the second part of Theorem 1 based on a mobility edge. For convenience, we start by repeating Definition 2 of a mobility gap.

Definition 2 (Mobility edge). *A Hamiltonian H is said to have a mobility edge at energy E_{mob} iff its time evolution satisfies for all times t, $\forall \rho \in \mathrm{span}\{|l\rangle\langle k| : E_l, E_k \leq E_{\mathrm{mob}}\}$*

$$|\mathrm{tr}\left(\rho[A(t), B]\right)| \leq \min(t, 1) c_{\mathrm{mob}} e^{-\mu \mathrm{d}(A,B)},$$

where c_{mob} is a t independent constant. That is, all transport is suppressed for states supported only on the low-energy sector below energy E_{mob}.

The first step for our proof will be to show that each eigenvectors only contributes to the correlation function by a term that is exponentially suppressed with the distance of the two observables. For this, we need the localisation assumption only for the eigenvector $|k\rangle$.

Lemma 11. *Let A, B be local observables and $|k\rangle$ a weakly localised eigenvector of H, i.e. satisfying*

$$\sup_{t\in[0,\infty)} |\langle k| [A(t), B] |k\rangle| \leq c_{\mathrm{mob}} e^{-\mu \mathrm{d}(A,B)}.$$

Then the contribution of any eigenvector $|l\rangle, l \neq k$, to the correlation function will be exponentially suppressed, in the sense that

$$|\langle k| A |l\rangle\langle l| B |k\rangle| \leq 2 c_{\mathrm{mob}} e^{-\mu \mathrm{d}(A,B)}.$$

Proof. The proof of this lemma again uses an energy filter with a Gaussian filter function (Def. 7) and works with $\langle k| B |k\rangle = 0$, which can always be achieved by using a shifted observable $\tilde{B} = B - \langle k| B |k\rangle$. For this proof the filter function will be multiplied by a complex factor $e^{it(E_k - E_l)}$ such that its Fourier transform approximately suppresses all transitions except the one from level k to level l. In a mild variant of the above filter function, we define $f : \mathbb{R} \to \mathbb{R}^+$ and the corresponding filter as

$$f(t) := \sqrt{\frac{\alpha}{\pi}} e^{it(E_k - E_l)} e^{-\alpha t^2},$$

$$I_f(A) := \int \mathrm{d}t f(t) A(t),$$

$$\langle r| I_f(A) |s\rangle = \langle r| A |s\rangle \, e^{-((E_s - E_r) - (E_l - E_k))^2/(4\alpha)}.$$

With this energy filter, we can proceed to prove the lemma.

$$|\langle k| A |l\rangle\langle l| B |k\rangle|$$

$$= \left| \langle k| I_f(A)B |k\rangle - \sum_{m\neq l} \langle k| A |m\rangle\langle m| B |k\rangle \, e^{-(E_m - E_l)^2/(4\alpha)} \right|$$

$$\leq |\langle k| [I_f(A), B] |k\rangle| + |\langle k| B I_f(A) |k\rangle| + 2^N e^{-\gamma^2/(4\alpha)}$$

$$\leq c_{\mathrm{mob}} e^{-d(A,B)} + 2^N \left(e^{-\gamma^2/(4\alpha)} + e^{-\zeta^2/(4\alpha)} \right).$$

Here, $\gamma > 0$ refers to the smallest gap and $\zeta > 0$ is the smallest degeneracy of the gaps for fixed E_k as defined in the main text in **AI** and **AIII** and in the last step we used weak dynamical localisation of the eigenvector $|k\rangle$. We can now conclude the proof by choosing α small enough such that

$$2^N \left(e^{-\gamma^2/(4\alpha)} + e^{-\zeta^2/(4\alpha)} \right) \leq c_{\mathrm{mob}} e^{-\mu \mathrm{d}(A,B)},$$

which means that we pick a filter function that is narrow enough in Fourier space. Setting $\xi := \min(\gamma, \zeta)$ it would even be enough to choose

$$\frac{\xi^2}{4} \left(\ln \frac{2^{2N+1}}{c_{\mathrm{mob}}} \right)^{-1} \geq \alpha,$$

independently of $\mathrm{d}(A, B)$. $\qquad\square$

With this, we are ready to prove part **b** of Theorem 1.

Theorem 1 (Clustering of correlations of eigenvectors). *The dynamical properties of a local Hamiltonian imply clustering correlations of its eigenvectors in the following way.*

> **b** *If the Hamiltonian has a mobility edge at energy E_{mob} and its spectrum fulfils assumptions **AI** and **AIII**, then all eigenvectors up to that energy E_{mob} cluster exponentially*
>
> $$|\langle k| AB |k\rangle - \langle k| A |k\rangle\langle k| B |k\rangle|$$
> $$\leq \left(12\pi\Theta(E_k + \kappa)c_{\mathrm{mob}} + \ln \frac{\pi\mu\mathrm{d}(A,B)e^{4+2\pi}}{\kappa^2} \right) \frac{e^{-\mu\mathrm{d}(A,B)}}{2\pi},$$
>
> *where $\Theta(E)$ is the number of states at energy E and κ is a constant that can be chosen arbitrarily to optimise the bound.*

*Proof of Theorem 1 part **b***. The proof runs along and builds upon the lines of thought of both Ref. [22] and of [28], and generalises both. The basic idea is again to start from the correlation function and to transform it into an expression depending on the commutator. For this, we fix a constant $\kappa > 0$ and define

$$P = \sum_{E_l \leq E_k + \kappa} |l\rangle\langle l|$$

to be the projector onto the subspace of energies smaller than $E_k + \kappa$ and write $P^\perp$ for $\mathbb{1} - P$. Decomposing the identity

with respect to P and $P^\perp$ and using a high-pass filter Γ_α for a suitable $\alpha > 0$ (Def. 8), we separate the correlator into terms

$$\langle k| AB |k\rangle = \langle k| (A - \Gamma_\alpha(A))PB |k\rangle$$
$$+ \langle k| (A - \Gamma_\alpha(A))P^\perp B |k\rangle$$
$$+ \langle k| BP\Gamma_\alpha(A) |k\rangle + \langle k| BP^\perp\Gamma_\alpha(A) |k\rangle$$
$$+ \langle k| [\Gamma_\alpha(A), B] |k\rangle .$$

Lets consider the first term on the right-hand side of Eq. (5) that contains P and expand the projector in the eigenbasis of H, which gives

$$| \langle k| (A - \Gamma_\alpha(A))PB |k\rangle |$$
$$\leq \sum_{E_l \leq E_k + \kappa} | \langle k| (A - \Gamma_\alpha(A)) |l\rangle\langle l| B |k\rangle |$$
$$\leq 2 \sum_{E_l \leq E_k + \kappa} | \langle k| A |l\rangle\langle l| B |k\rangle |.$$

Here, we have used that all matrix elements decrease under a high-pass filter (Lemma 8) for all $\alpha > 0$. As before, we will set $\langle k| B |k\rangle = 0$ w.l.o.g. and then proceed by bounding all the terms in the sum individually. Using Lemma 11 yields

$$| \langle k| (A - \Gamma_\alpha(A))PB |k\rangle | \leq 4\Theta(E_k + \kappa)c_{\mathrm{mob}} e^{-\mu\mathrm{d}(A,B)},$$

where $\Theta(E_k + \kappa)$ simply is the number of eigenvectors contained in P. The other term in Eq. (5) containing P can be bounded analogously, yielding

$$| \langle k| BP\Gamma_\alpha(A) |k\rangle | \leq 2\Theta(E_k + \kappa)c_{\mathrm{mob}} e^{-\mu\mathrm{d}(A,B)}.$$

The terms containing $P^\perp$ in Eq. (5) are bounded using the explicit form of the matrix elements of a high-pass filter (Lemma 8), which delivers

$$\langle k| BP^\perp\Gamma_\alpha(A) |k\rangle \leq \sum_{E_l > E_k + \kappa} \langle k| B |m\rangle\langle m| DA |k\rangle ,$$

where D is a diagonal matrix whose entries are bounded by $e^{-\kappa^2/(4\alpha)}/2$. A simple norm estimate concludes the estimation of this term

$$\langle k| BP^\perp\Gamma_\alpha(A) |k\rangle \leq \|D\| \leq \frac{e^{-\kappa^2/(4\alpha)}}{2}.$$

The other term containing $P^\perp$ can be estimated in the same way

$$\langle k| (A - \Gamma_\alpha(A)) P^\perp\Gamma_\alpha(A) |k\rangle \leq \frac{e^{-\kappa^2/(4\alpha)}}{2}.$$

The commutator term in Eq. (5) is bounded by using locality of a high-pass filter (Lemma 8). We are still free to choose a value for α in the high-pass filter. Taking

$$\alpha = \frac{\kappa^2}{4\mu\mathrm{d}(A, B)}$$

gives

$$\langle k | [\Gamma_\alpha(A), B] | k \rangle \leq \bigg(12\pi\Theta(E_k + \kappa)c_{\text{mob}}$$

$$+ 2\pi + 4 + \ln \frac{\pi\mu\mathrm{d}(A,B)}{\kappa^2} \bigg) \frac{e^{-\mu\mathrm{d}(A,B)}}{2\pi} \ .$$

This concludes the proof. $\qquad\square$

F. Proof of Corollary 2

It follows from the above results that in the case of 1D systems, the corresponding eigenstates satisfy an area law for a suitable entanglement entropy, which in turn implies that they can be well approximated by a matrix-product state of a low bond dimension. We will bound the bond dimension by using Theorem 1 from Ref. [32]. For convenience, we repeat our Corollary 2 from the main text.

Corollary 2 (Area laws and matrix-product states). *An eigenvector $|k\rangle$ of a localising Hamiltonian can be written as an MPS with fidelity $|\langle k|\mathrm{MPS}\rangle| \geq 1 - \epsilon$, where the bond dimension for a sufficiently large system size is given as follows.*

  *a If the Hamiltonian shows strong dynamical localisation and its spectrum fulfils assumptions **AI** and **AII**, then the statement holds for all eigenvectors $|k\rangle$ and the approximation has a bond dimension*

$$D = C\,(N/\epsilon)^{\frac{16}{\mu\log_2 e}} \ ,$$

  for some constant $C > 0$.

  *b If the Hamiltonian has a mobility edge at energy E_{mob}, and its spectrum fulfils assumptions **AI** and **AIII**, then the statement holds for all eigenvectors below this energy E_{mob} and the bond dimension is given by*

$$D = \mathrm{poly}\left(\Theta(E_k + \kappa), N\right) \ ,$$

  for any fixed κ which enters in the precise form of the polynomial.

Proof. As a first step of the proof, we reformulate our clustering of correlation results (Theorem 1) in terms of a correlation length according to Ref. [32]. For this, we need a decay of the correlator of the form $2^{-\mathrm{d}(A,B)/\xi}\ \forall\mathrm{d}(A,B) > l_0$ and start by rewriting our bound as

$$|\langle k| AB |k\rangle - \langle k| A |k\rangle\langle k| B |k\rangle| \leq Ce^{-\mu\mathrm{d}(A,B)} \qquad (11)$$

$$= C2^{-\mu\log_2 e\,\mathrm{d}(A,B)}.$$

Next we split the decaying term and use one part to get rid of the constant prefactor C and the other to preserve an exponential decay with a correlation length $\xi := 2/(\mu\log_2 e)$. This yields

$$|\langle k| AB |k\rangle - \langle k| A |k\rangle\langle k| B |k\rangle| \leq 2^{-\mathrm{d}(A,B)/\xi}$$
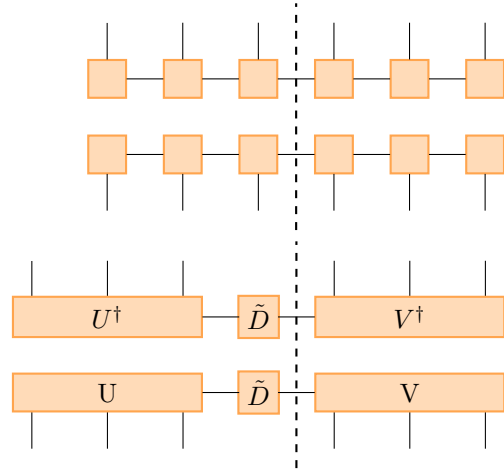$$\forall\,\mathrm{d}(A,B) \geq l_0,$$

FIG. 4. MPS description of an eigenstate $|E_k\rangle\langle E_k|$. Theorem 1 in Ref. [32] allows to bound the bond dimension across cuts by providing an upper bound to the smooth max entropy. The reduced state of the suitably normalised matrix product state is $\rho = U\tilde{D}^2 U^\dagger$.

where $l_0 = \xi/(\log_2 C)$. Following Ref. [32], we will use the exponential clustering to obtain a description of the eigenvector $|k\rangle$ in terms of MPS (see also Fig. 4). This is based on the so called smooth max entropy, which is defined as

$$H_{\max}^\delta(\rho_X) := \min_{\tilde{\rho}_X \in \mathcal{B}_\delta(\rho_X)} \log_2(\mathrm{rank}(\tilde{\rho}_X)) \ ,$$

with

$$\mathcal{B}_\delta(\rho_X) := \{\tilde{\rho}_X : D(\rho_X, \tilde{\rho}_X) < \delta\}$$

being a ball around ρ_X (see Appendix A of Ref. [32] for details). Here $D(A, B) := \|A - B\|_1/2$ denotes the trace distance. According to Theorem 1 in Ref. [32], we can obtain an upper bound for the smooth max entropy for any bipartite cut, as long as the system size is larger than $N \geq C_{\text{cut}}l_0/\xi$ with a constant $C_{\text{cut}} > 0$. Picking the approximation parameter in the smooth max entropy to be $\delta(l) = 2^{-l/(8\xi)}$, then Ref. [32] provides a bound of the form

$$H_{\max}^{\delta(l)} \leq c'l_0\xi^{c\xi} + l, \qquad (12)$$

for all $l \geq 8\xi$ with constants $c, c' > 0$. In the following, we will make Corollary 3 in Ref. [32] explicit, by deriving concrete bounds on the bond dimension of the matrix-product state chosen to approximate $|k\rangle$. Naturally, the first step for this is to express $|k\rangle$ as a matrix-product state vector with, a priori, exponentially large bond dimension (Fig. 4). For simplicity, we restrict the proof here to a linear system with open boundary conditions, but it can equivalently be reformulated for a periodic system on a ring [32]. The bond dimension of the matrix-product state will be bounded by truncating at each cut explicitly. For this, we start at one end of the chain and look at each cut separately. In each step, we apply a singular value decomposition (see also Fig. 4). Since we are truncating spectral values in each step, the positive operators will no

longer be normalised to unit trace and are states only up to normalisation. This results in a reduction on the left side of the cut of $\rho = U\tilde{D}^2 U^\dagger$, with U being unitary and $\tilde{D}$ diagonal. The goal is now to truncate the diagonal matrix $\tilde{D}$ to a fixed bond dimension D while creating only a small discarded weight [40]. Following Lemma 14 in Appendix B of Ref. [32], for any $\nu > 0$, we can choose the bond dimension as

$$D = 2^{H_{\max}^\nu(\rho)},$$

and create a discarded weight

$$\sum_{D+1}^{2^N} \lambda_k \leq 3\nu,$$

where λ_k are the eigenvalues of ρ or equivalently of $\tilde{D}^2$. Each time we create discarded weight, the fidelity with our original pure quantum state potentially is reduced by the discarded weight [33], upper bounded by 3ν. The truncation will result in a subnormalised state. Renormalising will, however, only increase the fidelity, allowing us to obtain a normalised MPS approximation. Thus, in total, the fidelity of our MPS approximation with bond dimension $D = 2^{H_{\max}^\nu(\rho)}$ is bounded by

$$|\langle \psi | \text{MPS} \rangle| \leq 1 - N 3\nu,$$

where N is the size of the linear 1D system. For a fixed global error $\epsilon > 0$, we therefore obtain an allowed local error

$$\nu = \frac{\epsilon}{3N} \ .$$

We now fix ν to take the role of $\delta(l)$ in (12). Plugging this bound into Eq. (12) and using $\delta(l) = 2^{-l/(8\xi)}$, we obtain

$$H_{\max}^{\frac{\epsilon}{3N}} \leq c' l_0 \xi^{c\xi} + 8\xi \log_2 \frac{3N}{\epsilon} \ ,$$

and thus obtain an approximation with

$$D = 2^{c' l_0 \xi^{c\xi}} \left(\frac{3N}{\epsilon} \right)^{8\xi}.$$

Coming back to our original clustering assumption in Eq. (11), this yields

$$D = C^{c' \xi^{c\xi+1}} \left(\frac{3N}{\epsilon} \right)^{8\xi},$$

with $\xi = 2/(\mu \log_2 e)$. Thus, we finally have a polynomial scaling in the constant $C > 0$ as well as the in the system size and in the inverse error of the approximation $1/\epsilon$. The rest of the proof follows directly by inserting the constants from Theorem 1. $\qquad\square$

G. Absence of thermalisation

In this subsection we make the link between dynamical localisation and the absence of thermalisation more explicit. On intuitive grounds, such a connection is much expected: if no transport happens, then the system will retain too much memory of the initial state for all times. This initial state dependence, then necessarily leads to an absence of thermalisation. Here, we will formulate the link between strong dynamical localisation in the sense of Definition 1 and the absence of thermalisation along the lines of Ref. [13] and consistent with Ref. [4].

To state this concisely, denote with ρ a Gibbs state of some inverse temperature $\beta > 0$. We denote with A both an observable as well as the support of A on the lattice. B is the region of distance no more than l from A. The following lemma shows that under the condition of strong localisation, there is an initial state ρ_0 different from ρ on A only, such that $\rho_0(t)$ will still be locally statistically distinguishable from ρ for all times $t > 0$ on the region B, in stark contrast to a presumed thermalisation. That is to say, in the sense of Ref. [13] the system retains too much of a memory of the initial condition to thermalise. Hence, the system must also violate the ETH.

Lemma 12 (Absence of thermalisation). *Let ρ be a Gibbs state. Then there exists a state ρ_0, different from ρ on A only, such that the trace distance of the time evolved state $\rho_0(t)$ and the Gibbs state ρ reduced to B is lower bounded by*

$$\|\rho_0(t)_B - \rho_B\|_1 \geq x - c_{\text{loc}} e^{-\mu l},$$

with $x := 2 - 2\lambda_{\min}(\rho_A)$, where ρ_A and ρ_B are the reductions to A and B, respectively.

Proof. For the distinguished region of the lattice, refer to ρ_A as its reduced state and $\rho_{\backslash A}$ its complement. Take as initial state $\rho_0 = \rho_{\backslash A} \otimes \xi_A$, where ξ_A is chosen such that

$$\|\rho_A - \xi_A\|_1 = 2 - 2\lambda_{\min}(\rho_A) = x.$$

Using that the trace norm is a unitarily invariant norm, it is clear that such a state ξ_A can always be found. Denote with A the observable satisfying $\|A\| \leq 1$ that saturates the trace distance. Then

$$x = \|\rho_A - \xi_A\|_1 = |\text{tr}(A(\rho_0 - \rho))|$$
$$= |\text{tr}(A(-t)(\rho_0(t) - \rho))|,$$

where we used that the Gibbs state is invariant under the time evolution of its Hamiltonian. Applying the triangle inequality, the condition of strong localisation, and the fact that $e^{itH_A^l} A e^{-itH_A^l}$ is an observable supported on B with operator norm upper bounded by unity, we get

$$x \leq |\text{tr}((e^{-itH} A e^{itH} - e^{-itH_A^l} A e^{itH_A^l})(\rho_0(t) - \rho))|$$
$$+ |\text{tr}((e^{-itH_A^l} A e^{itH_A^l})(\rho_0(t) - \rho))|$$
$$\leq \|(e^{-itH} A e^{itH} - e^{-itH_A^l} A e^{itH_A^l})\|$$
$$+ \|\rho_0(t)_B - \rho_B\|_1$$
$$\leq c_{\text{loc}} e^{-\mu l} + \|\rho_0(t)_B - \rho_B\|_1,$$

for l suitably chosen. From this the statement follows. $\qquad\square$