A simple model of many-body localization

Brian Swingle

Department of Physics, Harvard University, Cambridge MA 02138

(Dated: May 7, 2014)

We study a simple and tractable model of many-body localization. The main idea is to take a renormalization group perspective in which local entanglement is removed to reach a product state. The model is built from a random local unitary which implements a real space renormalization procedure and a fixed point Hamiltonian with random exponentially decaying interactions. We prove that every energy eigenstate is localized, that energy is not transported, and argue that despite being fine tuned, the model is stable to perturbations. We also show that every energy eigenstate obeys an area law for entanglement entropy and we consider the dynamics of entanglement entropy under perturbations. In the case of extensive perturbations we recover a logarithmic growth of entanglement observed in recent numerical simulations.

I. INTRODUCTION

Quantum systems can display the phenomenon of localization [1] where energy and charge are not transported under the dynamics of the system. Localization was first understood in the context of single particle physics with static disorder [1], but in many cases the inclusion of interactions or the coupling to a heat bath interrupts the localization physics and restores transport [2–4]. However, isolated local quantum systems may also enter a many-body localized phase where localization persists even with interactions [4]. Indeed, a many-body localized phase can exist when a local quantum system fails to thermalize and hence cannot serve as its own heat bath. Note also that there can in principle be many kinds of many-body localized phases, but here we focus exclusively on a many-body localized phase in which all energy eigenstates are localized [5]. Much interest in this topic has been sparked by a beautiful series of papers including Refs. [4–7]. Our particular interest is in understanding the interplay between entanglement and renormalization in a many-body localized phase.

The original work of Anderson suggested that localization could persist in the presence of interactions [1]. Basko et al. gave a picture of localization in Fock space as the physical basis for many-body localization [4]. Ref. [5] suggested that in systems with a finite Hilbert space, all energy eigenstates could be localized in a many-body localized phase. The phenomenology of entanglement dynamics in many-body localized phases has been explored in Refs. [8,9]. Many-body localization in the context of driven systems was investigated in Refs. [10,11]. Other interesting investigations include Refs. [12,15]. Independent of the present work [1], the dynamics of entanglement have been addressed in Refs. [16,18]. Although our model was developed independently of Refs. [16,18], it is easily seen that our fixed point Hamiltonian incorporates the dephasing dynamics first introduced in Ref. [16,17]. Finally, as this work was being finished, we learned of Ref. [19] where the idea of a local unitary is also used to elucidate the physics of localization. Ref. [19] also contains an extensive numerical analysis, so our model complements their analysis. To summarize, the physics we wish to describe is the existence of localized excited states, an entanglement area law for excited states, and the slow growth of entanglement under perturbations.

To this end we present a simple model of a many-body localized phase that can be straightforwardly analyzed. We show that our model has the following features. All many-body energy eigenstates are localized and obey an area law for entanglement entropy (defined below). Entanglement does not spread under a localized perturbation up to exponentially small corrections. Entanglement entropy displays a slow logarithmic growth for large regions at large times in the case of an extensive perturbation. We also argue that the many-body energy spectrum has the expected exponentially small level spacing, that energy is not transported, and that the system is stable to small local perturbations. Part of this stability argument is similar to that of Ref. [4]. We focus on entanglement because of the key role it has played in numerical studies and because it suggests a fruitful way to simulate many-body localized systems (using tensor networks [20,22] and DMRG [23,24]). Of course, in real experimental systems, various correlation functions will be much more accessible, but these can also be straightforwardly analyzed in our model.

Our perspective is strongly informed by renormalization group (RG) ideas and especially by the idea of real space renormalization using local unitaries (see for example Ref. [25]). A unitary $u$ is called local if it has the form $u = e^{-ik}$ for a local hermitian operator $k$ of bounded norm (not growing with system size). Local unitaries are of interest because they can be simply approximated by discrete circuits of few-spin unitary operators using the Trotter product formula. The key feature of such circuits in the case of local unitaries is that they have finite depth, that is they consist of only finitely many (not growing with system size) layers. Hence such circuits preserve the locality properties of states and operators on which they act.

Much work has focused on the idea that certain such bounded depth circuits can be interpreted as effecting a real space coarse-graining transformation. For example, in the context of "entanglement renormalization" [25], one imagines using local unitaries to successively remove short-range entanglement from a state. After each such step the degrees of

---

1 This model was developed in part from conversations with M.P.A. Fisher and A. Chandran at “Entanglement and Emergence II” at the Perimeter Institute. We were not aware at that time of the important related works of Refs. [16,18].
freedom may be safely coarse-grained and the process can be repeated. In this way, the entanglement in a quantum ground state, say, is displayed as a function of length scale. The physical picture which emerges is that we remove local entanglement and coarse-grain until we reach a product state, either because the system is short-range entangled or because the system has a finite size. There is growing evidence that this physical picture is a good one for ground states of local Hamiltonians. This evidence comes from numerical simulations of various quantum critical points in one and two dimensions [26][27], from the realization that certain kinds of ground states have exact representations in this language [28][29], and more exotically, from the proposal [30][31] that such a picture of entanglement renormalization also partially underlies holographic dualities [32][34] in the context of quantum gravity. In other words, a wide variety of ground states, including mean-field states, topologically ordered states, quantum critical points, and even states of gauge theories with holographic duals, can all be described in this unified language. It is furthermore satisfying that this physical picture is also associated with numerical tools, namely tensor networks, that can be used to simulate such ground states [24].

The picture just advocated works quite well so far for ground states, but the structure of entanglement in excited states is typically much more complex. However, a many-body localized phase is crucially different. There is a sense in which all energy eigenstates “look the same” (for example, all eigenstates might be localized [3]), so the story of entanglement renormalization in ground states should also apply to excited states in a many-body localized phase. This observation forms the basis for our model.

Following the above RG perspective, a key component of our model is a local unitary $U$ which transforms exponentially localized states into exactly localized product states. The other key component is a fixed point Hamiltonian $H_0$ which is diagonal in the local product basis and contains exponentially decaying interactions. Our model should thus be interpreted as a renormalization group fixed point of a many-body localized phase, although we do not explicitly specify a renormalization group procedure beyond the local unitary $U$. We note in passing that Refs. [35][36] have explored possibly related RG perspectives.

We should also immediately mention that the model we introduce is not without imperfections. For example, the Hamiltonian we analyze is not strictly local due to exponentially decaying interactions. Also, we do not generically expect that all energy eigenstates are localized or that a single local unitary can reduce every energy eigenstate to an exact product state. Thus our toy model is indeed fine tuned in the heuristic sense that all “irrelevant” perturbations are zero\(^2\), but it does capture many of the phenomena expected in a many-body localized phase.

The rest of the paper is organized as follows. First, we describe the basic model and establish that it is in a many-body localized phase in that energy is not transported. Second, we analyze the physics of entanglement in the state. Finally, we discuss the stability of the model and give a concluding discussion.

II. MODEL

Consider $N = L^d$ sites arranged on a hyper-cubic lattice in $d$ dimensions. Each site $r$ carries a spin-half and we label the Pauli operators $X_r, Y_r, Z_r$. We will use a renormalization group language to set up the model. First, we define a “fixed point” Hamiltonian in terms of projectors $P_r = (1 - Z_r)/2$ as

$$
H_0 = \sum_r J^i_r P_r + \sum_{r \neq r'} J^2_{rr'} P_r P_{r'} + \sum_{r \neq r' \neq r''} J^3_{rr'r''} P_r P_{r'} P_{r''} + ... \quad (2.1)
$$

The couplings $J^i_r, J^2_{rr'}, J^3_{rr'r''}$ are drawn from a probability distribution $P[J;i, r_1...r_l]$. The couplings are assumed to decay exponentially with the distances $|r_a - r_b|$ and with increasing number of spins involved. The length scale for the spatial decay is $\xi$, the localization length. Throughout we denote by $J_0$ a typical short-ranged energy scale in $H_0$.

The Hamiltonian $H_0$ describes a perfectly localized state by construction. Physically, we view it as the fixed point Hamiltonian obtained from a many-body localized system after some coarse-graining. Clearly the probability distributions $P$ must be “sufficiently random” for $H$ to truly describe a many-body localized state, otherwise adding small perturbations to $H$ could result in transport of energy under the perturbed dynamics, e.g. $P[J;1,r] = \delta(J - J_0)$ is clearly not suitable to describe a many-body localized state. We further discuss these stability issues below.

We next model the coarse-graining process. Introduce another local Hamiltonian $K$ which has strictly local interactions of bounded strength but is otherwise completely random. Define the unitary operator

$$
U = e^{-iK} \quad (2.2)
$$

which represents a fictitious time evolution for unit time under the Hamiltonian $K$. We emphasize that this $K$ is not the physical Hamiltonian. In fact, we wish to interpret $U$ as a renormalization group transformation which maps the physical exponentially localized many-body states to the ideal perfectly localized product states entering $H_0$. To be precise, we define the full Hamiltonian

$$
H = U H_0 U^{-1} \quad (2.3)
$$

where $H_0$ provides random energies and $U$ provides random localized wavefunctions. We emphasize that the spectra of $H$ and $H_0$ are identical. Also, although we do not consider it further here, we could include conserved charges by demanding that $K$ and $H_0$ commute with the conserved charge.
To show the locality of $H$, we use the machinery of Lieb-Robinson bounds \cite{[37]}. By assumption, $H_0$ consists of exponentially localized interactions. Thus we need only show that the effects of $U$ are also local up to exponential corrections. Since $K$ is a strictly local Hamiltonian built from Pauli operators with bounded coefficients, the Lieb-Robinson bound states \cite{[37]} that there are constants $\xi_{LR}$ (a length), $\nu_{LR}$ (a velocity), and $C$ such that for two commuting operators $A$ and $B$ separated by a distance $d$ we have

$$\| [A(t), B(0)] \| < C \| A \| \| B \| e^{-(d-\nu_{LR} t)/\xi_{LR}}. \quad (2.4)$$

where $A(t) = e^{iKt} A e^{-iKt}$ and where $\| O \|$ is the operator norm of $O$. This bound implies that for times of order one (e.g., for transformations by $U = e^{-iK}$), distant operators still commute up to exponential corrections. The length $\xi_{LR}$ plays the role of a localization length, and we henceforth restrict to $K$ such that the Lieb-Robinson bound is tight with $\xi_{LR} = \xi$. Thus the Hamiltonian $H$ consists of exponentially decaying interactions with characteristic length scale $\xi$ as desired. It should be noted that from one point of view our model is unrealistic in that it is not strictly local, that is it contains exponentially decaying (and not strictly local) interactions.

Let us also observe that the many-body spectrum of $H$, which is identical by construction to that of $H_0$, is generic in the sense that the level spacing is typically of order $2^{-N}$ where $N = L^d$ is the number of spins. This follows from the average density of states. For example, suppose $P[J, i > 1] = 0$ so that one has only random onsite energies with $P[|J|, 1] = \theta(J_0 - |J|) \theta(J_0) / J_0$. In this case the many-body spectrum is $E(n) = \sum_r J_r n_r$ with $n_r = 0, 1$. The density of states is

$$D(E) = \sum_n \delta(E - E(n)), \quad (2.5)$$

and the disorder averaged density of states is

$$\bar{D}(E) = \sum_n \delta(E - E(n)). \quad (2.6)$$

To compute the average we simply use that the onsite energies are uncorrelated, so the central limit theorem gives, for $\bar{N}(n) = \sum_r n_r \gg 1$,

$$p(E(\bar{N})) \propto \exp \left( - \frac{(E - \bar{N} J_0/2)^2}{2N (J_0^2/4)} \right). \quad (2.7)$$

The number of $n_r$ summing up to $\bar{N}$ is $\left( \begin{array}{c} N \\ \bar{N} \end{array} \right)$, and hence we have roughly

$$\bar{D}(E) \sim \int d\bar{N} \left( \begin{array}{c} N \\ \bar{N} \end{array} \right) \exp \left( - \frac{(E - \bar{N} J_0/2)^2}{2N (J_0^2/4)} \right). \quad (2.8)$$

The $\bar{N}$ integral is strongly peaked provided $2E/J_0 \sim N$ in which case $\bar{D}(E) \sim 2^N$ as claimed. Adding in further neighbor interactions does not substantially change the above reasoning.

The Hamiltonian $H$ describes a many-body localized system because energy is not transported. There are two results which illustrate this point. First, given any two many-body eigenstates that differ by spin flips in distant regions, the local physics of these states will be nearly identical away from the regions where the spin flips occurred. Indeed, the operator $UX_1 U^{-1}$ moves from one many-body energy eigenstate to another and is manifestly exponentially localized due to the localized structure of $U$. This means that adding a precise amount of energy to a state is a localized operation.

Second, suppose we start with a many-body eigenstate $|E\rangle$ and perturb the system locally by acting with a local operator $O(r_0)$. Then the amplitude to find the system in another many-body eigenstate associated with a distant eigenstate, e.g., $UX_1 U^{-1} |E\rangle$, is exponentially small as $|r - r_0| \to \infty$. This is true for all time since $U$ is localized and the dynamics of $H_0$ can only lead to dephasing and not distant spin flips. Hence the amplitude of a distant spin flip can be bounded at all times by a constant which decays exponentially with distance. Thus energy is not transported. For example, in a process where we start in the ground state, inject energy at $r_0$, and try extract energy at $r_1$, we only succeed with exponentially small probability.

III. ENTANGLEMENT

We now obtain results about the structure of entanglement in our model. Entanglement entropy is defined by splitting the Hilbert space into two pieces associated with a region $A$ and its complement $B$. The Hilbert space factorizes as $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ and the state of region $A$ may be obtained from the state of the whole system via a partial trace over $B$:

$$\rho_A = \text{tr}_B(\rho_{AB}). \quad (3.1)$$

The entanglement entropy $S(A)$ of $A$ is then the von Neumann entropy of $\rho_A$:

$$S(A) = -\text{tr}_A(\rho_A \ln \rho_A). \quad (3.2)$$

When $\rho_{AB}$ is a pure state, $S(A)$ indeed measures entanglement between $A$ and $B$.

A. Area law for energy eigenstates

We first address entanglement entropy for energy eigenstates. To begin, every eigenstate of $H_0$ is a product state and hence is unentangled. Every eigenstate of $H$ is obtained from an eigenstate of $H_0$ by acting with $U$: $|E\rangle = U \prod_r |n_r\rangle$ where $P_r |n_r\rangle = n_r |n_r\rangle$. Hence any entanglement present in $|E\rangle$ is due to the action of $U$. $U$ is generated by a fictitious time evolution with Hamiltonian $K$ for a time $t = 1$, and we may bound the entanglement generated by $U$ because $K$ is local.

Introduce the region $A$ whose entanglement we wish to compute and its complement $B$. $K$ may be decomposed as

$$K = K_A + K_B + \sum_\alpha k_\alpha \quad (3.3)$$
into terms acting only on $A$ or $B$ and $AB$ interaction terms of the form $\hat{O}_A = O_A^\alpha \otimes O_B^\beta$ where $O_R$ acts only on region $R$, $K_A$ and $K_B$ do not directly generate entanglement, so we must only consider the $AB$ interaction terms in $K$. The strict locality of $K$ implies that the number of such terms grows only with the boundary size $|\partial A|$ of $A$. The entangling power of each such term is rigorously bounded by

$$\frac{dS(A)}{dt} \leq c||O_A^\alpha||||O_B^\beta||$$

(3.4)

with $c$ a numerical constant [38][40]. Intuitively this is because the maximum possible entropy that a unitary of the form $e^{-i\hat{O}_A t}$ can add to region $A$ is bounded (by the log of the dimension $d_A$ of the Hilbert space on which it acts) with the maximum typically occurring when $t = t_{max} \sim ||\hat{O}_A||^{-1}$ (an inverse energy). Thus heuristically one has

$$\frac{dS}{dt} \sim \frac{\Delta S_{max}}{t_{max}} \sim \ln (d_A)||\hat{O}_A|| = \ln (d_A)||O_A^\alpha||||O_B^\beta||.$$ 

(3.5)

To treat the sum over $\alpha$ in $K$ we use the the Trotter product formula to show that

$$\frac{dS(A)}{dt} \leq \sum_{\alpha} c||O_A^\alpha||||O_B^\alpha||.$$ 

(3.6)

Since the total number of such operators is proportional to $|\partial A|$ and since all the interactions have bounded strength (meaning $||\hat{O}_A|| < \hat{k}$ for some fixed number $\hat{k}$), the total entropy generated by time evolution for time $t = 1$ is bounded

$$\Delta S(A) \leq c' |\partial A|.$$

(3.7)

Hence every many-body energy eigenstate obeys an area law for entanglement entropy.

B. Time dependent states

We can also consider time dependent states. Suppose first that we make a local perturbation starting from an energy eigenstate $|E\rangle$. As usual, we use $U$ to translate this perturbation into a state evolving under $H_0$. If the perturbation happens to push the system into another energy eigenstate then nothing will happen. If the perturbation produces a superposition of energy eigenstates then $H_0$ will rapidly decohere the superposition. However, as with energy transport, the maximum amplitude to produce distant excitations is bounded by an exponentially decaying envelope. If $\psi(n)$ are the amplitudes of the eigenstates $|n\rangle$ of $H_0$ then the maximum entropy of a region $A$ is the entropy of the density matrix

$$\rho_A(n_A, n_A') = \sum_{n_B} \psi(n_A n_B) \psi^*(n_A' n_B)$$

(3.8)

plus the maximum entropy $U$ can add. But by assumption the amplitudes $\psi$ factorize except near the localized perturbation and cannot give a contribution to the entropy which grows with region size. Hence the entropy of any region $A$ sufficiently larger than the size of the local perturbation will be fixed at its value in the initial eigenstate up to exponentially decaying corrections.

We can also consider extensive perturbations. Suppose we begin with a global product state. We use $U$ to translate this into a state which evolves under $H_0$. Because $U$ is local, the resulting initial state is short range correlated and has at most area law entanglement. The initial state will be a superposition of different eigenstates of $H_0$. Due to the exponentially decaying interactions in $H_0$, the initial state will slowly lose coherence and entanglement will be generated.

We give a simple model calculation which illustrates the effect. Consider two spins with Hamiltonian

$$H = J_1 P_1 + J_2 P_2 + J_{12} P_1 P_2$$

(3.9)

and suppose for simplicity that the spins begin in identical product states

$$|\psi(0)\rangle = (\alpha|0\rangle + \beta|1\rangle)^{\otimes 2}.$$ 

(3.10)

This state evolves into the state

$$|\psi(t)\rangle = e^{-iHt}|\psi(0)\rangle = \alpha^2|00\rangle + 2\alpha\beta e^{-iJ_{12}t}|01\rangle + \beta^2 e^{-i(J_1+J_2+J_{12})t}|11\rangle$$ 

(3.11)

after time $t$. For concreteness, suppose we have initially that $\alpha = \beta = 1/\sqrt{2}$. Then the reduced density matrix of spin one takes the form

$$\rho_1(t) = \frac{1}{2} \left( e^{-iJ_{12}t} \frac{1}{1+e^{-iJ_{12} t}} \frac{e^{iJ_{12} t}(1+e^{iJ_{12} t})}{2} \right).$$ 

(3.12)

Hence we see that, independent of $J_1$ and $J_2$, the state becomes maximally mixed when $J_{12} t = \pi$. Of course, in this simple model the system returns to a pure state after the same time, but in the Hamiltonian $H_0$ there are additionally many other spins which also interact with a given spin. These additional interactions lead to decoherence which implies that once purity is lost, it is not recovered (modulo recurrences on time scales of order $2^N$).

Returning to the main problem, consider again a region $A$ evolving from an initial product state for a time $t$. The physical picture is that at time $t$ entanglement has been generated between $A$ and spins a distance

$$\ell(t) \sim \xi \ln (J_0 t/\pi)$$ 

(3.13)

away. This is because, assuming exponentially decaying interactions and following the model calculation above, only these spins have had enough time to entangle with region $A$. Several types of behavior are then possible. Of course, there can be some initial transient behavior before the localization scale is reached, but afterwards, if the characteristic linear size $L$ of $A$ is much greater than $\ell(t)$, then we expect an entropy going like $S(A, t) \sim |\partial A|\ell(t)$. 


Eventually the entropy will saturate since we have the bound
\[ S(A) \leq N_A \ln 2. \]  
(3.14)
In fact, the entropy can never exceed the entropy of the “diagonal” state consisting of \( \rho_A \) with all off diagonal elements in the \( Z \) basis set to zero. Indeed, this process of setting off diagonal elements to zero amounts to a measurement of \( Z \) for every spin in \( A \),
\[ \rho_A \to \mathcal{E}(\rho_A) = \sum_{\{x_r=0,1\}} M(x)\rho_A M^\dagger(x), \]  
(3.15)
where
\[ M(x) = \left( \prod_r (P_r)^{x_r} (1 - P_r)^{1-x_r} \right). \]  
(3.16)
so that the set \( \{M(x)\} \) represents a projective measurement of all \( Z \) variables. Such a measurement always increases the entropy. Hence the diagonal ensemble upper bounds the entanglement entropy.

C. Justification of entanglement growth rate

We now give a more rigorous justification of the above results. First, since \( U \) can at most add an area law worth of entanglement to a state, for the purposes of asymptotic time dependence, we may as well work directly with a state evolving under \( H_0 \). \( H_0 \) has two very special properties that facilitate the analysis. First, the various terms in \( H_0 \) are all diagonal in a local product basis. Second, every term in \( H_0 \) commutes with every other term. This implies that the time evolution generated by \( H_0 \), call it \( W(t) = e^{-iH_0t} \), can be factorized as
\[ W = W_A W_B W_{AB}, \]  
(3.17)
where each term acts on the indicated subsystem and all terms commute.

If \( \rho_{AB}(0) \) is the initial state of \( AB \), then we wish to compute the time dependence of the state of \( A \) defined as
\[ \rho_A(t) = \text{tr}_B(W(t)\rho_{AB}(0)W^\dagger(t)). \]  
(3.18)
We immediately obtain
\[ \rho_A(t) = W_A \text{tr}_B(W_{AB}\rho_{AB}(0)W_{AB}^\dagger)W_A^\dagger \]  
(3.19)
where we have used the cyclic property of the trace to remove \( W_B \). Furthermore, the unitary \( W_A \) doesn’t change the spectrum of \( \rho_A \), so it cannot effect the entanglement entropy. Hence we may as well set \( W_A = W_B = 1 \), so that the whole entropy is manifestly dependent only on the \( AB \) interactions in \( W_{AB} \) and the initial state.

Suppose now that, unlike in our model, the interactions between \( A \) and \( B \) were of strictly finite range. Then we see that the \( W_{AB} \) acts on a finite Hilbert space near the boundary of \( A \) of dimension roughly \( 2^{R[\partial A]} \) where \( R \) is the range of the terms in the interaction Hamiltonian \( H_{AB} \). Since the amount of entropy such a unitary can add to a system is bounded by the logarithm of the dimension of the Hilbert space on which it acts, we see that such a \( W_{AB} \) with short-range interactions cannot add more than roughly \( R[\partial A] \) entropy. Hence a strict area law will be obeyed for sufficiently large regions at all times. In fact, in one dimension the situation is even worse. There, if the range \( R \) is finite, then \( W_{AB}(t) \) is quasiperiodic in \( t \) and the entanglement entropy actually suffers recurrences on time scales of order \( 1/J_0 \).

Thus we see that exponentially decaying interactions are essential if we want to have any sustained growth of entanglement for large regions. The virtue of the above argument is that it immediately provides an estimate of the entropy growth. We must ask what is the effective range, \( R_{\text{eff}} \), of \( H_{AB} \) at time \( t \). Since \( W_{AB} = e^{-iH_{AB}t} \) and since the terms in \( H_{AB} \) between spins a distance \( r \) apart are roughly of order \( J_0 e^{-r/\xi} \), the natural choice is to say that \( R_{\text{eff}}(t) \) is such that
\[ t J_0 e^{-R_{\text{eff}}(t)/\xi} \sim 1. \]  
(3.20)
Now because \( Z_r \) commutes with \( H_0 \), the diagonal elements of \( \rho_A \) in the \( Z \) basis do not change with time. As we already discussed, the entropy of the state with all the off-diagonal elements zero bounds the entropy of \( \rho_A \) itself. But more than this, since we expect dephasing to occur, the late time state should indeed be close to this \( Z \)-measured state. Let \( s \) be the entropy per spin in the dephased state (it can be much less than one and depends on the initial conditions). Then the entropy of region \( A \) should grow as
\[ S(A,t) = s|\partial A| R_{\text{eff}}(t). \]  
(3.21)
Since \( R_{\text{eff}}(t) = \xi \ln (J_0t) \) we indeed recover the claimed logarithmic growth. In fact, with the bounds on entanglement growth used above, we can prove that the entanglement cannot grow faster than this estimate. Furthermore, because the Hamiltonian is generic we expect it to saturate this bound on the rate of entanglement growth.

IV. Stability

The model we have presented is meant to represent a system deep within a many-body localized phase. However, as we already mentioned, a system should only be called many-body localized if it is stable to small perturbations. Thus in order for \( H \) to be described as a many-body localized system, it must be that \( H' = H + gV \) has all the same universal physics as \( H \) for all local perturbations \( V \), e.g. area laws for excited states and slow growth of entanglement, provided \( g \) is small enough. If, for example, we choose the couplings in \( H_0 \) and the generator \( K \) to be non-random, then we would certainly not be describing a many-body localized phases since perturbations would immediately lead to transport of energy. In this case we might say that the “fixed point” described by \( H \) is infinitely unstable, i.e. has infinitely many relevant perturbations, even though \( H \) is indeed localized by construction.
What follows in this section are heuristic arguments and some partial results towards a proof of stability. We first describe the basic intuition and then discuss a more technical approach based on adiabatic continuity.

Assuming that $H_0$ and $K$ are random as described above, we wish to analyze the effects of perturbations $gV$ with $V$ local and $g$ small. The perturbed Hamiltonian is

$$H' = H + gV,$$

or after transforming by the local unitary $U$,

$$H'_0 = H_0 + gU^{-1}VU = H_0 + gV_0. \quad (4.2)$$

The crucial point is that $V_0$ is still local up to exponentially decaying corrections. At this point, we certainly cannot prove that the many-body localized state is stable, and indeed, there are many possible subtleties, for example, a few many-body eigenstates might delocalize while the whole system effectively remains in a many-body localized phase. However, a some perturbative intuition that suggests the model is stable.

Indeed, since $V_0$ is a local operator, it will typically only connect states that differ by a finite energy (not decreasing with system size). If we assume that the terms in $V_0$ have bounded norm, then the matrix elements are of order $g$ or smaller and decay exponentially with distance. If $V_0 = \sum_r v_r$ with each $v_r$ a local operator and if $|n\rangle$ and $|m\rangle$ are two eigenstates of $H_0$, then matrix elements of the form $\langle n|v_r|m\rangle$ will vanish exponentially unless $n$ and $m$ differ only by changes near site $r$. However, in this case $E(n)$ will typically differ from $E(m)$ by a finite energy, say of order $J_0$, and hence the perturbing matrix elements will be quite small compared to the diagonal elements provided $g$ is small.

Of course, we can use $V_0$ to connect states which are closer in energy by going to higher orders in perturbation theory, but these processes will be exponentially suppressed due to the large number of off resonant intermediate states required. The picture which thus emerges is that we must go to an exponentially suppressed high order in perturbation theory or make use of an exponentially small interaction to connect states with nearly degenerate energies. This scenario is then essentially the picture advocated by Ref. [4] where one has localization in the many-body Fock space. It was argued in Ref. [4] that the growth in the number of terms at each order of perturbation theory is commensurate with an interpretation as an Anderson model in Fock space with bounded connectivity. More concretely, if free particle energy eigenstates are points and if two points share an edge provided they are connected by a matrix element of the interaction, then the resulting graph has an Anderson transition if the connectivity is finite. It would be interesting to carry out a more quantitative analysis of this scenario in our model, but in this work we turn to different but related point of view.

We will attack the problem from the point of view of adiabatic continuity. Ref. [11] has analyzed the quantum adiabatic algorithm for a different model and concluded that Anderson localization makes the adiabatic algorithm fail (due to an extremely small gap). However, our perspective will be somewhat different in that we will not require that ground states be mapped to ground states but only that energy eigenstates be mapped to energy eigenstates.

### A. Adiabatic continuity

Let us begin by recalling that for band insulators (not disordered), the ground state is localized and hence the linear response conductivity vanishes at zero temperature. Of course, at any finite temperature the system will conduct due to a finite density of delocalized quasiparticles. However, the important point is that the localization of the ground state persists as the Hamiltonian is smoothly changed. Hence the conductivity vanishes at zero temperature everywhere within the phase. On the other hand, highly excited energy eigenstates are not localized and an applied field leads to a non-zero response. From the point of view of adiabatic continuity, the ground state of, say, a band insulator, differs dramatically from a highly excited state in that the ground state is separated by a finite gap from other eigenstates while a highly excited state typically sits in a nearly continuous set of states with a level spacing going like $e^{-N}$. Furthermore, these nearly degenerate states can be easily connected to each other by local operators.

In the context of ground states of band insulators, the basic adiabatic argument runs as follows. Consider a family of Hamiltonians $H(g)$ each possessing a localized ground state $|\psi(g)\rangle$ and a gap $\Delta(g) \geq \tilde{\Delta}$. Suppose we start at the initial ground state $|\psi(0)\rangle$ and time evolve under the Hamiltonian $H(g(t))$ for some slowly varying function $g(t)$:

$$\frac{d|\phi(t)\rangle}{dt} = H(g(t))|\phi(t)\rangle \quad (4.3)$$

and

$$|\phi(t = 0)\rangle = |\psi(g = 0)\rangle. \quad (4.4)$$

Then because $H(g)$ is gapped for all $g$ with minimum gap $\tilde{\Delta}$, it follows that local properties of $|\psi(1)\rangle$ will be accurately reproduced provided $\frac{dg}{dt} \ll \tilde{\Delta}$. In general, to truly make the time evolved state $|\phi(t)\rangle$ close to $|\psi(1)\rangle$ in the sense of having high overlap, we must evolve for time growing with system size, but if we really want all global properties to be preserved then we can use a modified procedure known as quasi-adiabatic continuation [42] discussed below. Regardless, as far local properties are concerned the adiabatic argument works well. However, we see immediately that a key difference for excited states is the complete lack of a gap between nearby energy eigenstates. Since these eigenstates can be connected to each other via local operators, it follows that time evolution under a local Hamiltonian for any reasonable amount of time, e.g., less than times of order $e^{-N}$, will strongly mix various excited states. Hence we apparently cannot learn as much about excited states using adiabatic continuity.

The situation is quite different in a many-body localized phase. We have already shown that while the level spacing of the many-body spectrum goes like $e^{-N}$ near the center of the spectrum, states with nearly degenerate energies are typically
associated with very different configurations of spins. As argued in the perturbative discussion above, local operators typically only connect states with very different energies. Hence as far as local operators are concerned, energy eigenstates in a many-body localized state typically have an effective gap to other eigenstates. In other words, states $| E \rangle$ and $| E' \rangle$ with a sizable matrix element $\langle E | O | E' \rangle$ for some localized operator $O$ typically differ in energy by a relatively large amount, say, typically $| E - E' | \geq \Delta(O)$ where the effective gap $\Delta(O)$ depends on how localized the operator $O$ is.

We now repeat the dynamical evolution

$$\frac{d|\phi(t)\rangle}{dt} = H(g(t))|\phi(t)\rangle, \quad (4.5)$$

with

$$|\phi(t = 0)\rangle = |E\rangle \quad (4.6)$$

where $|E\rangle$ is an arbitrary many-body localized energy state. Then as we argued above, since $H(g(t))$ is local for all $t$ and since there is an effective gap for local perturbations, the adiabatic evolution maps eigenstates to eigenstates in a many-body localized phase. This should be compared with the result just for ground states in the case of conventional insulators.

Then using the Lieb-Robinson arguments above, it follows that the unitary $Q(t_{ad})$ which effects the adiabatic time evolution for time $t_{ad}$ is a local unitary up to exponentially decaying corrections. Hence the new energy eigenstates may be approximated by $Q(t_{ad})|E\rangle$, but since $|E\rangle$ is itself related to a product state by $U$, we see that the new energy eigenstates are also related to product states by the local unitary $Q(t_{ad})U$ and thus are localized. Hence the fact that energy eigenstates of the final Hamiltonian $H(1)$ are localized can be heuristically derived from the fact that the energy eigenstates of the initial Hamiltonian $H(0)$ are localized provided the effective gap $\Delta(O)$ for local perturbations doesn’t collapse.

### B. Quasi-adiabatic continuity

Having given various intuitive arguments for stability, we now want to formalize as much as possible the problem of proving stability. The main tool we use in the remainder of this section is Hastings’ quasi-adiabatic continuation [42]. Our considerations are also potentially related to Hastings’ recent work on adiabatic continuity in disordered systems [43].

To achieve high overlap between an adiabatically evolved state and the target, one must evolve the system for a time which grows with system size. The following simple example illustrates this need. Suppose we have a single spin with Hamiltonian $h(t)$ which we evolve for time $t_{ad}$. Let $|\phi(t)\rangle_1$ denote the result of adiabatic time evolution and let $|\psi(t)\rangle_1$ denote the instantaneous ground state of $h(t)$. Suppose that $h(t)$ changes slowly enough so that

$$\langle \phi(t) | \psi(t) \rangle_1 = 1 - \epsilon. \quad (4.7)$$

If we now perform the same adiabatic evolution in a many-body system of $N$ such spins, all with the same hamiltonian and not interacting, then we have

$$|\phi(t)\rangle_N = |\phi(t)\rangle_1^\otimes N \quad (4.8)$$

and

$$|\psi(t)\rangle_N = |\psi(t)\rangle_1^\otimes N. \quad (4.9)$$

Thus the overlap is

$$\langle \phi(t) | \psi(t) \rangle_N = (1 - \epsilon)^N \quad (4.10)$$

because we have an independent probability of error at each site. Hence to make $(1 - \epsilon)^N$ close to one for $N$ large, $\epsilon$ must be chosen to decrease with $N$.

Remarkably, it is possible to construct an exponentially localized hermitian operator which generates a fictitious time evolution that exactly reproduces the ground state in a gapped system. This method is called quasi-adiabatic evolution and was pioneered by Hastings [42]. Consider a Hamiltonian $H(s)$ which depends on a parameter $s$. We would like to find a local operator $J(s)$ such that the solution of

$$i\partial_s |\phi(s)\rangle = J(s) |\phi(s)\rangle \quad (4.11)$$

has $|\phi(s)\rangle$ equal to the exact ground state of $H(s)$ for all $s \in [0,1]$.

The ground state of $H(s)$ is denoted $|\psi_0(s)\rangle$. We assume the system has gap $\Delta(s) = E_1(s) - E_0(s)$ which is bounded from below by an $s$ and system size independent constant $\Delta(s) > 1$. Let $F(t)$ be a fast decaying function of $t$ with the following properties. First, its Fourier transform $\tilde{F}\omega$ satisfies

$$\tilde{F}\omega = -\frac{1}{\omega}, \quad \text{for } |\omega| \geq \Delta \text{ and second, } \tilde{F}\omega = 0. \quad (4.12)$$

If we now apply $J(s)$ to the ground state we find

$$-iJ(s)|\psi_0(s)\rangle = \sum_n \tilde{F}(E_n - E_0)|\psi_n(s)\rangle \times \langle \psi_n(s) | \partial_s H(s) |\psi_0(s)\rangle, \quad (4.13)$$

but this last expression, using the properties of $\tilde{F}$, is simply

$$\sum_{n \neq 0} \frac{-1}{E_n - E_0} |\psi_n(s)\rangle \langle \psi_n(s) | \partial_s H(s) |\psi_0(s)\rangle \times \langle \psi_n(s) | \partial_s H(s) |\psi_0(s)\rangle \quad (4.14)$$

The last equality above is conventional perturbation theory. Thus we have that

$$i\partial_s |\psi_0(s)\rangle = J(s) |\psi_0(s)\rangle \quad (4.15)$$

with $J(s)$ a quasi-local hermitian effective Hamiltonian which generates the quasi-adiabatic evolution.

The quasi-locality of $J(s)$ follows from the Lieb-Robinson bound and from the fast decay of $\tilde{F}(t)$ with $t$. In brief,
\( e^{iH(t)\Delta t} e^{-iH(s)\Delta t} \) is localized up to exponential corrections beyond the Lieb-Robinson light cone \( v_{LR} t \), and although we integrate over all \( t \), the integral is weighted with a fast decaying function \( F \). Large values of \( t \) are almost exponentially suppressed (faster than any power) and since small values of \( t \) lead to exponentially localized operators, we see that every term in \( \partial_s H \), even after the integral transform is applied, remains local up lengths of roughly \( v_{LR}/\Delta \) with nearly exponentially decaying corrections beyond. To formalize these statements, let \( \partial_s H(s) = \sum V_r(s), \) i.e. a sum over local terms \( V_r(s) \). Now consider \( J_r(s,T) \) defined by
\[
-\mathbf{i} J_r(s,T) = \int_{-T}^{T} dt F(t) e^{iH(s) t} V_r(s) e^{-iH(s)t}. \tag{4.16}
\]
The Lieb-Robinson bound states that \( J_r(s,T) \) is localized to within \( v_{LR} T \) of \( r \) up to an exponentially small error, but we also know from the fast decay of \( F \) that
\[
||J_r(s,T) - J_r(s,\infty)|| < F_{\infty}(\Delta T)||V_r(s)|| \tag{4.17}
\]
with \( F_{\infty}(x) \) another fast decaying function related to \( F \). For example, we could expect to have \( F_{\infty}(T\Delta) \sim e^{-T\Delta} \) or nearly exponential decay. Thus \( J_r(s,T) \) is localized up to exponential error, \( J_r(s,\infty) \) is exponentially close to \( J_r(s,\infty) \), and \( \sum_r J_r(s,\infty) \) is nothing but our original \( J(s) \).

It is interesting to instructive to apply the quasi-adiabatic technology to the non-interacting many spin problem from the beginning of Appendix A. In this case we see immediately that \( J(s) \) is strictly local and in fact is a sum of independent terms for each spin. Each such term is identical and by construction perfectly evolves ground states into ground states. Indeed, in this case it also perfectly evolves excited states into excited states since each spin has only two levels and unitary evolution preserves orthogonality.

### C. Quasi-adiabatic evolution for many-body localization

Since the evolution generated by \( J(s) \) is quasi-local, it follows that if one ground state is localized then so is every other ground state. Thus it would be very useful if we could apply quasi-adiabatic continuation to more general localized states. We can try to repeat the same story for a many-body localized phase, but the immediate problem is that the many-body localized Hamiltonian does not possess a gap (around the ground state or around excited states). However, as discussed above, there is a sense in which each energy eigenstate has a gap to other eigenstates which can be connected to it by a local operator. In this sense, all eigenstates of a many-body localized phase should be treated on the same footing. For example, in our model all eigenstates manifestly share the same properties and each could be the ground state of a local Hamiltonian.

Considering again a parameter dependent Hamiltonian \( H(s) \) with a spectrum in which every energy level is distinct and separated by a gap of at least \( \gamma \) from every other level. Let the normalized energy eigenstates be \( \psi_n(s) \). By definition we have
\[
H(s)|\psi_n(s)\rangle = E_n(s)|\psi_n(s)\rangle. \tag{4.18}
\]
Upon differentiating both sides with respect to \( s \) we find (denoting \( \partial_s \) with a ‘)\n\[
H'|\psi_n\rangle + H|\psi_n\rangle' = E'_n|\psi\rangle + E_n|\psi_n\rangle'. \tag{4.19}
\]
Since \( |\psi_n(s)\rangle \) is normalized for all \( s \), we have \( \langle \psi_n|(|\psi_n\rangle') = 0 \). Thus we can simplify the equation by projecting onto \( |\psi_n\rangle \) and the complement. Projecting onto \( |\psi_n\rangle \) gives
\[
E'_n = \langle \psi_n|H'|\psi_n\rangle. \tag{4.20}
\]
Projecting onto the complement gives
\[
(H - E_n)|\psi_n\rangle' = -(1 - |\psi_n\rangle\langle\psi_n|)H'|\psi_n\rangle. \tag{4.21}
\]

Since \( H - E_n \) is invertible on the the complement to \( |\psi_n\rangle \) we have an equation for \( |\psi_n\rangle' \)
\[
|\psi_n\rangle' = -(H - E_n)^{-1}(1 - |\psi_n\rangle\langle\psi_n|)H'|\psi_n\rangle. \tag{4.22}
\]

We now apply \( J(s) \) as defined above to a general energy eigenstate to obtain
\[
-i J(s)|\psi_m(s)\rangle = \sum \hat{F}(E_m - E_n)|\psi_m(s)\rangle \nonumber \times \langle \psi_m(s)|\partial_s H(s)|\psi_n(s)\rangle. \tag{4.23}
\]

If we choose \( \Delta \) in the definition of \( F \) to coincide with \( \gamma \), then we see that
\[
i\partial_s |\psi_n(s)\rangle = J(s)|\psi_n(s)\rangle. \tag{4.24}
\]
for all \( n \), not just the ground state.

Now, as we have repeatedly emphasized, there is no such gap \( \gamma \) in a many-body localized phase. However, there is typically an effective gap if we only consider states that can be connected to a given energy eigenstate by localized operators. This expectation is confirmed, for example, in Ref. [5], which demonstrated numerically the absence of level repulsion. Hence, because \( \partial_s H(s) \) is a sum of such localized operators, we can argue that although the filter function \( F \) cannot remove all states nearby in energy (not even close), it can effectively remove states which have a sizable matrix element of \( \partial_s H(s) \) with \( |\psi_n\rangle \). Of course, this may also fail in rare instances. We leave it to future work to make further progress on the issue of stability.

### D. Application to the model

Because our model is defined in terms of an arbitrary local unitary and a fixed point Hamiltonian, it immediately follows that there is a class of transformations under which the model is stable. For example, we may change the random local unitary to another random local unitary and the model clearly remains localized. Similarly, we may make small changes to energy level structure of the fixed point Hamiltonian and the phase will remain localized. For example, when changing the local unitary from \( U_1 \) to \( U_2 \), the transformation is effected \( U_2 U_1^{-1} \) which is still local. It is interesting to ask how
the quasi-adiabatic approach deals with such perturbations for which we know the exact answer. Recall the definition of the model Hamiltonian \( H \) in terms of \( H_0 \) and the unitary \( U \):
\[
H = U H_0 U^{-1}.
\]
(4.25)
Recall also that \( U \) was the exponential of a local operator \( K \):
\[
U = e^{-iK}.
\]
(4.26)
Now introduce a family of Hamiltonians \( H(s) \) defined by
\[
H(s) = e^{-isK} H_0 e^{isK} = U(s) H_0 U(s)^\dagger.
\]
(4.27)
We will test the machinery of quasi-adiabatic evolution on this family of Hamiltonians.

Since \( H(s = 0) = H_0 \) we clearly begin in maximally localized phase where all energy eigenstates are product states. We can calculate \( \partial_s H(s) \) as
\[
\partial_s H(s) = -i[K, H(s)] = -iU(s)[K, H_0]U(s)^\dagger.
\]
(4.28)
Thus we see that the quasi-adiabatic generator is
\[
U(s)^\dagger J(s) U(s) = \left[ \int dt F(t) e^{iH_0 t} [K, H_0] e^{-iH_0 t} \right].
\]
(4.29)
To compute matrix elements of this operator with respect to the exact many-body eigenstates of \( H(s) \), we use the fact that such eigenstates are related to eigenstates \( |n\rangle \) of \( H_0 \) by \( U(s) \). Hence we have
\[
\langle n| U(s)^\dagger J(s) U(s)|m\rangle = \langle n| U(s)^\dagger \int dt F(t) e^{iH_0 t} [K, H_0] e^{-iH_0 t} |m\rangle.
\]
(4.30)
Using the properties of \( F \) we see that for \( |E(n) - E(m)| \geq \Delta \), the matrix elements are just those of \( K \) itself. Hence the quasi-adiabatic generator coincides for such states. On the other hand, matrix elements of \( J(s) \) do not reduce to those of \( K \) for states with energies closer than \( \Delta \).

At this point, then, we see clearly the importance of localization physics. The matrix elements are
\[
\hat{F}(E(n) - E(m))(E(m) - E(n))\langle n|K|m\rangle.
\]
(4.31)
\( K \) is short-ranged and the \( |n\rangle \) are all product states, hence the matrix elements \( K(n,m) = \langle n|K|m\rangle \) can only be non-zero if \( n \) and \( m \) differ locally. Yet the local spectrum of such a many-body localized phase is discrete, with an effective gap \( \gamma \), hence almost all pairs of states with energy difference less than \( \gamma \) will give a zero matrix element of \( K \) (within a more general model, the matrix element might be exponentially small). The final conclusion is that the quasi-adiabatic generator \( J(s) \) is equivalent to \( K \) for almost all states. Since \( K \) generates the exact unitary which transforms between eigenvectors for different \( s \), we see that within our model the quasi-adiabatic method leads to the correct conclusion that localization persists as we vary the Hamiltonian, e.g., each \( H(s) \) is many-body localized.

E. Level crossings

We can also change the couplings entering the fixed point Hamiltonian without changing the operators which appear. Consider \( H_0(s) \) which is of the same form as \( H_0 \) but with variable couplings. \( H(s) \) is defined as \( U H_0(s) U^{-1} \) with \( U \) fixed. It follows that \( [\partial_s H(s), H(s)] = 0 \) and hence \( J(s) \) in this case is
\[
-J(s) \int dt F(t) e^{iH_0 t} \partial_s H e^{-iH_0 t} = \left( \int dt F(t) \right) \partial_s H(s).
\]
Since \( \partial_s H \) has all the same eigenvectors as \( H(s) \), it follows that \( J(s) \) only generates an uninteresting phase for each energy eigenstate. Hence in this case the resulting quasi-adiabatic transformation is of the form
\[
\sum_n e^{i\theta(n,s)} U|n\rangle \langle n|U^\dagger
\]
(4.32)
which manifestly preserves all eigenvectors up to a phase.

However, in a more generic physical system, one might expect avoided crossings over a very small energy window (see, for example, Ref. [41]). For example, imagine we build up the Hamiltonian from the extreme localized using the family of Hamiltonians \( H(s) \). As \( s \) is varied, there are bound to be very sharp avoided level crossings in the physical Hamiltonian. This is because the matrix elements of the perturbation, while they are exponentially suppressed, may still connect states with an even smaller energy splitting in the limit of very large systems. However, the quasi-adiabatic evolution will be fast on the scale of the induced mixing, so rather than follow the avoided crossing, the states will shoot through the crossing. Given two initial states, of lower and higher energy, the quasi-adiabatic evolution will map low to high and vice versa instead of low to low as would occur if we ran the evolution for a long enough time.

However, this simple picture suggests that the quasi-adiabatic evolution will still map energy eigenstates to energy eigenstates, it may just get them “out of order”. The locality of the quasi-adiabatic evolution combined with localized nature of all initial energy eigenstates means that all resulting evolved states are still localized. If the evolved states are indeed close to energy eigenstates of the target Hamiltonian, then we have succeeded in demonstrating the stability of many-body localization. Of course, what we have said is merely a sketch, since there are many closely spaced states and many-level crossings happening as the Hamiltonian is changed. It remains to be seen what part of the above sketch can be made completely rigorous.

V. DISCUSSION

In this paper we presented a simple model of a many-body localized phase. The model describes a system deep within a many-body localized and is analogous to a fixed point of a real
space renormalization scheme. We showed that the model has a number of phenomenological features expected of many-body localized states. All energy eigenstates are localized and obey an area law for entanglement entropy. Energy is not transported even at finite energy density. The growth of entanglement after a perturbation is also slow, either effectively not occurring at all or only growing logarithmically. Finally, we gave some arguments supporting the hypothesis that our model is stable and hence describes a many-body localized phase with no relevant perturbations.

It has already been observed in one dimension that the low entanglement scaling of energy eigenstates and the slow growth of entanglement with time makes tensor network states an ideal tool to investigate the physics of many-body localized states. Indeed, matrix product states and DMRG were used to great effect to study the entanglement dynamics of one dimensional localized phases [8,9]. Our renormalization group circuit picture naturally fits into the context of more general tensor network states like the MERA [20]. We have also proposed [30, 31] that MERA provides a kind of holographic description of ground states, so thinking along these lines, it would be interesting to find phases with features of a many-body localized state in the context of gravity. Perhaps the disordered black holes of Ref. [44] could be useful.

There are also numerous open questions. For example, it might be possible to rigorously prove or at least argue more completely that the model introduced above is stable. The assumption of a single unitary transformation $U$ is clearly too extreme in the generic case (see, for example, the results of Ref. [19]), but it is not clear what assumption should replace it. It should also be possible to give evidence for its stability using DMRG by checking the many-body spectrum for a variety of weak perturbations. Another interesting question relates to the possibility of having some delocalized energy eigenstates coexisting with localized eigenstates. Similarly, it would be interesting to try to understand a phase transition between a many-body localized phase and an ergodic phase where the system can serve as its own heat bath [4,7]. Finally, there is the question of whether the phenomena of many-body localization have any clear signature for experiments (beyond the absence of diffusion), for example, in cold atoms systems where a phonon heat bath can be excluded.

Acknowledgements: We thank M.P.A. Fisher and A. Chandran for enlightening discussions which helped inspire the present model and J. Sau for earlier conversations about logarithmically growing entanglement arising from exponentially small interactions. We also thank J. McGreevy, S. Gopalakrishnan, and M. Knap for discussions about many-body localization and Refs. [16,18]. We thank A. Pal, D. Huse, and D. Abanin for helpful comments on the manuscript. Significant progress on this work was made during “Entanglement and Emergence II” at the Perimeter Institute, and we thank the Perimeter Institute for hospitality. BGS is supported by a Simons Fellowship through Harvard University.

[1] P. W. Anderson, Phys. Rev. 109, 1492 (1958), URL [http://link.aps.org/doi/10.1103/PhysRev.109.1492]
[2] P. A. Lee and T. V. Ramakrishnan, Rev. Mod. Phys. 57, 287 (1985), URL [http://link.aps.org/doi/10.1103/RevModPhys.57.287]
[3] I. V. Gornyi, A. D. Mirlin, and D. G. Polyakov, Phys. Rev. Lett. 95, 206603 (2005), URL [http://link.aps.org/doi/10.1103/PhysRevLett.95.206603]
[4] D. M. Basko, I. L. Aleiner, and B. L. Altshuler, Annals of Physics 321, 1126 (2006), arXiv:cond-mat/0506617.
[5] V. Oganesyan and D. A. Huse, Phys. Rev. B 75, 155111 (2007), arXiv:cond-mat/0610854.
[6] V. Oganesyan, A. Pal, and D. A. Huse, Phys. Rev. B 80, 115104 (2009), 0905.4112.
[7] A. Pal and D. A. Huse, Phys. Rev. B 82, 174411 (2010), URL [http://link.aps.org/doi/10.1103/PhysRevB.82.174411]
[8] M. Znidarič, T. Prosen, and P. Prelovšek, Phys. Rev. B 77, 064426 (2008), 0706.2539.
[9] J. H. Bardarson, F. Pollmann, and J. E. Moore, Physical Review Letters 109, 017202 (2012), 1202.5532.
[10] S. Iyer, V. Oganesyan, G. Refael, and D. A. Huse, Phys. Rev. B 87, 134202 (2013), 1212.4159.
[11] L. D’Alessio and A. Polkovnikov, Annals of Physics 333, 19 (2013), 1210.2791.
[12] E. Canovi, D. Rossini, R. Fazio, G. E. Santoro, and A. Silva, New Journal of Physics 14, 095020 (2012), 1205.0370.
[13] A. De Luca and A. Scardicchio, EPL (Europhysics Letters) 101, 37003 (2013), 1206.2342.
[14] D. Delande, K. Sacha, M. Plodzien, S. K. Avazbaev, and J. Zakrzewski, New Journal of Physics 15, 045021 (2013), 1207.2001.
[15] D. A. Huse, R. Nandkishore, V. Oganesyan, A. Pal, and S. L. Sondhi, ArXiv e-prints (2013), 1304.1158.
[16] M. Serbyn, Z. Papić, and D. A. Abanin, Phys. Rev. Lett. 110, 260601 (2013), URL [http://link.aps.org/doi/10.1103/PhysRevLett.110.260601]
[17] D. A. Huse and V. Oganesyan, ArXiv e-prints (2013), 1305.4915.
[18] M. Serbyn, Z. Papić, and D. A. Abanin, ArXiv e-prints (2013), 1305.5554.
[19] B. Bauer and C. Nayak, ArXiv e-prints (2013), 1306.5753.
[20] G. Vidal, Phys. Rev. Lett. 101, 110501 (2008).
[21] F. Verstraete, J. Cirac, and V. Murg, Adv. Phys. 57, 143 (2008).
[22] Z.-C. Gu, M. Levin, and X.-G. Wen, Phys. Rev. B 78, 205116 (2008).
[23] S. R. White, Phys. Rev. Lett. 69, 2863 (1992), URL [http://link.aps.org/doi/10.1103/PhysRevLett.69.2863]
[24] U. Schollwöck, Rev. Mod. Phys. 77, 259 (2005).
[25] G. Vidal, Phys. Rev. Lett. 99, 220405 (2007).
[26] S. Montangero, M. Rizzi, V. Giovannetti, and R. Fazio, Phys. Rev. B 80, 113103 (2009), 0810.1414.
[27] R. N. C. Pfeifer, G. Evenbly, and G. Vidal, Phys. Rev. A 79, 040301 (2009), 0810.0580.
[28] Z.-C. Gu, M. Levin, B. Swingle, and X.-G. Wen, Phys. Rev. B 79, 085118 (2009), URL [http://link.aps.org/doi/10.1103/PhysRevB.79.085118]
[29] O. Buerschaper, M. Aguado, and G. Vidal, Phys. Rev. B 79, 085119 (2009), URL [http://link.aps.org/doi/10.
[30] B. Swingle, Phys. Rev. D 86, 065007 (2012), URL http://link.aps.org/doi/10.1103/PhysRevD.86.065007

[31] B. Swingle, ArXiv e-prints (2012), 1209.3304.

[32] J. M. Maldacena, Adv. Theor. Math. Phys. 2, 231 (1998).

[33] S. S. Gubser, I. R. Klebanov, and A. M. Polyakov, Phys. Lett. B 428, 105 (1998).

[34] E. Witten, Adv. Theor. Math. Phys. 2, 253 (1998).

[35] C. Monthus and T. Garel, Phys. Rev. B 81, 134202 (2010), 1001.2984.

[36] R. Vosk and E. Altman, Physical Review Letters 110, 067204 (2013), 1205.0026.

[37] E. Lieb and D. Robinson, Communications in Mathematical Physics 28, 251 (1972), ISSN 0010-3616, URL http://dx.doi.org/10.1007/BF01645779

[38] W. Dür, G. Vidal, J. I. Cirac, N. Linden, and S. Popescu, Phys. Rev. Lett. 87, 137901 (2001), URL http://link.aps.org/doi/10.1103/PhysRevLett.87.137901

[39] J. Eisert and T. J. Osborne, Phys. Rev. Lett. 97, 150404 (2006), URL http://link.aps.org/doi/10.1103/PhysRevLett.97.150404

[40] S. Bravyi, M. B. Hastings, and F. Verstraete, Phys. Rev. Lett. 97, 050401 (2006), URL http://link.aps.org/doi/10.1103/PhysRevLett.97.050401

[41] B. Altshuler, H. Krovi, and J. Roland, Proceedings of the National Academy of Sciences 107, 12446 (2010), http://www.pnas.org/content/107/28/12446.full.pdf+html, URL http://www.pnas.org/content/107/28/12446.abstract

[42] M. B. Hastings and X.-G. Wen, Phys. Rev. B 72, 045141 (2005), URL http://link.aps.org/doi/10.1103/PhysRevB.72.045141

[43] M. B. Hastings, ArXiv e-prints (2010), 1001.5280.

[44] A. Adams and S. Yaida, ArXiv e-prints (2011), 1102.2892.