

Many-body localization landscape

Shankar Balasubramanian,^{1,2} Yunxiang Liao³ , and Victor Galitski³¹*Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*²*Joint Quantum Institute, Department of Physics, University of Maryland, College Park, Maryland 20742-4111, USA*³*Joint Quantum Institute and Condensed Matter Theory Center, Department of Physics, University of Maryland, College Park, Maryland 20742-4111, USA*

(Received 24 September 2019; revised manuscript received 7 December 2019; published 6 January 2020)

We generalize the notion of “localization landscape,” introduced by M. Filoche and S. Mayboroda [*Proc. Natl. Acad. Sci. USA* **109**, 14761 (2012)] for the single-particle Schrödinger operator, to a wide class of interacting many-body Hamiltonians. The many-body localization landscape (MBLL) is defined on a graph in the Fock space, whose nodes represent the basis vectors in the Fock space and edges correspond to transitions between the nodes connected by the hopping term in the Hamiltonian. It is shown that in analogy to the single-particle case, the inverse MBLL plays the role of an effective potential in the Fock space. We construct a generalized discrete Agmon metric and prove Agmon inequalities on the Fock-state graph to obtain bounds on the exponential decay of the many-body wave functions in the Fock space. The corresponding construction is motivated by the semiclassical WKB approximation, but the bounds are exact and fully quantum mechanical. We then prove a series of locality theorems which establish where in the Fock space we expect eigenstates to localize. Using these results as well as the locator expansion, we establish evidence for the existence of many-body localized states for a wide class of lattice models *in any physical dimension* in at least a part of their Hilbert space. The key to this argument is the observation that in sharp contrast to the conventional locator expansion for the Green’s function, the locator expansion for the landscape function contains no resonances. For short-range hopping, which limits the connectivity of the Fock-state graph, the locator series is proven to be convergent and bounded by a simple geometric series. This, in combination with the discrete Agmon-type inequalities and the locality theorems, shows that localization for a fraction of the Hilbert space survives weak interactions and weak hopping at least for some realizations of disorder, but cannot prove or rule out localization of the entire Hilbert space. We qualitatively discuss potential breakdown of the locator expansion in the MBLL for long-range hopping and the appearance of a mobility edge in higher-dimensional theories.

DOI: [10.1103/PhysRevB.101.014201](https://doi.org/10.1103/PhysRevB.101.014201)

I. INTRODUCTION

There have been two major recent developments in the study of Anderson localization, a mature field pioneered by Anderson [1]. First, the breakthrough paper by Basko, Aleiner, and Altshuler [2–4] and followup theoretical and experimental research have demonstrated that single-particle localization can survive interactions in many-particle disordered systems [5–19]. This localization in the Fock space [2,20] was called many-body localization (MBL), which remains a fast-developing field, with a number of open questions and controversies still unresolved. A second development came as a surprise from joint work between the mathematics and physics community, in which Mayboroda and collaborators have developed a new, intuitive approach to single-particle Anderson localization [21–24]. Namely, it was proven that a simple structure $u(\mathbf{x})$, defined via $[-\nabla^2 + V(\mathbf{x})]u(\mathbf{x}) = 1$ (here V is the disorder potential) and called the localization landscape (LL) [21,22,25–27], when inverted represents an effective potential, where localized states up to a certain energy window *actually* reside. More specifically, it was shown that in contrast to the bare disorder potential, whose explicit form does not necessarily shed much light on the structure of localization, the LL carries this information in a straightforward

way. Valleys of the inverse localization landscape (which includes both the effect of disorder and hopping/kinetic energy in a nonperturbative way) form boundaries that confine the quantum particle [21,22]. A series of bounds and estimates have been proven to show exponential localization of the wave function in the wells of the inverse LL [23,24]. The LL has also been extended to more general single-particle models [28].

As far as the first major development goes, regarding MBL, there are currently four pieces of evidence for its existence. First, numerous exact diagonalization studies of finite-size systems, both spin chains with random disorder and one-dimensional interacting fermions in a quasiperiodic potential (e.g., the interacting Aubry-André model [29–31]), contain signatures of a transition [9–11], or possibly crossover, from the localized, nonergodic phase to a metallic, ergodic one. The numerical metrics used to characterize the many-body phases include calculating many-body inverse-participation ratio [32] and level statistics [6] or its proxies. However, these careful exact diagonalization studies can suffer from strong finite-size effects, which limit the usefulness of a numerical method in providing unambiguous evidence. Furthermore, recent and disconcerting data analysis by Šuntajs and collaborators [33] provided tentative evidence, based on data collapse

in a model widely believed to be many-body localized, that favors “many-body quantum chaos” (i.e., an ergodic phase with Wigner-Dyson level statistics), although this interpretation remains controversial. While the exact diagonalization studies can neither serve as a conclusive proof of MBL as a stable phase [34], nor rule it out, they do seem to provide a faithful description of physics actually seen in cold-atom experiments (at least in one dimension). The experimental studies, in particular by Schreiber *et al.* and Choi *et al.* (in both one and two dimensions) [14,15], Kondov *et al.* [16], and Smith *et al.* [17], represent a second piece of evidence. However, given the subtle mathematical nature of the question regarding the existence of MBL and a number of additional complications in experiments (the finite lifetime of many-body states, presence of the trap potential, etc.), experiments can only provide strong indication that MBL exists, but not a proof.

The third piece of evidence is an actual proof by Imbrie in one dimension [7]. This work is one of a few papers at a mathematical level of rigor addressing the MBL problem. However, it includes a physically reasonable but mathematically *ad hoc* assumption about the absence of level attraction. As was pointed out by Polkovnikov [35], coupling to a bath may in principle lead to level attraction. Indeed, there are examples in the literature where phenomena of this type have been reported [36,37]. However, it is unclear at this stage whether interactions may lead to an effective bath that would enable level attraction. It appears unlikely, but until this issue is mathematically settled, Imbrie’s proof is incomplete. Finally, the fourth piece of evidence for MBL comes from the original MBL work by Basko, Aleiner, and Altshuler [2–4]. The work is based on the locator expansion, which is also a method Anderson used in his pioneering paper [1]. The key idea of the locator expansion is simple. One starts with a strongly disordered model, with hopping “turned” off (which is trivially localized), and performs a series expansion in the hopping to obtain the Green’s functions. The method is straightforward, but suffers from divergences associated with resonances in the expansion, which does not necessarily converge. A renormalization scheme exists to resum the series, but it is not mathematically rigorous (unless the hopping takes place on a Cayley graph and specifically a Bethe lattice, in which rigorous statements have been known since the early work by Abou-Chacra, Anderson, and Thouless [38]). All in all, there exists overwhelming evidence for the existence of MBL, but a conclusive mathematical proof for physically relevant lattices is still lacking.

In contrast, the localization landscape of Mayboroda and collaborators [21–24] is a set of mathematically rigorous statements and bounds which physically correspond to single-particle Anderson localization. The key to the LL construction is a set of universal Agmon-type bounds proven in Refs. [39,40]. Agmon’s estimates have been known since 1979 and represent a set of *exact* inequalities for the eigenstates of the single-particle Schrödinger operator (which can be generalized to other operators). The construction of the estimates is motivated by a semiclassical WKB analysis, but Agmon’s estimates hold beyond the usual WKB regime. Particular Agmon’s bounds are potential specific and have been developed for a wide variety of problems such as

the double-well potential [41–43], a particle in a magnetic field [44,45], the Klein-Gordon equation [46,47], etc. A remarkable achievement of Mayboroda *et al.* [23,24,44] is that they have proven a set of universal Agmon’s estimates for the eigenstates of the Schrödinger operator in a random potential (while potential does not have to be random for the construction to hold, the bounds appear to be most useful when it is), but with the role of an effective potential played by the inverse landscape function $1/u(\mathbf{x})$ rather than the bare potential itself. Most importantly, these bounds predict exponential decay of eigenfunctions localized in the valleys of the LL. This statement is both mathematically rigorous and supported by numerical simulations, which provide an intuitive illustration of the structure of the spectrum. Note that the method works in an arbitrary dimension, and may contain a route to describing the delocalizing transition in three and higher dimensions via a percolation in the LL [48], although it has not been rigorously described yet.

This paper brings together many-body localization and localization landscape by constructing a many-body localization landscape (MBLL) on a graph in Fock space. We prove that the MBLL serves as an effective potential where many-body states are localized. First, we define the MBLL for a wide class of interacting spin Hamiltonians on an arbitrary-dimensional lattice or more generally on a graph with nearest-neighbor interactions (or equivalently hoppings) t . Next, we generalize Mayboroda’s arguments to an undirected discrete graph in the Fock space. The vertices of the graph correspond to many-body states and the edges to transitions enabled by the hopping term in the Hamiltonian. We show that the inverse many-body localization landscape $1/u_\alpha$ can be viewed as an effective potential in the Fock space, by deriving the following equation:

$$-t \sum_{\beta \in N(\alpha)} u_\beta u_\alpha \left(\frac{\psi_\beta}{u_\beta} - \frac{\psi_\alpha}{u_\alpha} \right) + \psi_\alpha = E' u_\alpha \psi_\alpha, \quad (1)$$

which represents a discrete Fock-space generalization of Eq. (5) in Ref. [23] by Mayboroda *et al.* involving a “weighted” Laplacian [α and β represent nodes of the Fock-space graph corresponding to the many-body states $|\alpha\rangle$ and $|\beta\rangle$, $N(\alpha)$ is the set of neighbors of node α , and E' is a shifted many-body energy as explained in the text]. The next step involves a generalization of Agmon’s bounds for this equation. It is accomplished by following semiclassical WKB-type intuition and a subsequent path-integral-like construction on the Fock-space graph. In analogy with the single-particle Agmon’s bound, a many-body Agmon’s bound can be proven, which is exact, and rigorously establishes (given certain properties of the landscape, as discussed in the text) exponential decay of the eigenfunctions in the valleys of the MBLL.

It is important to note that since exponential decay occurs in the valleys of the MBLL (or the mountains of the inverse MBLL), the Agmon bounds do not rule out the possibility of the wave function spreading out across all of the MBLL mountains. To better quantify this, we prove a series of locality theorems; these theorems roughly state that eigenstates in the global Fock-space graph are close to eigenstates defined in a particular well of the inverse MBLL if the corresponding

energies are close to each other. Therefore, eigenstates at a particular energy primarily localize in the set of wells whose corresponding spectra resonate at that energy. This condition can severely restrict the possible support of the eigenstate, assuming that we choose a reasonable realization of disorder.

While these exponential bounds are exact, they assume the existence of valleys in the MBLL, which is *a priori* not guaranteed. We then perturbatively explore the topology of the many-body localization landscape using an Anderson-type locator expansion. Importantly, to establish the existence of the valleys, we do not need information about the entire spectrum (which is provided by the Green's function in the usual locator expansion) and the knowledge of the MBLL suffices. Note that finding the latter is much simpler, as it involves inverting a matrix (as opposed to fully diagonalizing it, which in effect is required for the full Green's function and is more computationally costly). Another critical observation is that the locator expansion for MBLL *does not have the problem of resonances*, which has been the major roadblock in studies of the localization physics using conventional methods. This simplification allows us to circumvent the need for renormalization in the locator expansion, bound it by a geometric series, and hence prove convergence for a wide class of models. If we start with a strongly disordered Hamiltonian with no hopping, localization at the bare level is trivial, and the convergence of the series in conjunction with the locality theorem, essentially proves that *some* states remain many-body localized for particular realizations of disorder. However, this analysis cannot provide useful bounds for all states in the Hilbert space. It is also unclear at this stage what “fraction” of the Hilbert space can be proven to remain localized per this method (this seems to remain an open problem for the single-particle localization landscape as well). We discuss this critical point toward the end of the paper along with the structure of the Fock space graph corresponding to physical lattices of different dimension. Simple estimates for graph connectivity (which enters the many-body Agmon bounds) point to a route that can help explore mobility-edge physics in higher-dimensional MBL models.

II. MODEL AND THE FOCK-SPACE GRAPH

The canonical model that is used to describe localization is the Anderson model, which consists of noninteracting particles in a random potential. We are interested in understanding the effects of adding an interaction. The simplest model which incorporates nontrivial interaction is given by the Hamiltonian

$$\mathcal{H} = -t \sum_{\langle i,j \rangle} (\sigma_i^+ \sigma_j^- + \text{H.c.}) + \sum_i \epsilon_i n_i + V \sum_{\langle i,j \rangle} n_i n_j, \quad (2)$$

where σ^+ and σ^- are spin- $\frac{1}{2}$ raising and lowering operators, and $n = 2\sigma^z + 1$, which we refer to as the “density.”

The first term denotes hopping between nearest neighbors on graph \mathcal{G}_P with edge set E_P and vertex set V_P . For instance, this graph can be a one-dimensional lattice or a square lattice, but for the sake of generality we assume it can be arbitrary. We assume that $t \geq 0$; however, if the graph is bipartite, the sign of t is unimportant and can always be made negative by the mapping $\sigma^+ \rightarrow -\sigma^+$ and $\sigma^- \rightarrow -\sigma^-$ on one of the sub-graphs. This latter case where $t \leq 0$ is more conventionally

studied in the context of MBL. Call $|V_P| = L$ the number of vertices in \mathcal{G}_P (or correspondingly, the number of sites in the lattice). This is because such interactions may be effectively treated as a second hopping in our analysis.

The second term denotes the disorder potential. In the case of the Anderson model, $\epsilon_j \sim U[0, W]$, where W is a constant, but the specific distribution of disorder is unimportant for our construction. The last term denotes density-density interactions that occur between nearest neighbors in \mathcal{G}_P ; we may generalize the density-density interactions to quartic interactions $c_\alpha^\dagger c_\beta^\dagger c_\gamma c_\delta$, so long as the interaction strength is negative (although there exist graphs such that a similar particle-hole symmetry as described in the previous paragraph renders the sign unimportant).

An interpretation that will be useful in the subsequent section involves the visualization of this Hamiltonian in Fock space. Since the Hamiltonian in Eq. (2) is particle conserving, we restrict the Hilbert space to the set of states with N occupied sites. A tuple of occupation numbers at each site uniquely describes a basis vector in the Hilbert space. Associate each of these $M = \binom{L}{N}$ basis vectors to a vertex in a different graph \mathcal{G}_F . Edges connecting two such vertices in \mathcal{G}_F occur when the two configurations are connected by the hopping term. The resulting “hypergraph” is essentially an embedding of the physical graph (or lattice) \mathcal{G}_P in Fock space, and indicates the structure of the hopping. This is a visual that allows us to better understand MBL in Fock space, as well as help develop intuition for bounds we present throughout the rest of the paper. Figure 1 is an example of the Fock-space graph \mathcal{G}_F for a small system with $N = 3$ particles; the physical graph \mathcal{G}_P is a one-dimensional chain with $L = 5$ sites and periodic boundary conditions.

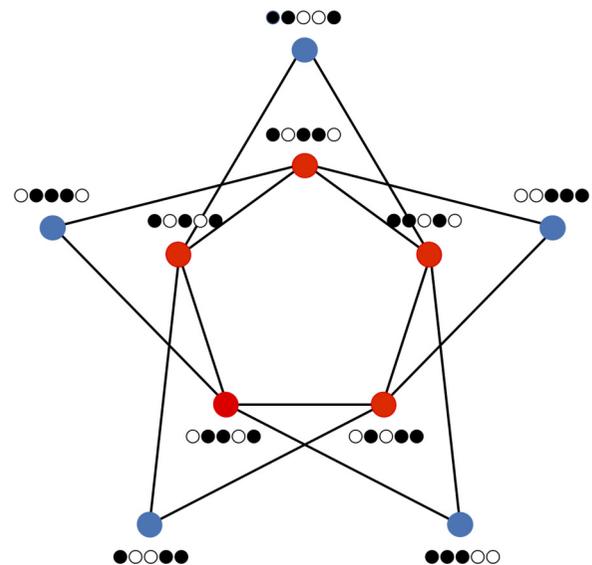


FIG. 1. The Fock-space graph \mathcal{G}_F for a one-dimensional chain with $L = 5$ and $N = 3$. For convenience, the occupation of each state is denoted next to each node; blackened circles correspond to occupied sites and unfilled circles correspond to empty sites. Notice that we assume periodic boundary conditions on the physical lattice in this construction, although this assumption is not needed in general.

Henceforth, we will use the latin alphabet (i, j) to denote nodes in the real-space (“physical”) graph \mathcal{G}_P and the greek alphabet (α, β) to denote nodes in the Fock-space graph \mathcal{G}_F .

III. DEFINING THE MBL

In this section, we briefly visit the premiere results of Mayboroda *et al.* [21,23,24,44] regarding single-particle localization via uniformly upper bounding the eigenstates by the so-called landscape function. We then show that the landscape formalism naturally extends to the many-body case with interactions. This allows us to conclude similar bounds to those derived by Mayboroda *et al.*

A. One-particle landscape

Filоче and Mayboroda [21] considered the Hamiltonian $\mathcal{H} = -\nabla^2 + V(\mathbf{x})$ to solve the eigenvalue problem $\mathcal{H}\psi(\mathbf{x}) = E\psi(\mathbf{x})$, $\mathbf{x} \in \Omega \subset \mathbb{R}^d$, and $\psi(\mathbf{x})|_{\partial\Omega} = 0$. Rather than solving the above Schrödinger equation, the crucial insight is to solve $\mathcal{H}u(\mathbf{x}) = 1$, where $u(\mathbf{x})$ is the landscape function. The resulting eigenstate satisfies $|\psi(\mathbf{x})| \leq Eu(\mathbf{x}) \sup_{\mathbf{x}} |\psi(\mathbf{x})|$. Thus, the landscape function can indicate whether states are localized or extended. Moreover, it was also argued that $1/u(\mathbf{x})$ acts as an effective confining potential [23,24], so at a given energy the eigenfunction can exponentially decay in classically forbidden regions. Our goal is to extend this intuition to many-body systems.

Along with continuum models, discrete single-particle systems were also analyzed. Here, the wave function is discretized into a vector of scalar entries, the Laplacian term of the Hamiltonian is discretized so that $(\nabla^2)\psi_i = \psi_{i+1} + \psi_{i-1} - 2\psi_i$, and the potential is treated as an onsite term. It was shown that solving the matrix equation $\mathcal{H}u = 1$ gives a landscape function that satisfies the same bounds as in the continuous case [22].

B. Many-body landscape

We now construct a similar landscape function for the Hamiltonian in Eq. (2). To do so, we notice that in the Fock-space representation, the Hamiltonian consists of a sum of two parts. The first part is the diagonal contribution due to the onsite and interaction terms. The second part is $-t$ times the adjacency matrix of graph \mathcal{G}_F . Thus, this Hamiltonian behaves similarly to the single-particle case albeit with long-range hopping. It thus suffices to show that a landscape function can be defined given the Hamiltonian in the Fock space.

As before, we restrict the Hilbert space to states with a fixed number of occupied sites; label the corresponding basis states by $|\alpha\rangle$ (occupation number representation). For an arbitrary superposition of such states, call v_α the coefficient of state $|\alpha\rangle$. Then, the Hamiltonian acting on state $|v\rangle = \sum_\alpha v_\alpha |\alpha\rangle$ gives

$$\langle \alpha | \mathcal{H} | v \rangle = -t \sum_{\beta \in N(\alpha)} v_\beta + \left(\sum_{(i,j)} V n_i^{(\alpha)} n_j^{(\alpha)} + \sum_i \epsilon_i n_i^{(\alpha)} \right) v_\alpha, \quad (3)$$

where $n_i^{(\alpha)}$ denotes the occupation number of site i for state α , and $N(\alpha)$ is the set of neighbors of α in graph \mathcal{G}_F .

For the purposes of the following proofs, we also need to impose Dirichlet zero-boundary conditions in the sense of the graph \mathcal{G}_F . This can be accomplished by augmenting \mathcal{G}_F to \mathcal{G}'_F and connecting an additional node α' to each node α . We enforce the condition $v_{\alpha'} = 0$, which results in the above equation not changing. At the end, we perform a restriction of \mathcal{G}'_F to \mathcal{G}_F without changing any of the main results. In this section, we will also define $N'(\alpha)$ to be the neighbors of α in \mathcal{G}'_F , and it is clear that $|N'(\alpha)| = |N(\alpha)| + 1$.

The following lemma gives conditions for when we should expect v_α to be positive. The proof method is analogous to that of the maximum principle for harmonic functions in a bounded domain.

Lemma 1. Suppose $\langle \alpha | \mathcal{H} + K | v \rangle \geq 0$ where K is a positive constant for all α and $v_{\alpha'} = 0$ for all α' . Then, if

$$K - t[|N(\alpha)| + 1] + \sum_{(i,j)} V n_i^{(\alpha)} n_j^{(\alpha)} + \sum_i \epsilon_i n_i^{(\alpha)} \geq 0 \quad (4)$$

for all α , we have that $v_\alpha \geq 0$. Furthermore, if there exists a β for which $\langle \beta | \mathcal{H} + K | v \rangle > 0$, then $v_\alpha > 0$ for all α .

Proof. Suppose that, for the sake of contradiction, the global minimum $\text{argmin}_{\alpha'} v_\alpha$ does not occur within the set of α' nodes. With an abuse of notation, call v_α the global minimum. Then, it follows that v_α is a local minimum and $v_\beta \geq v_\alpha$, where $\beta \in N'(\alpha)$. From this, we find that the following weaker condition must be satisfied:

$$\sum_{\beta \in N'(\alpha)} (v_\beta - v_\alpha) \geq 0. \quad (5)$$

This condition enforces that v_α is a local minimum with respect to its neighboring nodes in \mathcal{G}'_F . Next, since $\langle \alpha | \mathcal{H} + K | v \rangle \geq 0$, it follows that

$$t \sum_{\beta \in N'(\alpha)} v_\beta \leq K v_\alpha + \left(\sum_{(i,j)} V n_i^{(\alpha)} n_j^{(\alpha)} + \sum_i \epsilon_i n_i^{(\alpha)} \right) v_\alpha. \quad (6)$$

Thus, combining with Eq. (5), we find

$$0 \leq t \sum_{\beta \in N'(\alpha)} (v_\beta - v_\alpha) \quad (7)$$

$$\leq K v_\alpha - t |N'(\alpha)| v_\alpha + \left(\sum_{(i,j)} V n_i^{(\alpha)} n_j^{(\alpha)} + \sum_i \epsilon_i n_i^{(\alpha)} \right) v_\alpha. \quad (8)$$

To remind, note that $|N(\alpha)| + 1$ is the number of neighbors of node α in \mathcal{G}'_F , which is one greater than the number of neighbors in \mathcal{G}_F due to enforcement of the Dirichlet boundary condition. Simple rearrangement gives

$$0 \leq \left(K - t |N'(\alpha)| + \sum_{(i,j)} V n_i^{(\alpha)} n_j^{(\alpha)} + \sum_i \epsilon_i n_i^{(\alpha)} \right) v_\alpha. \quad (9)$$

The quantity in the parentheses is always positive by assumption, so $v_\alpha \geq 0$. However, this contradicts the assumption that v_α is a global minimum since it is at least equal to the value of $v_{\alpha'}$ for all α' . Thus, v_α only acquires a minimum in the nodes α' , and it follows that $v_\alpha \geq 0$ for all nodes in \mathcal{G}'_F .

Proving strict positivity follows directly. Suppose $v_\alpha = 0$ for some α . Then, reading from Eq. (6), we find

$$t \sum_{\beta \in N'(\alpha)} v_\beta \leq K v_\alpha + \left(\sum_{(i,j)} V n_i^{(\alpha)} n_j^{(\alpha)} + \sum_i \epsilon_i n_i^{(\alpha)} \right) v_\alpha = 0. \quad (10)$$

Since we proved that $v_\alpha \geq 0$ for all α , we have that $v_\beta = 0$ for $\beta \in N'(\alpha)$. We can represent this as the diffusion of an action on \mathcal{G}'_F : we set v_α equal to zero in node α and spread this action to all neighboring nodes. We then inductively repeat. Since \mathcal{G}'_F is irreducible, there exists an iteration in which all nodes will equal zero, thus, all $v_\alpha = 0$. Therefore, $\langle \alpha | \mathcal{H} + K | v \rangle = 0$. Hence, taking the contrapositive of this logic combined with $v_\alpha \geq 0$ finishes the proof. ■

Notice that the constant K equals

$$K = - \min_\alpha \left[\sum_{(i,j)} V n_i^{(\alpha)} n_j^{(\alpha)} + \sum_i \epsilon_i n_i^{(\alpha)} - t |N'(\alpha)| \right]. \quad (11)$$

This corresponds to a constant shift of the Hamiltonian, which does not affect the eigenstates. Note that Lyra *et al.* [22] had to perform a similar shift to impose positive onsite potentials. Next, we show that this conclusion implies that the elements of $(\mathcal{H} + K)^{-1}$ are positive:

Lemma 2. Consider the restriction of matrix $\mathcal{H} + KI \triangleq \mathcal{H} + K$ to states in \mathcal{G}_F , where I denotes the identity matrix. The inverse of the restriction of $\mathcal{H} + K$ is entrywise positive.

Proof. Suppose $\langle \alpha | \mathcal{H} + K | v \rangle = f_\alpha$ for $f_\alpha \geq 0$. By Lemma 1, $v_\alpha \geq 0$. Inverting this equation and solving for v_α gives

$$v_\alpha = \sum_\beta (\mathcal{H} + K)_{\alpha\beta}^{-1} f_\beta. \quad (12)$$

Calling $f_\beta = \delta_{\gamma\beta}$, then $0 \leq v_\alpha = (\mathcal{H} + K)_{\alpha\gamma}^{-1}$. Repeating for all α and γ finishes the proof. Strict positivity follows from the fact that the above choice of f satisfies $f_\alpha > 0$ for at least one value of α . ■

In this next lemma, we show that the spectrum of $\mathcal{H} + K$ is strictly positive.

Lemma 3. The matrix $\mathcal{H} + K$ has positive eigenvalues.

Proof. From Lemma 1, if $(\mathcal{H} + K)|v\rangle = 0$, then $v_\alpha \geq 0$. Suppose the null space of $\mathcal{H} + K$ is not trivial; then it must contain some state $|v\rangle \neq 0$. Thus, both $|v\rangle$ and $-|v\rangle$ are in the null space, which contradicts Lemma 1 since v_α cannot be negative; it follows that the null space of $\mathcal{H} + K$ is trivial. This implies that there is no eigenvalue equal to zero.

It is trivial to show that Lemma 1 is satisfied for any choice of $K' > K$, and $\mathcal{H} + K'$ does not have any zero eigenvalues for $K' > K$. Suppose, for the sake of contradiction, that there exists a negative eigenvalue $-\lambda$ of $\mathcal{H} + K'$. Then, setting $K'' = K' + \lambda \geq K'$, it is clear that $\mathcal{H} + K''$ has a zero eigenvalue. This contradicts the fact that $\mathcal{H} + K'$ does not have any zero eigenvalues for $K' > K$. This completes the proof. ■

We next define the MBLL to be the solution to

$$(\mathcal{H} + K)|u\rangle = |1\rangle, \quad (13)$$

where $|1\rangle$ denotes the vector of all ones. Written out more explicitly, the defining equation for the MBLL is given by

$$-t \sum_{\beta \in N(\alpha)} u_\beta + \left(\sum_{(i,j)} V n_i^{(\alpha)} n_j^{(\alpha)} + \sum_i \epsilon_i n_i^{(\alpha)} \right) u_\alpha = 1. \quad (14)$$

In an identical manner to the proof in Lyra *et al.* [22], we arrive at the following bound on the eigenstates via the landscape function:

Theorem 1. Assume $(\mathcal{H} + K)|\psi\rangle = (E + K)|\psi\rangle$ with $\psi_{\alpha'} = 0$, where E is the many-body energy, and K is defined in Eq. (11). It follows that for all α ,

$$\frac{|\psi_\alpha|}{\max_\beta |\psi_\beta|} \leq (E + K) u_\alpha. \quad (15)$$

Proof. We use the two lemmas that have been derived to write

$$\begin{aligned} |\psi_\alpha| &= (E + K) \left| \sum_\beta (\mathcal{H} + K)_{\alpha\beta}^{-1} \psi_\beta \right| \\ &\leq (E + K) \max_\beta |\psi_\beta| \sum_\beta (\mathcal{H} + K)_{\alpha\beta}^{-1} \\ &= (E + K) \max_\beta |\psi_\beta| u_\alpha, \end{aligned} \quad (16)$$

where in the first line we use Lemma 3, in the second line we use the fact that $(\mathcal{H} + K)_{\alpha\beta}^{-1}$ is positive (from Lemma 2), and in the third line we use the definition of u_α . ■

Surprisingly, Mayboroda's single-particle bound carries over in the many-body case even in the presence of interactions. However, the visualization of this bound is more complex. Rather than viewing the landscape function on a standard plot to determine domains of localization, one must view the landscape function with respect to the Fock-space graph \mathcal{G}_F . Localization occurs in *connected* components of \mathcal{G}_F with large values of u which are surrounded by valleys with small values of u .

Also notice that our bounds depend on the choice of K . While K is arbitrary in principle, it is lower bounded. Choosing too large of K results in landscape functions that are not good representatives of the localization regions for the eigenstates. However, it is not clear whether there exists an optimal value of K that gives the best possible upper bounds.

From this point on, we denote $E' = E + K$ for simplicity and readability. We will also change notation by calling $N(\alpha)$ the set of neighbors of node α in \mathcal{G}_F .

IV. MBLL AS AN EFFECTIVE POTENTIAL

The bound derived in the last section gives evidence that the MBLL informs us about possible localization in the Fock space. This is consistent with the general qualitative picture of many-body localization, as discussed by Basko, Aleiner, and Altshuler [2–4], who proposed studying MBL in terms of the localization of eigenstates in Fock space. However, to understand whether eigenfunctions exponentially decay outside of localized regions, we must borrow intuition from the single-particle case. Here, Mayboroda *et al.* [23,24] showed that $1/u$ behaves like an effective potential that confines the

eigenstates. This effective potential is often more informative than considering the original onsite potential, as it incorporates both disorder and the kinetic energy in a nonperturbative way and more accurately identifies regions where eigenstates localize. We want to understand whether this analogy extends to the MBL, as this gives significant insight into when and where in the Fock space we would expect strong localization to occur. The result we show is that the rate of decay of eigenfunctions depends on the connectivity of a particular portion of \mathcal{G}_F : if a node has fewer neighbors, the resulting states have large contiguous regions of occupied sites and will more strongly localize.

The method we use to analyze this decay was first developed by Agmon [39,40] several decades ago. It essentially formalizes a semiclassical WKB analysis of wave-function tunneling to exact upper bounds on the magnitude of the wave function, which are known as *Agmon inequalities*.

A. Effective Schrödinger equation via MBL

We will first proceed to rewrite the eigenvalue equation in terms of the MBL. Let us start by noting that the eigenstates $|\psi\rangle = \sum_{\alpha} \psi_{\alpha} |\alpha\rangle$ of the many-body Hamiltonian satisfy

$$-t \sum_{\beta \in N(\alpha)} \psi_{\beta} + \left(\sum_{(i,j)} V n_i^{(\alpha)} n_j^{(\alpha)} + \sum_i \epsilon_i n_i^{(\alpha)} \right) \psi_{\alpha} = E \psi_{\alpha}. \quad (17)$$

For brevity, we call the term in the parentheses V_{α} . Next, we perform the substitution $\psi_{\alpha} = \Psi_{\alpha} u_{\alpha}$. We may then write

$$-t \sum_{\beta \in N(\alpha)} \Psi_{\beta} u_{\beta} + V_{\alpha} \Psi_{\alpha} u_{\alpha} = E \Psi_{\alpha} u_{\alpha}, \quad (18)$$

which can be rearranged as

$$\left[-t \sum_{\beta \in N(\alpha)} u_{\beta} + (V_{\alpha} + K) u_{\alpha} \right] \Psi_{\alpha} - t \sum_{\beta \in N(\alpha)} u_{\beta} (\Psi_{\beta} - \Psi_{\alpha}) = E' \Psi_{\alpha} u_{\alpha}. \quad (19)$$

The quantity in the brackets equals 1 by definition of the MBL, and if we rewrite the equation in terms of ψ_{α} , we obtain the equation

$$-t \sum_{\beta \in N(\alpha)} u_{\beta} u_{\alpha} \left(\frac{\psi_{\beta}}{u_{\beta}} - \frac{\psi_{\alpha}}{u_{\alpha}} \right) + \psi_{\alpha} = E' u_{\alpha} \psi_{\alpha}. \quad (20)$$

This is a more revealing form, as the first term is equal to t times a weighted Laplacian matrix L acting on the state $|\psi/u\rangle = \sum_{\alpha} (\psi_{\alpha}/u_{\alpha}) |\alpha\rangle$. A Laplacian matrix with weighted edges $w_{\alpha\beta}$ has elements equal to

$$L_{\alpha\beta} = \begin{cases} -w_{\alpha\beta}, & \text{for } \alpha \sim \beta \\ \sum_{\alpha \in N(\beta)} w_{\alpha\beta}, & \text{for } \alpha = \beta \end{cases} \quad (21)$$

where $\alpha \sim \beta$ means that nodes α and β are connected in \mathcal{G}_F . This quantity is a discrete generalization of the differential operator for a graph whose distance metric is given by the weights $w_{\alpha\beta}$. In the above equation, the weights of the Laplacian are $w_{\alpha\beta} = u_{\alpha} u_{\beta}$, and so are symmetric with respect to interchanging α and β . Moreover, they are also positive as per

the analysis of the previous section. Write the left-hand side of Eq. (20) as $M|\psi\rangle$, where $M_{\alpha\beta} = tL_{\alpha\beta}/u_{\beta} + \delta_{\alpha\beta}$. Next, we apply $\langle \psi/u |$ to the matrix version of the above equation:

$$\begin{aligned} \left\langle \frac{\psi}{u} \middle| M \middle| \psi \right\rangle &= t \sum_{\langle \alpha, \beta \rangle} u_{\alpha} u_{\beta} \left(\frac{\psi_{\alpha}}{u_{\alpha}} - \frac{\psi_{\beta}}{u_{\beta}} \right)^2 + \sum_{\alpha} \frac{\psi_{\alpha}^2}{u_{\alpha}} \\ &= E' \sum_{\alpha} \psi_{\alpha}^2, \end{aligned} \quad (22)$$

where for the first term we use the well-known identity

$$\langle x | L | x \rangle = \sum_{\langle \alpha, \beta \rangle} w_{\alpha\beta} (x_{\alpha} - x_{\beta})^2, \quad (23)$$

which holds for any weighted Laplacian with $w_{\alpha\beta} = w_{\beta\alpha}$. Equation (22) indicates that $1/u$ also acts as a confining potential. The first term is the expected kinetic energy of the particle, which has been renormalized by the MBL. The second term denotes an expectation of a ‘‘potential energy,’’ if $1/u$ is treated as an effective potential. However, this does not necessarily tell us any information of the behavior of eigenfunctions under this effective potential. To analyze the decay of eigenfunctions, we will need to generalize Agmon estimates to the Fock-space graph as discussed at the beginning of the section.

B. Agmon estimates for eigenfunction decay

The intuition behind the derivation of Agmon estimates is to consider the behavior of an eigenstate modulated by an exponentially large amplitude and show that this new function is bounded in a certain sense. This immediately shows that the eigenstate exponentially decays in particular regions. As we will show, this exponential amplitude is obtained by minimizing an effective action on \mathcal{G}_F , and also depends on the connectivity of a particular node in the graph.

We first prove some identities obtained through brute-force computation.

Lemma 4. Given a choice of $S_{\alpha} \in \mathbb{R}$, the following equation holds:

$$\sum_{\alpha} \psi_{\alpha}^2 e^{2S_{\alpha}} \left[\frac{1}{u_{\alpha}} - E' - \frac{t}{2} \sum_{\beta \in N(\alpha)} (e^{S_{\beta} - S_{\alpha}} - 1)^2 \right] \leq 0. \quad (24)$$

Proof. We first compute the seemingly unrelated quantity $\langle \frac{\psi}{u} e^{2S} | M | \psi \rangle$. From manipulating the first line of Eq. (22), this is equal to

$$\begin{aligned} \left\langle \frac{\psi}{u} e^{2S} \middle| M \middle| \psi \right\rangle &= t \sum_{\langle \alpha, \beta \rangle} u_{\alpha} u_{\beta} \left(\frac{\psi_{\alpha}}{u_{\alpha}} e^{2S_{\alpha}} - \frac{\psi_{\beta}}{u_{\beta}} e^{2S_{\beta}} \right) \\ &\quad \times \left(\frac{\psi_{\alpha}}{u_{\alpha}} - \frac{\psi_{\beta}}{u_{\beta}} \right) + \sum_{\alpha} \frac{\psi_{\alpha}^2}{u_{\alpha}} e^{2S_{\alpha}}. \end{aligned} \quad (25)$$

Furthermore, manipulating the second line of Eq. (22), in particular utilizing the relationship between $M|\psi\rangle$ and the many-body energy E , the above quantity can be rewritten in terms of E' :

$$\left\langle \frac{\psi}{u} e^{2S} \middle| M \middle| \psi \right\rangle = E' \sum_{\alpha} \psi_{\alpha}^2 e^{2S_{\alpha}}. \quad (26)$$

By equating both of these equations and manipulating with simple algebra, we find the following identity:

$$t \sum_{(\alpha, \beta)} u_\alpha u_\beta \left(\frac{\psi_\alpha}{u_\alpha} e^{S_\alpha} - \frac{\psi_\beta}{u_\beta} e^{S_\beta} \right)^2 + \sum_\alpha \frac{\psi_\alpha^2}{u_\alpha} e^{2S_\alpha} - t \sum_{(\alpha, \beta)} \psi_\alpha \psi_\beta (e^{S_\alpha} - e^{S_\beta})^2 = E' \sum_\alpha \psi_\alpha^2 e^{2S_\alpha}. \quad (27)$$

Using the fact that u_α is positive, the first term is positive and the identity may be written as the inequality

$$\sum_\alpha \psi_\alpha^2 e^{2S_\alpha} \left(\frac{1}{u_\alpha} - E' \right) - t \sum_{(\alpha, \beta)} \psi_\alpha \psi_\beta (e^{S_\alpha} - e^{S_\beta})^2 \leq 0. \quad (28)$$

Finally, we note the inequality $\psi_\alpha \psi_\beta \leq \frac{1}{2}(\psi_\alpha^2 + \psi_\beta^2)$ true for any ψ and the fact that the second sum in the above equation is symmetric in the indices α and β . Applying this inequality and symmetrizing the second sum with respect to α and β , we find Eq. (24), where we add an additional factor of 2 from splitting the sum over edges to a double sum over nodes and avoiding double counting pairs of nodes. ■

Notice that the previous Lemma 4 is general: it holds for an arbitrary choice of S . Next, we choose S so that the cumulative amplitude of the wave function in the classically forbidden region is “small.” By classically forbidden, we mean regions for which the many-body energy is smaller than the value of the effective potential $1/u$. We make rigorous this notion in the following lemma.

Lemma 5. There exists a choice S_α for all α such that for any $1 > \epsilon > 0$, the energy inequality can be refined to

$$\sum_\alpha \psi_\alpha^2 e^{2S_\alpha} \left(\frac{1}{u_\alpha} - E' \right) \leq (1 - \epsilon) \sum_\alpha \psi_\alpha^2 e^{2S_\alpha} \left(\frac{1}{u_\alpha} - E' \right)_+, \quad (29)$$

where $(x)_+ = \max(x, 0)$.

Proof. First, define \bar{U} to be the set of nodes α which reside in the classically allowed region, where $\frac{1}{u_\alpha} - E' < 0$. Suppose that S_α satisfies

$$S_\alpha = \inf_{\substack{p(0) \in \bar{U} \\ p(-1) = \alpha}} \sum_k |\mathcal{L}_{p(k-1), p(k)}|, \quad (30)$$

where the infimum is over paths from a point in a classically allowed region to α . By path, we mean a list of nodes such that each consecutive pair has an edge in \mathcal{G}_F ; this is denoted by the vector p where $p(i)$ is the i th element in the path. The notation $p(-1)$ denotes the last element in the path. We also assume that $|\mathcal{L}_{\alpha, \beta}| = |\mathcal{L}_{\beta, \alpha}|$. It straightforwardly follows that, for $\beta \in N(\alpha)$,

$$S_\beta \leq S_\alpha + |\mathcal{L}_{\alpha, \beta}|, \quad (31)$$

and similarly $S_\beta \geq S_\alpha - |\mathcal{L}_{\beta, \alpha}|$. Therefore, $|S_\beta - S_\alpha| \leq |\mathcal{L}_{\alpha, \beta}| = |\mathcal{L}_{\beta, \alpha}|$. This essentially states that S satisfies a triangle inequality, and therefore is a valid metric on \mathcal{G}_F . Then, we may write

$$\begin{aligned} \frac{t}{2} \sum_{\beta \in N(\alpha)} (e^{S_\beta - S_\alpha} - 1)^2 &\leq \frac{t}{2} \sum_{\beta \in N(\alpha)} (e^{|S_\beta - S_\alpha|} - 1)^2 \\ &\leq \frac{t}{2} \sum_{\beta \in N(\alpha)} (e^{|\mathcal{L}_{\alpha, \beta}|} - 1)^2. \end{aligned} \quad (32)$$

Next, we make the following choice for $|\mathcal{L}_{\alpha, \beta}| = |\mathcal{L}_{\beta, \alpha}|$:

$$|\mathcal{L}_{\alpha, \beta}| = \min_{k=\alpha, \beta} \left[\log \left(1 + \sqrt{\frac{2(1-\epsilon)}{t|N(k)|} \left(\frac{1}{u_k} - E' \right)_+} \right) \right]. \quad (33)$$

With this choice, we may further bound Eq. (32) by

$$\begin{aligned} \frac{t}{2} \sum_{\beta \in N(\alpha)} (e^{|\mathcal{L}_{\alpha, \beta}|} - 1)^2 &= (1 - \epsilon) \sum_{\beta \in N(\alpha)} \min_{k=\alpha, \beta} \frac{(u_k^{-1} - E')_+}{|N(k)|} \\ &\leq (1 - \epsilon) \left(\frac{1}{u_\alpha} - E' \right)_+. \end{aligned} \quad (34)$$

Substituting this inequality into the last term on the left-hand side of Eq. (24), we find the desired refinement of Eq. (24):

$$\sum_\alpha \psi_\alpha^2 e^{2S_\alpha} \left(\frac{1}{u_\alpha} - E' \right) \leq (1 - \epsilon) \sum_\alpha \psi_\alpha^2 e^{2S_\alpha} \left(\frac{1}{u_\alpha} - E' \right)_+. \quad (35)$$

The choice of variables S and \mathcal{L} suggests that the function S_α can be viewed as a saddle-point approximation on an effective action S which is a sum of Lagrangian terms along a path from a node in a classically allowed region to α . This saddle-point value of the action will dictate the decay rate of the eigenfunction as defined on \mathcal{G}_F . The statement that was just derived roughly states that the sum of $\psi_\alpha^2 e^{2S_\alpha}$ over the classically allowed region \bar{U} is greater than some constant times the sum of $\psi_\alpha^2 e^{2S_\alpha}$ over the classically disallowed region U . Since $S_\alpha = 0$ for $\alpha \in \bar{U}$, the sum of the eigenfunctions in U multiplied by a term exponential in the action S is less than the sum of eigenfunctions in \bar{U} without this exponential amplification. Thus, in a sense, the value of an eigenfunction in U must be exponentially small for its cumulative sum in U to be small. This intuition is made explicit in the following theorem.

Theorem 2. Define $\bar{u} = \max_{\beta \in \bar{U}} u_\beta$. For all nodes α in the classically forbidden region, an eigenstate with energy E' satisfies

$$|\psi_\alpha| \leq \sqrt{\frac{E' \bar{u} u_\alpha - u_\alpha}{\bar{u} - E' \bar{u} u_\alpha}} \left(\frac{e^{-S_\alpha}}{\sqrt{\epsilon}} \right), \quad (36)$$

with S_α defined in Eq. (30) and $1 > \epsilon > 0$ a positive constant.

Proof. Based on the result in the previous lemma, we may split the sum over nodes in U and nodes in \bar{U} . This gives

$$\sum_{\alpha \in \bar{U}} \psi_\alpha^2 e^{2S_\alpha} \left(\frac{1}{u_\alpha} - E' \right) + \epsilon \sum_{\alpha \in U} \psi_\alpha^2 e^{2S_\alpha} \left(\frac{1}{u_\alpha} - E' \right) \leq 0. \quad (37)$$

Since $S_\alpha = 0$ for $\alpha \in \bar{U}$, we obtain the following chain of inequalities:

$$\begin{aligned} \epsilon \sum_{\alpha \in U} \psi_\alpha^2 e^{2S_\alpha} \left(\frac{1}{u_\alpha} - E' \right) &\leq \sum_{\alpha \in \bar{U}} \psi_\alpha^2 \left(E' - \frac{1}{u_\alpha} \right) \\ &\leq \left(\sum_{\alpha \in \bar{U}} \psi_\alpha^2 \right) \max_{\alpha \in \bar{U}} \left(E' - \frac{1}{u_\alpha} \right) \\ &\leq \max_{\alpha \in \bar{U}} \left(E' - \frac{1}{u_\alpha} \right). \end{aligned} \quad (38)$$

Each term in the sum on the left-hand side of the above equation must be strictly less than or equal to the right-hand side. Call $\bar{u} = \max_{\beta \in \bar{U}} u_\beta$. This results in the bound

$$|\psi_\alpha| \leq \sqrt{\frac{E' - \frac{1}{\bar{u}}}{\frac{1}{u_\alpha} - E'}} \left(\frac{e^{-S_\alpha}}{\sqrt{\epsilon}} \right) = \sqrt{\frac{E' \bar{u} u_\alpha - u_\alpha}{\bar{u} - E' \bar{u} u_\alpha}} \left(\frac{e^{-S_\alpha}}{\sqrt{\epsilon}} \right), \quad (39)$$

where $\alpha \in U$. \blacksquare

Therefore, the eigenfunction decays exponentially in the minimum value of the action. The coefficient before the exponential has an intuitive meaning: it is the square root of the ratio of the height of the deepest classical well to the depth of penetration through the classically forbidden region, measured with respect to the many-body energy E' . For instance, if the height to the deepest classical well is small, which occurs when the eigenstate is toward the bottom of the spectrum, then the coefficient scales as a constant, and decreases as the penetration depth into the classically forbidden region increases. In the standard WKB formalism, one finds the following semiclassical decay for the wave function:

$$\psi(x) \propto \frac{1}{\sqrt{p(x)}} \exp\left(-\int p(x) dx\right). \quad (40)$$

Here, we find a similar expression, assuming that p denotes a “graph momentum,” equal to $p_\alpha \sim \sqrt{(1/u_\alpha - E')/|N(\alpha)|}$. The quantity within the exponential is a sum of Lagrangian terms $\mathcal{L} \sim \log(1 + p)$ along a path; this approximately reduces to a discrete WKB formula when the “graph momentum” is small.

Next, we would like to understand the nature of the exponential decay in the action. Each Lagrangian term scales like $\sim \sqrt{(1/u_\alpha - E')/|N(\alpha)|}$, assuming that $|N(\alpha)|$ scales with system size in Eq. (32). Thus, the decay rate is crucially dictated by the size of $N(\alpha)$.

For simplicity, let us assume \mathcal{G}_P is a one-dimensional chain with nearest-neighbor hopping. Then, $|N(\alpha)|$ has the following behavior:

$$\Theta(1) < |N(\alpha)| < \Theta(N). \quad (41)$$

We adopt the $\Theta(\cdot)$ notation to mean that if $g(N) = \Theta(f(N))$, then $c_1 f(N) < g(N) < c_2 f(N)$ for large N and constants c_1 and c_2 . For states with a large domain of occupied sites, $|N(\alpha)|$ will scale like a constant. In the worst case, if occupied and filled sites alternate, $|N(\alpha)|$ will scale linearly in N . Thus, in the worst case, the action may only increase by $O(1/\sqrt{N})$ as one hops farther out in \mathcal{G}_F .

Suppose we deal with these worst case states where $|N(\alpha)| \propto N$. Further fix the many-body energy E' , and denote \mathcal{H}_1 to the space of states living in the classically allowed regions. As we hop along \mathcal{G}_F and into the classically forbidden region, the barrier height is at least a constant (we make an implicit assumption that we are considering typical paths in which the probability of encountering a node with barrier height scaling inversely with system size is small). Suppose the minimizing path in Eq. (30) to some node has length n into the classically disallowed region; then, the action will scale

like

$$S \gtrsim \sum_{i=1}^n \log\left(1 + \sqrt{\frac{k_i}{N}}\right) \sim \frac{n}{N^{1/2}}, \quad (42)$$

where k_i are unimportant constants. Then, if $n = \Omega(N^{1/2})$, the action will begin to scale as a function of the system size. Note that this bound is weak because the Lagrangian depends on $1/u - E'$, which increases along the minimizing path. When this energy difference scales as $\Theta(N)$, we obtain exponential decay, which conceivably occurs for a fraction of the Hilbert space.

With this intuition, we define three regions. Each region covers nodes of the Fock-space graph \mathcal{G}_F , but should be interpreted as the Hilbert space (i.e., the span of states associated with nodes in the region), denoted by \mathcal{H} :

(i) \mathcal{H}_1 : Fixing a value of the many-body energy E' , this region covers classically allowed regions of \mathcal{G}_F where the wave function behaves like a constant.

(ii) \mathcal{H}_2 : This region is an annulus of thickness $N^{1/2}$ around \mathcal{H}_1 regions. It corresponds to the intermediate region in which there exist portions that do not decay at a rate scaling with the system size. In the thermodynamic limit, we will not observe decay in these portions. However, there may exist several portions of \mathcal{H}_2 where we do find decay scaling with system size.

(iii) \mathcal{H}_3 : This region is outside of \mathcal{H}_1 and \mathcal{H}_2 , in which eigenstates almost surely decay at a rate that scales with some increasing function in the system size. For instance, if $n = \Theta(N^{1/2} \log N)$, then the eigenstates obey a power-law decay. Eigenstates will decay exponentially when $1/u - E' = \Theta(N)$.

These regions are illustrated in Fig. 2. In addition, since the degree of each node is at most N , the dimension \mathcal{H}_2 is roughly $N^{\sqrt{N}}$, which is of zero measure with respect to the dimension of the full Hilbert space, roughly equal to $e^{NH(\nu)}$ [here, we make the assumption that N is a fraction of L , so that $\binom{L}{N} \sim e^{NH(\nu)}$ where ν is the filling fraction and $H(\nu) = -\nu \log \nu - (1 - \nu) \log(1 - \nu)$]. For \mathcal{H}_1 regions centered about states with degree that is $o(N)$, the estimates derived above can be refined by substituting N with the degree. Thus, based on the bound in Eq. (41), there can exist some regions of the full Hilbert space that exhibit strong exponential localization.

It is worth noting what changes as the underlying dimension increases. For instance, if \mathcal{G}_P is a d -dimensional lattice, then the topology of \mathcal{G}_F becomes far more connected. In particular, the degree distribution follows

$$\Theta\left(N^{1-\frac{1}{d}}\right) < |N(\alpha)| < \Theta(N), \quad (43)$$

with minimum degree nodes having the property that linearly sized domains are occupied, and maximum degree nodes having a large number of constant-size domains. Note that the minimum degree corresponds to an area law, while the maximum degree corresponds to a volume law. The bounds on decay rate developed in the preceding paragraph follow immediately by substituting N with the degree. The higher connectivity present in these graphs could point to the existence of a mobility edge in higher dimensions at least in the Anderson model, as will be further elaborated in Sec. V.

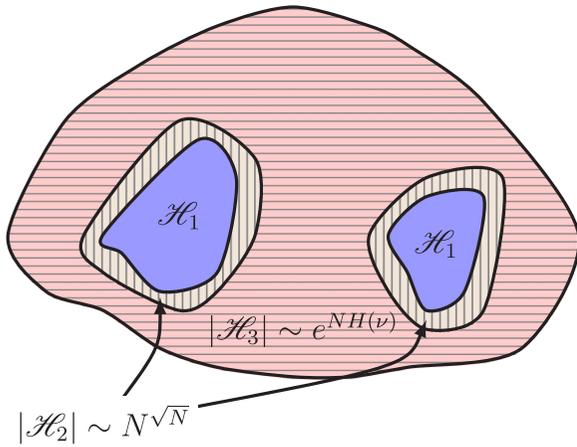


FIG. 2. The breakup of the full Hilbert space into the three regions described in the text. One can visualize the Fock-space graph \mathcal{G}_F as embedded into this Hilbert space. The blue region (no texture) is the classically allowed area formed from a fixed many-body energy. As shown, for sufficiently small energies, the classical regions will form islands in the full Hilbert space. In the brown region (vertical lines), eigenstates decay, but at a rate does not increase with increasing system size. The number of states in these regions is of order $N^{N^{1/2}}$. In the red region (horizontal lines), the eigenstates decay at a rate scaling with an increasing function of system size; eigenstates decay exponentially for a constant fraction of the red region.

C. Predicting regions of localization

In the previous section, we have addressed the question of what occurs in classically forbidden regions. Now, we explore how the eigenstates behave in the classically allowed regions. In the single-particle case, it was argued that so long as these regions are far apart and not resonant, the eigenstates (toward the bottom of the spectrum) tend to localize at only one of these regions, and have exponentially small amplitudes on the others [24]. Having a similar result in the many-body case would support any assertion of MBL.

To proceed, it is first important for us to precisely define the regions of interest. As before, denote \bar{U} the set of classically allowed regions which satisfy $1/u_\alpha - E' \leq 0$. These regions/wells form sets of disjoint connected regions in \mathcal{G}_F , which we label by the index ℓ . Let us fix a particular such well $\bar{U}(\ell)$; compute

$$\bar{S} = \inf\{S(\alpha, \beta) : \alpha \in \bar{U}(\ell), \beta \in \bar{U}(\ell')\} \quad (44)$$

for $\ell' \neq \ell$. The notation $S(\alpha, \beta)$ denotes the action of a path from node α to β . Essentially, this quantity measures the minimum value of the action along any path connecting well ℓ to any other well. Call $W(\ell)$ the set of nodes α such that $\inf\{S(\alpha, \beta) : \beta \in \bar{U}(\ell)\} \leq \bar{S}/2$. This defines a neighborhood around well ℓ consisting of points for which ℓ is the closest such well. Repeating for all other wells, it is clear that $\bigcup_\ell W(\ell)$ is a subset of all nodes in \mathcal{G}_F and $W(\ell)$ is disjoint from $W(\ell')$ [we may arbitrarily assign tie-breaker nodes shared by $W(\ell)$ and $W(\ell')$ at the boundary to assure disjointness].

Let us further define $|\phi^{(\ell)}\rangle$ to be the eigenstates of the eigenvalue problem restricted to nodes in $W(\ell)$ (with zero

boundary conditions on all other nodes). In the statement of Locality Theorem I, we drop the label ℓ in $|\phi^{(\ell)}\rangle$. We will first need the following lemma in order to proceed:

Lemma 6. Consider the subgraph induced by the nodes in $W(\ell)$, which we call \mathcal{G}_W , and solve the eigenvalue problem on \mathcal{G}_W with zero boundary conditions applied to all nodes not in $W(\ell)$. Call an eigenstate of this problem $|\psi^{(W)}\rangle$ for which the energy of such a state is E'_W upper bounded by E' . Call u the landscape for the original eigenvalue problem on the entirety of \mathcal{G}_F , and $u^{(W)}$ the landscape for the eigenvalue problem on the subgraph induced by $W(\ell)$. Then,

$$|\psi_\alpha^{(W)}| \leq \sqrt{\frac{E'_W - \frac{1}{\bar{u}^{(W)}}}{\frac{1}{u_\alpha^{(W)}} - E'_W}} \left(\frac{e^{-S_\alpha}}{\sqrt{\epsilon}} \right) \leq O\left(\sqrt{\frac{N}{\sigma}}\right) e^{-S_\alpha} \quad (45)$$

for some $\epsilon > 0$ and S_α the action of the eigenvalue problem on all of \mathcal{G}_F with energy E' . The notation $\bar{u}^{(W)}$ indicates the maximum value of $u^{(W)}$ in classically allowed regions of $W(\ell)$ where $1/u^{(W)} - E'_W < 0$. Furthermore, the second inequality follows from $E'_W - \frac{1}{\bar{u}^{(W)}} = O(N)$ and $\frac{1}{u_\alpha^{(W)}} - E'_W > \sigma$ in classically disallowed regions for some $\sigma > 0$.

This lemma essentially justifies that splicing the Fock space into wells and solving eigenvalue problems in each well will only cause a stronger localization of eigenfunctions; this is seen by proving that the Agmon decay estimates are stronger for the well eigenstates than for the global eigenstates. We refer the reader to the Appendix for a proof. With this, we may now state the following locality theorem:

Theorem 3. (Locality Theorem I): Suppose $|\phi\rangle$ is an eigenstate with eigenvalue μ of the eigenvalue problem on the subgraph induced by nodes in $W(\ell)$. Denote by $P_\psi(\mu - \delta, \mu + \delta)$ the projection operator onto the basis of eigenstates (in the full graph) with eigenvalues between $\mu - \delta$ and $\mu + \delta$; that is,

$$P_\psi(\mu - \delta, \mu + \delta) = \sum_{a \in \bar{A}} |\psi^{(a)}\rangle \langle \psi^{(a)}|, \quad (46)$$

where A is the set of eigenstates with eigenvalues λ_a satisfying $|\lambda_a - \mu| \geq \delta$. Then,

$$\|[I - P_\psi(\mu - \delta, \mu + \delta)]|\phi\rangle\|_2 \leq O\left(\sqrt{\frac{N}{\sigma}}\right) \frac{N}{\delta} e^{-\bar{S}/2}, \quad (47)$$

where σ is defined in the previous lemma.

Again, we direct the reader to the Appendix for a full proof. In short, this theorem states that solving the eigenvalue problem in the well $W(\ell)$ gives an eigenstate that is very close to the span of eigenstates in the entire domain with energies sufficiently close. To show that the well eigenstates are themselves similar to the full eigenstates requires us to prove an additional statement. Before presenting the theorem, we note the change in notation and denote $|\phi^{(\ell, \alpha)}\rangle$ to be the α th eigenstate of the eigenvalue problem in well $W(\ell)$.

Theorem 4. (Locality Theorem II): Suppose $|\psi\rangle$ is an eigenstate with eigenvalue μ of the eigenvalue problem on the entire graph \mathcal{G}_F . Denote by $P_\phi(\mu - \delta, \mu + \delta)$ the projection operator onto the basis of the union of all eigenstates in each

well $W(\ell)$ with eigenvalues between $\mu - \delta$ and $\mu + \delta$; that is,

$$P_\phi(\mu - \delta, \mu + \delta) = \sum_{(\ell, a) \in A'} |\phi^{(\ell, a)} \langle \phi^{(\ell, a)} |, \quad (48)$$

where A' is the set of pairs (ℓ, a) of eigenstates and corresponding wells with eigenvalues $\lambda_{\ell, a}$ satisfying $|\lambda_{\ell, a} - \mu| \geq \delta$. Then,

$$\| [I - P_\phi(\mu - \delta, \mu + \delta)] |\psi\rangle \|_2 \leq O\left(\sqrt{\frac{N}{\sigma}}\right) \frac{N}{\delta} e^{-\bar{C}/2}. \quad (49)$$

Once again, the proof is in the Appendix. This statement shows that a global eigenstate $|\psi\rangle$ is close to the span of well eigenstates with similar eigenvalues. Note that this (coupled with the last theorem) presents a bijective correspondence between well and global eigenstates. In particular, if the spectrum associated with a well is not resonant with that of any another well, then, any global eigenvalue that is close enough to an eigenvalue of the first well will imply that the global eigenstate will localize in that well. We call these ‘‘locality theorems’’ as they argue that eigenstates do not uniformly spread over all possible classically allowed regions, but rather tend to only localize at a small handful. Furthermore, both these theorems have useful bounds if the minimum Agmon distance between wells is of order $\log(N)$, otherwise, the right-hand side of both theorems will scale directly with system size.

One can convert both of these theorems into statements comparing the global spectrum with the well spectrum. Specifically, call $C(\lambda) = \#\{\lambda_\alpha : \lambda_\alpha < \lambda\}$ and $C_0(\lambda) = \#\{\lambda_{\ell, \alpha} : \lambda_{\ell, \alpha} < \lambda\}$, which counts the number of eigenvalues in the global and local problems, respectively, below a particular value λ . From an identical analysis as in Ref. [24], we then arrive at the following corollary:

Corollary 1. Suppose we choose \bar{C} and δ such that

$$O\left(\sqrt{\frac{N}{\sigma}}\right) \frac{N\bar{C}}{\delta} < e^{\bar{C}/2}. \quad (50)$$

Then, it follows that

$$\begin{aligned} \min(\bar{C}, C_0(\mu - \delta)) &\leq C(\mu), \\ \min(\bar{C}, C(\mu - \delta)) &\leq C_0(\mu). \end{aligned} \quad (51)$$

Proof. See Ref. [24], Corollary 5.2. ■

This corollary states that the global and the well spectrum can be δ -close to each other up to an upper counting limit of $\bar{C}(\delta)$. This gives the intuition that the eigenvalues of the global and well problems more closely agree toward the bottom of the spectrum (since δ can be chosen to be smaller without significantly decreasing the size of \bar{C}): this indicates that global eigenstates are likely to only resonate with a single well toward the bottom of the spectrum.

V. LOCATOR EXPANSION FOR MBL

In the previous two sections, we have defined a many-body generalization of the LL, u , and showed that it satisfies the same identities as derived by Lyra, Mayboroda, and Filoche [21,22] in the single-particle case. We further showed that $1/u$ behaves as an effective potential and derived

decay rates that relate to the height of an effective potential barrier and the connectivity of a particular node in the Fock-space graph. However, the effective potential is in general an analytically intractable object to calculate because it requires one to invert the Hamiltonian. We would thus like to calculate corrections to the effective potential perturbatively, much like Anderson has done using the famous locator expansion (LE) [1]. In this section, we show that naive perturbation theory does not suffer from the resonances that plagued Anderson’s LE. This allows us to state that the Agmon estimates derived in the previous section are robust to small interaction and hopping, and MBL for low-energy states should continue to persist.

We start by splitting the Hamiltonian into three terms corresponding to the hopping, onsite potential, and interaction terms: $\mathcal{H} = T + U + V$. The landscape is given by

$$|u\rangle = \frac{1}{U + (T + V)} |1\rangle, \quad (52)$$

where the suggestive grouping differentiates the nonperturbative term and the perturbations. We may write this recursively, by utilizing the identity $A^{-1} = B^{-1} + B^{-1}(B - A)A^{-1}$ for $A = U + (T + V)$ and $B = U$:

$$|u\rangle = \frac{1}{U} \left[I + (T + V) \frac{1}{U + (T + V)} \right] |1\rangle, \quad (53)$$

and iteration (assuming inverses are well defined) results in the following infinite series:

$$\begin{aligned} |u\rangle &= U^{-1} \sum_{n=0}^{\infty} [(T + V)U^{-1}]^n |1\rangle \\ &= U^{-1} \sum_{\beta} \sum_{n=0}^{\infty} [(T + V)U^{-1}]^n |\beta\rangle, \end{aligned} \quad (54)$$

where in the second line we use the fact that $|1\rangle$ is the all ones vector. The sum is over all basis states in the Fock space (which are eigenstates of U). Next, we call $\mathcal{E}_\alpha = \sum_i n_i^{(\alpha)} \epsilon_i$, which is the unperturbed eigenvalues of U . Inserting n complete sets of eigenstates labeled by $|\alpha_i\rangle \langle \alpha_i|$, we find

$$\begin{aligned} \langle \alpha | u \rangle &= \sum_{n=0}^{\infty} \sum_{\alpha_1, \alpha_2, \dots, \alpha_{n+1}, \beta} \langle \alpha | U^{-1} | \alpha_1 \rangle \\ &\quad \times \langle \alpha_1 | (T + V) U^{-1} | \alpha_2 \rangle \langle \alpha_2 | (T + V) U^{-1} | \alpha_3 \rangle \\ &\quad \times \dots \langle \alpha_n | (T + V) U^{-1} | \alpha_{n+1} \rangle \langle \alpha_{n+1} | \beta \rangle. \end{aligned} \quad (55)$$

This can be written in terms of \mathcal{E} , resulting in the following expression:

$$\begin{aligned} \langle \alpha | u \rangle &= \frac{1}{\mathcal{E}_\alpha} \sum_{n=0}^{\infty} \sum_{\alpha_1 = \alpha, \alpha_2, \dots, \alpha_{n+1}} \frac{\langle \alpha_1 | (T + V) | \alpha_2 \rangle}{\mathcal{E}_{\alpha_2}} \\ &\quad \times \frac{\langle \alpha_2 | (T + V) | \alpha_3 \rangle}{\mathcal{E}_{\alpha_3}} \dots \frac{\langle \alpha_n | (T + V) | \alpha_{n+1} \rangle}{\mathcal{E}_{\alpha_{n+1}}}. \end{aligned} \quad (56)$$

This series has a graphical representation: we start at a node α and define two propagators: the first carries an amplitude $-\frac{1}{\mathcal{E}_\alpha}$ and occurs along the directed edge from node α to any of its neighboring nodes. The second is a self-propagator carrying an amplitude $\frac{V}{\mathcal{E}_\alpha} (\sum_{\langle i, j \rangle} n_i^{(\alpha)} n_j^{(\alpha)})$. Thus, the n th-order term is

the sum over all possible paths of products of n propagators along a given path.

In the conventional LE, quantities like $E - \mathcal{E}_\alpha$ appear in the denominators, which lead to an unregulated perturbation series; however, such resonances are crucially not present in this expansion. This allows us to show that for sufficiently small perturbation, this series converges:

Theorem 5. Suppose the onsite potentials have positive support; i.e., $\epsilon_i > \delta$ for δ a positive constant. Then, the perturbation series for $\langle \alpha | u \rangle$ converges for sufficiently small perturbation if \mathcal{G}_P is a cubic lattice in d dimensions.

Proof. The proof is straightforward. First, note that the eigenstates are not affected by a constant shift in the onsite potential and thus the onsite potential can be chosen to have positive support. This implies that $\mathcal{E}_\alpha \geq \delta N$ for δ the minimum value in the support of the onsite potential. The n th-order term in the infinite series is given by

$$\begin{aligned}
 C_n^{(\alpha)} &= \frac{1}{\mathcal{E}_\alpha} \sum_{\alpha_2, \dots, \alpha_{n+1}} \frac{\langle \alpha | (T + V) | \alpha_2 \rangle}{\mathcal{E}_{\alpha_2}} \\
 &\quad \times \frac{\langle \alpha_2 | (T + V) | \alpha_3 \rangle}{\mathcal{E}_{\alpha_3}} \dots \frac{\langle \alpha_n | (T + V) | \alpha_{n+1} \rangle}{\mathcal{E}_{\alpha_{n+1}}} \\
 &\leq \frac{1}{\mathcal{E}_\alpha} \frac{1}{(\delta N)^n} \sum_{\alpha_2, \dots, \alpha_{n+1}} \langle \alpha | (T + V) | \alpha_2 \rangle \\
 &\quad \times \langle \alpha_2 | (T + V) | \alpha_3 \rangle \dots \langle \alpha_n | (T + V) | \alpha_{n+1} \rangle. \quad (57)
 \end{aligned}$$

The sum is interpreted as a sum over paths of length n in \mathcal{G} starting at α . Furthermore, we allow self-paths to occur, which result from $\langle \alpha | V | \alpha \rangle$ being nonzero. Additionally, since this term can be of order N (since it is equal to $V \sum_{(i,j)} n_i^{(\alpha)} n_j^{(\alpha)}$), we create N copies of each node, such that the matrix element due to transitioning to any neighboring node is constant. For the case of a lattice, the number of paths in this enlarged graph of length n can therefore be bounded by $(cN)^n$, where c is a constant; this follows from the fact that the maximum degree of a node in \mathcal{G} for such graphs is $O(N)$. The transition amplitude along an edge is upper bounded in absolute value by a constant, which we call λ . Therefore, $C_n^{(\alpha)} \leq \frac{1}{\mathcal{E}_\alpha} \frac{\lambda^n (cN)^n}{(\delta N)^n} = \frac{1}{\mathcal{E}_\alpha} \left(\frac{c\lambda}{\delta}\right)^n$. Choosing λ small enough so that $c\lambda/\delta < 1$, $C_n^{(\alpha)}$ decays geometrically in n and the series therefore converges. ■

Persistence of MBL with interactions

While the MBLL illuminates the structure of eigen-states and their exponential decay away from the classically allowed regions in the Fock space for a given many-body energy, it does not provide direct information about what the many-body energies are. Of course, at the bare level before performing the locator expansion (i.e., with only a disorder potential present $\{\epsilon_j\}$), the many-body spectrum is known: $\mathcal{E}_\alpha = \sum_{j=1}^L \epsilon_j n_j^{(\alpha)}$ (where $n_j^{(\alpha)}$ is the occupation number of site j in the many-body state $|\alpha\rangle$). The corresponding landscape is just its inverse $\langle \alpha | u \rangle \propto \frac{1}{\mathcal{E}_\alpha}$ and, in particular, it scales as $1/N$. The convergence of the landscape under the locator expansion implies that its structure does not change much under perturbation theory, the original localizing traps remain, and the underlying scaling of the effective potential $1/u$ with N persists. However, determining how the many-body energies shift under

even a small perturbation is not immediately obvious. Finding self-energy corrections via perturbation theory and in general determining the full spectrum are more difficult problems than perturbing the landscape, because of the resonances complicating the former.

It is worth mentioning here the intriguing results of Mayboroda *et al.* on connecting the spectrum of the Schrödinger operator and the structure of the single-particle landscape. As reported in Ref. [49] and Mayboroda’s talk at the 2018 International Congress of Mathematicians [50], there is strong evidence that the local minima of the single-particle landscape determine the spectrum to great accuracy. In particular, it was proposed that $E_j \approx (1 + \frac{d}{4})(\min u^{-1})_j$, which relates the j th energy level to the depth of the j th deepest well in the inverse landscape (here, d is the dimension of space). This heuristic formula was numerically shown to work with unprecedented precision for the lower part of spectrum and some heuristic analytical arguments were presented in Ref. [49] to support it. These arguments have recently been made rigorous and a new “landscape law” for the density of states was proposed [51], but it is unclear whether the bounds generalize to discrete lattices or graphs. However, if there were a one-to-one correspondence between the local minima of the MBLL and many-body energies, this would imply (in conjunction with the convergence of the locator expansion) many-body localization of the states associated with the MBLL minima (note that there are in general “more” states than there are minima in the inverse landscape).

Apart from this observation, a more convincing argument is based on physical reasoning as follows. Under a weak enough perturbation λ , the greatest shift we can expect in many-body energies is $O(\lambda N)$. For λ small, this shift is small and the size of the classical region does not expand much from the size in the $t = V = 0$ limit. Since states are localized at the $t = V = 0$ limit, we conclude that MBL still persists for small t and V , at least for the low-energy portion of the spectrum. Utilizing the particle-hole symmetry if \mathcal{G}_P is bipartite also informs us that states are localized at high energies (by this we mean low-energy states of $-H$); this is analogous to the argument used in Ref. [22].

We note here that by MBL, we imply the existence of a subset of states in the Hilbert space which decay in the Fock space exponentially with the system size in the presence of weak interaction. This is a weaker definition than that used by Basko, Aleiner, and Altshuler [2], who called a state $|\Psi_\alpha\rangle$ in the Fock space localized if the creation of a particle-hole pair (or equivalently flipping two spins) results in a state $\sigma_i^+ \sigma_j^- |\Psi_\alpha\rangle$ with a support on a finite, i.e., $O(1)$, number of states in the Fock space. This is contrasted with an ergodic state, whose support is on $O(|\mathcal{H}|)$ states, i.e., spanning most of the Hilbert space. Our landscape construction corresponds to an in-between situation where the support (corresponding roughly to a size of classically allowed regions) may, in principle, be of size $\Omega(1)$, spanning a part of the Hilbert space that scales with system size. Yet, the Basko-Aleiner-Altshuler definition of MBL is weaker than the definition used by Oganesyan and Huse [5] and many subsequent publications, where localization of the entire Hilbert space is required. This may occur in a lattice model in one dimension, but most certainly cannot happen in three and higher dimensions,

nor in continuous models. Our construction does not rule out the possibility that these stronger versions of MBL hold in a particular model, but can only point to the existence of states exponentially localized in a low-energy part of the Hilbert space (for an arbitrary-dimensional lattice or graph, \mathcal{G}_P), where the landscape function is most informative.

Furthermore, the MBLL does not rule out that states may delocalize at the middle of the band, which may correspond to a percolationlike transition on \mathcal{G}_F as a function of energy. This potential tendency to delocalize increases as the graph becomes more connected. This is because choosing an energy toward the middle of the spectrum causes a considerable fraction of nodes to be in the classically allowed region. Once this fraction is as large as the critical fraction for percolation in \mathcal{G}_F , states will become extended throughout the entire Hilbert space. As \mathcal{G}_F becomes more connected, the percolation transition occurs at a lower filling fraction. Since \mathcal{G}_F becomes more connected as the dimension increases, this may point to evidence of the formation of a mobility edge for high-dimensional systems.

One caveat to note is when one would expect the perturbative series described in this section to diverge. This occurs when the physical graph \mathcal{G}_P is highly connected. Note that if \mathcal{G}_P is locally connected (for instance in a lattice), then \mathcal{G}_F has a maximum degree that is linear in N . The moment \mathcal{G}_P becomes more strongly connected by longer-range hopping, \mathcal{G}_F will have nodes with degree that scale as $\Omega(N)$, and the resulting bound presented in Theorem 5 of this section will not hold. This is consistent with the intuition that systems with long-range hopping do not exhibit MBL.

VI. CONCLUSIONS

We construct a generalization of the single-particle localization landscape to interacting many-body systems. The many-body localization landscape is a useful representation of interacting many-body systems on a general class of lattices. The MBLL is given by $(\mathcal{H} + K)|u\rangle = |1\rangle$ in Fock space, and thus is more appropriately defined with respect to a graph in the Fock space, which we call \mathcal{G}_F . We show that the MBLL satisfies $|\psi_\alpha| \leq E(\max_\beta |\psi_\beta|)u_\alpha$, and thus the basins of $1/u$ can be interpreted as “traps” (classically allowed regions) with respect to \mathcal{G}_F that confine many-body eigenstates.

We then describe the behavior of eigenfunctions outside the classically allowed regions by generalizing Agmon decay estimates. These decay estimates are defined with respect to paths in the Fock-space graph \mathcal{G}_F and can be interpreted as the minimization of an effective action. We show that the eigenfunctions begin to decay slowly once they pass through the classically forbidden region, but it is only after roughly $\Theta(N^{1/2})$ steps into the forbidden region that the decay rate begins to scale with the system size. When the depth into the forbidden region is $\Theta(N)$, the decay rate becomes exponential in the path length.

A locality theorem is then proved, which roughly states that the eigenstate in a particular well, as defined by the MBLL, is sufficiently close to the span of eigenstates in the entire Fock space whose energies are sufficiently close; a similar statement about global eigenstates sufficiently close to the span of well eigenstates with close energies is also proved.

This implies that eigenstates tend to localize in a small number of individual wells as opposed to being spread out in the Hilbert space among many wells.

We then develop a perturbative expansion for the MBLL similar to the locator expansion; i.e., we treat the disorder as the leading term in the Hamiltonian and perform a series expansion in both the hopping and interaction. Unlike in the locator expansion, where one encounters resonances, we show that the perturbative expansion for the MBLL converges for small hopping and interaction. This implies that the landscape only changes by a multiplicative constant from the unperturbed landscape. Assuming the corresponding energy shift is not too large, the size of the classically allowed regions should not change by much with small interaction and the localization properties should be similar to those of the noninteracting problem.

The convergent locator expansion combined with the locality theorem allows us to prove that at least for particular kinds of disorder distributions, localization does persist in the presence of weak interactions for a part of the Hilbert space. Namely, we can choose an initial disorder distribution with a desired level statistics, such that the wells do not percolate up to an energy threshold. The corresponding landscape without hopping and interactions is just the disorder potential itself and the convergent locator series implies that it is weakly corrected as long as the disorder potential is dominant compared to the hopping and interactions. The locality theorem implies that the set of tightly localized states in the disorder-only model remain exponentially localized in a small number of nonpercolating wells in the presence of small interactions. This rigorously establishes a weak version of many-body localization for a class of models.

We note that the landscape construction crucially does not shed light on whether full MBL, by which we mean localization of all states in the Hilbert space and the complete breakdown of ergodicity, may exist in a model. What it does show, however, is that there exist a rich variety of possible localization structures in the Fock space, which include the possibility of Basko-Aleiner-Altshuler-type MBL, full MBL of the entire spectrum, and a continuum of structures in-between. In the language of the MBLL, there is a question of quantifying the size of the “classically allowed regions,” where the many-body wave functions do not decay; this is analogous to the size of support in Fock space of a state with a particle-hole excitation as described in Ref. [2]. This region can consist of a finite number of states [2], or scale with the system size while still covering a measure zero set in the entire Hilbert space. Furthermore, the decay properties beyond these regions may differ, e.g., power law vs exponential, based on the region in the Fock-space graph. As we have described throughout this paper, different parts of the Hilbert space may in principle have different properties. This points to a variety of possible phases and suggests that a finer classification of MBL needs to be developed to describe these structures.

We also point out some limitations of the MBLL construction we presented in this paper. For one, we cannot fully determine the energy shifts due to interaction, which is crucial to determining the size of classically allowed regions. This computation requires one to calculate self-energies, a method

which is only known to be rigorous on a Bethe lattice. However, Mayboroda and collaborators have several conjectures (numerically and theoretically verified to some degree) stating the rather astounding notion that the set of local minima of the inverse landscape can determine the spectrum to high accuracy [49,51]. In particular with this result, we would not only be able to approximate energy shifts, but also determine level statistics of the spectrum, which are commonly used in identifying the MBL transition.

Second, it is not known whether the Agmon estimates derived are tight; it may be possible to improve these bounds by eliminating or refining the presence of $\sqrt{1/|N(\alpha)|}$ in the “graph momentum.” Doing so would allow one to make stronger statements of rapid exponential decay for a larger fraction of states in the Fock space, which would improve understanding of the mechanism for MBL.

In the perturbative series we performed, where we treat the disorder as large, the landscape essentially only contains the onsite disorder potential. The landscape thus gives us more information when the hopping becomes considerably large. Moreover, there are rigorous mathematical proofs for the existence of Anderson localization in 1D noninteracting systems [52,53]. Thus, a more natural way to use perturbation theory would be to consider the noninteracting Anderson model as the starting point and calculate the MBLL perturbatively in the interaction only.

On top of these avenues for exploration, there are several other modifications that may also lead to interesting effects. For one, we argued that the sign of t should not matter due to particle-hole symmetry if the lattice is bipartite, but it would be important to understand whether this constraint can be relaxed to frustrated lattices that are not bipartite. It is conceivable that frustrated disordered lattices may favor delocalization, an interesting scenario. Second, we assumed in our treatment that the interaction term is repulsive, but a similar theory can be developed for attractive interactions. Such “negative- U ” Hubbard models with disorder are of great interest for understanding the superconductor-to-insulator transition in disordered films, where a crossover is expected from a superconductor to a regime of localized Cooper pairs to that of the more conventional electronic insulator [54]. Third, we briefly mentioned that other nonlocal interactions can be used, so long as the interaction strength is negative (or there exists a similar particle-hole symmetry so that the sign of the interaction does not matter), it would similarly be important to understand for what class of interactions the landscape picture would break down. In particular, it may be possible to develop the MBLL theory in the case of the Dirac Sachdev-Ye-Kitaev (SYK) model [55], in which case the MBLL may contain signatures of ergodicity.

The current theory developed shows that the many-body localization landscape is a relatively simple yet powerful object that reveals the rich nature of many-body localization and possibly other interacting many-body problems. Furthermore, the computational price for calculating the many-body localization landscape is smaller than that required for exact diagonalization; therefore, the exact, non-perturbative landscape can be computed numerically for larger system sizes than what has been possible with exact diagonalization.

ACKNOWLEDGMENTS

The authors are grateful to Y.-T. Hsu for useful discussions and a fruitful collaboration on a related numerical project to be published elsewhere. The authors are also grateful to S. Mayboroda and C. Beenakker for insightful and illuminating discussions. The authors also thank both anonymous referees for their detailed feedback. S.B. was supported by NSF Grant No. DMR-1613029 while at the Physics Frontier Center in JQI. This research was also supported by US-ARO (Contract No. W911NF1310172) and DARPA DRINQS program (Y.L.), DOE-BES (Grant No. DESC0001911) (V.G.), and the Simons Foundation. V.G. is especially grateful to the Simons Foundation for hospitality during the MPS Annual Meetings, where this research project was conceived.

APPENDIX: PROOFS OF LOCALITY THEOREMS

We explicitly provide proofs for the Locality Theorems stated in Sec. IV C. We will substantially make use of the test state $|\zeta\rangle = \sum_{a \in A} \gamma_a |\psi^{(a)}\rangle$, which is written in the eigenbasis (and we enumerate eigenstates using the Latin alphabet), where A is specified in the theorems stated below.

1. Proof of Lemma 6

To prove this statement, it suffices to show that $1/u^{(W)} - E'_W$ for the subgraph problem is greater than $1/u - E'$ for the entire problem; therefore, the action for the subgraph problem will be larger than that for the entire problem. First, we know that

$$-t \sum_{\beta \in N(\alpha)} u_\beta + \left(\sum_{i=1}^L V n_i^{(\alpha)} n_{i+1}^{(\alpha)} + \epsilon_i n_i^{(\alpha)} + K \right) u_\alpha = 1 \quad (\text{A1})$$

for the landscape on the entire domain, and

$$-t \sum_{\beta \in N_W(\alpha)} u_\beta^{(W)} + \left(\sum_{i=1}^L V n_i^{(\alpha)} n_{i+1}^{(\alpha)} + \epsilon_i n_i^{(\alpha)} + K \right) u_\alpha^{(W)} = 1, \quad (\text{A2})$$

where $N_W(\alpha)$ denotes the set of neighbors of α in the subgraph induced by the well nodes. Note that $N_W(\alpha)$ is a subset of $N(\alpha)$. Subtracting the two equations above for all $\alpha \in W(\ell)$ gives

$$\begin{aligned} & -t \sum_{\beta \in N_W(\alpha)} (u_\beta - u_\beta^{(W)}) + \left(\sum_{i=1}^L V n_i^{(\alpha)} n_{i+1}^{(\alpha)} + \epsilon_i n_i^{(\alpha)} + K \right) \\ & \times (u_\alpha - u_\alpha^{(W)}) = t \sum_{\beta \in N(\alpha) \setminus N_W(\alpha)} u_\beta. \end{aligned} \quad (\text{A3})$$

From Lemma 1, the right-hand side of this equation is positive, and by the same lemma, this implies that $(u_\alpha - u_\alpha^{(W)}) \geq 0$. Therefore, $1/u_\alpha \leq 1/u_\alpha^{(W)}$ and by the assumption that $E'_W \leq E'$, we have that $1/u_W - E'_W \geq 1/u - E'$. Thus, utilizing the bounds presented in the previous subsection along with this result, we prove this lemma.

2. Proof of Theorem 3

We need to make note of some additional notation. We call \mathcal{G}_W the subgraph induced by nodes in the well, and we

call $\partial\mathcal{G}_W$ the subgraph induced by the set of edges connecting nodes in $W(\ell)$ to nodes outside of $W(\ell)$.

Throughout this proof, we will be using labels such as $|fg\rangle$ to mean $\sum_{\alpha} f_{\alpha} g_{\alpha} |\alpha\rangle$. These do not represent physical quantum states because they are not normalized. First, recall that operators such as M defined at the beginning of this section operate on the entirety of \mathcal{G}_F . For this proof, we will assume that they only operate on \mathcal{G}_W . Since $|\phi\rangle$ has eigenvalue μ , we know that $M|\phi\rangle = \mu|u^{(W)}\phi\rangle$. Let us instead consider the state $|\eta\phi\rangle$, for some η to be presented momentarily. Then, $M|\eta\phi\rangle = \mu|u\eta\phi\rangle + |ur\rangle$ for some residual state $|r\rangle$. We note that u here is the landscape defined for the global \mathcal{G}_F . Applying the test state $\langle\zeta|u|$ to this equation gives us the following expression for $\langle\zeta|ur\rangle = \langle\zeta|r\rangle$:

$$\begin{aligned} \langle\zeta|r\rangle &= t \sum_{(\alpha,\beta)\in\mathcal{G}_W} u_{\alpha}u_{\beta} \left(\frac{\zeta_{\alpha}}{u_{\alpha}} - \frac{\zeta_{\beta}}{u_{\beta}} \right) \left(\frac{\phi_{\alpha}\eta_{\alpha}}{u_{\alpha}} - \frac{\phi_{\beta}\eta_{\beta}}{u_{\beta}} \right) \\ &\quad + \sum_{\alpha} \zeta_{\alpha}\phi_{\alpha}\eta_{\alpha} \left(\frac{1}{u_{\alpha}} - \mu \right). \end{aligned} \quad (\text{A4})$$

This expression may be written as

$$\begin{aligned} \langle\zeta|r\rangle &= t \sum_{(\alpha,\beta)\in\mathcal{G}_W\cup\partial\mathcal{G}_W} u_{\alpha}u_{\beta} \left(\frac{\zeta_{\alpha}}{u_{\alpha}} - \frac{\zeta_{\beta}}{u_{\beta}} \right) \left(\frac{\phi_{\alpha}\eta_{\alpha}}{u_{\alpha}} - \frac{\phi_{\beta}\eta_{\beta}}{u_{\beta}} \right) \\ &\quad - t \sum_{(\alpha,\beta)\in\partial\mathcal{G}_W} u_{\alpha}u_{\beta} \left(\frac{\zeta_{\alpha}}{u_{\alpha}} - \frac{\zeta_{\beta}}{u_{\beta}} \right) \left(\frac{\phi_{\alpha}\eta_{\alpha}}{u_{\alpha}} - \frac{\phi_{\beta}\eta_{\beta}}{u_{\beta}} \right) \\ &\quad + \sum_{\alpha} \zeta_{\alpha}\phi_{\alpha}\eta_{\alpha} \left(\frac{1}{u_{\alpha}} - \mu \right). \end{aligned} \quad (\text{A5})$$

Expanding $|\zeta\rangle$ via its definition into the basis of eigenstates in the set A results in

$$\begin{aligned} \langle\zeta|r\rangle &= t \sum_{a\in A} \gamma_a \sum_{(\alpha,\beta)\in\mathcal{G}_W\cup\partial\mathcal{G}_W} u_{\alpha}u_{\beta} \left(\frac{\psi_{\alpha}^{(a)}}{u_{\alpha}} - \frac{\psi_{\beta}^{(a)}}{u_{\beta}} \right) \\ &\quad \times \left(\frac{\phi_{\alpha}\eta_{\alpha}}{u_{\alpha}} - \frac{\phi_{\beta}\eta_{\beta}}{u_{\beta}} \right) \\ &\quad + \sum_{a\in A} \gamma_a \sum_{\alpha\in W} \psi_{\alpha}^{(a)} \phi_{\alpha}\eta_{\alpha} \left(\frac{1}{u_{\alpha}} - \lambda_a \right) \\ &\quad - t \sum_{a\in A} \gamma_a \sum_{(\alpha,\beta)\in\partial\mathcal{G}_W} u_{\alpha}u_{\beta} \left(\frac{\psi_{\alpha}^{(a)}}{u_{\alpha}} - \frac{\psi_{\beta}^{(a)}}{u_{\beta}} \right) \\ &\quad \times \left(\frac{\phi_{\alpha}\eta_{\alpha}}{u_{\alpha}} - \frac{\phi_{\beta}\eta_{\beta}}{u_{\beta}} \right) \\ &\quad + \sum_{a\in A} \gamma_a \sum_{\alpha\in W} \psi_{\alpha}^{(a)} \phi_{\alpha}\eta_{\alpha} (\lambda_a - \mu). \end{aligned} \quad (\text{A6})$$

The first two terms can be combined to equal zero from the eigenvalue equation $M|\psi_{\alpha}^{(a)}\rangle = \lambda_a|u\psi_{\alpha}^{(a)}\rangle$. The addition of the edges $\partial\mathcal{G}_W$ in the summation was necessary so that the boundary nodes satisfy the appropriate eigenvalue equation for the entire domain. The residual term (the third term) can be made to be zero by forcing $\eta = 0$ on all nodes in $\partial\mathcal{G}_W$. Thus, with this choice of η , we are left with the fourth term:

$$\langle\zeta|r\rangle = \sum_{a\in A} \gamma_a \sum_{\alpha} \psi_{\alpha}^{(a)} \phi_{\alpha}\eta_{\alpha} (\lambda_a - \mu), \quad (\text{A7})$$

where the remaining terms in Eq. (A4) disappear from application of the eigenvalue equation for $|\psi^{(a)}\rangle$. Denoting the inner product $\langle\psi^{(a)}|\eta\phi\rangle = \nu_a$, we find

$$\langle\zeta|r\rangle = \sum_{a\in A} \gamma_a \nu_a (\lambda_a - \mu). \quad (\text{A8})$$

Next, we derive a bound on $\langle\zeta|r\rangle$. We start by noting that the following identity is true:

$$\begin{aligned} &\left\langle \frac{\eta\zeta}{u} \middle| M \middle| \phi \right\rangle - \mu \langle \eta\zeta | \phi \rangle \\ &= t \sum_{(\alpha,\beta)\in\mathcal{G}_W} u_{\alpha}u_{\beta} \left(\frac{\zeta_{\alpha}\eta_{\alpha}}{u_{\alpha}} - \frac{\zeta_{\beta}\eta_{\beta}}{u_{\beta}} \right) \left(\frac{\phi_{\alpha}}{u_{\alpha}} - \frac{\phi_{\beta}}{u_{\beta}} \right) \\ &\quad + \sum_{\alpha\in W} \zeta_{\alpha}\phi_{\alpha}\eta_{\alpha} \left(\frac{1}{u_{\alpha}} - \mu \right). \end{aligned} \quad (\text{A9})$$

This expression in fact equals zero. To show this, let us rewrite the eigenvalue problem for $|\phi\rangle$ in terms of the *global* landscape $|u\rangle$. Following the analysis of Sec. IV A and replacing the local landscape by the global landscape, we find that Eq. (20) becomes

$$-t \sum_{\beta\in N(\alpha)} u_{\beta}u_{\alpha} \left(\frac{\phi_{\beta}}{u_{\beta}} - \frac{\phi_{\alpha}}{u_{\alpha}} \right) + \phi_{\alpha} = \mu u_{\alpha}\phi_{\alpha}, \quad (\text{A10})$$

so long as α is located in the bulk of W and not at the boundary ∂W . Then, since $\eta_{\alpha} = 0$ at the boundary, the quantity $\langle \frac{\eta\zeta}{u} | M | \phi \rangle - \mu \langle \eta\zeta | \phi \rangle$ does not depend on the boundary behavior. In the bulk, $M|\phi\rangle = \mu|u\phi\rangle$, so the quantity equals zero, as desired.

Subtracting Eq. (A9) from Eq. (A4) and simple algebraic manipulation gives the following expression:

$$\langle\zeta|r\rangle = t \sum_{(\alpha,\beta)} (\eta_{\beta} - \eta_{\alpha}) (\phi_{\alpha}\zeta_{\beta} - \phi_{\beta}\zeta_{\alpha}). \quad (\text{A11})$$

Next, we choose η_{α} to equal 1 if $\inf\{S(\alpha, \beta) : \beta \in U(\ell)\} \leq \bar{S}/2 - \delta'$ and zero otherwise. We also have the additional restriction that $\eta_{\alpha} = 0$ for nodes at the boundary ∂W . Therefore, the sum above has nonzero contribution only if the edge (α, β) connects a node where $\eta = 0$ and a node where $\eta = 1$. For δ' chosen small enough, the sum will only have contribution in the subgraph $\partial\mathcal{G}_W$. We may then write

$$\langle\zeta|r\rangle = t \langle\zeta|P|\phi\rangle, \quad (\text{A12})$$

where P denotes a matrix with elements $P_{\alpha\beta} = \eta_{\beta} - \eta_{\alpha}$. Next, we perform a restriction of both $|\phi\rangle$ and $|\zeta\rangle$ as well as P to nodes α for which $P_{\alpha\beta} \neq 0$ for some β . By the Cauchy-Schwartz inequality, we find

$$|\langle\zeta|r\rangle|^2 \leq t^2 \sigma_{\max}^2(P) \langle\phi|\phi\rangle \langle\zeta|\zeta\rangle, \quad (\text{A13})$$

where $\sigma_{\max}(P)$ denotes the maximum singular value of P , or alternately the 2-norm. First, we have $\langle\zeta|\zeta\rangle \leq \sum_{a\in A} \gamma_a^2$. Next, we refer to Lemma 6. Since $|\phi\rangle$ is an eigenstate on \mathcal{G}_W , then from the lemma and Eq. (38)

$$\epsilon \sum_{\alpha\in U} \phi_{\alpha}^2 e^{2S_{\alpha}} \leq \frac{\max_{\alpha\in\bar{U}} \Delta_{\alpha}^{(W)}}{\min_{\alpha\in U} -\Delta_{\alpha}^{(W)}} \leq O\left(\frac{N}{\sigma}\right), \quad (\text{A14})$$

where $\Delta_\alpha = 1/u_\alpha - E'$ and $\Delta_\alpha^{(W)} = 1/u_\alpha^{(W)} - E'_W$. Furthermore, U and \bar{U} denote classically disallowed and classically allowed regions, respectively. Since the action of all nodes $\alpha \in \partial W$ is at least $\bar{S}/2 - \delta'$, then we have (assuming δ' is constant and can be ignored)

$$\sum_{\alpha \in \partial W} \phi_\alpha^2 \leq \frac{e^{-\bar{S}}}{\epsilon} \frac{\max_{\alpha \in \bar{U}} \Delta_\alpha^{(W)}}{\min_{\alpha \in U} -\Delta_\alpha^{(W)}} \leq O\left(\frac{N}{\sigma}\right) e^{-\bar{S}}. \quad (\text{A15})$$

Next, we note the well-known identity (which is a special case of Holder's inequality)

$$\sigma_{\max}(P) \leq \sqrt{\|P\|_1 \|P\|_\infty} \leq kN, \quad (\text{A16})$$

where the second inequality follows from the fact that the 1-norm and the ∞ -norm are upper bounded by the degree of any node in \mathcal{G}_F , and k is some constant. Recall that the 1-norm of matrix P is the maximum column sum of $|P|$, while the ∞ -norm is the maximum row sum of $|P|$. Utilizing all of the bounds derived, Eq. (A13) becomes

$$\begin{aligned} |\langle \zeta | r \rangle|^2 &\leq \frac{k^2 t^2 N^2}{\epsilon} \frac{\max_{\alpha \in \bar{U}} \Delta_\alpha^{(W)}}{\min_{\alpha \in U} -\Delta_\alpha^{(W)}} e^{-\bar{S}} \left(\sum_{a \in A} \gamma_a^2 \right) \\ &\leq O\left(\frac{N^3}{\sigma}\right) e^{-\bar{S}}. \end{aligned} \quad (\text{A17})$$

Next, return to Eq. (A8) and choose $\gamma_a = v_a \text{sgn}(\lambda_a - \mu)$ for all $a \in A$. Then, combining Eq. (A8) with the above, we find that

$$\left(\sum_{a \in A} v_a^2 |\lambda_a - \mu| \right)^2 \leq O\left(\frac{N^3}{\sigma}\right) e^{-\bar{S}} \left(\sum_{a \in A} v_a^2 \right), \quad (\text{A18})$$

where we have suppressed the unnecessary constants for convenience. Because $|\lambda_a - \mu| \geq \delta$, we obtain the inequality

$$\sum_{a \in A} v_a^2 \leq \frac{O(N^3)}{\delta^2 \sigma} e^{-\bar{S}}. \quad (\text{A19})$$

However, the left-hand side of this expression is just equal to the norm squared of the orthogonal complement of $|\eta\phi\rangle$ onto the span of the eigenvectors in A . In terms of notation described in the theorem statement, this may be written as

$$\| [I - P_\psi(\mu - \delta, \mu + \delta)] |\eta\phi\rangle \|_2 \leq O\left(\sqrt{\frac{N}{\sigma}}\right) \frac{N}{\delta} e^{-\bar{S}/2}. \quad (\text{A20})$$

Next, we note that $\langle (1 - \eta)\phi | (1 - \eta)\phi \rangle$ satisfies the following bound:

$$\langle (1 - \eta)\phi | (1 - \eta)\phi \rangle \leq \frac{\max_{\alpha \in \bar{U}} \Delta_\alpha}{\min_{\alpha \in U} -\Delta_\alpha} e^{-\bar{S}} \leq O\left(\frac{N}{\sigma}\right) e^{-\bar{S}}, \quad (\text{A21})$$

which follows both from the fact that $|(1 - \eta)\phi\rangle$ has support on nodes with action greater than $\bar{S}/2 - \delta'$ as well as Eq. (38) and Lemma 6. Because $I - P_\psi(\mu - \delta, \mu + \delta)$ has matrix norm equal to one, we have

$$\| [I - P_\psi(\mu - \delta, \mu + \delta)] |(1 - \eta)\phi\rangle \|_2 \leq O\left(\sqrt{\frac{N}{\sigma}}\right) e^{-\bar{S}/2}. \quad (\text{A22})$$

Adding this to Eq. (A20), and using the triangle inequality, we find

$$\| [I - P_\psi(\mu - \delta, \mu + \delta)] |\phi\rangle \|_2 \leq O\left(\sqrt{\frac{N}{\sigma}}\right) \frac{N}{\delta} e^{-\bar{S}/2}, \quad (\text{A23})$$

where again we have absorbed all constants or lower-order contributions into the $O(\sqrt{\frac{N}{\sigma}})$ term.

3. Proof of Theorem 4

The proof is very similar to that of the previous theorem, but there are several key changes. Most of the changes involve swapping ϕ and ψ , and thus we will be very brief with the proof. First, define $|\psi\rangle$ and note that it satisfies the eigenvalue equation. Consider the restriction of $|\psi\rangle$, given by $|\eta\psi\rangle$, which satisfies $M|\psi\rangle = \mu|u\eta\psi\rangle + |ur\rangle$, again for some residual vector $|r\rangle$.

Next, define the test state $|\zeta\rangle = \sum_{(\ell, a) \in A'} \gamma_{\ell, a} |\phi^{(\ell, a)}\rangle$; the difference here is that we take a superposition of different eigenstates from different wells if their corresponding eigenvalue is close to μ . Then, Eq. (A7) becomes

$$\langle \zeta | r \rangle = \sum_{(\ell, a) \in A'} \gamma_{\ell, a} \sum_{\alpha} \phi_{\alpha}^{(\ell, a)} \psi_{\alpha} \eta_{\alpha} (\lambda_{\ell, a} - \mu). \quad (\text{A24})$$

Before, to arrive at this equation, we needed to set η at the boundary nodes equal to zero; here, a similar manipulation holds except η will equal zero on nodes at the boundary of each well ℓ appearing in A' . Moreover, u is the landscape for the global problem, not the well problem for $|\phi^{(\ell, a)}\rangle$, but this can be amended by using the fact that $|\phi\rangle$ solves the eigenvalue problem with the global landscape for nodes in the bulk of the well as well as $\eta = 0$ in the boundary of the wells. This is identical to the reasoning used for setting Eq. (A9) to zero.

With this, Eq. (A8) then becomes

$$\langle \zeta | r \rangle = \sum_{(\ell, a) \in A'} \gamma_{\ell, a} v_{\ell, a} (\lambda_{\ell, a} - \mu), \quad (\text{A25})$$

where $v_{\ell, a} = \langle \phi^{(\ell, a)} | \eta\psi \rangle$. After some basic manipulation as in the previous theorem, we have the following expression:

$$\langle \zeta | r \rangle = t \sum_{\langle \alpha, \beta \rangle} (\eta_{\beta} - \eta_{\alpha}) (\psi_{\alpha} \zeta_{\beta} - \psi_{\beta} \zeta_{\alpha}). \quad (\text{A26})$$

Next, since this sum is over the entire Fock-space graph, we select $\eta_{\alpha} = 1$ if $S_{\alpha} \leq \bar{S}/2 - \delta'$ and $\eta_{\alpha} = 0$ otherwise. Roughly speaking, we draw thick boundaries around each of the wells, and it is within these regions where $\eta = 0$. As before, for sufficiently small δ' , the region in which $\eta = 0$ will include all nodes that we require $\eta = 0$ at. As a result, the sum above is only nonzero for special edges $\langle \alpha, \beta \rangle \in \bigcup_{\ell} \partial \mathcal{G}_W(\ell)$ where $\eta_{\alpha} = 0$ and $\eta_{\beta} = 1$ or vice versa; furthermore, the nodes for which η is nonzero have action at least equal to $\bar{S}/2 - \delta'$. The analysis then proceeds identically to the previous theorem (with ϕ replaced with ψ), and use of Cauchy-Schwartz along with the other identities gives

$$\begin{aligned} |\langle \zeta | r \rangle|^2 &\leq \frac{k^2 t^2 N^2}{\epsilon} \frac{\max_{\alpha \in \bar{U}} \Delta_{\alpha}}{\min_{\alpha \in U} -\Delta_{\alpha}} e^{-\bar{S}} \left(\sum_{(\ell, a) \in A'} \gamma_{\ell, a}^2 \right) \\ &\leq O(N^3) e^{-\bar{S}}. \end{aligned} \quad (\text{A27})$$

In this case, the analysis is simpler because we do not need to make use of Lemma 6. Next, we choose $\gamma_{\ell,a} = v_{\ell,a} \text{sgn}(\lambda_{\ell,a} - \mu)$, which leads to the equation

$$\sum_{(\ell,a) \in A'} v_{\ell,a}^2 \leq \frac{O(N^3)}{\delta^2 \sigma} e^{-\bar{s}}, \quad (\text{A28})$$

leading to

$$\| [I - P_\phi(\mu - \delta, \mu + \delta)] \eta \psi \rangle \|_2 \leq O\left(\sqrt{\frac{N}{\sigma}}\right) \frac{N}{\delta} e^{-\bar{s}/2}. \quad (\text{A29})$$

Finally, we have a similar bound for $|(1 - \eta)\psi\rangle$ by utilizing the same logic as in the last theorem. Thus, we are done.

-
- [1] P. W. Anderson, *Phys. Rev.* **109**, 1492 (1958).
- [2] D. M. Basko, I. L. Aleiner, and B. L. Altshuler, *Ann. Phys.* **321**, 1126 (2006).
- [3] D. M. Basko, I. L. Aleiner, and B. L. Altshuler, *Phys. Rev. B* **76**, 052203 (2007).
- [4] D. M. Basko, I. L. Aleiner, and B. L. Altshuler, in *Problems of Condensed Matter Physics*, edited by A. L. Ivanov and S. G. Tikhodeev (Oxford University Press, Oxford, England, 2008), pp. 50–69.
- [5] V. Oganessian and D. A. Huse, *Phys. Rev. B* **75**, 155111 (2007).
- [6] A. Pal and D. A. Huse, *Phys. Rev. B* **82**, 174411 (2010).
- [7] J. Z. Imbrie, *J. Stat. Phys.* **163**, 998 (2016).
- [8] M. Žnidarič, T. Prosen, and P. Prelovšek, *Phys. Rev. B* **77**, 064426 (2008).
- [9] R. Vosk, D. A. Huse, and E. Altman, *Phys. Rev. X* **5**, 031032 (2015).
- [10] A. C. Potter, R. Vasseur, and S. A. Parameswaran, *Phys. Rev. X* **5**, 031033 (2015).
- [11] M. Serbyn and J. E. Moore, *Phys. Rev. B* **93**, 041424(R) (2016).
- [12] M. Serbyn, Z. Papić, and D. A. Abanin, *Phys. Rev. Lett.* **111**, 127201 (2013).
- [13] V. Ros, M. Müller, and A. Scardicchio, *Nucl. Phys. B* **891**, 420 (2015).
- [14] M. Schreiber, S. S. Hodgman, P. Bordia, H. P. Lüschen, M. H. Fischer, R. Vosk, E. Altman, U. Schneider, and I. Bloch, *Science* **349**, 842 (2015).
- [15] J.-y. Choi, S. Hild, J. Zeiher, P. Schauß, A. Rubio-Abadal, T. Yefsah, V. Khemani, D. A. Huse, I. Bloch, and C. Gross, *Science* **352**, 1547 (2016).
- [16] S. S. Kondov, W. R. McGehee, W. Xu, and B. DeMarco, *Phys. Rev. Lett.* **114**, 083002 (2015).
- [17] J. Smith, A. Lee, P. Richerme, B. Neyenhuis, P. W. Hess, P. Hauke, M. Heyl, D. A. Huse, and C. Monroe, *Nat. Phys.* **12**, 907 (2016).
- [18] R. Nandkishore and D. A. Huse, *Annu. Rev. Condens. Matter Phys.* **6**, 15 (2015).
- [19] D. A. Abanin, E. Altman, I. Bloch, and M. Serbyn, *Rev. Mod. Phys.* **91**, 021001 (2019).
- [20] I. V. Gornyi, A. D. Mirlin, and D. G. Polyakov, *Phys. Rev. Lett.* **95**, 206603 (2005).
- [21] M. Filoche and S. Mayboroda, *Proc. Natl. Acad. Sci. USA* **109**, 14761 (2012).
- [22] M. Lyra, S. Mayboroda, and M. Filoche, *Europhys. Lett.* **109**, 47001 (2015).
- [23] D. N. Arnold, G. David, D. Jerison, S. Mayboroda, and M. Filoche, *Phys. Rev. Lett.* **116**, 056602 (2016).
- [24] D. N. Arnold, G. David, M. Filoche, D. Jerison, and S. Mayboroda, *Commun. Partial Diff. Equations* **44**, 1186 (2019).
- [25] M. Filoche, M. Piccardo, Y.-R. Wu, C.-K. Li, C. Weisbuch, and S. Mayboroda, *Phys. Rev. B* **95**, 144204 (2017).
- [26] M. Piccardo, C.-K. Li, Y.-R. Wu, J. S. Speck, B. Bonef, R. M. Farrell, M. Filoche, L. Martinelli, J. Peretti, and C. Weisbuch, *Phys. Rev. B* **95**, 144205 (2017).
- [27] C.-K. Li, M. Piccardo, L.-S. Lu, S. Mayboroda, L. Martinelli, J. Peretti, J. S. Speck, C. Weisbuch, M. Filoche, and Y.-R. Wu, *Phys. Rev. B* **95**, 144206 (2017).
- [28] J. Buhl, I. Cinzori, I. Ginnett, M. Landry, Y. Li, and X. Liu, *arXiv:1907.00808*.
- [29] S. Iyer, V. Oganessian, G. Refael, and D. A. Huse, *Phys. Rev. B* **87**, 134202 (2013).
- [30] V. P. Michal, B. L. Altshuler, and G. V. Shlyapnikov, *Phys. Rev. Lett.* **113**, 045304 (2014).
- [31] M. Žnidarič and M. Ljubotina, *Proc. Natl. Acad. Sci. USA* **115**, 4595 (2018).
- [32] X. Li, X. Li, and S. Das Sarma, *Phys. Rev. B* **96**, 085119 (2017).
- [33] J. Šuntajs, J. Bonča, T. Prosen, and L. Vidmar, *arXiv:1905.06345*.
- [34] W. De Roeck and F. Huveneers, *Phys. Rev. B* **95**, 155129 (2017).
- [35] A. Polkovnikov (private communication).
- [36] N. R. Bernier, L. D. Tóth, A. K. Feofanov, and T. J. Kippenberg, *Phys. Rev. A* **98**, 023841 (2018).
- [37] M. Harder, Y. Yang, B. M. Yao, C. H. Yu, J. W. Rao, Y. S. Gui, R. L. Stamps, and C. M. Hu, *Phys. Rev. Lett.* **121**, 137203 (2018).
- [38] R. Abou-Chakra, D. Thouless, and P. Anderson, *J. Phys. C: Solid State Phys.* **6**, 1734 (1973).
- [39] S. Agmon, *Lectures on Exponential Decay of Solutions of Second-order Elliptic Equations: Bounds on Eigenfunctions of n-body Schrödinger Operations (Mathematical Notes Volume 29)* (Princeton University Press, Princeton, NJ, 2014).
- [40] S. Agmon, *Schrödinger Operators* (Springer, New York, 1985), pp. 1–38.
- [41] B. Simon, *Annales de l'I.H.P. Physique théorique* **38**, 295 (1983).
- [42] B. Helffer and J. Sjostrand, *Commun. Partial Diff. Equations* **9**, 337 (1984).
- [43] B. Simon, *Bull. Am. Math. Soc.* **8**, 323 (1983).
- [44] S. Mayboroda and B. Poggi, *Trans. Am. Math. Soc.* **372**, 4313 (2019).
- [45] B. Helffer, *Spectral and Scattering Theory and Applications* (Mathematical Society of Japan, Tokyo, 1994), pp. 113–141.
- [46] B. Helffer and B. Parrisé, *Annales de l'I.H.P. Physique théorique* **60**, 147 (1994).
- [47] X. P. Wang, *Annales de l'I.H.P. Physique théorique* **43**, 269 (1985).

- [48] S. Mayboroda, Invited talk at the Institute for Advanced Study, Princeton, NJ (March 8, 2016); available online: <https://www.youtube.com/watch?v=4OxJqrf8S-k>
- [49] D. Arnold, D. Guy, M. Filoche, D. Jerison, and S. Mayboroda, *SIAM J. Sci. Comput.* **41**, B69 (2019).
- [50] S. Mayboroda, Invited talk at the International Congress of Mathematicians (ICM 2018), Rio de Janeiro, Brazil (August 8, 2018); available online: <https://www.youtube.com/watch?v=FhPsWJL9eNQ>
- [51] G. David, M. Filoche, and S. Mayboroda, [arXiv:1909.10558](https://arxiv.org/abs/1909.10558).
- [52] I. Goldsheid, S. Molchanov, and L. Pastur, *Funct. Anal. Appl.* **11**, 8 (1977).
- [53] H. Kunz and B. Souillard, *Commun. Math. Phys.* **78**, 201 (1980).
- [54] A. Gangopadhyay, V. Galitski, and M. Müller, *Phys. Rev. Lett.* **111**, 026801 (2013).
- [55] S. Sachdev and J. Ye, *Phys. Rev. Lett.* **70**, 3339 (1993).