Contemporary Mathematics

The lattice Anderson model with discrete disorder

John Z. Imbrie

Abstract. Consider the Anderson tight-binding model on the lattice in any dimension, with single-site potential having a discrete distribution taking N values, with N large. We discuss recent work elucidating mechanisms by which randomness localizes eigenfunctions, smooths out eigenvalue distributions, and produces eigenvalue separation.

1. Introduction

Quantum systems with disorder present particular mathematical challenges. In order to understand the phenomenon of localization in such systems, it is important to get control over eigenvalue statistics, especially the density of states. The Anderson model [**And58**] describes a particle on a lattice with a random potential. If the distribution of the potential has some regularity, then one may control the density of states using the Wegner argument [**Weg81**]. But this breaks down if the potential has a discrete distribution.

The case of a binary distribution has been called the Anderson-Bernoulli model. Localization in one dimension has been proven on the lattice [**CKM87, SVW98**] and in the continuum [**DSS02**]. Further results on the lattice in one dimension include improved regularity of the density of states for weak disorder [**Bou12, Bou14**]. In higher dimensions, results on localization have been confined to the continuum case [**BK05**], where localization was established near the bottom of the spectrum, using a quantitative form of the unique continuation principle. This led to several generalizations in other cases involving singular potentials [**GK07, GHK07, GK13, KT16**]. Recently, the author has obtained results on localization and eigenvalue statistics for all energies for the lattice Anderson model with discrete disorder distribution taking N values, with N large [**Imb17**]. Here we present some of the key ideas in this work.

In the course of the analysis, it was essential to obtain a degree of energy level separation. This is a particular challenge in the case of a discrete distribution, because degeneracies appear with positive probability in small subsystems. Thus it is essential to demonstrate that as larger neighborhoods of a subsystem are taken into account, the disorder has a tendency to break these degeneracies. However,

2010 *Mathematics Subject Classification.* Primary 82D30, 82B44.

©0000 (copyright holder)

due to the exponential decay of influence of distant random variables, the gaps between eigenvalues may be extremely small.

It should be evident that controlling the statistical separation of eigenvalues is important tool for the mathematical study of disordered quantum systems. The author's proof of many-body localization [**Imb16b**] was dependent on an assumption on level-separation for that system. The usual Minami estimate [**Min96**] is not available for discrete disorder distributions. So in order to handle the level-separation problem, we use a method of iterated Schur complementation. A sequence of reductions to lower-dimensional problems with associated effective Hamiltonians allows a precise examination of the dependence on the disorder. The reductions are accomplished by eliminating sites or regions where the spectrum of the effective Hamiltonian is disjoint from the energy interval of interest. The regions that remain may be called resonant regions or resonant blocks. One may then examine the way the spectrum of the effective Hamiltonian for a resonant block depends on the disorder. We use the disorder in the vicinity of a resonant block to produce eigenvalue movement and separation.

2. Model and Results

We work on a rectangular sublattice $\Lambda \subset \mathbb{Z}^d$. The Anderson model Hamiltonian is

$$H = -\gamma\Delta + v, \tag{2.1}$$

where Δ is the lattice Laplacian and v is multiplication by the lattice potential v_x, $x \in \Lambda$. We assume that $\{v_x\}_{x\in\Lambda}$ is a collection of iid random variables, each with a uniform distribution on $\{0, \frac{1}{N-1}, \frac{2}{N-1}, \dots, 1\}$, with N an integer much greater than 1. This generalizes the Anderson-Bernouilli model, which corresponds to $N = 2$. The diagonal entries of H are $2d\gamma + v_x$, and the off-diagonal part of H is γJ, where

$$J_{xy} = \begin{cases} 1, & \text{if } |x-y| = 1; \\ 0, & \text{otherwise,} \end{cases} \tag{2.2}$$

and $|x| = \sum_{i=1}^d |x_i|$.

In stating our results, we use a parameter $p > d$, a sufficiently large constant that serves as the exponent for log-Hölder continuity of the density of states. We take N sufficiently large, depending on p. Then we take $\gamma \le \varepsilon^{20}$, where $\varepsilon \equiv \frac{1}{N-1}$. Let $I_\delta(E) = [E-\delta, E+\delta]$, and let $\mathcal{N}(I)$ denote the number of eigenvalues of H in I. Let $\{E_\beta, \varphi_\beta\}_{\beta=1,\dots,|\Lambda|}$ denote the eigenvalues and associated normalized eigenvectors of H.

Theorem 2.1. *Choose a sufficiently large p. Then for N sufficiently large (depending on p) and γ sufficiently small (depending on N),*

$$\mathbb{E}\,\mathcal{N}\big(I_\delta(E)\big) \le |\Lambda|(\log_\gamma \delta)^{-p}. \tag{2.3}$$

for any rectangle Λ and any $\delta \in [\gamma^{\mathrm{diam}(\Lambda)/2}, 1]$. Furthermore, the eigenfunction correlator satisfies

$$\mathbb{E}\sum_\beta |\varphi_\beta(x)\varphi_\beta(y)| \le \big(|x-y| \vee 1\big)^{-(p/2-d-1)}. \tag{2.4}$$

Finally, the minimum eigenvalue spacing satisfies

$$P\Big(\min_{\beta\neq\tilde\beta}|E_\beta - E_{\tilde\beta}| < \delta\Big) \leq |\Lambda|^2(\log_\gamma \delta)^{-(p/2-1)}, \tag{2.5}$$

for any rectangle Λ *and any* $\delta \in [\gamma^{\mathrm{diam}(\Lambda)/2}, 1]$.

Although we only establish power-law decay of the eigenfunction correlator, we do obtain exponential decay of the eigenfunctions, in the sense that

$$P\bigg(\max_{y:\,|y-x|\geq R}\sum_\beta |\varphi_\beta(x)\varphi_\beta(y)|\gamma^{-|x-y|/5} > 1\bigg) \leq R^{-(p/2-4d-1)}. \tag{2.6}$$

Note that (2.5) implies that the probability of an exact degeneracy in the spectrum of H decreases as a large power of $\mathrm{diam}(\Lambda)$.

3. Resonant Blocks and Random-Walk Expansions

3.1. Schur Complement. The following lemma from [**IM16**] shows how to reduce to an effective Hamiltonian on a lower-dimensional subspace. The Schur complement gives an accurate picture of the spectrum in some neighborhood about a target energy E.

LEMMA 3.1. *Let* K *be a* $(p+q)\times(p+q)$ *symmetric matrix in block form,* $K = \left(\begin{smallmatrix} A & B \\ C & D\end{smallmatrix}\right)$, *with* A *a* $p\times p$ *matrix,* D *a* $q\times q$ *matrix, and* $C = B^T$. *Assume that* $\|(D-E)^{-1}\| \leq \tilde\varepsilon^{-1}, \|B\| \leq \tilde\gamma, \|C\| \leq \tilde\gamma$. *Define the Schur complement with respect to* λ*:*

$$F_\lambda \equiv A - B(D-\lambda)^{-1}C. \tag{3.1}$$

Let $\tilde\varepsilon$ *and* $\tilde\gamma/\tilde\varepsilon$ *be small, and* $|\lambda - E| \leq \tilde\varepsilon/2$. *Then*

(1) If φ *is an eigenvector for* F_λ *with eigenvalue* λ, *then* $\big(\varphi, -(D-\lambda)^{-1}C\varphi\big)$ *is an eigenvector for* K *with eigenvalue* λ, *and all eigenvectors of* K *with eigenvalue* λ *are of this form.*

(2)

$$\|F_\lambda - F_E\| \leq 2\Big(\frac{\tilde\gamma}{\tilde\varepsilon}\Big)^2|\lambda - E|. \tag{3.2}$$

(3) The spectrum of K *in* $[E-\tilde\varepsilon/2, E+\tilde\varepsilon/2]$ *is in close agreement with that of* F_E *in the following sense. If* $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_m$ *are the eigenvalues of* K *in* $[E-\tilde\varepsilon/2, E+\tilde\varepsilon/2]$, *then there are corresponding eigenvalues* $\tilde\lambda_1 \leq \tilde\lambda_2 \leq \ldots \leq \tilde\lambda_m$ *of* F_E, *and* $|\lambda_i - \tilde\lambda_i| \leq 2(\tilde\gamma/\tilde\varepsilon)^2|\lambda_i - E|$.

PROOF. The first statement follows from the fact that $\Big(\left(\begin{smallmatrix} A & B \\ C & D\end{smallmatrix}\right) - \lambda\Big)\left(\begin{smallmatrix}\varphi\\ \tilde\varphi\end{smallmatrix}\right) = 0$ if and only if $C\varphi + (D-\lambda)\tilde\varphi = 0$ and $(F_\lambda - \lambda)\varphi = 0$. The second statement follows by writing

$$F_E - F_\lambda = B(D-E)^{-1}(\lambda - E)(D-\lambda)^{-1}C. \tag{3.3}$$

and using the assumed bounds on each factor. The last statement follows from Weyl's inequality and (3.2). □

The lemma provides an algorithm for finding the eigenvalues of K near E. Again using Weyl's inequality and (3.2), we see that the eigenvalues of F_λ are Lipschitz continuous functions of λ, with a small Lipschitz constant. Therefore, the eigenvalues of K near E may be found using a fixed point argument, to solve the condition $\lambda \in \mathrm{spec}\, F_\lambda$.

3.2. First Step. The spectrum of $H = -\gamma\Delta + v$ is contained in the interval $[0, 1+4d\gamma]$. Let us work in the vicinity of some energy E in this interval. Each v_x is a multiple of $\frac{1}{N-1}$ in $[0,1]$ in $[0,1]$, so the spacing between these values is $\varepsilon \equiv \frac{1}{N-1}$. Let us define $\varepsilon_1 \equiv \varepsilon/3$. We say that a site x is resonant to E if $v_x + 2d\gamma \in I_{\varepsilon_1}(E)$. It should be clear that at most one potential value v_x will permit x to be resonant to E. Thus the probability that x is resonant to E is bounded by $\frac{1}{N} < \varepsilon$. Since N is large, the set of resonant sites will typically be a dilute subset of Λ. We denote this subset by $R^{(1)}$.

It is natural to use the set of resonant sites as the basis for a block decomposition of the Hamiltonian:

$$H = \begin{pmatrix} A^{(1)} & B^{(1)} \\ C^{(1)} & D^{(1)} \end{pmatrix}, \tag{3.4}$$

with $A^{(1)}$ denoting the restriction of H to the subspace with indices in $R^{(1)}$, and $D^{(1)}$ denoting the restriction to the space with indices in $\Lambda \setminus R^{(1)}$. This leads to the Schur complement

$$F_\lambda^{(1)} \equiv A^{(1)} - B^{(1)}(D^{(1)} - \lambda)^{-1}C^{(1)}. \tag{3.5}$$

We may decompose $D^{(1)}$ into diagonal and off-diagonal parts:

$$D^{(1)} = W^{(1)} - V^{(1)}, \tag{3.6}$$

with $W_{xy}^{(1)} = (2d\gamma + v_x)\delta_{xy}$ and $V_{xy}^{(1)} = \gamma J_{xy}$. Let us take $\lambda - E \le \varepsilon_1/2$, so that $(2d\gamma + v_x - \lambda)^{-1} \le 2/\varepsilon_1$ for $x \in \Lambda \setminus R^{(1)}$. Then the Neumann series

$$(D^{(1)} - \lambda)^{-1} = (W^{(1)} - \lambda)^{-1} + (W^{(1)} - \lambda)^{-1}V^{(1)}(W^{(1)} - \lambda)^{-1} + \dots \tag{3.7}$$

converges for γ small. Since $V^{(1)}$ induces nearest-neighbor steps, we may expand the matrix products to obtain a random-walk expansion for $\left[B^{(1)}(D^{(1)} - \lambda)^{-1}C^{(1)}\right]_{xy}$ as a sum of graphs from x to y, with bounds that decay exponentially in the length of the walk. Thus the Schur complement matrix $F_\lambda^{(1)}$ has long-range (but exponentially decaying) interactions.

To set the stage for the next step, we introduce a sequence of length scales,

$$L_k = L_0 2^k, k = 1, 2, \dots, \tag{3.8}$$

where L_0 to be a large integer, whose choice will depend on the value of p. We form connected components of $R^{(1)}$ by declaring that x and y are connected if $|x - y| \le L_1^\alpha$, with $\alpha \equiv \frac{3}{2}$. Let B_1 denote one of these components. Then let $\bar{B}_1$ denote the set of lattice points within a distance $2L_1$ of B_1. Each component or block B_1 will be considered an indivisible unit (the analog of a site) in subsequent steps. We need to distinguish between small blocks satisfying $\mathrm{diam}(B_1) \le L_1$ and the rest. The small blocks are called *isolated* because the distance to other blocks is much greater than their diameter. For each isolated component, we define a localized version of $F_\lambda^{(1)}$:

$$[\tilde{F}_\lambda^{(1)}(B_1)]_{xy} \equiv A_{xy}^{(1)} - \sum_{g_1 : x \to y,\ g_1 \subseteq \bar{B}_1} \prod_{i=1}^{m} \frac{1}{2d\gamma + v_{x_i} - \lambda} \prod_{j=0}^{m} V_{x_j x_{j+1}}^{(1)}. \tag{3.9}$$

We have written out the random-walk expansion explicitly as a sum of walks g_1 from x to y, but here the walks are required to visit only sites within $\bar{B}_1$. The point

is to specify for the next step what constitutes the "diagonal" part of the effective Hamiltonian $F_\lambda^{(1)}$.

3.3. k^{th} Step. As in the first step, we need to define a notion of resonance. But this will only apply to the isolated blocks, because the others are still too singular to participate in the next random-walk expansion (they are too large). We define the energy window width for the k^{th} step:

$$\varepsilon_k \equiv \gamma^{1.6L_k}, \text{ for } k > 1. \tag{3.10}$$

This is designed to be a bit larger than anticipated bounds on the terms discarded in (3.9). In the k^{th} step, we are working with blocks B_{k-1} that are the connected components of the $(k-1)^{\text{st}}$ resonant set $R^{(k-1)}$. Here connectivity is defined using connections of length L_{k-1}^α. For each block we have a matrix $\tilde{F}_\lambda^{(k-1)}(B_{k-1})$ that is obtained by restricting the set of graphs that define $F_\lambda^{(k-1)}$ to those that remain within $\bar{B}_{k-1}$. Here the set $\bar{B}_{k-1}$ is essentially the neighborhood of width $2L_{k-1}$ about B_{k-1}, but its boundary is constructed so as to maintain a distance $L_j^{\sqrt{\alpha}}$ from the sets $\bar{B}_j$ from prior scales. These constructions are modeled on those in [**FS83**].

Definition 3.2. Let B_{k-1} be a component of $R^{(k-1)}$ with $k \geq 2$. We say that B_{k-1} is **resonant in step** k if it is isolated in step k and if

$$\operatorname{dist}\bigl(\operatorname{spec} \tilde{F}_E^{(k-1)}(B_{k-1}), E_k\bigr) \leq \varepsilon_k. \tag{3.11}$$

The new resonant set $R^{(k)}$ is obtained by deleting from $R^{(k-1)}$ all of its components that are isolated but not resonant in step k. As in the first step, $R^{(k)}$ determines the block decomposition

$$F_\lambda^{(k-1)} = \begin{pmatrix} A^{(k)} & B^{(k)} \\ C^{(k)} & D^{(k)} \end{pmatrix}. \tag{3.12}$$

Then we may define the next Schur complement matrix

$$F_\lambda^{(k)} = A^{(k)} - B^{(k)}(D^{(k)} - \lambda)^{-1} C^{(k)}. \tag{3.13}$$

In order to generate the random-walk expansion in the k^{th} step, we restrict to the neighborhood $|\lambda - E| \leq \varepsilon_k/2$ and write

$$D^{(k)} = W^{(k)} - V^{(k)}, \tag{3.14}$$

where $W^{(k)}$ is a block diagonal matrix formed out of blocks $\tilde{F}_\lambda^{(k-1)}(B_{k-1})$. The "off-diagonal" matrix $V^{(k)}$ contains the long graphs not included in $\tilde{F}_\lambda^{(k-1)}(B_{k-1})$. Then we expand $(D^{(k)} - \lambda)^{-1}$ in a Neumann series:

$$(D^{(k)} - \lambda)^{-1} = (W^{(k)} - \lambda)^{-1} + (W^{(k)} - \lambda)^{-1} V^{(k)} (W^{(k)} - \lambda)^{-1} + \dots. \tag{3.15}$$

Working inductively, we can use bounds on the random-walk expansion to prove a Lipschitz bound

$$\|\tilde{F}_\lambda^{(k-1)}(B_{k-1}) - \tilde{F}_E^{(k-1)}(B_{k-1})\| \leq \gamma |\lambda - E_k|. \tag{3.16}$$

This is less than $\varepsilon_k/6$, because $|\lambda - E| \leq \varepsilon_k/2$. This series involves only nonresonant blocks satisfying $\operatorname{dist}\bigl(\operatorname{spec} \tilde{F}_E^{(k-1)}(B_{k-1}), E\bigr) \geq \varepsilon_k$, and so $\|(W^{(k)} - \lambda)^{-1}\| \leq 3\varepsilon_k^{-1}$. Although this is a large factor, the graphs contributing to $V^{(k)}$ extend at least a distance $4L_{k-1} = 2L_k$ (from B_{k-1} to $\bar{B}_{k-1}^{\text{c}}$ and back). Our inductive bounds on the random-walk expansion imply decay like $\gamma^{.85R}$ for graphs of length R. Thus

$\|V^{(k)}\| \le \gamma^{1.7L_k}$, which is sufficient to control a factor of $\varepsilon_k^{-1} = \gamma^{-1.6L_k}$ (compare with the condition $\tilde{\gamma}/\tilde{\varepsilon} \ll 1$ in Lemma 3.1). See [**Imb17**] for further details on the random-walk expansion bounds.

4. Density of States Smoothing from Environmental Disorder

The basic mechanism for demonstrating some degree of smoothness in the density of states is the reduction of the size of $R^{(k)}$ as k grows and ε_k shrinks. This goes hand in hand with a decrease in $\mathcal{N}\big(I_\delta(E)\big)$ as δ decreases. Consider the fate of a block B_{k-1} in the k^{th} step. There are four possibilities. The block may fail to be resonant in step k, in which case it is not part of $R^{(k)}$. It may join up with other blocks in forming a component B_k of $R^{(k)}$. It may survive to $R^{(k)}$ on account of its not being isolated. Finally, it may be isolated and resonant in step $k-1$ and remain resonant in step k. In the first case, the requisite reduction in the size of $R^{(k)}$ does occur. In the second and third cases, there is no reduction, but the probability is small, due to the large size of the blocks involved. The fourth case requires careful analysis, because we need to argue for the rarity of the set of potentials in $\bar{B}_k \setminus \bar{B}_{k-1}$ that permit the continued resonance of B_{k-1} in step k.

4.1. Influential Sites. We begin the analysis by establishing the existence of sites with significant influence on the relevant eigenvalues. Let H_X denote the matrix obtained by restricting H to $X \subseteq \Lambda$. It is useful to note that $\tilde{F}_\lambda^{(k-1)}(B_{k-1})$ is the same as what we would have obtained for $F_\lambda^{(k-1)}(B_{k-1})$, if we had begun with $H_{\bar{B}_{k-1}}$. (This is the effect of restricting graphs to $\bar{B}_{k-1}$.) Thus, any solution to $\lambda \in \operatorname{spec} \tilde{F}_\lambda^{(k-1)}(B_{k-1})$ in $I_{\varepsilon_k/2}(E_k)$ corresponds to an eigenvalue of $H_{\bar{B}_{k-1}}$, via repeated application of Lemma 3.1. Likewise, the eigenfunctions φ_λ corresponding to each such solution map to an eigenfunction ψ_λ of $H_{\bar{B}_{k-1}}$. The lemma shows that we may write this map as $\psi_\lambda = G_\lambda^{(k-1)}\varphi_\lambda$, where $G_\lambda^{(k-1)}$ is given by a product of operators $\left(\begin{smallmatrix} I \\ -(D^{(j)}-\lambda)^{-1}C^{(j)} \end{smallmatrix}\right)$ mapping functions on $R^{(j)}$ to functions on $R^{(j-1)}$ (working all the time in $\bar{B}_{k-1}$).

LEMMA 4.1. *Let B_{k-1} be isolated in step k, i.e.* $\operatorname{diam}(B_{k-1}) \le L_{k-1}$. *Let $\psi = G_\lambda^{(k-1)}\varphi$, with $\big(\tilde{F}_\lambda^{(k-1)}(B_{k-1}) - \lambda\big)\varphi = 0$, $\|\varphi\| = 1$, and $\lambda \in I_{\varepsilon_k/2}(E)$. Then $(H_{\bar{B}_{k-1}} - \lambda)\psi = 0$. For any y with* $\operatorname{dist}(y, \bar{B}_{k-1}) = 1$, *define the influence of y as*

$$\mathcal{I}_\psi(y) = \left| \sum_{x \in \bar{B}_{k-1},\, |x-y|=1} \psi(x) \right|. \tag{4.1}$$

If $\operatorname{diam}(\bar{B}_{k-1}) < \operatorname{diam}(\Lambda)$, *then for γ small, there exists at least one $y \in \Lambda \setminus \bar{B}_{k-1}$ with $\mathcal{I}_\psi(y) \ge \gamma^{3.1L_{k-1}}$.*

PROOF. Since φ is normalized, there must be a point $\bar{x}$ of B_{k-1} such that $|\varphi(\bar{x})| \ge |B_{k-1}|^{-1/2}$. Let us use $\bar{x}$ as the origin of coordinates in Λ. Then let us assume that the z-coordinate runs toward a boundary face of Λ that contains no points of $\bar{B}_{k-1}$ – this is possible because $\operatorname{diam}(\bar{B}_{k-1}) < \operatorname{diam}(\Lambda)$. The construction of $\bar{B}_{k-1}$ ensures that it extends no further than a distance $2.05L_{k-1}$ from B_{k-1}. Let $z_{\max}$ denote the maximal z-coordinate for points in $\bar{B}_{k-1}$. Then $z_{\max} \le 3.05L_{k-1}$.

Working toward a proof by contradiction, let us suppose that there is no site $y \in \Lambda \setminus \bar{B}_{k-1}$ with $\mathcal{I}_\psi(y) \ge \gamma^{3.1L_{k-1}}$. Then for each x_0 in the top layer at $z = z_{\max}$

we have $|\psi(x_0)| < \gamma^{3.1L_{k-1}}$. (Each site y with d^{th} coordinate $z_{\max}+1$ is in Λ and is adjacent to no more than one site of $\bar{B}_{k-1}$, so the sum in (4.1) reduces to a single term.)

Let x be a site of $\bar{B}_{k-1}$ that is immediately below a top-layer site x_0 of $\bar{B}_{k-1}$. Let $y_1, \dots, y_{2d-1}$ denote the other neighbors of x_0. Then

$$-\gamma\Big(\psi(x) + \sum_{i-1}^{2d-1} \psi(y_i)\Big) + (2d\gamma + v_{x_0} - \lambda)\psi(x_0) = 0, \tag{4.2}$$

where we put $\psi(x) = 0$ for $x \notin \bar{B}_{k-1}$. Observing that $\lambda \in [0, 1+4d\gamma]$, $v_{x_0} \in [0,1]$, we have that

$$|2d\gamma + v_{x_0} - \lambda| \le 1 + 2d\gamma. \tag{4.3}$$

Hence

$$|\psi(x)| \le [\tfrac{1}{\gamma}(1+2d\gamma) + 2d - 1]\gamma^{3.1L_{k-1}} = (\tfrac{1}{\gamma} + 4d - 1)\gamma^{3.1L_{k-1}} \le \tfrac{2}{\gamma}\gamma^{3.1L_{k-1}}. \tag{4.4}$$

The remaining sites x in the second layer lie below a site $y \notin \bar{B}_{k-1}$, which then must satisfy $\mathcal{I}_\psi(y) < \gamma^{3.1L_{k-1}}$. We have already established that the other neighbors of y satisfy $|\psi(y_i)| \le \gamma^{3.1L_{k-1}}$. Therefore, $|\psi(x)| \le 2d\gamma^{3.1L_{k-1}}$ (otherwise, even after a cancellation with the other neighbors, $\mathcal{I}_\psi(y)$ would be too large). Thus (4.4) holds for all sites in the second layer (for γ small).

We continue this argument on successive layers, obtaining a bound

$$|\psi(x)| \le \big(\tfrac{2}{\gamma}\big)^{z_{\max}-z}\gamma^{3.1L_{k-1}} \le (2\gamma)^{.05L_{k-1}} \tag{4.5}$$

for the layer with d-coordinate $z \ge 0$. Thus we learn that $|\psi(\bar{x})| = |\varphi(\bar{x})| \le (2\gamma)^{.05L_{k-1}}$. This contradicts our assumption that $|\varphi(\bar{x})| \ge |B_{k-1}|^{-1/2} \ge (2L_{k-1}+1)^{-d/2}$. Hence there must be at least one influential site $y \in \Lambda \setminus \bar{B}_{k-1}$ satisfying $\mathcal{I}_\psi(y) \ge \gamma^{3.1L_{k-1}}$. □

Note that this argument does not work for every lattice. For example, on the Kagome lattice, some points are connected to more than one point at the next level down. See Fig. 1. Indeed, on the Kagome lattice there may be compactly supported eigenfunctions, which would be immune to the influence of disorder on neighboring sites. In contrast, on the rectangular lattice, eigenfunctions cannot exhibit uniform decay that is faster than an exponential [**CS83**].

4.2. Local Degeneracy of the Spectrum. We will need to follow the behavior of the number of eigenvalues in small windows around various energies. Define for each k and each component of $R^{(k)}$

$$\hat{n}_k(B_k) = \text{the number of eigenvalues of } \tilde{F}_E^{(k)}(B_k) \text{ in } I_{\varepsilon_{k+1}}(E). \tag{4.6}$$

Here we count eigenvalues with multiplicity.

PROPOSITION 4.2. *Let L_0 be sufficiently large. Take $\varepsilon = \frac{1}{N-1}$ to be sufficiently small, depending on L_0, and take $\gamma \le \varepsilon^{20}$. Assume that B_{k-1} is isolated and resonant in step k with respect to energy E, and that $\mathrm{diam}(\bar{B}_{k-1}) < \mathrm{diam}(\Lambda)$. Assume that B_{k-1} remains isolated in step $k+1$, so that $B_k = B_{k-1}$. Fix all v_y for $y \in \bar{B}_{k-1}$. These determine a particular $\bar{y}$ such that $\mathrm{dist}(\bar{y}, \bar{B}_{k-1}) = 1$. Fix all remaining $v_y \in \bar{B}_k$, $y \ne \bar{y}$. Then*

$$\hat{n}_k(B_k) \le \hat{n}_{k-1}(B_{k-1}). \tag{4.7}$$

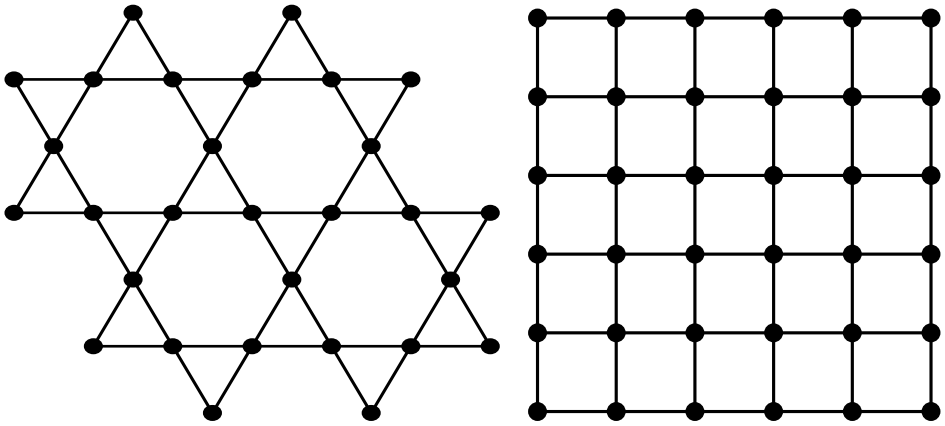

FIGURE 1. On the Kagome lattice (left), some points are connected to more than one point with smaller z-coordinate. Each point on the rectangular lattice (right) is singly connected to the layer below.

Furthermore,

$$\hat{n}_k(B_k) < \hat{n}_{k-1}(B_{k-1}) \tag{4.8}$$

for all but one value of $v_{\bar{y}}$.

We outline the main ideas behind the proof. To begin, we need to produce a useful representation for

$$\Delta F_\lambda \equiv \tilde{F}_\lambda^{(k-1)}(B_{k-1}) - \tilde{F}_\lambda^{(k)}(B_k), \tag{4.9}$$

which represents the change in the effective Hamiltonian as we expand the domain from $\bar{B}_{k-1}$ to $\bar{B}_k$. This will lead to a choice of $\bar{y}$ as a site adjacent to $\bar{B}_{k-1}$ with maximum influence on the set of eigenvalues in $I_{\varepsilon_k/2}(E)$, based on the leading term in this representation. Lemma 4.1 provides a lower bound on this influence. We show that higher-order terms are either smaller than the leading term, or independent of $v_{\bar{y}}$. Then it will be possible to demonstrate sufficient eigenvalue movement to obtain (4.8).

Let λ_0 be the closest eigenvalue of $\tilde{F}_E^{(k-1)}(B_{k-1})$ to E. We can assume that $\lambda_0 \in I_{\varepsilon_k/9}(E)$, because otherwise we would have that $\hat{n}(B_k) = 0$, due to the rapid contraction of energy windows. Let λ be the solution to $\lambda \in \operatorname{spec} \tilde{F}_\lambda^{(k-1)}(B_{k-1})$ that is closest to λ_0. Using the Lipschitz continuity of $\tilde{F}_\lambda^{(k-1)}(B_{k-1})$, there is a solution satisfying $|\lambda - E| \le \varepsilon_k/6$. Define

$$\hat{n} = \text{the number of eigenvalues of } \tilde{F}_\lambda^{(k-1)}(B_{k-1}) \text{ in } I_{\varepsilon_k/2}(E_{k+1}). \tag{4.10}$$

Then we have that

$$\hat{n} \le \hat{n}_{k-1}(B_{k-1}), \tag{4.11}$$

because eigenvalues of $\tilde{F}_E^{(k-1)}(B_{k-1})$ outside of $I_{\varepsilon_k}(E)$ cannot migrate into $I_{\varepsilon_k/2}(E)$ with the change $E \to \lambda$.

Using resolvent identities, one can show that

$$\Delta F_\lambda = \tilde{G}_\lambda^{(k-1)\mathrm{tr}} \Gamma \big(H_{\bar{B}_k \setminus B_k} - \lambda\big)^{-1} \Gamma \tilde{G}_\lambda^{(k-1)}, \tag{4.12}$$

where

$$\Gamma_{xy} = \begin{cases} \gamma, & \text{if } |x-y| = 1 \text{ with one in } \bar{B}_{k-1} \text{ and the other in } \bar{B}_k; \\ 0, & \text{otherwise.} \end{cases} \tag{4.13}$$

Here $\tilde{G}_\lambda^{(k-1)} : L^2(B_{k-1}) \to L^2(\bar{B}_{k-1} \setminus B_{k-1})$ is the operator providing the extension of the eigenfunction from B_{k-1} to $\bar{B}_{k-1} \setminus B_{k-1}$. Thus $G_\lambda^{(k-1)} = \left(\begin{smallmatrix} I \\ \tilde{G}_\lambda^{(k-1)} \end{smallmatrix} \right)$. We may interpret (4.12) graphically: the operators $\tilde{G}_\lambda^{(k-1)}$ involve graphs extending from B_{k-1} to its boundary, then Γ produces a step into $\bar{B}_k \setminus \bar{B}_{k-1}$, and $\big(H_{\bar{B}_k \setminus B_k} - \lambda\big)^{-1}$ connects the two ends of the graph with walks in $\bar{B}_k \setminus B_k$. All these graphical expansions are under control, with bounds $\sim \gamma^{.85R}$ for graphs of length R.

Let us use a basis of normalized eigenvectors $\{\varphi_1, \dots, \varphi_{\hat{n}}, \varphi_{\hat{n}+1}, \dots, \varphi_n\}$ corresponding to eigenvalues $\lambda_1, \dots, \lambda_{\hat{n}}, \lambda_{\hat{n}+1}, \dots, \lambda_n$. of $\tilde{F}_\lambda^{(k-1)}(B_{k-1})$. Here $\lambda = \lambda_1, \lambda_2, \dots, \lambda_{\hat{n}}$ are the eigenvalues in $I_{\varepsilon_k/2}(E_{k+1})$, and $n = |B_{k-1}|$ is the number of sites in B_{k-1}. In this basis,

$$\Delta F_{\lambda,\beta\tilde{\beta}} = \big\langle \varphi_\beta, \tilde{G}_\lambda^{(k-1)\mathrm{tr}} \Gamma \big(H_{\bar{B}_k \setminus B_k} - \lambda\big)^{-1} \Gamma \tilde{G}_\lambda^{(k-1)} \varphi_{\tilde{\beta}} \big\rangle. \tag{4.14}$$

Define for any $y \in \bar{B}_k \setminus \bar{B}_{k-1}$ with $\mathrm{dist}(y, \bar{B}_{k-1}) = 1$

$$\chi_y(x) = \begin{cases} 1, \text{ if } |x-y| = 1 \text{ and } x \in \bar{B}_{k-1}; \\ 0, \text{ otherwise.} \end{cases} \tag{4.15}$$

This is the indicator function for the set of sites in $\bar{B}_{k-1}$ that are adjacent to y. Then put

$$a_\beta(y) \equiv \langle \chi_y, \tilde{G}_\lambda^{(k-1)} \varphi_\beta \rangle. \tag{4.16}$$

The vector $\mathbf{a}^{(\mathrm{r})}(y) = (a_1(y), \dots, a_{\hat{n}}(y))$ – in particular its length-squared $|\mathbf{a}^{(\mathrm{r})}(y)|^2 = \sum_{\beta=1}^{\hat{n}} a_\beta(y)^2$ – is a measure of the influence of v_y on the family of eigenvalues $\{\lambda_1, \dots, \lambda_{\hat{n}}\}$. Choose $\bar{y} \in \Lambda$ with $\mathrm{dist}(\bar{y}, \bar{B}_{k-1}) = 1$ to be a site that maximizes $|\mathbf{a}^{(\mathrm{r})}(y)|$ from amongst all neighbors of $\bar{B}_{k-1}$. Lemma 4.1 implies that $|a_1(y)| \geq \gamma^{3.1 L_{k-1}}$ for at least one y adjacent to $\bar{B}_{k-1}$. Hence $|\mathbf{a}^{(\mathrm{r})}(\bar{y})| \geq \gamma^{3.1 L_{k-1}}$.

Let us write

$$\Delta F_{\lambda,\beta\tilde{\beta}} = \sum_{xy} a_\beta(x) K(x,y) a_{\tilde{\beta}}(y), \tag{4.17}$$

where

$$K(x,y) \equiv \gamma^2 \big(H_{\bar{B}_k \setminus B_k} - \lambda\big)^{-1}_{xy} = K_0(x,y) + K_1(x,y) + K_2(x,y). \tag{4.18}$$

for $x, y \in \bar{B}_k \setminus \bar{B}_{k-1}$ and adjacent to $\bar{B}_{k-1}$. Here K_0 is the sum of graphs for $\big(H_{\bar{B}_k \setminus \bar{B}_{k-1}} - \lambda\big)^{-1}$ that do not include the site $\bar{y}$ (which means that K_0 is independent of $v_{\bar{y}}$). For K_1, only the trivial graph of at $\bar{y}$ is included; thus

$$K_1(x,y) = \frac{\gamma^2 \delta_{x\bar{y}} \delta_{\bar{y}y}}{v_{\bar{y}} + 2d\gamma - \lambda}. \tag{4.19}$$

The remaining graphs make up K_2; they must contain $\bar{y}$ and have at least three steps.

To proceed further, we would need to make another Schur complement, to focus on the modes $\{\varphi_1, \dots, \varphi_{\hat{n}}\}$. But let us avoid this complication by assuming that $n = \hat{n}$. The following proposition gives the estimates needed for the proof of Proposition 4.2.

PROPOSITION 4.3. *Let L_0 be sufficiently large. Take $\varepsilon = \frac{1}{N-1}$ to be sufficiently small, depending on L_0, and take $\gamma \le \varepsilon^{20}$. Assume that B_{k-1} is isolated and resonant in step k with respect to energy E_k, and that B_{k-1} remains isolated in step $k+1$, so that $B_{k-1} = B_k$. Given E, there is a decomposition*

(4.20)
$$\big(\tilde{F}_\lambda^{(k-1)}(B_{k-1}) - \tilde{F}_E^{(k)}(B_k)\big)_{\beta\tilde{\beta}} = \frac{\gamma^2}{v_{\bar{y}} + 2d\gamma - \lambda} a_\beta^{(r)}(\bar{y}) a_{\tilde{\beta}}^{(r)}(\bar{y}) + \mathcal{C}_{\beta\tilde{\beta}} + \mathcal{R}(v_{\bar{y}})_{\beta\tilde{\beta}},$$

where $\mathcal{C}$ is independent of $v_{\bar{y}}$, and

$$\|\mathcal{C}\| \le \gamma\varepsilon_k \tag{4.21}$$

$$\|\mathcal{R}(v_{\bar{y}})\| \le \gamma^{2.5} |\mathbf{a}^{(r)}(\bar{y})|^2. \tag{4.22}$$

PROOF. We put

(4.23)
$$\tilde{F}_\lambda^{(k-1)}(B_{k-1}) - \tilde{F}_E^{(k)}(B_k) = \big(\tilde{F}_\lambda^{(k-1)}(B_{k-1}) - \tilde{F}_\lambda^{(k)}(B_k)\big) - \big(\tilde{F}_\lambda^{(k)}(B_k) - \tilde{F}_E^{(k)}(B_k)\big).$$

Consider the first term in (4.23), which corresponds to (4.17). The K_1 term appears explicitly in (4.20). The K_0 term may be denoted $\mathcal{C}^{(1)}$; it is independent of $v_{\bar{y}}$, and has a minimum decay distance $4L_{k-1} = 2L_k$ from the trip from B_{k-1} to $\bar{B}_k \setminus \bar{B}_{k-1}$ and back. Thus $\|\mathcal{C}^{(1)}\| \le \gamma^{.85 \cdot 2L_k} = \gamma^{1.7L_k} \le \frac{1}{2}\gamma\varepsilon_k$, since $\varepsilon_k = \gamma^{1.6L_k}$. Denoting the K_2 term by $\mathcal{R}^{(1)}(v_{\bar{y}})$, recall that $\bar{y}$ is defined as the site that maximizes $|\mathbf{a}^{(\mathrm{r})}(y)|$, so $|\mathbf{a}^{(\mathrm{r})}(y)| \le |\mathbf{a}^{(\mathrm{r})}(\bar{y})|$ for all y adjacent to $\bar{B}_{k-1}$. Therefore,

$$\|\mathcal{R}^{(1)}(v_{\bar{y}})\| \le \sum_{xy} \|\mathbf{a}^{(\mathrm{r})}(x)\mathbf{a}^{(\mathrm{r})}(y)^{\mathrm{tr}}\| |K_2(x,y)| \le \tfrac{1}{2}\gamma^{2.5} |\mathbf{a}^{(\mathrm{r})}(\bar{y})|^2. \tag{4.24}$$

We have used the fact that the norm of an outer product matrix $\mathbf{u}\mathbf{w}^{\mathrm{tr}}$ is bounded by $|\mathbf{u}||\mathbf{w}|$. Also, we have used $\sum_{xy} |K_2(x,y)| \le \frac{1}{2}\gamma^{2.5}$, because the relevant graphs are tied to a single point, $\bar{y}$, and have minimum length 3 (the leading term with length 2 is in K_0). This is a critical bound that controls remainders in terms of $|\mathbf{a}^{(\mathrm{r})}(\bar{y})|^2$, which governs the lower bound on the leading term.

The second term of (4.23) may be written as a sum of $\mathcal{C}^{(2)} + \mathcal{R}^{(2)}(v_{\bar{y}})$, with $\mathcal{C}^{(2)} = \tilde{F}_\lambda^{(k-1)}(B_{k-1}) - \tilde{F}_E^{(k-1)}(B_{k-1})$ and $\mathcal{R}^{(2)}(v_{\bar{y}}) = \Delta F_E - \Delta F_\lambda$. Note that $\mathcal{C}^{(2)}$ is independent of the potentials in $\bar{B}_k \setminus \bar{B}_{k-1}$. It is bounded by $\frac{1}{2}\gamma\varepsilon_k$, by the Lipschitz continuity estimate for $\tilde{F}_\lambda^{(k-1)}(B_{k-1})$, since $|\lambda - E| \le \varepsilon_k/2$. The term $\mathcal{R}^{(2)}(v_{\bar{y}})$ also involves a difference $|\lambda - E|$, and as above the graphs involved span a distance $4L_{k-1}$. Hence it is bounded by $\gamma^{1.7L_k}\varepsilon_k \le \gamma^{3.3L_k} \le \frac{1}{2}\gamma^{2.5}|\mathbf{a}^{(\mathrm{r})}(\bar{y})|^2$, because $|\mathbf{a}^{(\mathrm{r})}(\bar{y})|^2 \ge (\gamma^{3.1L_{k-1}})^2 = \gamma^{3.1L_k}$. If we write $\mathcal{C} = \mathcal{C}^{(1)} + \mathcal{C}^{(2)}$, $\mathcal{R}(v_{\bar{y}}) = \mathcal{R}^{(1)}(v_{\bar{y}}) + \mathcal{R}^{(2)}(v_{\bar{y}})$, and combine the bounds proven above, we obtain (4.21), (4.22). □

PROOF OF PROPOSITION 4.2. We will again simplify the analysis by assuming that $n = \hat{n}$. Then the inequality $\hat{n}_k(B_k) \le \hat{n}$ is true by construction, as $\hat{n}$ is the dimension of the matrix $\tilde{F}_E^{(k)}(B_k)$. Combining this fact with (4.11), we obtain (4.7).

With $\lambda \in [0, 1 + 4d\gamma]$, take any v_1, v_2 in the set of allowed potential values $\{0, \frac{1}{N-1}, \frac{2}{N-1}, \dots, 1\}$. A straightforward calculation shows that for small γ,

$$\left| \frac{1}{v_1 + 2d\gamma - \lambda} - \frac{1}{v_2 + 2d\gamma - \lambda} \right| \geq \frac{1}{N}. \tag{4.25}$$

Consequently, there are N distinct values of $(v + 2d\gamma - \lambda)^{-1}$ as v varies, and the minimum gap between these values is $\frac{1}{N}$.

Consider the case $\hat{n} = 1$. Then the matrix $a^{(\mathrm{r})}_{\beta}(\bar{y}) a^{(\mathrm{r})}_{\tilde{\beta}}(\bar{y})$ reduces to a number $|a^{(\mathrm{r})}_1(\bar{y})|^2 \geq \gamma^{3.1L_k}$. Likewise $\tilde{F}^{(k)}_E(B_k)$ and $\tilde{F}^{(k-1)}_\lambda(B_{k-1})$ are numbers, with the latter independent of $v_{\bar{y}}$. It is evident from Proposition 4.3 and (4.25) that the set of values that $\tilde{F}^{(k)}_E(B_k)$ takes as $v_{\bar{y}}$ varies is spaced apart by at least $(\gamma^2/N)|a^{(\mathrm{r})}_1(\bar{y})|^2 \gg \varepsilon_{k+1}$. Hence there is at most one value of $v_{\bar{y}}$ such that $\tilde{F}^{(k)}_E(B_k)$ lies in $I_{\varepsilon_{k+1}}(E)$.

Now consider the case $\hat{n} > 1$. The spread of a Hermitian matrix is defined as the difference between its largest and smallest eigenvalues. Weyl's inequality implies that

$$\mathrm{spread}(M_1 + M_2) \geq |\mathrm{spread}(M_1) - \mathrm{spread}(M_2)|. \tag{4.26}$$

Put $M_1 = \mathbf{a}^{(\mathrm{r})}(\bar{y}) \mathbf{a}^{(\mathrm{r})}(\bar{y})^{\mathrm{tr}}$. This is a rank-one matrix, so its spread is equal to its nonzero eigenvalue $|\mathbf{a}^{(\mathrm{r})}(\bar{y})|^2 \geq \gamma^{3.1L_k} \gg \varepsilon_{k+1}$.

From (4.20) we have that

$$\tilde{F}^{(k)}_E(B_k) = \tilde{F}^{(k-1)}_\lambda(B_{k-1}) - \frac{\gamma^2}{v_{\bar{y}} + 2d\gamma - \lambda} M_1 - \mathcal{C} - \mathcal{R}(v_{\bar{y}}). \tag{4.27}$$

Let $M_2 = f^{(k-1)}_\lambda - \mathcal{C}$. It should be clear that

$$\left| \mathrm{spread}(M_2) - \mathrm{spread}\left(\frac{\gamma^2}{v_{\bar{y}} + 2d\gamma - \lambda} M_1 \right) \right| \leq \frac{\gamma^2}{3N} |\mathbf{a}^{(\mathrm{r})}(\bar{y})|^2 \tag{4.28}$$

for at most one value of $v_{\bar{y}}$. (The values of the second spread are spaced out by at least $(\gamma^2/N)|a^{(\mathrm{r})}_1(\bar{y})|^2$, so cancellation as in (4.2) can happen at most once.) The bound on $\|\mathcal{R}(v_{\bar{y}})\|$ is much smaller than this, so

$$\mathrm{spread}\big(\tilde{F}^{(k)}_E(B_k)\big) \geq \Big(\frac{\gamma^2}{3N} - 2\gamma^{2.5} \Big) |\mathbf{a}^{(\mathrm{r})}(\bar{y})|^2 \geq \frac{\gamma^2}{4N} \gamma^{3.1L_k} \gg \varepsilon_{k+1} = \gamma^{3.2L_k}, \tag{4.29}$$

for all but one value of $v_{\bar{y}}$. Hence for all but one value of $v_{\bar{y}}$, at least one eigenvalue of $\tilde{F}^{(k)}_E(B_k)$ must fall outside of $I_{\varepsilon_{k+1}}(E)$. Thus we have demonstrated that $\hat{n}_k(B_k) < \hat{n}$ for all but one value of $v_{\bar{y}}$. Combining this with (4.11), we obtain (4.8). □

Remark 4.4. From (4.29) we see the importance of the fact that ε_{k+1}, the next energy window, fits inside the proven spread for $\tilde{F}^{(k)}_E(B_k)$, originating from the lower bound on the influence of $\bar{y}$. We have already seen that ε_{k+1} must be larger than any of the terms not included in $\tilde{F}^{(k)}_E(B_k)$, specifically graphs extending farther than $2L_k$ from B_k. It is inevitable that there is a gap between the decay rates that can be established for upper bounds and for lower bounds. Since the upper bound for step $k+1$ must fit inside the lower bound for step k, we are forced to use a sequence of length scales L_k that increase geometrically. Consequently, we are only able to make use of a single random potential per annular region on each scale. We obtain a Cantor-like spreading of the eigenvalue distribution that

develops with the logarithm of the length scale; this leads to log-Hölder continuity of the density of states as in (2.3).

4.3. Percolation Estimates. Proposition 4.2 gives the key result on the manner in which $\hat{n}_k(B_k)$, the local degeneracy of the spectrum, has a strong tendency to decrease with increasing k. The fact that $\hat{n}_k(B_k)$ fails to decrease for at most one value of $v_{\bar{y}}$ can be translated into probabilistic terms to demonstrate that the resonant regions $R^{(k)}$ get more and more dilute as k increases. The resonant blocks B_k can be considered components of a multi-scale percolation problem. Accurate control of this problem is a prerequisite for obtaining our main results. Here we discuss the problem in broad terms; the details may be found in [**Imb17**].

The first step blocks B_1 are actually the connected components of a site percolation problem, where the probability that a site is occupied is bounded by $\frac{1}{N} < \varepsilon$. Unlike the usual site percolation problem on $\mathbb{Z}^d$, connectedness is defined based on linking sites up to a distance L_1^α. Nevertheless, one can estimate the connectivity function (the probability that x and y are in the same percolation cluster) by summing over sequences of occupied sites, each separated from the previous one by a distance no greater than L_1^α, and including a factor of ε for each site. In the k^{th} step, components from the $(k-1)^{\text{st}}$ step are either resonant or not. The resonant ones are joined using connections up to a distance L_k^α. Thus we have a competition between the increasing distance over which connections are made and the decreasing probability that components remain resonant.

The simplest situation occurs when a component satisfies $\hat{n}_{k-1}(B_{k-1}) = 1$, *i.e.* the associated local Hamiltonian $\tilde{F}_E^{(k-1)}(B_{k-1})$ has only a single eigenvalue in $I_{\varepsilon_k}(E)$. Then if no new connection is made, then Proposition 4.2 implies that the probability that the component remains resonant is bounded by $\frac{1}{N-1} = \varepsilon$. (The probability in question is actually a conditional probability. Since we condition on the B_{k-1}, $\bar{B}_{k-1}$, and the set of potentials in $\bar{B}_{k-1}$, this eliminates one of the possible values for $v_{\bar{y}}$, as it would introduce a component B_1 adjacent to $\bar{B}_{k-1}$, which is incompatible with the definition of $\bar{B}_{k-1}$. Hence the denominator is $N-1$, not N. A similar situation occurs for ordinary site percolation: if B is a component, then neighbors of B are necessarily unoccupied.) If a single-site component B_1 remains resonant through step k, then the probability is bounded by $\varepsilon^k \leq L_k^{-q}$, provided ε is taken sufficiently small, depending on L_0 and q. (Recall that $L_k \equiv L_0 2^k$.) Thus we see that the procedure naturally leads to a decrease in probability as a large inverse power of the length scale. This is sufficient also to handle sums over $\sim L_k^{\alpha d}$ sites when making connections, provided $q > \alpha d$. Thus we should expect the connectivity function to decay as a large power $p = q - \alpha d$, when q is large enough. However, there is a complication in the multi-scale situation when two components, each with probability L_{k-1}^{-p}, are connected at a distance L_k^α. Then we require $(L_{k-1}^{-p})^2 \leq L_k^{-\alpha p}$, which means that α must be less than 2. On the other hand, to preserve exponential decay of the random-walk expansion, we require that $\alpha > 1$, so that the fraction of decay lost due to gaps at resonant blocks is summable. These limitations on α go back to [**FS83**].

Finally, let us consider the complication that arises when a component has multiple eigenvalues in the window, *i.e.* $\hat{n}_{k-1}(B_{k-1}) > 1$. (This is only an issue if no new connections are made, since new connections entail sufficient smallness, as discussed above.) Then we may fail to get the anticipated factor of ε when

going to the next scale. However, Proposition 4.2 shows that if this happens, $\hat{n}_k(B_k) < \hat{n}_{k-1}(B_{k-1}) = 1$. Thus there can be a delay in the accumulation of factors of ε while the degeneracy works its way down to 1. However, an n-fold degeneracy can only occur for a component of size n or more. Thus we have n factors of ε available at the start, and these are sufficient to make up for what is missing due to degeneracies. As mentioned earlier, the density of states problem is intertwined with the problem of degeneracy breaking or level separation.

Let us conclude this section by stating a bound on the connectivity function for the percolation problem in step k. It is a straightforward application of the percolation estimates of [**Imb17**]. Let $\mathcal{Q}_{xy}^{(k)}$ denote the probability that x and y are in the same component B_k of $R^{(k)}$. Then for any sufficiently large p, we may choose N sufficiently large (and γ sufficiently small) so that

$$\mathcal{Q}_{xy}^{(k)} \leq \varepsilon^{1/4}\big(|x-y| \vee L_{k-1}\big)^{-p}. \tag{4.30}$$

5. Proof of Main Results

5.1. Density of States. We are ready to prove statement (2.3) of Theorem 2.1, which is that $\mathbb{E}\mathcal{N}(I_\delta(E)) \leq |\Lambda|(\log_\gamma \delta)^{-p}$. As we are primarily interested in small δ, let us consider the case where

$$\varepsilon_{k+1}/3 < \delta \leq \varepsilon_k/3 \tag{5.1}$$

for some $k \geq 2$. In Theorem 2.1 we restrict to $\delta > \gamma^{\mathrm{diam}(\Lambda)/2}$; this is to ensure that value of k determined by (5.1) is such that that $\mathrm{diam}(\bar{B}_{k-1}) < \mathrm{diam}(\Lambda)$, as required by Proposition 4.2. (Once this is no longer true, we run out of random variables to continue the smoothing process.)

By repeated application of Lemma 3.1, we can see that the number of eigenvalues in $I_{\varepsilon_k/3}(E)$ is bounded by $|R^{(k)}|$, the dimension of the space on which $F_E^{(k)}$ acts. The probability that $x \in R^{(k)}$ may be bounded by L_{k-1}^{-p} by taking $y = x$ in the connectivity function (4.30). Then using (5.1) and $\varepsilon_{k+1} = \gamma^{6.4L_{k-1}}$, we obtain

$$\begin{aligned} \mathbb{E}\mathcal{N}\big(I_\delta(E)\big) &\leq \mathbb{E}\mathcal{N}\big(I_{\varepsilon_k/3}(E)\big) \leq \sum_x P\big(x \in R^{(k)}\big) \\ &\leq |\Lambda|\varepsilon^{1/4}L_{k-1}^{-p} = |\Lambda|\varepsilon^{1/4}\big(\tfrac{1}{6.4}\log_\gamma \varepsilon_{k+1}\big)^{-p} \leq |\Lambda|(\log_\gamma \delta)^{-p}. \end{aligned} \tag{5.2}$$

5.2. Energy-Following Procedure. Before proceeding to the other statements in Theorem 2.1, we need to discuss a new tool, which we call the energy-following procedure (EFP). There is a problem that arises from the fact that the eigenfunction correlator involves a sum over all eigenvalues, not just those that lie in a predetermined interval. Likewise, the statement on the minimum eigenvalue spacing applies to all eigenvalues of H. One cannot simply choose eigenvalues as the center points for energy windows, because they depend on the potentials throughout Λ, and we would lose the independence that is a crucial aspect of our percolation estimates. Nor can one consider a complete system of overlapping energy windows, because our windows are too narrow (and the probabilities are not small enough) to control the sum over all cases. (This strategy only works if the density of states is Hölder continuous with exponent greater than $\frac{1}{2}$ – see [**CGK09**].) The energy-following procedure was first introduced in [**IM16**] to deal with these issues in the context of the block Anderson model.

To deal with this issue, we construct all the eigenvalues through a sequence of successive approximations that are generated as the length-scale induction proceeds. Thus we allow the center point for energy windows to depend on k. To initiate the procedure, we choose a site $x \in \Lambda$ and let $E_1 = 2d\gamma + v_x$, the diagonal entry of H at x. Then we choose E_2 to be one of the solutions to $\lambda \in \operatorname{spec} \tilde{F}^{(1)}_\lambda(B_{x,1})$ in $I_{\varepsilon_1/3}$, where $B_{x,1}$ is the component of $R^{(1)}$ containing x. Continuing in this fashion, we obtain in each step a better approximate eigenvalue. Once L_k exceeds the diameter of Λ, the approximate eigenvalue becomes an exact eigenvalue of H. One can show that the EFP is complete, in the sense that every eigenvalue of H can be obtained by some sequence of choices $E_1, E_2, \ldots$. The key feature of the procedure is its *quasilocality*. This is the property that each E_k depends only on the potential in a neighborhood of x whose radius is some multiple of L_k (and subsequent corrections to the energy are exponentially small in L_k, with probability $\sim 1 - L_k^{-p}$). The way the energies are chosen, each $B_{x,k}$ is automatically resonant, but we preserve the conditional independence of events defined in terms of the potentials outside $\bar{B}_{x,k}$. Thus with minor modifications (in particular p becomes $p/2 - 1$), we obtain percolation estimates such as (4.30). The fact that the initial basis can be deformed into the eigenvector basis by means of quasilocal rotations has played a major role in the theory of many-body localization, see [**IRS17**] for a review. This was first established in the Anderson model in [**Imb16a**], and in the many-body context in [**Imb16b**], under an assumption on level statistics.

5.3. Eigenfunction Correlator and Level Spacing. Here we give a brief discussion of the proofs of the remaining statements in Theorem 2.1. We run the EFP up to the final step $\bar{k}$, at which point $\bar{B}_{x,\bar{k}} = \Lambda$. In the EFP each eigenfunction can be written as $G^{(\bar{k}-1)}_{E_{\bar{k}}}\varphi^{(\bar{k}-1)}$, for some eigenvector $\varphi^{(\bar{k}-1)}$ of $F^{(\bar{k}-1)}_{E_{\bar{k}}}$ (here $E_{\bar{k}}$ is the corresponding eigenvalue). Then we have

(5.3)
$$\sum_\beta |\varphi_\beta(y_1)\varphi_\beta(y_2)| \le \sum_x \sum_{\beta \text{ reachable from } x} \big|\big(G^{(\bar{k}-1)}_{E_{\bar{k}}}\varphi^{(\bar{k}-1)}_\beta\big)(y_1)\big(G^{(\bar{k}-1)}_{E_{\bar{k}}}\varphi^{(\bar{k}-1)}_\beta\big)(y_2)\big|.$$

In the second sum on the right-hand side, we are summing over all possible choices in the energy-following procedure, starting at x. (Each β may be reachable from more than one x.) Since $G^{(\bar{k}-1)}_{E_{\bar{k}}}$ decays exponentially, we may rewrite this as

$$\sum_\beta |\varphi_\beta(y_1)\varphi_\beta(y_2)| \le \sum_{x,z_1,z_2} \gamma^{.85|y_1-z_1|}\gamma^{.85|y_2-z_2|}N_{x,z_1,z_2}, \tag{5.4}$$

where $N_{x,y,z}$ denotes the number of eigenvalues of H that can be reached via the EFP with starting point x, and with a resonant region $B_{x,\bar{k}-1}$ that includes y and z. In order to bound $\mathbb{E}N_{x,y,z}$, we need to control the sum over choices in the EFP. These are never more numerous than the size n of the blocks involved, and a factor ε^n is available from the percolation estimates. Thus $\mathbb{E}N_{x,y,z}$ can be controlled as if it were a 3-point connectivity function for the percolation problem, and as in (4.30) we obtain decay as a large inverse power of $\operatorname{diam}(\{x,y,z\})$. As a result, we can take the expectation of (5.4) and perform the sums over x, y, z to obtain

$$\mathbb{E}\sum_\beta |\varphi_\beta(x)\varphi_\beta(y)| \le \big(|x-y| \vee 1\big)^{-(p/2-d-1)}, \tag{5.5}$$

which is (2.4).

We proceed to the level-spacing statement in Theorem 2.1. We may write

$$P\Big(\min_{\beta\neq\tilde\beta}|E_\beta - E_{\tilde\beta}| < \delta\Big) \leq \tfrac{1}{2}\sum_x \mathbb{E}N_x(\delta), \tag{5.6}$$

where $N_x(\delta)$ is the number of eigenvalues λ_0 of H that can be reached via the EFP starting at x, and which have another eigenvalue in $I_\delta(\lambda_0)$. Then (2.5) follows from the estimate

$$\mathbb{E}N_x(\delta) \leq 2|\Lambda|(\log_\gamma \delta)^{-(p/2-1)}, \tag{5.7}$$

where we take $\delta \in [\gamma^{\mathrm{diam}(\Lambda)}, \gamma]$. Let us define k by the inequality

$$\varepsilon_{k+1}/4 < \delta \leq \varepsilon_k/4, \tag{5.8}$$

and again we consider only the case $k \geq 2$. When we run the EFP starting at x, we have the condition that another eigenvalue is within δ of the final eigenvalue $E_{\bar k}$. This implies that one of two things happen. In case 1, the multiplicity $\hat n_{k-1}(B_{x,k-1})$ is greater than 1. Then the percolation estimates produce a bound $L_{k-1}^{-(p/2-1)}$ due to the continuous failure of the multiplicity to drop to 1. (In the EFP, we follow an eigenvalue, so the multiplicity cannot drop below 1, but Proposition 4.2 still leads to a factor of ε if it fails to drop below 2.) In case 2, there is a block B_{k-1} other than $B_{x,k-1}$ satisfying

$$\mathrm{dist}\big(\mathrm{spec}\,\tilde F_{E_k}^{(k-1)}(B_{k-1}), E_k\big) \leq \varepsilon_k, \tag{5.9}$$

so it is resonant in step k and survives to $R^{(k)}$. Thus as in the proof of the density of states bound (2.3), we obtain a bound of $|\Lambda|L_{k-1}^{-p}$ on the expectation for case 2. Combining the two cases, we obtain (5.7) to complete the proof.

References

[And58] P. Anderson, *Absence of diffusion in certain random lattices*, Phys. Rev. **109** (1958), 1492–1505.

[BK05] J. Bourgain and C. E. Kenig, *On localization in the continuous Anderson-Bernoulli model in higher dimension*, Invent. Math. **161** (2005), 389–426.

[Bou12] J. Bourgain, *On the Furstenberg measure and density of states for the Anderson-Bernoulli model at small disorder*, JAMA **117** (2012), 273–295.

[Bou14] ———, *An application of group expansion to the Anderson-Bernoulli model*, Geom. Funct. Anal. **24** (2014), 49–62.

[CGK09] J.-M. Combes, F. Germinet, and A. Klein, *Generalized eigenvalue-counting estimates for the Anderson model*, J. Stat. Phys. **135** (2009), 201–216.

[CKM87] R. Carmona, A. Klein, and F. Martinelli, *Anderson localization for Bernoulli and other singular potentials*, Commun. Math. Phys. **108** (1987), 41–66.

[CS83] W. Craig and B. Simon, *Log Hölder continuity of the integrated density of states for stochastic Jacobi matrices*, Commun. Math. Phys. **90** (1983), 207–218.

[DSS02] D. Damanik, R. Sims, and G. Stolz, *Localization for one-dimensional, continuum, Bernoulli-Anderson models*, Duke Math. J. **114** (2002), 59–100.

[FS83] J. Fröhlich and T. Spencer, *Absence of diffusion in the Anderson tight binding model for large disorder or low energy*, Commun. Math. Phys. **88** (1983), 151–184.

[GHK07] F. Germinet, P. D. Hislop, and A. Klein, *Localization for the Schrödinger operator with a Poisson random potential*, J. Eur. Math. Soc. **9** (2007), 577–607.

[GK07] F. Germinet and A. Klein, *Localization for some Cantor-Anderson Schrödinger operators*, Adventures in Mathematical Physics (F. Germinet and P. D. Hislop, eds.), Contem. Math., vol. 447, American Mathematical Society, Providence, RI, 2007, pp. 103–112.

[GK13] ———, *A comprehensive proof of localization for continuous Anderson models with singular random potentials*, J. Eur. Math. Soc. **15** (2013), 53–143.

[IM16] J. Z. Imbrie and R. Mavi, *Level spacing for non-monotone Anderson models*, J. Stat. Phys. **162** (2016), 1451–1484.

[Imb16a] J. Z. Imbrie, *Multi-scale Jacobi method for Anderson localization*, Commun. Math. Phys. **341** (2016), 491–521.

[Imb16b] ———, *On many-body localization for quantum spin chains*, J. Stat. Phys. **163** (2016), 998–1048.

[Imb17] ———, *Localization and eigenvalue statistics for the lattice anderson model with discrete disorder*, arXiv:1705.01916.

[IRS17] J. Z. Imbrie, V. Ros, and A. Scardicchio, *Local integrals of motion in many-body localized systems*, Ann. Phys. (Berlin) (2017), 1600278.

[KT16] A. Klein and C. Tsang, *Quantitative unique continuation principle for Schrödinger operators with singular potentials*, Proc. Am. Math. Soc. **144** (2016), 665–679.

[Min96] N. Minami, *Local fluctuation of the spectrum of a multidimensional Anderson tight binding model*, Commun. Math. Phys. **177** (1996), 709–725.

[SVW98] C. Shubin, R. Vakilian, and T. Wolff, *Some harmonic analysis questions suggested by Anderson-Bernoulli models*, Geom. Funct. Anal. **8** (1998), 932–964.

[Weg81] F. Wegner, *Bounds on the density of states in disordered systems*, Z. Phys. B **44** (1981), 9–15.

Department of Mathematics, University of Virginia, Charlottesville, VA 22904
E-mail address: imbrie@virginia.edu