Ann. Henri Poincaré 13 (2012), 525–541

1424-0637/12/030525-17
published online August 17, 2011
DOI 10.1007/s00023-011-0132-1

Annales Henri Poincaré

Lifshits Tails in the Hierarchical Anderson Model

Simon Kuttruf and Peter Müller

Abstract. We prove that the homogeneous hierarchical Anderson model exhibits a Lifshits tail at the upper edge of its spectrum. The Lifshits exponent is given in terms of the spectral dimension of the homogeneous hierarchical structure. Our approach is based on Dirichlet–Neumann bracketing for the hierarchical Laplacian and a large-deviation argument.

1. Introduction

Hierarchical models have a long tradition in statistical physics. Dyson [5,6] introduced them as an auxiliary tool in his study of phase transitions in the one-dimensional Ising ferromagnet with long-range interactions. An important feature of hierarchical models is that they preserve their structure under renormalisation-group transformations. Bleher and Sinai [2,3] exploited this to determine critical properties of hierarchical spin models.

Bovier [4] seems to be the first who studied the hierarchical Anderson model, that is, the Anderson model on a countably infinite configuration space with kinetic energy given by the hierarchical Laplacian. He also pursued a renormalisation-group approach and showed analyticity properties of the density of states. Under very mild conditions, Molchanov [15,16] established that the hierarchical Anderson model with Cauchy-distributed random variables has only pure point spectrum. In particular, his result does not require a homogeneous hierarchical structure. More recently, Kritchevski [11–13] continued the investigations of the hierarchical Anderson model. In [11,12] he removed the requirement for a Cauchy distribution and proved Anderson localisation at all energies and for general single-site distributions. However, his proof only works for homogeneous hierarchical structures with spectral dimension $d_s \leqslant 4$ (the spectral dimension will be introduced in Definition 2.4 below). In [13] he

This work has been partially supported by Sfb/Tr 12 of the German Research Foundation (DFG). Simon Kuttruf dedicates this work to Mark.

proved Poisson statistics of rescaled eigenvalue distributions for the homogeneous hierarchical Anderson model with $d_s < 1$. If the hierarchical Laplacian is modified as to contain also suitable negative hopping rates, then Monthus and Garel [17] argue in favour of a localisation–delocalisation transition in the hierarchical Anderson model. Fyodorov, Ossipov and Rodriguez found analytical and numerical evidence for a localisation–delocalisation transition for random matrices with a hierarchical structure of the correlations between the matrix entries [7].

In this paper, we prove the occurrence of a Lifshits tail at the upper spectral edge of the hierarchical Anderson model. More precisely, we only deal with homogeneous hierarchical structures for which one can define a spectral dimension d_s (but we do not impose any restriction on the value of d_s). In particular, we find that the Lifshits exponent of the integrated density of states of the hierarchical Anderson model coincides with $d_s/2$. Technically, we follow the approach that was successfully employed for Poisson and alloy-type random Schrödinger operators by Kirsch and Martinelli [9], and for the Anderson model on the lattice by Simon [21], see also the recent survey by Kirsch [8]. The method requires Dirichlet–Neumann bracketing for finite-volume operators which we establish for suitable finite-volume restrictions of the hierarchical Laplacian.

The quantity $d_s/2$, which we find for the Lifshits exponent in Theorem 2.6, is also referred to as the van Hove exponent, since it governs the van Hove "singularities" of the integrated density of states in the absence of disorder, see also Lemma 2.5. Amazingly, the equality of the Lifshits and van Hove exponent is known to hold for very different types of random hopping models. First and foremost we mention the standard alloy-type random Schrödinger operators in $\mathbb{R}^d$ or $\mathbb{Z}^d, d \in \mathbb{N}$. In that case $d_s = d$, and the Lifshits exponent equals $d/2$, see e.g. [9,21] and references therein. But it also holds for the integrated density of states of the Dirichlet Laplacian on percolation subgraphs. This was first shown for bond percolation on $\mathbb{Z}^d$ in [10,18] and then generalised to a large class of Cayley graphs in [1], see also the recent review [19]. From this perspective, Theorem 2.6 is interesting because it establishes the equality of the Lifshits and van Hove exponent for a random perturbation of the rather peculiar hierarchical Laplacian. The deeper reason behind it in all mentioned models is the intuition explained in Remark 2.7.

The paper is organised as follows. After introducing our notation and presenting the main result, Theorem 2.6, in Sect. 2, we turn to Dirichlet–Neumann bracketing for hierarchical finite-volume operators in Sect. 3. Finally, Sect. 4 is devoted to the proof of Theorem 2.6, and the appendix compiles the ergodic structure of the hierarchical Anderson model.

2. Model and Result

We consider a countably infinite configuration space $\mathbb{X}$ and the quantum Hamiltonian

$$H^\omega := \Delta + V^\omega \tag{2.1}$$

acting on the Hilbert space $\ell^2(\mathbb{X})$ of complex-valued, square-summable sequences over $\mathbb{X}$. The Hamiltonian describes diagonal disorder through its potential energy, which acts as the multiplication operator

$$(V^\omega\psi)(x) := \omega_x\,\psi(x) \quad \text{for all } \psi \in \ell^2(\mathbb{X}) \text{ and all } x \in \mathbb{X}. \tag{2.2}$$

Here, $\omega := (\omega_x)_{x\in\mathbb{X}}$ is a family of independent and identically distributed (i.i.d.), real-valued random variables. We think of them as being canonically realised in the probability space $\Omega := \mathbb{R}^{\mathbb{X}}$, equipped with the product Borel σ-algebra $\bigotimes_{x\in\mathbb{X}}\mathcal{B}_{\mathbb{R}}$ and the product probability measure $\mathbb{P} := \bigotimes_{x\in\mathbb{X}}\mathbb{P}_0$. We assume throughout that the single-site distribution $\mathbb{P}_0$ is compactly supported, $\operatorname{supp}\mathbb{P}_0 \subseteq [v_-, v_+]$ for some $v_-, v_+ \in \mathbb{R}$, in order to avoid irrelevant technical complications in dealing with unbounded operators. For the proof of the Lifshits tail in Theorem 2.6 we suppose in addition that $\mathbb{P}_0$ is not concentrated at one single point, i.e.

$$\mathbb{P}_0(\{v\}) < 1 \quad \text{for every } v \in \mathbb{R}, \tag{2.3}$$

and that its upper tail decays no faster than any power, i.e. there exist real constants $C, \mu > 0$ such that

$$\mathbb{P}_0([v_+ - \varepsilon, v_+]) \geqslant C\varepsilon^\mu \tag{2.4}$$

for every $\varepsilon > 0$ sufficiently small.

The operator Δ in (2.1) is the hierarchical Laplacian (2.7) and refers to a *hierarchical structure* on $\mathbb{X}$, which we need to explain first. A hierarchical structure on $\mathbb{X}$ is a sequence of partitions $(\mathcal{P}_r)_{r\in\mathbb{N}_0}$ of $\mathbb{X}$ together with a sequence $(n_r)_{r\in\mathbb{N}}$ of natural numbers such that the properties (H1)–(H3) below hold. By definition, each partition subdivides $\mathbb{X}$ into mutually disjoint subsets, which we call *clusters*. The clusters of $\mathcal{P}_r$ are referred to as *clusters of rank* r.

(H1) $\mathcal{P}_0$ is the trivial partition, the clusters of which consist precisely of the single elements of $\mathbb{X}$.
(H2) Every cluster of rank $r \in \mathbb{N}$ is a union of n_r distinct clusters of rank $r-1$.
(H3) Given $x, y \in \mathbb{X}$ there is a cluster of some rank containing both x and y.

We denote by $Q_r(x)$ the unique cluster of rank r containing $x \in \mathbb{X}$, and we write $|A|$ for the number of elements of a finite set A. By (H2) the number of elements $|Q_r(x)| = \prod_{r'=1}^{r} n_{r'}$ of this cluster does not depend on $x \in \mathbb{X}$. Thus, we will simply write $|Q_r|$ for the cluster size. The elements of $\mathbb{X}$ can be enumerated

$$\mathbb{N}_0 \to \mathbb{X}, \qquad k \mapsto x_k, \tag{2.5}$$

in such a way that x_{k_1} and x_{k_2} belong to the same cluster of rank r if and only if there exists $M \in \mathbb{N}_0$ with $k_1, k_2 \in \{M|Q_r|, \ldots, (M+1)|Q_r| - 1\}$, see Fig. 1.

A hierarchical structure is called *homogeneous of degree* $n \in \mathbb{N}$, if

$$n_r = n \quad \text{for all } r \in \mathbb{N}. \tag{2.6}$$

In this case, the size of any cluster of rank r is given by $|Q_r| = n^r$.

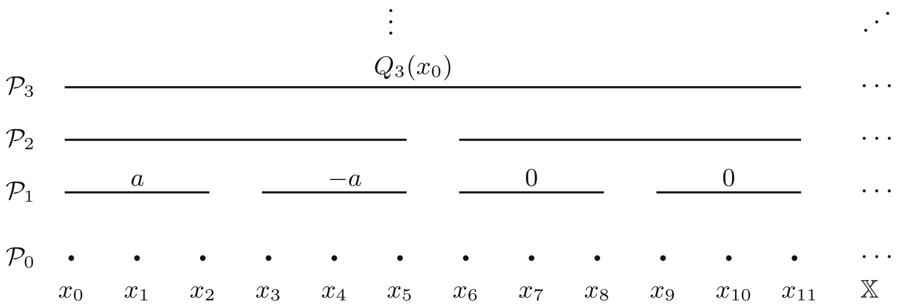

FIGURE 1. Sketch of a hierarchical structure for $n_1 = 3, n_2 = 2, n_3 = 2$ up to the rank-3-cluster containing x_0. A function taking on the values $\pm a \in \mathbb{C}\backslash\{0\}$, when restricted to the clusters $Q_1(x_0), Q_1(x_3)$, and being equal to zero everywhere else is an eigenfunction of the hierarchical Laplacian Δ corresponding to the eigenvalue λ_1

Given a sequence of probability weights $(p_s)_{s\in\mathbb{N}}, 0 < p_s < 1$ for all $s \in \mathbb{N}$ and $\sum_{s=1}^{\infty} p_s = 1$, the *hierarchical Laplacian* is defined as the weighted sum

$$\Delta := \sum_{s=1}^{\infty} p_s \, \mathbf{E}_s \tag{2.7}$$

of the *cluster averaging operators* $\mathbf{E}_s \colon \ell^2(\mathbb{X}) \to \ell^2(\mathbb{X})$,

$$(\mathbf{E}_s\psi)(x) := \frac{1}{|Q_s|} \sum_{y\in Q_s(x)} \psi(y), \tag{2.8}$$

where $\psi \in \ell^2(\mathbb{X}), x \in \mathbb{X}$ and $s \in \mathbb{N}_0$. For convenience, we also introduce $p_0 := 0$. Note that $\mathbf{E}_0 = \mathbf{1}$, the identity operator, and that $p_s \neq 0$ for every $s \in \mathbb{N}$. The random operator (2.1) is referred to as the *hierarchical Anderson Hamiltonian* or the *hierarchical Anderson model*. We use the additional specification "homogeneous" if (2.6) holds.

The basic spectral theorem for the self-adjoint hierarchical Laplacian Δ is

Theorem 2.1. *The spectral decomposition of the hierarchical Laplacian Δ reads*

$$\Delta = \sum_{r=0}^{\infty} \lambda_r \left(\mathbf{E}_r - \mathbf{E}_{r+1}\right), \tag{2.9}$$

where

$$\lambda_r := \sum_{s=0}^{r} p_s, \quad r \in \mathbb{N}_0, \tag{2.10}$$

are its eigenvalues (of infinite multiplicity) and $\mathbf{E}_r - \mathbf{E}_{r+1}$ is the orthogonal projection onto the eigenspace corresponding to λ_r. In particular, these eigenvalues and their accumulation point $\lambda_\infty := 1$ belong to the essential spectrum of Δ.

We refer to [11, Thm. 1.1] for a proof. The spectral value $\lambda_\infty = 1$ is never an eigenvalue since $p_s > 0$ for every $s \in \mathbb{N}$. Any eigenfunction $\psi_r \in \ell^2(\mathbb{X})$ corresponding to the eigenvalue λ_r is constant on every cluster of rank r. In addition, the sum of these constants over all rank-r-clusters which belong to the same rank-$(r+1)$-cluster is always equal to zero, see Fig. 1 for a sketch.

Using [12, Lemma 1.2] together with ergodicity of H^ω, see Lemmas A.1 and A.3, we conclude the standard

Lemma 2.2. *There exists a non-random compact subset $\Sigma \subset \mathbb{R}$ such that* $\operatorname{spec} H^\omega = \Sigma$ *for $\mathbb{P}$-a.e. $\omega \in \Omega$. The location of the deterministic spectrum Σ obeys*

$$\operatorname{spec}\Delta + \operatorname{supp}\mathbb{P}_0 \subseteq \Sigma \subseteq \big(\operatorname{spec}\Delta + \operatorname{ch}(\operatorname{supp}\mathbb{P}_0)\big) \cap \big([0,1] + \operatorname{supp}\mathbb{P}_0\big), \quad (2.11)$$

where $\operatorname{ch} B$ *denotes the convex hull of a set $B \subset \mathbb{R}$. In particular, we have* $\sup(\inf)\Sigma = \sup(\inf)\{\operatorname{spec}\Delta + \operatorname{supp}\mathbb{P}_0\}$*, and if* $\operatorname{supp}\mathbb{P}_0$ *is even connected, then also* $\Sigma = \operatorname{spec}\Delta + \operatorname{supp}\mathbb{P}_0$*.*

Remark 2.3. The preceding lemma strengthens Lemma 1.2 in [12], in as much as non-randomness of $\operatorname{spec} H^\omega$ is established irrespective of the connectedness of $\operatorname{supp}\mathbb{P}_0$.

For homogeneous hierarchical structures, we will focus on the special case where the decay rate of $(p_s)_{s\in\mathbb{N}}$ is linked to the degree n of the structure.

Definition 2.4. Consider a homogeneous hierarchical structure of degree $n \geqslant 2$. Suppose that there exist constants $C_1, C_2 > 0$ and $\rho > 1$ such that

$$C_1 \rho^{-r} \leqslant p_r \leqslant C_2 \rho^{-r} \quad (2.12)$$

for all $r \in \mathbb{N}$ large enough. Then the *spectral dimension* of this model is defined as

$$d_s \equiv d_s(n, \rho) := 2\,\frac{\ln n}{\ln \rho}. \quad (2.13)$$

In other words, this amounts to $C_1 n^{-2r/d_s} \leqslant p_r \leqslant C_2 n^{-2r/d_s}$ for large r. One motivation for the definition of d_s will be given by Lemma 2.5 below. To this end we introduce the *integrated density of states* $N_0 : \mathbb{R} \to [0,1]$ of Δ, which is defined by

$$E \mapsto N_0(E) := \langle \delta_{x_0}, \chi_{]-\infty,E]}(\Delta)\delta_{x_0}\rangle \quad (2.14)$$

for some $x_0 \in \mathbb{X}$. Here, χ_B stands for the indicator function of a set B, $\langle\cdot,\cdot\rangle$ denotes the canonical scalar product of the Hilbert space $\ell^2(\mathbb{X})$ and δ_x the canonical basis vector associated with $x \in \mathbb{X}$, i.e. $\delta_x(y) = 0$ for every $y \in \mathbb{X}\backslash\{x\}$ and $\delta_x(x) = 1$. We remark that N_0 is a right-continuous distribution function, which is normalised according to $N_0(1) = 1$. Moreover, it does not depend on the choice of x_0. This can be seen from the more explicit expression

$$N_0(E) = \sum_{r\in\mathbb{N}_0:\, \lambda_r \leqslant E} \left(\frac{1}{|Q_r|} - \frac{1}{|Q_{r+1}|}\right) = 1 - \frac{1}{|Q_{r(E)+1}|}, \quad (2.15)$$

which follows from the spectral representation (2.9). The second equality in (2.15) makes only sense for $E \in [0,1[$ with $r(E) := \max\{r' \in \mathbb{N}_0 : \lambda_{r'} \leqslant E\}$.

Lemma 2.5. *Let Δ be the hierarchical Laplacian of a homogeneous hierarchical structure with spectral dimension d_s. Then the integrated density of states N_0 of Δ exhibits the upper-edge asymptotics*

$$\lim_{E \searrow 0} \frac{\ln [1 - N_0(1-E)]}{\ln E} = \frac{d_s}{2}. \tag{2.16}$$

We refer to, e.g., [11, Prop. 1.3] or [16] for a proof.

Next, we turn to the central quantity of this paper. The *integrated density of states* $N : \mathbb{R} \to [0,1]$ of the hierarchical Anderson model is defined by

$$E \mapsto N(E) := \mathbb{E}\big[\langle \delta_{x_0}, \chi_{]-\infty,E]}(H^\omega)\delta_{x_0}\rangle\big], \tag{2.17}$$

where $\mathbb{E}$ denotes the probabilistic expectation associated with $\mathbb{P}$. The integrated density of states is a right-continuous distribution function and independent of the choice of $x_0 \in \mathbb{X}$. The set of growth points of N coincides $\mathbb{P}$-a.s. with the compact deterministic spectrum described in (2.11). At the lower spectral edge $\inf \operatorname{supp} \mathbb{P}_0$ the asymptotics of N is solely determined by the single-site distribution $\mathbb{P}_0$ of the potential, because the second-lowest eigenvalue $\lambda_1 = p_1 > 0$ of the hierarchical Laplacian is separated from $\lambda_0 = 0$ by a finite gap. The interesting case is the upper spectral edge $1 + \sup \operatorname{supp} \mathbb{P}_0$, because the eigenvalues of the hierarchical Laplacian accumulate at $\sup \operatorname{spec} \Delta = 1$. The main result of the present paper, Theorem 2.6, concerns this case.

Theorem 2.6. *Let N be the integrated density of states of a homogeneous hierarchical Anderson model with the properties* (2.3), (2.4) *and* (2.12). *Then, N has a Lifshits tail at the upper spectral edge $1 + v_+$ in the sense that*

$$\lim_{E \searrow 0} \frac{\ln \big| \ln \big[1 - N(1 + v_+ - E)\big]\big|}{\ln E} = -\frac{d_s}{2}. \tag{2.18}$$

We note that $N(1 + v_+) = 1$ and recall that d_s denotes the spectral dimension (2.13) *of the homogeneous hierarchical model.*

The proof of Theorem 2.6 can be found in Sect. 4.

Remark 2.7. The intuition behind Theorem 2.6 is that eigenvalues $1 + v_+ - E$ close to the upper edge of the spectrum, i.e. for $E \ll 1$, must maximise both the kinetic and the potential energy. In order to achieve a large kinetic energy of the order $1 - E =: 1 - \sum_{r=\kappa(E)}^{\infty} p_r$, the associated eigenfunction must be approximately constant over a cluster of high rank $\kappa(E) \sim -d_s(\ln E)/(2 \ln n) \gg 1$, that is, with volume $|Q_{\kappa(E)}| = n^{\kappa(E)} \sim E^{-d_s/2}$. To ensure a large potential energy at the same time then requires most coupling constants of this cluster to take values close to $v_+ = \sup \operatorname{supp} \mathbb{P}_0$. This is a large-deviation event with approximate probability $\exp\{-|Q_{\kappa(E)}|\} \sim \exp\{-E^{-d_s/2}\}$ that sets the scale for $N(1 + v_+ - E)$. We refer to the end of Sect. 1 for further comments on the fact that the Lifshits exponent equals $d_s/2$.

3. Dirichlet–Neumann Bracketing

Technically, many proofs of Lifshits tails rely on Dirichlet–Neumann bracketing. This allows for a two-sided estimate of the integrated density of states in terms of finite-volume operators. We will also follow this route. Thus, it is a main point of this paper to find a pair of suitable finite-volume restrictions of the hierarchical Laplacian for which Dirichlet–Neumann bracketing works.

Definition 3.1. For $x_0 \in \mathbb{X}$ fixed and $\kappa \in \mathbb{N}_0$ we consider the finite cluster $Q_\kappa \equiv Q_\kappa(x_0)$ and introduce the *Neumann*, resp. *Dirichlet finite-volume restrictions*

$$\Delta_{\mathrm{N},Q_\kappa} := \sum_{s=1}^{\kappa} p_s \mathbf{E}_s\Big|_{\ell^2(Q_\kappa)}, \qquad \Delta_{\mathrm{D},Q_\kappa} := \Delta_{\mathrm{N},Q_\kappa} + \sum_{s=\kappa+1}^{\infty} p_s\, \mathbf{1} \tag{3.1}$$

of the hierarchical Laplacian to the finite-dimensional subspace $\ell^2(Q_\kappa)$. For $\mathrm{X} \in \{\mathrm{N}, \mathrm{D}\}$ we then set

$$H^\omega_{\mathrm{X},Q_\kappa} := \Delta_{\mathrm{X},Q_\kappa} + V^\omega. \tag{3.2}$$

To simplify notation we write $\Delta_{\mathrm{X},\kappa} \equiv \Delta_{\mathrm{X},Q_\kappa}$ and $H^\omega_{\mathrm{X},\kappa} \equiv H^\omega_{\mathrm{X},Q_\kappa}$, if there is no danger of confusion.

The desired property is stated in

Lemma 3.2 (Dirichlet–Neumann decoupling). *Consider a fixed finite cluster Q_κ of rank $\kappa \in \mathbb{N}$. Let $r \in \mathbb{N}_0, r < \kappa$, and assume that the cluster Q_κ is the union of m disjoint clusters $Q_r^1, \ldots, Q_r^m$ of lower rank r. Writing $H^{\omega,j}_{\mathrm{X},r} \equiv H^\omega_{\mathrm{X},Q_r^j}$ for $\mathrm{X} \in \{\mathrm{N}, \mathrm{D}\}$, we have in the sense of quadratic forms*

$$H^\omega_{\mathrm{N},\kappa} \geqslant \bigoplus_{j=1}^m H^{\omega,j}_{\mathrm{N},r} \quad \text{and} \quad H^\omega_{\mathrm{D},\kappa} \leqslant \bigoplus_{j=1}^m H^{\omega,j}_{\mathrm{D},r}. \tag{3.3}$$

Proof. The subspace $\ell^2(Q_r^j)$ is left invariant under $\mathbf{E}_s\big|_{\ell^2(Q_\kappa)}$ for every $s \in \{1, \ldots, r\}$ and every $j \in \{1, \ldots, m\}$. Thus, we have

$$H^\omega_{\mathrm{N},\kappa} = \left(\bigoplus_{j=1}^m H^{\omega,j}_{\mathrm{N},r}\right) + \sum_{s=r+1}^{\kappa} p_s \mathbf{E}_s\Big|_{\ell^2(Q_\kappa)} \geqslant \bigoplus_{j=1}^m H^{\omega,j}_{\mathrm{N},r} \tag{3.4}$$

and

$$H^\omega_{\mathrm{D},\kappa} = \left(\bigoplus_{j=1}^m H^{\omega,j}_{\mathrm{D},r}\right) - \sum_{s=r+1}^{\kappa} p_s \big(\mathbf{1} - \mathbf{E}_s\big)\Big|_{\ell^2(Q_\kappa)} \leqslant \bigoplus_{j=1}^m H^{\omega,j}_{\mathrm{D},r}. \tag{3.5}$$

□

Being an operator on a finite dimensional space, $H^\omega_{\mathrm{X},\kappa}$ has discrete (random) eigenvalues

$$e^\omega_{\mathrm{X},\kappa}(1) \leqslant e^\omega_{\mathrm{X},\kappa}(2) \leqslant \cdots \leqslant e^\omega_{\mathrm{X},\kappa}(|Q_\kappa|), \tag{3.6}$$

which are counted according to their multiplicities. We define the corresponding normalised eigenvalue counting function $N^{\omega}_{\mathrm{X},\kappa}\colon \mathbb{R} \to [0,1]$ by

$$E \mapsto N^{\omega}_{\mathrm{X},\kappa}(E) := \frac{1}{|Q_\kappa|} \sum_{j=1}^{|Q_\kappa|} \chi_{]-\infty,E]}\big(e^{\omega}_{\mathrm{X},\kappa}(j)\big)\,, \quad \mathrm{X} \in \{\mathrm{N},\mathrm{D}\}. \tag{3.7}$$

In the macroscopic limit, this quantity is self-averaging and justifies the interpretation of N as an integrated density of states.

Lemma 3.3. *For* $\mathrm{X} \in \{\mathrm{N},\mathrm{D}\}$ *there exists a set* $\Omega_0 \subseteq \Omega$ *of full probability,* $\mathbb{P}[\Omega_0] = 1$*, such that for every* $\omega \in \Omega_0$ *we have*

$$\lim_{\kappa\to\infty} N^{\omega}_{\mathrm{X},\kappa}(E) = N(E) \tag{3.8}$$

at each continuity point E *of* N.

Remark 3.4. The above lemma extends Thm. 3.3 in [13] to the Dirichlet case. Whereas we rely on ergodicity, the (longer) argument in [13] follows a different route which is based on the law of large numbers instead.

Proof of Lemma 3.3. *Case* $\mathrm{X} = \mathrm{N}$. Given $Q \subseteq \mathbb{X}$, we write tr_Q for the trace over $\ell^2(Q)$ and identify a function on Q with the corresponding multiplication operator by this function on $\ell^2(Q)$. The covariance and ergodicity of the hierarchical Anderson model, Lemmas A.3 and A.2, yield for every continuous function $\varphi \in C_c(\mathbb{R})$ with compact support

$$\begin{aligned}\frac{1}{|Q_\kappa|}\,\mathrm{tr}_{\mathbb{X}}\left[\chi_{Q_\kappa(x_0)}\,\varphi(H^{\omega})\right] &= \frac{1}{|Q_\kappa|}\sum_{x\in Q_\kappa(x_0)} \left\langle \delta_{x_0}, \varphi\big(H^{\tau_x(\omega)}\big)\delta_{x_0}\right\rangle \\ &\xrightarrow{\kappa\to\infty} \int_\Omega \mathrm{d}\mathbb{P}(\omega')\,\langle \delta_{x_0}, \varphi(H^{\omega'})\delta_{x_0}\rangle \end{aligned} \tag{3.9}$$

for $\mathbb{P}$-a.e. $\omega \in \Omega$. Next, we introduce the unrestricted but truncated operator $K^{\omega}_{\kappa} := \sum_{s=1}^{\kappa} p_s \mathbf{E}_s + V^{\omega}$ on $\ell^2(\mathbb{X})$. Since K^{ω}_{κ} converges to H^{ω} in operator norm as $\kappa \to \infty$ for $\mathbb{P}$-a.e. $\omega \in \Omega$, we conclude that also

$$\frac{1}{|Q_\kappa|}\,\mathrm{tr}_{\mathbb{X}}\left[\chi_{Q_\kappa(x_0)}\,\varphi(K^{\omega}_{\kappa})\right] \xrightarrow{\kappa\to\infty} \int_\Omega \mathrm{d}\mathbb{P}(\omega')\,\langle \delta_{x_0}, \varphi(H^{\omega'})\delta_{x_0}\rangle \tag{3.10}$$

for $\mathbb{P}$-a.e. $\omega \in \Omega$ and every given $\varphi \in C_c(\mathbb{R})$. This implies $\mathbb{P}$-a.s. vague convergence of the corresponding probability measures on $\mathbb{R}$, see e.g. the proof of Thm. 5.5 in [8]. Hence, there exists a set $\Omega_0 \subseteq \Omega$ of full probability, $\mathbb{P}[\Omega_0] = 1$, such that for every $\omega \in \Omega_0$ the corresponding distribution functions converge

$$N^{\omega}_{\mathrm{N},\kappa}(E) = \frac{1}{|Q_\kappa|}\,\mathrm{tr}_{\mathbb{X}}\left[\chi_{Q_\kappa(x_0)}\,\chi_{]-\infty,E]}(K^{\omega}_{\kappa})\right] \xrightarrow{\kappa\to\infty} N(E) \tag{3.11}$$

for every continuity point $E \in \mathbb{R}$ of N. We note that the left equality above relies on $\ell^2(Q_\kappa(x_0))$ being an invariant subspace of K^{ω}_{κ}.

Case $\mathrm{X} = \mathrm{D}$. We repeat the argument of the Neumann case with

$$K^{\omega}_{\kappa} := \sum_{s=1}^{\kappa} p_s \mathbf{E}_s + \sum_{s=\kappa+1}^{\infty} p_s\,\mathbf{1} + V^{\omega} \tag{3.12}$$

acting on $\ell^2(\mathbb{X})$. ∎

Lemma 3.5 (Dirichlet–Neumann bracketing). *For every cluster Q_r of rank $r \in \mathbb{N}_0$ and for every $E \in \mathbb{R}$ the integrated density of states N obeys the two-sided estimate*

$$\mathbb{E}\big[N^{\omega}_{\mathrm{D},r}(E)\big] \leqslant N(E) \leqslant \mathbb{E}\big[N^{\omega}_{\mathrm{N},r}(E)\big]. \tag{3.13}$$

Proof. We fix $E \in \mathbb{R}, r \in \mathbb{N}_0$ and $x_0 \in \mathbb{X}$. Consider a cluster $Q_\kappa \equiv Q_\kappa(x_0)$ of rank $\kappa \in \mathbb{N}, \kappa > r$, such that Q_κ is the union of m disjoint rank-r-clusters $Q_r^1, \ldots, Q_r^m$ for some $m \in \mathbb{N}$. Using Lemma 3.2 and the min–max principle, we conclude that

$$\begin{aligned}\mathbb{E}\left[\frac{1}{|Q_\kappa|}\mathrm{tr}_{Q_\kappa}\chi_{]-\infty,E]}(H^{\omega}_{\mathrm{N},\kappa})\right] &\leqslant \mathbb{E}\left[\frac{1}{m|Q_r^1|}\,\mathrm{tr}_{Q_\kappa}\left(\bigoplus_{j=1}^{m}\chi_{]-\infty,E]}\big(H^{\omega,j}_{\mathrm{N},r}\big)\right)\right]\\ &= \frac{1}{m|Q_r^1|}\,\sum_{j=1}^{m}\mathbb{E}\left[\mathrm{tr}_{Q_r^j}\,\chi_{]-\infty,E]}\big(H^{\omega,j}_{\mathrm{N},r}\big)\right]\\ &= \frac{1}{|Q_r^1|}\,\mathbb{E}\left[\mathrm{tr}_{Q_r^1}\chi_{]-\infty,E]}\big(H^{\omega,1}_{\mathrm{N},r}\big)\right].\end{aligned} \tag{3.14}$$

Here, the last equality relies on the r-cluster permutation invariance of $\Delta_{\mathrm{N},r}$ and the identical distribution of the random variables. By dominated convergence and Lemma 3.3, the left-hand side of (3.14) converges to $\mathbb{E}[N(E)] = N(E)$ as $\kappa \to \infty$ provided E is a continuity point. In this case we obtain $N(E) \leqslant \mathbb{E}[N^{\omega}_{\mathrm{N},r}(E)]$. If E happens to be a discontinuity point of N—for which we have no convergence statement in Lemma 3.3—simply replace E by a monotone decreasing sequence of continuity points $E_l \searrow E$ and use right-continuity. The lower bound follows by the same line of reasoning. □

4. Proof of Theorem 2.6

The strategy outlined in Remark 2.7 suggests to estimate the maximal eigenvalue of random operators on finite clusters. For a general self-adjoint operator, an upper bound on the maximal eigenvalue is provided by Temple's inequality, which we recall from [20, Thm. XIII.5] (with A replaced by $-A$).

Lemma 4.1. *Let A be a self-adjoint operator in a Hilbert space and let $E_{\max}(A) := \sup\operatorname{spec} A$ be an isolated eigenvalue of A. We write $E_1(A) := \sup\big\{\operatorname{spec}(A)\backslash\{E_{\max}(A)\}\big\}$ and assume the existence of a vector ψ in the domain of A such that $\langle\psi,\psi\rangle = 1$ and*

$$\langle\psi, A\psi\rangle > E_1(A). \tag{4.1}$$

Then we have the estimate

$$E_{\max}(A) \leqslant \langle\psi, A\psi\rangle + \frac{\langle\psi, A^2\psi\rangle - \langle\psi, A\psi\rangle^2}{\langle\psi, A\psi\rangle - E_1(A)}. \tag{4.2}$$

Proof of Theorem 2.6. We assume w.l.o.g. that $v_+ = 0$. This can always be achieved by adding the constant term $-v_+\mathbf{1}$ to H^ω. In order to prove the assertion, we will construct an upper and a lower bound on $1 - N(1 - E)$, which, after taking double logarithms, will asymptotically coincide as $E \searrow 0$.

In what follows we choose some fixed cluster $Q_\kappa \equiv Q_\kappa(x_0)$, $x_0 \in \mathbb{X}$, of finite rank $\kappa \in \mathbb{N}$ and let $0 < E \ll 1$ be arbitrary but fixed.

(a) *Upper bound.* We recall the notation (3.6) and estimate the expectation of the finite-volume eigenvalue counting function $N^\omega_{\mathrm{D},\kappa}(E')$ for every $E' \in \mathbb{R}$ in terms of the maximal eigenvalue $E_{\max}(H^\omega_{\mathrm{D},\kappa}) \equiv e^\omega_{\mathrm{D},\kappa}(|Q_\kappa|)$ of $H^\omega_{\mathrm{D},\kappa}$ by

$$\begin{aligned}\mathbb{E}\left[N^\omega_{\mathrm{D},\kappa}(E')\right] &= \frac{1}{|Q_\kappa|}\,\mathbb{E}\Bigg[\sum_{j=1}^{|Q_\kappa|}\chi_{]-\infty,E']}\big(e^\omega_{\mathrm{D},\kappa}(j)\big)\Bigg]\\ &\geqslant \mathbb{E}\Big[\chi_{]-\infty,E']}\big(E_{\max}(H^\omega_{\mathrm{D},\kappa})\big)\Big] = \mathbb{P}\big[E_{\max}(H^\omega_{\mathrm{D},\kappa}) \leqslant E'\big].\end{aligned} \tag{4.3}$$

Setting $E' = 1 - E$ and applying Lemma 3.5, we conclude that

$$1 - N(1-E) \leqslant 1 - \mathbb{E}\left[N^\omega_{\mathrm{D},\kappa}(1-E)\right] \leqslant \mathbb{P}\big[E_{\max}(H^\omega_{\mathrm{D},\kappa}) > 1-E\big]. \tag{4.4}$$

To proceed further, we need an upper bound on $E_{\max}(H^\omega_{\mathrm{D},\kappa})$. Temple's inequality cannot be applied directly: even the normalised trial function $\psi_0 = |Q_\kappa|^{-1/2} \in \ell^2(Q_\kappa)$, which is the eigenfunction of $\Delta_{\mathrm{D},\kappa}$ corresponding to the maximal eigenvalue $E_{\max}(\Delta_{\mathrm{D},\kappa}) = 1$, will not satisfy the condition (4.1) for κ large enough. This problem is circumvented by introducing the auxiliary operator

$$\widetilde{H}^\omega_{\mathrm{D},\kappa} := \Delta_{\mathrm{D},\kappa} + V^\omega_\kappa \geqslant \Delta_{\mathrm{D},\kappa} + V^\omega = H^\omega_{\mathrm{D},\kappa} \tag{4.5}$$

on $\ell^2(Q_\kappa)$ with the new, higher potential $V^\omega_\kappa(x) := \max\{\omega_x, -p_\kappa/3\}$ for all $x \in \mathbb{X}$. This operator satisfies

$$\langle\psi_0, \widetilde{H}^\omega_{\mathrm{D},\kappa}\psi_0\rangle \geqslant 1 - \frac{p_\kappa}{3} > 1 - p_\kappa = E_1(\Delta_{\mathrm{D},\kappa}) \geqslant E_1(\widetilde{H}^\omega_{\mathrm{D},\kappa}) \tag{4.6}$$

and may thus be employed in Temple's inequality. In order to simplify the right-hand side of (4.2), we use the estimates $\langle\psi_0, \widetilde{H}^\omega_{\mathrm{D},\kappa}\psi_0\rangle - E_1(\widetilde{H}^\omega_{\mathrm{D},\kappa}) \geqslant 2p_\kappa/3$ and $(V^\omega_\kappa)^2 \leqslant (p_\kappa/3)|V^\omega_\kappa|$, and arrive at

$$E_{\max}(H^\omega_{\mathrm{D},\kappa}) \leqslant 1 + \frac{1}{2|Q_\kappa|}\sum_{x\in Q_\kappa} V^\omega_\kappa(x). \tag{4.7}$$

Together with (4.4), this implies

$$1 - N(1-E) \leqslant \mathbb{P}\Bigg[\frac{1}{|Q_\kappa|}\sum_{x\in Q_\kappa} V^\omega_\kappa(x) > -2E\Bigg]. \tag{4.8}$$

By hypothesis of Theorem 2.6, the spectral dimension d_s exists and the estimate (2.12) is valid. So we choose

$$\kappa = k(E) := \max\big\{r \in \mathbb{N} : |Q_r| \leqslant (\alpha E)^{-d_s/2}\big\}, \tag{4.9}$$

where $\alpha > 0$ is a free parameter to be determined below. Since, by definition, $n = \rho^{d_s/2}$ for some $\rho > 1$ and since $|Q_r| = n^r$, the inequality in (4.9) is equivalent to $\rho^{-r} \geqslant \alpha E$. Therefore, (2.12) guarantees that for $E > 0$ small enough, we have the estimate $p_{k(E)} > C_1 \rho^{-k(E)} \geqslant C_1 E\alpha$ with some constant $C_1 > 0$. This estimate and (4.8) imply

$$1 - N(1-E) \leqslant \mathbb{P}\Big[\frac{1}{|Q_{k(E)}|} \sum_{x \in Q_{k(E)}} V^{\omega}_{k(E)}(x) > -\frac{2p_{k(E)}}{\alpha\, C_1}\Big] =: \mathbb{P}\big[\mathcal{A}_{k(E)}\big] \qquad (4.10)$$

for all $E > 0$ small enough.

For every $r \in \mathbb{N}$, and assuming $\alpha > 6/C_1$, it follows that the condition

$$\frac{1}{|Q_r|}\Big|\Big\{x \in Q_r : V^{\omega}_r(x) > -\frac{p_r}{3}\Big\}\Big| \geqslant 1 - \frac{6}{\alpha C_1} =: z > 0 \qquad (4.11)$$

is necessary for the event $\mathcal{A}_r$ to occur. Hence

$$\mathbb{P}[\mathcal{A}_r] \leqslant \mathbb{P}\Big[\frac{1}{|Q_r|}\Big|\Big\{x \in Q_r : V^{\omega}_r(x) > -\frac{p_r}{3}\Big\}\Big| \geqslant z\Big]. \qquad (4.12)$$

To eliminate the dependence of this probability on p_r, we pick $\gamma \in]\inf \operatorname{supp} \mathbb{P}_0\,, 0[$, introduce the i.i.d. random variables $\eta^{\omega}_x := \chi_{]\gamma,0]}(\omega_x)$ for $x \in \mathbb{X}$ and note that

$$\Big|\Big\{x \in Q_r : V^{\omega}_r(x) > -\frac{p_r}{3}\Big\}\Big| = \Big|\Big\{x \in Q_r : \omega_x > -\frac{p_r}{3}\Big\}\Big| \leqslant \sum_{x \in Q_r} \eta^{\omega}_x, \qquad (4.13)$$

where we assumed $\gamma \leqslant -p_r/3$ in the last step. This implies

$$\mathbb{P}[\mathcal{A}_r] \leqslant \mathbb{P}\Big[\frac{1}{|Q_r|}\sum_{x \in Q_r} \eta^{\omega}_x \geqslant z\Big] \qquad (4.14)$$

for all $r \in \mathbb{N}$ large enough, since $(p_r)_r$ is a null sequence. The latter probability can be handled with the standard estimate $\mathbb{P}(X \geqslant \delta) \leqslant e^{-t\delta}\, \mathbb{E}(e^{tX})$ for any $t \geqslant 0$ and any random variable X. Thus there exists $r_0 \in \mathbb{N}$ such that for every $r \in \mathbb{N}$ with $r \geqslant r_0$ we have

$$\mathbb{P}[\mathcal{A}_r] \leqslant \mathrm{e}^{-tz|Q_r|}\; \mathbb{E}\big[\,\mathrm{e}^{t\sum_{x \in Q_r} \eta^{\omega}_x}\,\big] = \mathrm{e}^{-tz|Q_r|}\,\Big(\mathbb{E}_0\big[\,\mathrm{e}^{t\eta^{\omega}_{x_0}}\,\big]\Big)^{|Q_r|} = \mathrm{e}^{-|Q_r| f(t)} \qquad (4.15)$$

for any $t \geqslant 0$, where $f(t) := tz - \ln \mathbb{E}_0[\mathrm{e}^{t\eta^{\omega}_{x_0}}]$ and $\mathbb{E}_0$ is the expectation associated with the single-site distribution $\mathbb{P}_0$. In view of (4.10) this means that there exists $E_u > 0$ such that for every $E \in]0, E_u]$ the estimate

$$1 - N(1-E) \leqslant \mathrm{e}^{-|Q_{k(E)}| f(t)} \tag{4.16}$$

holds for all $t \geqslant 0$.

Now we choose γ close enough to 0 such that $q := \mathbb{E}_0(\eta_x^\omega) = \mathbb{P}_0(\,]\gamma, 0]\,) \in]0,1[$. This is always possible in view of conditions (2.3) and (2.4). By adjusting the free parameter $\alpha > 6/C_1$ large enough, we ensure in addition that $q < z$. Then we have $f(0) = 0$ and

$$f'(0) = z - \frac{\mathbb{E}_0\big[\eta_{x_0}^\omega\, \mathrm{e}^{t\eta_{x_0}^\omega}\big]}{\mathbb{E}_0\big[\,\mathrm{e}^{t\eta_{x_0}^\omega}\,\big]}\Bigg|_{t=0} = z - q > 0. \tag{4.17}$$

Hence, there is a $t_0 > 0$ such that $f(t_0) > 0$. We remark that neither t_0 nor $f(t_0)$ depend on E. Definition (4.9) implies

$$|Q_{k(E)}| = \frac{1}{n}|Q_{k(E)+1}| \geqslant \frac{1}{n}(\alpha E)^{-d_s/2} \tag{4.18}$$

for the homogeneous model. This estimate and (4.16) then yield the desired upper bound

$$1 - N(1-E) \leqslant \exp\big\{ - C_u\, E^{-d_s/2}\big\} \tag{4.19}$$

for all $E \in]0, E_u]$ with the constant $C_u := f(t_0)\, \alpha^{-d_s/2}/n > 0$.

(b) *Lower bound.* This time we use the upper bound of Lemma 3.5 and estimate

$$\begin{aligned} 1 - N(1-E) &\geqslant 1 - \mathbb{E}\big[N_{\mathrm{N},\kappa}^\omega(1-E)\big] \\ &= \frac{1}{|Q_\kappa|}\, \mathbb{E}\Big[\big|\big\{\text{eigenvalues of } H_{\mathrm{N},\kappa}^\omega > 1-E\big\}\big|\Big] \\ &\geqslant \frac{1}{|Q_\kappa|}\, \mathbb{P}\big[E_{\max}(H_{\mathrm{N},\kappa}^\omega) > 1-E\big], \end{aligned} \tag{4.20}$$

where $E_{\max}(H_{\mathrm{N},\kappa}^\omega) \equiv e_{\mathrm{N},\kappa}^\omega(|Q_\kappa|) = \sup_{0\neq\varphi\in\ell^2(Q_\kappa)} \langle\varphi, H_{\mathrm{N},\kappa}^\omega\varphi\rangle / \langle\varphi,\varphi\rangle$ denotes the maximal eigenvalue of $H_{\mathrm{N},\kappa}^\omega$. The choice $\varphi = \psi_0 = |Q_\kappa|^{-1/2}$ for the trial function yields

$$E_{\max}(H_{\mathrm{N},\kappa}^\omega) \geqslant \sum_{s=1}^{\kappa} p_s + \frac{1}{|Q_\kappa|} \sum_{x\in Q_\kappa} \omega_x. \tag{4.21}$$

Therefore we get together with (4.20)

$$1 - N(1-E) \geqslant \frac{1}{|Q_\kappa|}\, \mathbb{P}\Bigg[\frac{1}{|Q_\kappa|} \sum_{x\in Q_\kappa} \omega_x > -E + \sum_{s=\kappa+1}^{\infty} p_s\Bigg] \tag{4.22}$$

So far, $\kappa \in \mathbb{N}$ was fixed arbitrary. Now we choose

$$\kappa = K(E) := \min\Big\{ r \in \mathbb{N} : \sum_{s=r+1}^{\infty} p_s < \frac{E}{2}\Big\} \tag{4.23}$$

and conclude

$$1 - N(1-E) \geqslant \frac{1}{|Q_{K(E)}|}\, \mathbb{P}\Bigg[\frac{1}{|Q_{K(E)}|} \sum_{x\in Q_{K(E)}} \omega_x > -\frac{E}{2}\Bigg]. \tag{4.24}$$

Since the random variables are assumed to be i.i.d., we obtain

$$\mathbb{P}\Bigg[\frac{1}{|Q_{K(E)}|}\sum_{x\in Q_{K(E)}}\omega_x > -\frac{E}{2}\Bigg] \geqslant \mathbb{P}\Bigg[\forall x\in Q_{K(E)} : \omega_x > -\frac{E}{2}\Bigg]$$
$$= \Big(\mathbb{P}_0[\omega_{x_0} > -E/2]\Big)^{|Q_{K(E)}|}$$
$$= \mathrm{e}^{-g(E)|Q_{K(E)}|} \tag{4.25}$$

with $g(E) := -\ln\mathbb{P}_0[\omega_{x_0} > -E/2]$ for every $E > 0$. The function g is well defined because of Assumption (2.4) on the tails of the single-site distribution. In fact (2.4) implies the estimate

$$0 \leqslant g(E) \leqslant m|\ln(E/2)| - \ln C \tag{4.26}$$

for all sufficiently small $E > 0$ with some E-independent constants $C, m > 0$.

We have by definition (4.23) that $\sum_{s=K(E)}^{\infty} p_s \geqslant E/2$. In conjunction with the upper bound in (2.12), this yields the existence of $E_l > 0$ such that $\rho^{K(E)} \leqslant C_l^{2/d_s}E^{-1}$ for all $E \in]0, E_l]$ with the constant $C_l := [2C_2\rho/(\rho-1)]^{d_s/2}$. Since $\rho^{d_s/2} = n$, we conclude

$$|Q_{K(E)}| \leqslant C_l E^{-d_s/2} \tag{4.27}$$

for all $E \in]0, E_l]$. Collecting (4.24), (4.25) and (4.27), we finally arrive at the desired lower bound

$$1 - N(1-E) \geqslant C_l^{-1}E^{d_s/2}\exp\{-C_l E^{-d_s/2}g(E)\} \tag{4.28}$$

for all $E \in]0, E_l]$.

(c) *Limit* $E \searrow 0$. The lower bound (4.28) can be simplified by observing (4.26), enlarging the constant C_l and diminishing E_l: there exist constants $\widetilde{C}_l, \widetilde{E}_l > 0$ such that

$$1 - N(1-E) \geqslant \widetilde{C}_l^{-1}\exp\{-\widetilde{C}_l E^{-d_s/2}|\ln E|\} \tag{4.29}$$

for all $E \in]0, \widetilde{E}_l]$. This clearly implies

$$\liminf_{E\searrow 0}\frac{\ln\big|\ln\big[1-N(1-E)\big]\big|}{\ln E} \geqslant -\frac{d_s}{2}. \tag{4.30}$$

On the other hand, we deduce from the upper bound (4.19) that

$$\limsup_{E\searrow 0}\frac{\ln\big|\ln\big[1-N(1-E)\big]\big|}{\ln E} \leqslant -\frac{d_s}{2}, \tag{4.31}$$

and Theorem 2.6 is proven. □

Appendix A. Ergodicity

Here we briefly compile the ergodic structure of the hierarchical Anderson model, which we have not found in the literature. Due to the use of ergodicity, our Lemmas 2.2 and 3.3 give slightly stronger results, respectively, have shorter proofs, than corresponding statements in [11–13], who argued without it.

In what follows, we enumerate $\mathbb{X}$ as in (2.5) and arrange the elements in increasing order from left to right, see Fig. 1. Also, we think of $\mathbb{X}$ as being isomorphic to the space

$$\mathbb{X} \cong \left\{ (\xi_r)_{r\in\mathbb{N}} : \xi_r \in \{0,1,\ldots,n_r-1\} \text{ for all } r\in\mathbb{N} \text{ and } \sum_{r=1}^{\infty} \xi_r < \infty \right\}, \quad \text{(A.1)}$$

which consists of sequences with only finitely many non-zero elements. The identification $x = (\xi_r)_{r\in\mathbb{N}}$, which underlies (A.1), works as follows: ξ_1 determines the position of x in $Q_1(x)$, where $\xi_1 = 0$ corresponds to the left-most position in Fig. 1, $\xi_1 = 1$ to the second position from the left, and so on. Similarly, ξ_r encodes the position of $Q_{r-1}(x)$ in $Q_r(x)$ for every $r \geqslant 2$, where, again, cluster positions are counted from the left, starting with 0. For example, in Fig. 1 we have $x_5 = (2, 1, 0, 0, \ldots)$. Every $x \in \mathbb{X}$ eventually belongs to $Q_{r_0}(x_0)$, the left-most cluster of rank r_0, for some sufficiently large r_0. Therefore, any sequence $(\xi_r)_{r\in\mathbb{N}}$ has only finitely many non-zero elements. Moreover, we have the representation $x_k = (\xi_r)_{r\in\mathbb{N}}$, if and only if $k = \sum_{r=1}^{\infty} \xi_r |Q_{r-1}|$.

The countable space $\mathbb{X}$ can be equipped with an Abelian group structure, which we write in an additive way

$$x + y := \big((\xi_r + \eta_r) \mod n_r\big)_{r\in\mathbb{N}} \quad \text{(A.2)}$$

for all $x = (\xi_r)_{r\in\mathbb{N}}$ and $y = (\eta_r)_{r\in\mathbb{N}}$ in $\mathbb{X}$. The identity of the group is given by x_0 and the inverse of x by $-x := (n_r - \xi_r)_{r\in\mathbb{N}}$.

The discrete Abelian group $\mathbb{X}$ acts on Ω according to $\mathbb{X}\times\Omega \to \Omega$, $(x, \omega) \mapsto \tau_x(\omega)$, where

$$\tau_x(\omega) := (\omega_{x+y})_{y\in\mathbb{X}} \quad \text{(A.3)}$$

for every $x \in \mathbb{X}$ and every $\omega := (\omega_y)_{y\in\mathbb{X}} \in \Omega$. Since $\mathbb{P} = \bigotimes_{x\in\mathbb{X}} \mathbb{P}_0$, every τ_x is measure-preserving, and we have

Lemma A.1. *The group of transformations $\{\tau_x\}_{x\in\mathbb{X}}$ is ergodic with respect to $\mathbb{P}$.*

Proof. Let $A \in \bigotimes_{x\in\mathbb{X}} \mathcal{B}_{\mathbb{R}}$ be invariant under τ_x for every $x \in \mathbb{X}$. We show $\mathbb{P}[A]$ is either 0 or 1. Let $\varepsilon > 0$ be given. Since every product-measurable set can be approximated arbitrarily well by cylinder sets, there exists $\kappa \in \mathbb{N}$ and $Z = \times_{x\in\mathbb{X}} Z_x$ with $Z_x \in \mathcal{B}_{\mathbb{R}}$ for every $x \in \mathbb{X}$ and $Z_x = \mathbb{R}$ for every $x \in \{x_k \in \mathbb{X} : k \geqslant \kappa\}$, and such that

$$\mathbb{P}[A \vartriangle Z] \leqslant \varepsilon. \quad \text{(A.4)}$$

Here $\vartriangle$ denotes the symmetric difference. Pick $r_0 \in \mathbb{N}$ such that $x_\kappa \in Q_{r_0}(x_0)$ and define $w := (\delta_{r,r_0+1})_{r\in\mathbb{N}} \in \mathbb{X}$. Hence, $\big(w + Q_{r_0}(x_0)\big) \cap Q_{r_0}(x_0) = \varnothing$, which in turn implies the crucial identity

$$\mathbb{P}\big[\tau_w(Z) \cap Z\big] = \mathbb{P}\big[\tau_w(Z)\big]\ \mathbb{P}[Z] = \big(\mathbb{P}[Z]\big)^2 \quad \text{(A.5)}$$

because $\tau_w(Z) = \bigtimes_{x\in\mathbb{X}} Z_{w+x}$. Using invariance of A under τ_w and (A.5), we conclude

$$\begin{aligned}
0 \leqslant \mathbb{P}[A] - \big(\mathbb{P}[A]\big)^2 &= \mathbb{P}\big[A \cap \tau_w(A)\big] - \mathbb{P}\big[Z \cap \tau_w(Z)\big] + \big(\mathbb{P}[Z]\big)^2 - \big(\mathbb{P}[A]\big)^2 \\
&\leqslant \mathbb{P}\big[\big(A \cap \tau_w(A)\big) \vartriangle \big(Z \cap \tau_w(Z)\big)\big] + 2\,\mathbb{P}[A \vartriangle Z] \\
&\leqslant \mathbb{P}\big[(A \vartriangle Z) \cup \big(\tau_w(A) \vartriangle \tau_w(Z)\big)\big] + 2\,\mathbb{P}[A \vartriangle Z] \\
&\leqslant 4\varepsilon. \qquad\qquad (A.6)
\end{aligned}$$

In the second line of (A.6), we estimated $\mathbb{P}[C] - \mathbb{P}[D] \leqslant \mathbb{P}[C \vartriangle D]$ for events C and D, which follows from $C = (C\backslash D) \cup (C \cap D) \subseteq (C \vartriangle D) \cup D$. The inequality in the third line of (A.6) is based upon the inclusion

$$\begin{aligned}
(C_1 \cap D_1)\backslash(C_2 \cap D_2) &= (C_1 \cap C_2^c \cap D_1) \cup (D_1 \cap D_2^c \cap C_1) \\
&\subseteq (C_1\backslash C_2) \cup (D_1\backslash D_2)
\end{aligned}$$

and its mirror $1 \longleftrightarrow 2$. In order to get to the last line of (A.6) we used $\tau_w(C) \vartriangle \tau_w(D) = \tau_w(C \vartriangle D)$, the fact that τ_w is measure preserving and (A.4). Since $\varepsilon > 0$ is arbitrary in (A.6), this completes the proof. □

The (Følner) sequence of growing clusters $\big(Q_r(x_0)\big)_{r\in\mathbb{N}}$ exhausts the (amenable) ergodic group $\mathbb{X}$ and fulfils Shulman's temperedness condition [14, Def. 1.1]. Therefore we can apply the general pointwise ergodic theorem of Lindenstrauss [14, Thm. 1.2] and conclude

Lemma A.2. (Birkhoff ergodic theorem) *For every $\mathbb{P}$-integrable random variable $h : \Omega \to \mathbb{C}$ we have*

$$\lim_{r\to\infty} \frac{1}{|Q_r|} \sum_{x\in Q_r(x_0)} h\big(\tau_x(\omega)\big) = \mathbb{E}[h] \qquad (A.7)$$

for $\mathbb{P}$-almost every $\omega \in \Omega$.

The reason for introducing the particular group structure (A.2) is its compatibility with the hierarchical structure on $\mathbb{X}$. This underlies

Lemma A.3. *The hierarchical Anderson Hamiltonian* (2.1) *transforms covariantly under the group action on Ω,*

$$H^{\tau_x(\omega)} = U_x^* H^\omega U_x \qquad (A.8)$$

for every $x \in \mathbb{X}$ and every $\omega \in \Omega$. Here we have introduced the unitary representation of the group $\mathbb{X}$ on $\ell^2(\mathbb{X})$, given by $(U_x\psi)(y) := \psi(y - x)$ for all $x, y \in \mathbb{X}$ and all $\psi \in \ell^2(\mathbb{X})$.

Proof. Clearly we have

$$\begin{aligned}
(U_x^* V^\omega U_x \psi)(y) &= (V^\omega U_x\psi)(x + y) = \omega_{x+y}(U_x\psi)(x + y) \\
&= \big(\tau_x(\omega)\big)_y \psi(y) = (V^{\tau_x(\omega)}\psi)(y), \qquad (A.9)
\end{aligned}$$

so that it remains to verify the invariance of the hierarchical Laplacian

$$\begin{aligned}(U_x^* \Delta U_x \psi)(y) &= \sum_{s=1}^{\infty} p_s (\mathbf{E}_s U_x \psi)(x+y) = \sum_{s=1}^{\infty} p_s \frac{1}{|Q_s|} \sum_{v \in Q_s(x+y)} (U_x \psi)(v) \\ &= \sum_{s=1}^{\infty} p_s \frac{1}{|Q_s|} \sum_{v \in x+Q_s(y)} \psi(v-x) \\ &= \sum_{s=1}^{\infty} p_s \frac{1}{|Q_s|} \sum_{w \in Q_s(y)} \psi(w) = (\Delta\psi)(y). \end{aligned} \tag{A.10}$$

The equality in the second line of (A.10) rests on the identity

$$\begin{aligned} Q_s(x+y) &= \Big\{ (\zeta_r)_{r\in\mathbb{N}} : \ \zeta_r \in \{0, \ldots, n_r - 1\} \quad \text{for } r \leqslant s, \\ &\qquad\qquad \zeta_r = (\xi_r + \eta_r) \bmod n_r \quad \text{for } r > s \Big\} \\ &= x + Q_s(y) \end{aligned} \tag{A.11}$$

for all $s \in \mathbb{N}, x = (\xi_r)_{r\in\mathbb{N}}$ and $y = (\eta_r)_{r\in\mathbb{N}}$ in $\mathbb{X}$. □

References

[1] Antunović, T., Veselić, I.: Equality of Lifshitz and van Hove exponents on amenable Cayley graphs. J. Math. Pures Appl. **92**, 342–362 (2009)

[2] Bleher, P.M., Sinai, Ja.G.: Investigation of the critical point in models of the type of Dyson's hierarchical models. Commun. Math. Phys. **33**, 23–42 (1973)

[3] Bleher, P.M., Sinai, Ya.G.: Critical indices for Dyson's asymptotically-hierarchical models. Commun. Math. Phys. **45**, 247–278 (1975)

[4] Bovier, A.: The density of states in the Anderson model at weak disorder: a renormalization group analysis of the hierarchical model. J. Stat. Phys. **59**, 745–779 (1990)

[5] Dyson, F.J.: Existence of a phase-transition in a one-dimensional Ising ferromagnet. Commun. Math. Phys. **12**, 91–107 (1969)

[6] Dyson, F.J.: An Ising ferromagnet with discontinuous long-range order. Commun. Math. Phys. **21**, 269–283 (1971)

[7] Fyodorov, Y.V., Ossipov, A., Rodriguez, A.: The Anderson localization transition and eigenfunction multifractality in an ensemble of ultrametric random matrices. J. Stat. Mech. **L12001**, 1–5 (2009)

[8] Kirsch, W.: An invitation to random Schrödinger operators—with an appendix by F. Klopp. In: Panor. Synthèses. Random Schrödinger operators, vol. 25, pp. 1–119. Society of Mathematics France, Paris (2008)

[9] Kirsch, W., Martinelli, F.: Large deviations and Lifshitz singularity of the integrated density of states of random Hamiltonians. Commun. Math. Phys. **89**, 27–40 (1983)

[10] Kirsch, W., Müller, P.: Spectral properties of the Laplacian on bond-percolation graphs. Math. Z. **252**, 899–916 (2006)

[11] Kritchevski, E.: Spectral localization in the hierarchical Anderson model. Proc. Am. Math. Soc. **135**, 1431–1440 (2007)

[12] Kritchevski, E.: Hierarchical Anderson Model. In: Probability and mathematical physics. CRM Proceedings and Lecture Notes, vol. 42, pp. 309–322. American Mathematical Society, Providence (2007)

[13] Kritchevski, E.: Poisson statistics of eigenvalues in the hierarchical Anderson model. Ann. Henri Poincaré **9**, 685–709 (2008)

[14] Lindenstrauss, E.: Pointwise theorems for amenable groups. Invent. Math. **146**, 259–295 (2001)

[15] Molchanov, S.A.: Lectures on random media. In: Lectures on probability theory. Lecture Notes in Mathematics, vol. 1581, pp. 242–411. Springer, Berlin (1994)

[16] Molchanov, S. : Hierarchical random matrices and operators. Application to Anderson model. In: Gupta, A.K., Girko, V.L. (eds.) Multidimensional Statistical Analysis and Theory of Random Matrices, pp. 179–194. VSP, Utrecht (1996)

[17] Monthus, C., Garel, T.: A critical Dyson hierarchical model for the Anderson localization transition. J. Stat. Mech. **P05005**, 1–27 (2011)

[18] Müller, P., Stollmann, P.: Spectral asymptotics of the Laplacian on supercritical bond-percolation graphs. J. Funct. Anal. **252**, 233–246 (2007)

[19] Müller, P., Stollmann, P.: Percolation Hamiltonians. In: Lenz, D., Sobieczky, F., Woess, W. (eds.) Random walks, boundaries and spectra, Proceedings, Graz—St. Kathrein 2009, Progress in Probability, vol. 64, pp. 235–258. Springer, Basel (2011)

[20] Reed, M., Simon, B.: Methods of Modern Mathematical Physics IV: Analysis of Operators. Academic Press, San Diego (1978)

[21] Simon, B.: Lifschitz tails for the Anderson model. J. Stat. Phys. **38**, 65–76 (1985)

Simon Kuttruf and Peter Müller
Mathematisches Institut
Universität München
Theresienstraße 39
80333 Munich, Germany
e-mail: mueller@lmu.de

Communicated by Claude Alain Pillet.

Received: January 24, 2011.

Accepted: June 28, 2011.