HIERARCHICAL ANDERSON MODEL.

EVGENIJ KRITCHEVSKI

Dedicated to Stanislav Molchanov on the occasion of his 65'th birthday.

1. INTRODUCTION

In this article, we will review the spectral localization problem in the hierarchical Anderson model. We will present the original result of Molchanov (see Theorem 1.3 below) on localization at arbitrary disorder in any spectral dimension, when the random perturbations are i.i.d. with a Cauchy distribution, as well as generalization to the case of mixed Cauchy distributions. We will also prove the new result (see Theorem 1.6 below), that in spectral dimension d ≤ 4, localization holds for very general distributions of the perturbation. This is an improvement of the main theorem of [K1]. Our notation will be consistent with [K1] and this article should be a complementary reading to [K1] and [M2]. For additional information about the hierarchical structures and the hierarchical Anderson model we refer the reader to [D, BS, Bo, K1, K2, M1, M2].

We first recall the definition of the model and its basic properties. Let X be an infinite countable set. Throughout the paper, δ_x will denote the Kronecker delta function at $x \in X$. A partition $\mathcal{P}$ of X is a collection of its disjoint subsets whose union is equal to X. Let $\mathbf{n} = (n_r)_{r\geq 0}$ be a sequence of positive integers and $\mathbf{P} = (\mathcal{P}_r)_{r\geq 0}$ a sequence of partitions of X. The elements of $\mathcal{P}_r$ are called "clusters" of rank r. We say that $(X, \mathbf{P}, \mathbf{n})$ is a hierarchical structure if the following hold:

(1) $n_0 = 1$ and every $Q \in \mathcal{P}_0$ has exactly one element.
(2) For $r \geq 1$, every $Q \in \mathcal{P}_r$ is a disjoint union of n_r clusters in $\mathcal{P}_{r-1}$.
(3) Given $x, y \in X$, there is a cluster Q of some rank containing both x and y.

Let us state some immediate consequences of this definition. Every cluster of rank $r \geq 0$ has size $N_r := \prod_{s=0}^{r} n_s$. Given $x \in X$ and $r \geq 0$, there is a unique cluster of rank r containing x. We denote this cluster by $Q_r(x)$. The map

$$d(x, y) := \min\{r : y \in Q_r(x)\},$$

is a metric on X and $Q_r(x) = \{y : d(x, y) \leq r\}$. Note that $Q_r(x) = Q_r(y)$ whenever $d(x, y) \leq r$. Given an integer $n \geq 2$, a hierarchical structure is called *homogeneous of degree* n if $n_r = n$ for all $r \geq 1$. Figure 1 illustrates the first three levels of hierarchy in a hierarchical structure with $n_1 = 3, n_2 = 2, n_3 = 3$.

The free Laplacian on the hierarchical structure $(X, \mathbf{P}, \mathbf{n})$ is defined as follows. For each $r \geq 0$, let $E_r : l^2(X) \to l^2(X)$ be the averaging operator

$$(E_r\psi)(x) := \frac{1}{N_r} \sum_{d(x,y)\leq r} \psi(y).$$

Date: Friday, April 21, 2006.
2000 *Mathematics Subject Classification.* Primary 47B80, 47A55, 93A13.
This work was supported in part by an FQRNT grant.

$$\overbrace{\overbrace{\overbrace{\underbrace{x_1}_{Q_0(x_1)}, x_2, x_3}^{Q_1(x_3)}, \overbrace{x_4, x_5, x_6}}^{Q_2(x_3)=Q_2(x_5)}, \overbrace{\overbrace{x_7, x_8, x_9}, \overbrace{x_{10}, x_{11}, x_{12}}^{Q_1(x_8)}}^{Q_2(x_9)}}^{Q_3(x_1)}, \overbrace{x_{13}, x_{14}, x_{15}}, \dots$$

here $d(x_1, x_1) = 0, d(x_1, x_3) = 1, d(x_1, x_6) = 2, d(x_1, x_8) = 3$, etc

FIGURE 1. A hierarchical structure

Let $\mathbf{p} = (p_r)_{r\geq 1}$ be a sequence of positive numbers such that $\sum_{r=1}^{\infty} p_r = 1$. In the sequel we set $p_0 := 0$ and

$$\lambda_r := \sum_{s=0}^{r} p_s, \qquad r = 0, 1, \cdots, \infty.$$

The hierarchical Laplacian Δ on $l^2(X)$ is defined by

$$\Delta := \sum_{r=0}^{\infty} p_r E_r.$$

Clearly, Δ is a bounded self-adjoint operator and $0 \leq \Delta \leq 1$.

A hierarchical model is a hierarchical structure $(X, \mathbf{P}, \mathbf{n})$ together with the hierarchical Laplacian Δ. The basic spectral theorem for Δ is the following.

Theorem 1.1. (1) *The spectrum of Δ is given by*

$$\sigma(\Delta) = \{\lambda_r \,:\, r = 0, \cdots, \infty\}.$$

Each λ_r, $r < \infty$, is an eigenvalue of Δ of infinite multiplicity. The point $\lambda_\infty = 1$ is not an eigenvalue.
(2) *$E_r - E_{r+1}$ is the orthogonal projection onto the eigenspace of λ_r and*

$$\Delta = \sum_{r=0}^{\infty} \lambda_r (E_r - E_{r+1}).$$

(3) *For every $x \in X$, the spectral measure for δ_x and Δ is given by*

$$\mu = \sum_{r=0}^{\infty} \left(\frac{1}{N_r} - \frac{1}{N_{r+1}} \right) \delta(\lambda_r),$$

where $\delta(\lambda_r)$ stands for the Dirac unit mass at λ_r. Note that μ does not depend on x.

A proof can be found in [K1] or [M2].

We now define the hierarchical Anderson model associated to $(X, \mathbf{P}, \mathbf{n})$ and the hierarchical Laplacian Δ. Consider the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ where $\Omega := \mathbb{R}^X$, $\mathcal{F}$ is the usual Borel σ-algebra in Ω, and $\mathbb{P}$ is a given probability measure on $(\Omega, \mathcal{F})$. For $\omega \in \Omega$, we set

$$V_\omega := \sum_{x\in X} \omega(x) \langle \delta_x | \cdot \rangle \delta_x.$$

V_ω is a self-adjoint (possibly unbounded) multiplication operator on $l^2(X)$. Let

$$H_\omega := \Delta + V_\omega, \qquad \omega \in \Omega.$$

The family of self-adjoint operators $\{H_\omega\}_{\omega\in\Omega}$ indexed by the events of the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is called the *hierarchical Anderson model.* The main problem is to understand the spectral properties of H_ω. The first natural question is to describe the spectrum $\sigma(H_\omega)$ of H_ω as a set. Note that, unlike the spectrum of the standard discrete Laplacian on $\mathbb{Z}^{\mathrm{d}}$, the spectrum $\sigma(\Delta)$ of the hierarchical Laplacian is a disconnected set.

Lemma 1.2. *Suppose that the random variables $\{\omega(x) : x \in X\}$ are i.i.d. with distribution ν, i.e. $\mathbb{P} = \otimes_{x\in X}\nu$. Let S be the support of ν and $ch(S)$ the convex hull of S. Then, for $\mathbb{P}$-a.e. $\omega \in \Omega$,*

$$\sigma(\Delta) + S \subset \sigma(H_\omega) \subset (\sigma(\Delta) + ch(S)) \cap ([0, 1] + S).$$

In particular, if S is connected, then $\sigma(H_\omega) = \sigma(\Delta) + S$, for $\mathbb{P}$-a.e. $\omega \in \Omega$.

The proof of Lemma 1.2 (given in the appendix) is a straight forward adaptation of the proof of the corresponding well known result for the standard Anderson model on $\mathbb{Z}^{\mathrm{d}}$, where the spectrum is almost surely equal to $[0, 2\mathrm{d}]+\mathrm{supp}(\nu)$ (cf for example Proposition V.3.5. in [CL]).

We denote by $\sigma_{\mathrm{ac}}(H_\omega)$ the absolutely continuous part of the spectrum of H_ω and by $\sigma_{\mathrm{cont}}(H_\omega)$ the continuous part. In [M2], Molchanov proved the following localization result:

Theorem 1.3. *Suppose that*

$$\sum_{r=1}^{\infty} p_r r^{1+\varepsilon} < \infty, \tag{1.1}$$

for some $\varepsilon > 0$. If the random variables $\{\omega(x) : x \in X\}$ are i.i.d. with a Cauchy distribution,

$$k_{a,h}(e)de := \frac{1}{\pi}\frac{h}{(e-a)^2 + h^2}de,$$

for some parameters $a \in \mathbb{R}$, $h > 0$, then $\sigma_{\mathrm{cont}}(H_\omega) = \emptyset$ for $\mathbb{P}$-a.e. ω.

A proof will be given in section 3. A generalization of Molchanov's theorem will be stated and proven in section 4.

Molchanov's theorem is remarkable for several reasons. First of all, it is a result valid for any disorder. Given a coupling constant $c > 0$, the random operator $\Delta + cV_\omega$ can be rewritten as $\Delta + \sum_{x\in X} c\omega(x)\langle\delta_x|\cdot\rangle\delta_x$. If $\omega(x)$ has a Cauchy distribution $k_{a,h}$, then $c\omega(x)$ has a Cauchy distribution $k_{ca,ch}$. Hence, under the hypotheses of the theorem, $\Delta + cV_\omega$ has pure point spectrum for $\mathbb{P}$-a.e. ω. The next observation is that the condition (1.1) is very weak, it particular it does not at all depend on n_r. To appreciate the generality of (1.1), one has to look at the density of states and at the spectral dimension of Δ. Let $x_0 \in X$ be given and consider the increasing sequence of clusters $Q_r(x_0)$, $r \geq 0$. Let P_r be the orthogonal projection onto the N_r-dimensional subspace

$$l^2(Q_r(x_0)) := \left\{\psi \in l^2(X) : \psi(x) = 0 \text{ for } x \notin Q_r(x_0)\right\}.$$

Let $e_1^{(r)} \leq e_2^{(r)} \leq \cdots \leq e_{N_r}^{(r)}$ be the eigenvalues of the restricted Laplacian $P_r\Delta P_r$ acting on $l^2(Q_r(x_0))$ and

$$\nu_r := \frac{1}{N_r}\sum_{s=1}^{r}\delta(e_s^{(r)}),$$

the corresponding counting measure.

Proposition 1.4. *The weak-* limit* $\lim_{r\to\infty} \nu_r$ *exists and is equal to* μ.

A detailed proof can be found in [K1]. In the original paper [M2], an analogous result is proven, but with a different definition of ν_r. The content of the proposition is that the spectral measure μ can be naturally interpreted as the integrated density of states of the operator Δ.

If

$$\lim_{t\downarrow 0} \frac{\log \mu([1-t,1])}{\log t} = \mathrm{d}/2,$$

then the number d is called the *spectral dimension* of Δ. This definition is motivated by analogy with the edge asymptotics of the density of states of the standard discrete Laplacian on $\mathbb{Z}^{\mathrm{d}}$, for which the spectral and spatial dimensions coincide.

The relation $\sum_{y\in X}\langle \delta_x|\Delta\delta_y\rangle = 1$ yields that Δ generates a random walk on X. We recall that the random walk on $\mathbb{Z}^{\mathrm{d}}$ generated by the standard discrete Laplacian is recurrent if $\mathrm{d} = 1, 2$ and transient if $\mathrm{d} > 2$. The corresponding result for the hierarchical Laplacian is:

Proposition 1.5. *Consider a homogeneous hierarchical structure of degree* $n \geq 2$. *Suppose that there exist constants* $C_1 > 0, C_2 > 0$ *and* $\rho > 1$ *such that*

$$C_1\rho^{-r} \leq p_r \leq C_2\rho^{-r},$$

for r *big enough. Then:*
(1) *The spectral dimension of this model is*

$$\mathrm{d}(n,\rho) = 2\frac{\log n}{\log \rho}.$$

Hence $0 < \mathrm{d}(n,\rho) \leq 2$ *iff* $n \leq \rho$.
(2) *The random walk generated by* Δ *is recurrent if* $0 < \mathrm{d}(n,\rho) \leq 2$ *and transient if* $\mathrm{d}(n,\rho) > 2$.

See [K1], [M2] for a proof. We can now see that Theorem 1.3 holds in arbitrary spectral dimension. Indeed, let $n \geq 2$ be given and let d be a positive real number. Then there is a unique $\rho > 1$ with $\mathrm{d} = \mathrm{d}(\mathrm{n}, \rho)$. Set $p_r = C\rho^{-r}$, where C is a normalization constant making $\sum_{r=1}^{\infty} p_r = 1$. Then p_r satisfies (1.1). Therefore, we can construct homogeneous hierarchical Anderson models of arbitrary spectral dimension for which Theorem 1.3 holds.

Since Cauchy random variables play a very special role in the theory of random Schrödinger operators, it is natural to ask whether one can extend Theorem 1.3 to distributions other than Cauchy. A partial answer to this question is that one can prove localization for very general distributions of $\omega(x)$, at any disorder, provided that one imposes stronger decay conditions on p_r than (1.1). We will prove the following new result:

Theorem 1.6. *Assume that there exists a sequence* $u_r > 0$ *with* $\sum_{r=1}^{\infty} u_r^{-1} < \infty$ *and*

$$\sum_{r=1}^{\infty} p_r u_r \sqrt{N_r} < \infty. \tag{1.2}$$

Then:
(1) *For all* $\omega \in \Omega$, $\sigma_{\mathrm{ac}}(H_\omega) = \emptyset$.
(2) *If* $\mathbb{P}$ *is conditionally a.c. then* $\sigma_{\mathrm{cont}}(H_\omega) = \emptyset$ *for* $\mathbb{P}$*-a.e.* ω.

A proof will be given in section 3.

Remark 1. The condition (1.2) is more demanding than (1.1). The required decay of p_r imposes an upper bound on the spectral dimension of Δ. Theorem 1.6 and Proposition 1.5 allow us to construct hierarchical models with spectral dimension $\mathrm{d} \leq 4$ that exhibit localization at arbitrary disorder. If $(X, \mathbf{P}, \mathbf{n})$ is a homogeneous hierarchical structure of degree $n \geq 2$ and $p_r = C\rho^{-r}$ with $\rho > \sqrt{n}$, then the hypothesis (1.2) is fulfilled for $u_r = r^{1+\varepsilon}$. Given $0 < \mathrm{d} < 4$ one can adjust $\rho > \sqrt{n}$ to make $d(n, \rho) = \mathrm{d}$. If $p_r = Cr^{-2-\varepsilon}n^{-r/2}$, then the model has spectral dimension $\mathrm{d} = 4$ and (1.2) is verified for $u_r = r^{1+\varepsilon/2}$. One can also construct trivial models with $\mathrm{d} = 0$ by taking p_r to decrease faster than ρ^{-r} for any ρ. We emphasize that homogeneity of the hierarchical structure is not required for Theorem 1.6.

Remark 2. Theorem 1.6 is an improvement of the main localization theorem of [K1]. In [K1], instead of (1.2), it is assumed that $\sum_{r=1}^{\infty} p_r N_{r-1} u_{r-1} u_r < \infty$, which imposes the spectral dimension to be ≤ 2.

2. Localization criterion

In this section we will formulate and prove a sufficient condition for H_ω to have $\sigma_{\mathrm{ac}}(H_\omega) = \emptyset$ for all $\omega \in \Omega$, and a sufficient condition for H_ω to have $\sigma_{\mathrm{cont}}(H_\omega) = \emptyset$ for $\mathbb{P}$-a.e. ω. Both theorems 1.3 and 1.6 will follow from this localization criterion and will be proven in the next section.

Consider the truncated hierarchical Laplacian

$$\Delta_r := \sum_{s=0}^{r} p_s E_s, \qquad r \geq 0,$$

and set

$$H_{\omega,r} := V_\omega + \Delta_r, \qquad r \geq 0.$$

Fix $\omega \in \Omega$. Note that for any $Q_r \in \mathcal{P}_r$, the subspace $l^2(Q_r)$ is invariant for $H_{\omega,r}$. This is the main reason for working with the truncated hierarchical Laplacian. Let $\sigma(\omega, Q_r)$ be the set of the eigenvalues of the restricted operator $H_{\omega,r} \restriction l^2(Q_r)$ and $\sigma_\omega := \bigcup \sigma(\omega, Q_r)$ where the union is over all clusters of all ranks. Clearly, σ_ω is a countable subset of $\mathbb{R}$, and hence of zero Lebesgue measure. For $z \in \mathbb{C} \backslash \sigma_\omega$, $r \geq 0$, and $x, y \in X$, we set

$$G_{\omega,r}(x, y; z) := \langle \delta_x | (H_{\omega,r} - z)^{-1} \delta_y \rangle.$$

For $z \in \mathbb{C} \backslash \sigma_\omega$, $r \geq 0$ and $t \in X$, let $g_{\omega,r}(t; z)$ be the average of $G_{\omega,r}(\cdot, t; z)$ over the cluster $Q_r(t)$, i.e.

$$g_{\omega,r}(t; z) := \frac{1}{N_r} \sum_{d(t',t) \leq r} G_{\omega,r}(t', t; z).$$

Since the joint spectral measure for δ_t, $\delta_{t'}$ and $H_{\omega,r}$ is real, $G_{\omega,r}(t', t; z) = G_{\omega,r}(t, t'; z)$ and

$$g_{\omega,r}(t; z) = \frac{1}{N_r} \sum_{d(t',t) \leq r} G_{\omega,r}(t, t'; z) = \frac{1}{N_r} \langle \delta_t | (H_{\omega,r} - z)^{-1} \mathbf{1}_{Q_r(t)} \rangle. \tag{2.1}$$

For $Q_r \in \mathcal{P}_r$, let

$$\gamma_\omega(Q_r; z) := \frac{1}{N_r} \langle \mathbf{1}_{Q_r} | (H_{\omega,r} - z)^{-1} \mathbf{1}_{Q_r} \rangle.$$

We denote by m the Lebesgue measure on $\mathbb{R}$. For $x \in X$ and $\omega \in \Omega$, we denote by μ_x^ω the spectral measure for $\Delta + V_\omega$ and δ_x, by $\mu_{x,\mathrm{cont}}^\omega$ the continuous part of

μ_x^ω and by $\mu_{x,\mathrm{ac}}^\omega$ the a.c. part. The main spectral localization criterion for the hierarchical Anderson model is the following

Theorem 2.1. *Assume that there exists a sequence $u_r > 0$ with $\sum_{r=1}^\infty u_r^{-1} < \infty$ and*

$$\sum_{r=1}^{\infty} p_r u_r < \infty. \tag{2.2}$$

Let $x \in X$ and a Borel set $B \subset \mathbb{R}$ be given. Then:
(1) *If for a fixed $\omega \in \Omega$,*

$$\sum_{s=1}^{\infty} p_s \left|\gamma_\omega(Q_s(x); e)\right| < \infty, \tag{2.3}$$

for m-a.e. $e \in B$, then $\mu_{x,\mathrm{ac}}^\omega(B) = 0$.
(2) *If $\mathbb{P}$ has a conditional density along the x'th coordinate, and* (2.3) *holds for $\mathbb{P} \otimes m$-a.e. $(\omega, e) \in \Omega \times B$, then $\mu_{x,\mathrm{cont}}^\omega(B) = 0$ for $\mathbb{P}$-a.e. $\omega \in \Omega$.*
(3) *Assume that for $\mathbb{P} \otimes m$-a.e. $(\omega, e) \in \Omega \times B$, there is a finite constant $C_{\omega,e,x}$ such that*

$$|\gamma_\omega(Q_r(x); e)| \le C_{\omega,e,x} u_r, \tag{2.4}$$

for all $r \ge 0$. If $\mathbb{P}$ has a conditional density along the x'th coordinate, then $\mu_{x,\mathrm{cont}}^\omega(B) = 0$ for $\mathbb{P}$-a.e. $\omega \in \Omega$.

To facilitate the exposition, we will first prove three simple propositions (2.2, 2.3 and 2.4). Then we will review the Simon-Wolff localiztion criterion (Theorem 2.5) and give the proof of Theorem 2.1.

Proposition 2.2. *Let $\omega \in \Omega$, $x, y \in X$, $z \in \mathbb{C}\backslash\sigma_\omega$ and $r \ge 0$ be given. Then*

$$G_{\omega,r}(x, y; z) = G_{\omega,0}(x, y; z) - \sum_{s=d(x,y)}^{r} p_s N_{s-1} g_{\omega,s-1}(x; z) g_{\omega,s}(y; z), \tag{2.5}$$

and

$$g_{\omega,r}(x; z) = \frac{1}{N_r} G_{\omega,0}(x, x; z) \prod_{s=1}^{r} (1 - p_s \gamma_\omega(Q_s(x); z)). \tag{2.6}$$

Proof. The formula (2.5) holds for $r = 0$ since $p_0 = 0$. For $s \ge 1$, the resolvent identity yields

$$(H_{\omega,s} - z)^{-1}\delta_y - (H_{\omega,s-1} - z)^{-1}\delta_y = -(H_{\omega,s-1} - z)^{-1} p_s E_s (H_{\omega,s} - z)^{-1}\delta_y.$$

Observe that $E_s(H_{\omega,s} - z)^{-1}\delta_y = g_{\omega,s}(y; z)\mathbf{1}_{Q_s(y)}$. Taking $\langle\delta_x|\cdot\rangle$ in the above equation yields

$$G_{\omega,s}(x, y; z) - G_{\omega,s-1}(x, y; z) = -p_s g_{\omega,s}(y; z)\langle\delta_x|(H_{\omega,s-1} - z)^{-1}\mathbf{1}_{Q_s(y)}\rangle. \tag{2.7}$$

Note that by (2.1),

$$\langle\delta_x|(H_{\omega,s-1} - z)^{-1}\mathbf{1}_{Q_s(y)}\rangle = \begin{cases} N_{s-1} g_{\omega,s-1}(x; z), & \text{if } d(x, y) \le s, \\ 0, & \text{if } d(x, y) > s. \end{cases}$$

The formula (2.5) follows after adding (2.7) for $s = 1, 2, \cdots, r$.

The proof of (2.6) is similar. The resolvent identity yields

$$\langle\delta_x|(H_{\omega,r} - z)^{-1}\mathbf{1}_{Q_r(x)}\rangle = (1 - p_r\gamma_\omega(Q_r(x); z))\langle\delta_x|(H_{\omega,r-1} - z)^{-1}\mathbf{1}_{Q_{r-1}(x)}\rangle,$$

and (2.6) follows after iterating the above formula. □

Proposition 2.3. *Let $u_r > 0$ be a sequence with $\sum_{r=1}^{\infty} u_r^{-1} < \infty$. Let $\omega \in \Omega$ and $x \in X$ be fixed. Then for m-a.e. $e \in \mathbb{R}$, there exists a finite constant C_e such that*

$$\left(\sum_{d(x,y)\leq s} |g_{\omega,s}(y;e)|^2 \right)^{1/2} \leq C_e u_s, \tag{2.8}$$

for all $s \geq 0$.

Proof. Since $l^2(Q_r(x))$ is an N_r-dimensional invariant subspace for $H_{\omega,r}$ and since $\left\| \mathbf{1}_{Q_r(x)} \right\|_2 = \sqrt{N_r}$, we have by Lemma 5.1(see appendix) that for $M_r > 0$,

$$m\left(\left\{ e \in \mathbb{R}\backslash\sigma_\omega : \left\| (H_{\omega,r} - e)^{-1} \mathbf{1}_{Q_r(x)} \right\|_2^2 \geq M_r \right\} \right) \leq \frac{4N_r}{\sqrt{M_r}}.$$

Let $M_r := (u_r N_r)^2$. Then $\sum_{r=1}^{\infty} N_r M_r^{-1/2} < \infty$. By the Borel-Cantelli lemma, for m-a.e. $e \in \mathbb{R}\backslash\sigma_\omega$, there exists a finite constant C_e such that

$$\left\| (H_{\omega,r} - e)^{-1} \mathbf{1}_{Q_r(x)} \right\|_2 < C_e \sqrt{M_r}, \tag{2.9}$$

for all $r \geq 0$. Observe that

$$\begin{aligned} \left(\sum_{d(x,y)\leq s} |g_{\omega,s}(y;e)|^2 \right)^{1/2} &= \left(\sum_{d(x,y)\leq s} \left| \frac{1}{N_s} \langle \delta_y | (H_{\omega,s} - e)^{-1} \mathbf{1}_{Q_s(y)} \rangle \right|^2 \right)^{1/2} \\ &= \frac{1}{N_s} \left(\sum_{d(x,y)\leq s} \left| \langle \delta_y | (H_{\omega,s} - e)^{-1} \mathbf{1}_{Q_s(x)} \rangle \right|^2 \right)^{1/2} \\ &= \frac{1}{N_s} \left\| (H_{\omega,s} - e)^{-1} \mathbf{1}_{Q_s(x)} \right\|_2 . \end{aligned}$$

Inequality (2.9) yields (2.8). □

The key step in proving Theorem 2.1 is:

Proposition 2.4. *Assume that there exists a sequence $u_r > 0$ with $\sum_{r=1}^{\infty} u_r^{-1} < \infty$ and $\sum_{r=1}^{\infty} p_r u_r < \infty$. Let $B \subset \mathbb{R}$ be a Borel set and let $\omega \in \Omega$, $x \in X$ be fixed. Suppose that*

$$\sum_{s=1}^{\infty} p_s \left| \gamma_\omega(Q_s(x);e) \right| < \infty, \tag{2.10}$$

for m-a.e. $e \in B$. Then, for m-a.e. $e \in B$,

$$\sup_{r\geq 0} \sum_{y\in X} |G_{\omega,r}(x,y;e)|^2 < \infty. \tag{2.11}$$

Proof. For all $e \in \mathbb{C}\backslash\sigma_\omega$, the representation formula (2.5) and Cauchy-Schwarz inequality yield

$$\left(\sum_{y\in X} |G_{\omega,r}(x,y;e)|^2\right)^{1/2} \leq |G_{\omega,0}(x,x;e)| + \sum_{s=1}^{r} p_s N_{s-1} |g_{\omega,s-1}(x;e)| \left(\sum_{d(x,y)\leq s} |g_{\omega,s}(y;e)|^2\right)^{1/2}.$$

Hence, by Proposition 2.3, for m-a.e. $e \in \mathbb{R}$,

$$\left(\sum_{y\in X} |G_{\omega,r}(x,y;e)|^2\right)^{1/2} \leq |G_{\omega,0}(x,x;e)| + C_e \sum_{s=1}^{r} p_s u_s N_{s-1} |g_{\omega,s-1}(x;e)|. \tag{2.12}$$

Hypothesis (2.10) implies that for m-a.e. $e \in B$, the product $\prod_{s=1}^{\infty}(1-p_s\gamma_\omega(Q_s(x);e))$ converges. For such e, it follows from (2.6) that there is a finite constant $C_e^{'}$ such that

$$N_s |g_{\omega,s}(x;e)| \leq C_e^{'}, \tag{2.13}$$

for all $s \geq 0$.

Combination of (2.12) with (2.13) yields the estimate

$$\left(\sum_{y\in X} |G_{\omega,r}(x,y;e)|^2\right)^{1/2} \leq |G_{\omega,0}(x,x;e)| + C_e C_e^{'} \sum_{s=1}^{r} p_s u_s,$$

and the result follows. □

Before proving Theorem 2.1, let us recall the Simon-Wolff localization criterion. Define the function $G_{\omega,x} : \mathbb{R} \to [0,+\infty]$ by

$$G_{\omega,x}(e) := \int_{\mathbb{R}} \frac{d\mu_x^\omega(\lambda)}{(e-\lambda)^2} = \lim_{\epsilon\downarrow 0} \left\|(\Delta + V_\omega - e - i\epsilon)^{-1}\delta_x\right\|^2.$$

By the Theorem of de la Vallé Poussin,

$$d\mu_{x,\mathrm{ac}}^\omega(e) = \pi^{-1}\left(\lim_{\epsilon\downarrow 0} \varepsilon \left\|(\Delta + V_\omega - e - i\epsilon)^{-1}\delta_x\right\|^2\right) de.$$

Hence, if for a fixed $\omega \in \Omega$ we have that $G_{\omega,x}(e) < \infty$ for m-a.e. $e \in \mathbb{R}$, then $\mu_{x,\mathrm{ac}}^\omega = 0$.

The Simon-Wolff localization criterion is summarized in:

Theorem 2.5. *Assume that $\mathbb{P}$ has a conditional density along the x'th coordinate. Let $B \subset \mathbb{R}$ be a Borel set such that $G_{\omega,x}(e) < \infty$ for $\mathbb{P}\otimes m$-a.e. $(\omega,e) \in \Omega\times B$. Then $\mu_{x,\mathrm{cont}}^\omega(B) = 0$ for $\mathbb{P}$-a.e. $\omega \in \Omega$.*

Theorem 2.5 is a well known consequence of the rank-1 Simon-Wolff theorem [SW] and the conditional Fubini's theorem.

Now we can prove the main localization criterion.

Proof of Theorem 2.1. Fix $\omega \in \Omega$ and fix $e \in \mathbb{R}\backslash\sigma_\omega$ for which (2.2) holds. Proposition 2.4 yields that then we also have the bound (2.11). By monotone convergence

$$\int_{\mathbb{R}} \frac{d\mu_x^\omega(\lambda)}{(e-\lambda)^2} = \lim_{\varepsilon \downarrow 0} \int_{\mathbb{R}} \frac{d\mu_x^\omega(\lambda)}{(e-\lambda)^2 + \varepsilon^2} = \sup_{\varepsilon > 0} \int_{\mathbb{R}} \frac{d\mu_x^\omega(\lambda)}{(e-\lambda)^2 + \varepsilon^2}.$$

Since for any $z \in \mathbb{C}\backslash\mathbb{R}$,

$$\lim_{r \to \infty} \left\| (H_{\omega,r} - z)^{-1} - (H_\omega - z)^{-1} \right\| = 0,$$

we have that the weak-* limit $\lim_{r\to\infty} \mu_{x,r}^\omega$ equals μ_x^ω, where $\mu_{x,r}^\omega$ is the spectral measure for $H_{\omega,r}$ and δ_x. Therefore

$$\begin{aligned}
\int_{\mathbb{R}} \frac{d\mu_x^\omega(\lambda)}{(e-\lambda)^2} = \sup_{\varepsilon>0} \lim_{r\to\infty} \int_{\mathbb{R}} \frac{d\mu_x^{\omega,r}(\lambda)}{(e-\lambda)^2+\varepsilon^2} &\le \sup_{\varepsilon>0, r\ge 1} \int_{\mathbb{R}} \frac{d\mu_x^{\omega,r}(\lambda)}{(e-\lambda)^2+\varepsilon^2} \\
&= \sup_{r\ge 1} \int_{\mathbb{R}} \frac{d\mu_x^{\omega,r}(\lambda)}{(e-\lambda)^2} \\
&= \sup_{r\ge 1} \left\| (H_{\omega,r} - e)^{-1} \delta_x \right\|^2 \\
&= \sup_{r \ge 1} \sum_{y \in X} |G_{\omega,r}(x,y)|^2 < \infty.
\end{aligned}$$

In the final equality we used the fact that $\{\delta_y : y \in X\}$ is an orthonormal basis for $l^2(X)$. Since $m(\sigma_\omega) = 0$ and since the bound (2.11) holds for m-a.e. $e \in \mathbb{R}\backslash\sigma_\omega$, we have that for every fixed $\omega \in \Omega$, $G_{\omega,x}(e) < \infty$ for m-a.e. $e \in \mathbb{R}$. This proves part (1). Part (2) follows from the fact that $G_{\omega,x}(e) < \infty$ for $\mathbb{P} \otimes m$-a.e. $(\omega, e) \in \Omega \times \mathbb{R}$ and the Simon-Wolff criterion. Part (3) is an immediate consequence of part (2) and assumption (2.2). □

3. Proofs of localization theorems

In this section we show how to apply the localization criterion (Theorem 2.1) to prove theorems 1.6 and 1.3.

Proof of Theorem 1.6. Since $\gamma_\omega(Q_s(x); e) = \sum_{d(x,y)\le s} g_{\omega,s}(y; e)$, Cauchy-Schwarz inequality and Proposition 2.3 yield that for m-a.e. $e \in \mathbb{R}$,

$$|\gamma_\omega(Q_s(x); e)| \le C_e u_s \sqrt{N_s}. \tag{3.1}$$

Hence $\sum_{s\ge 0} p_s |\gamma_\omega(Q_s(x); e)| < \infty$ and the result follows from parts (1) and (2) of Theorem 2.1. □.

The proof of Molchanov's theorem 1.3 requires more work. Let us introduce some notation. It is convenient to form the complex number $v = a + ih \in \mathbb{C}_+$ and to denote by k_v the Cauchy distribution $k_{a,h}$. If ν is a Borel probability measure on the Riemann sphere $\mathbb{S} := \mathbb{C} \cup \{\infty\}$, and $\tau : \mathbb{S} \to \mathbb{S}$ is a Borel measurable map, then $\tau\mu$ will denote the induced Borel probability measure on $\mathbb{S}$:

$$(\tau\nu)(B) := \nu(\tau^{-1}(B)),$$

for Borel sets $B \subset \mathbb{S}$. If $n \ge 2$ is a integer, and ν is a Borel probability measure on $\mathbb{S}$, we let

$$A_n \nu := \tau(\underbrace{\nu \star \cdots \star \nu}_{n \text{ times}}),$$

where $\tau(z) = z/n$, and $\star$ is the usual convolution of measures. Hence, if $Y_1, \dots Y_n$ are i.i.d. random variables on $\mathbb{S}$ with distribution ν, then $A_n\nu$ is the distribution of $(Y_1 + \cdots + Y_n)/n$. The following proposition, whose proof is left an an easy exercise for the reader, summarizes some of the well known classical facts about Cauchy distributions.

Proposition 3.1. *Cauchy distributions have the following properties:*

(1) *If $v_1, v_1 \in \mathbb{C}_+$, then*
$$k_{v_1} \star k_{v_2} = k_{v_1+v_2}.$$
(2) *If $\tau(z) = -\overline{z}$ or if $\tau(z) = (az+b)/(cz+d)$, where $a, b, c, d \in \mathbb{R}$, $ad - bc > 0$, then*
$$\tau k_v = k_{\tau v},$$
for all $v \in \mathbb{C}_+$.
(3) *For all $v \in \mathbb{C}_+$ and integers $n \geq 2$,*
$$A_n k_v = k_v.$$

We will use fractional linear transformations of the special form
$$\tau_p(z) = \frac{z}{1 + pz},$$
where $p \geq 0$. Note that $\tau_{p+p'} = \tau_p \circ \tau_{p'}$. The transformations τ_p are important in the hierarchical Anderson model because of the following recursive relation, proven by Molchanov in [M2].

Proposition 3.2. *Let $\omega \in \Omega$ and $z \in \mathbb{C}\backslash\sigma_\omega$ be given. Then for all $x \in X$,*
$$\gamma_\omega(Q_0(x); z) = \frac{1}{\omega(x) - z}. \tag{3.2}$$
Moreover, for all $r \geq 1$ and $Q_r \in \mathcal{P}_r$, we have
$$\gamma_\omega(Q_r; z) = \tau_{p_r}\left(\frac{1}{n_r}\sum_{Q^j_{r-1}\subset Q_r} \gamma_\omega(Q^j_{r-1}; z)\right), \tag{3.3}$$
where the sum contains n_r terms, one for each rank $r - 1$ subcluster Q^j_{r-1} of Q_r.

Proof. The formula (3.2) is clear since $G_{\omega,0}(x, y; z) := (\omega(x) - z)^1 \langle\delta_x|\delta_y\rangle$. For $r \geq 1$, the resolvent identity yields
$$(H_{\omega,r} - z)^{-1}\mathbf{1}_{Q_r} - (H_{\omega,r-1} - z)^{-1}\mathbf{1}_{Q_r} = -p_r\gamma_\omega(Q_r; z)(H_{\omega,r-1} - z)^{-1}\mathbf{1}_{Q_r}.$$
Taking $\langle \mathbf{1}_{Q_r}|\cdot\rangle$ in the above equation yields
$$N_r\gamma_\omega(Q_r; z) - \sum_{Q^j_{r-1}\subset Q_r} N_{r-1}\gamma_\omega(Q^j_{r-1}; z) = -p_r\gamma_\omega(Q_r; z)\sum_{Q^j_{r-1}\subset Q_r} N_{r-1}\gamma_\omega(Q^j_{r-1}; z).$$
The formula (3.3) follows after dividing by N_r in the above equation and then solving for $\gamma_\omega(Q_r; z)$. □

Proof of Theorem 1.3.
Since for every fixed ω, $m(\sigma_\omega) = 0$, we have that for $\mathbb{P} \otimes m$-a.e. $(\omega, e) \in \Omega \times \mathbb{R}$, $e \notin \sigma_w$. Hence, there is a Borel mesurable set $E \subset \mathbb{R}$ with the properties

(1) $m(\mathbb{R}\backslash E) = 0$, and

(2) for every fixed $e \in E$, there is a set $\Omega_e \in \mathcal{F}$ with $\mathbb{P}(\Omega_e) = 1$ and such that $e \notin \sigma_w$ for all $\omega \in \Omega_e$.

Let us fix $e \in E$. Then for all $\omega \in \Omega_e$, and hence for $\mathbb{P}$-a.e. $\omega \in \Omega$, all the representation formulas in propositions 2.2 and 3.2 are valid. It follows from (3.2) and from Proposition 3.1 that $\{\gamma_\omega(Q_0(x); e) : x \in X\}$ are i.i.d. Cauchy random variables with distribution k_{v_0}, where $v_0(e) = -(e - a + ih)^{-1}$. Moreover, (3.3) and Proposition 3.1 yield that for $r \geq 1$, $\{\gamma_\omega(Q_r; e) : Q_r \in P_r\}$ are i.i.d. Cauchy random variables with distribution k_{v_r}, where $v_r(e) = \tau_{\lambda_r} v_0(e)$. Since $\lambda_r \to \lambda_\infty = 1$ as $r \to \infty$, the closure of the orbit

$$V(e) := \bigcup_{r \geq 0} \{v_r(e)\},$$

is equal to $V(e) \cup \{\tau_1 v_0(e)\}$, a compact set in $\mathbb{C}_+$. Since $\sup_{r\geq 0} |v_r(e)| < \infty$, there is a constant $K(e) < \infty$, such that

$$\mathbb{P}\left(|\gamma_\omega(Q_r; e)| > u\right) \leq \frac{K(e)}{u},$$

for all real $u > 0$, integer $r \geq 0$ and $Q_r \in P_r$. Let us now fix $x \in X$ and take $u_r = r^{1+\varepsilon}$. By the Borel-Cantelli lemma, for $\mathbb{P}$-a.e. $\omega \in \Omega_e$, there exists a finite constant $L_\omega(e)$ such that

$$|\gamma_\omega(Q_r(x); e)| \leq L_\omega(e) u_r, \tag{3.4}$$

for all $r \geq 0$. Part (3) of Theorem 2.1 yields that $\mu^\omega_{x,\mathrm{cont}} = 0$ for $\mathbb{P}$-a.e. $\omega \in \Omega$. Since $x \in X$ is arbitrary, the result follows. □.

Note that, in the proof given above, (3.4) is a significant improvement of (3.1) since the factor $\sqrt{N_r}$ is no longer present.

4. A generalization of Molchanov's theorem

In this section we generalize Theorem 1.3 to convex combinations of Cauchy distributions. Let $\mathcal{M}$ denote the set of Borel probability measures on $\mathbb{C}_+$. Given $\nu \in \mathcal{M}$ and a Borel set $B \subset \mathbb{R}$, we set

$$k_\nu(B) := \int k_v(B) d\nu(v).$$

Then k_ν is a Borel probability measure on $\mathbb{R}$ and for every bounded Borel function $f : \mathbb{R} \to \mathbb{C}$,

$$\int_{\mathbb{R}} f dk_\nu = \int \left(\int_{\mathbb{R}} f dk_v \right) d\nu(v).$$

We call k_ν a *mixed Cauchy distribution.* Note that the usual Cauchy distribution k_{a+ih} is a special case of the above general defintion: $k_{a+ih} = k_\nu$ where $\nu = \delta(a+ih)$. Let S be the set of mixed Cauchy distributions k_ν, such that the distance from $\mathbb{R}$ to the support of ν is strictly positive.

Theorem 4.1. *Assume that there exists a sequence $u_r > 0$ with $\sum_{r=1}^\infty u_r^{-1} < \infty$ and*

$$\sum_{r=1}^\infty p_r u_r < \infty.$$

If the random variables $\{\omega(x) : x \in X\}$ are i.i.d. with a mixed Cauchy distrubution $k_\nu \in S$, then $\sigma_{\mathrm{cont}}(H_\omega) = \emptyset$ for $\mathbb{P}$-a.e. ω.

Remark 1. Theorem 4.1 is valid at arbitrary disorder. Indeed, if $\omega(x)$ has a distribution $k_\nu \in S$, and $c > 0$, then $c\omega$ has again a distribution of the form $k_{\nu'} \in S$. Precisely, $\nu' = \tau\nu$, where $\tau(z) = cz$.

Remark 2. If f is any compactly supported continuous probability density on $\mathbb{R}$, and $\varepsilon > 0$, then there is a mixed Cauchy distribution $k_\nu \in S$ s.t.

$$\sup_{e\in\mathbb{R}} |f(e) - dk_\nu/de| < \varepsilon.$$

To see this, recall that the harmonic function $F(v) = \int_{\mathbb{R}} f(e)dk_v(e)$, $v \in \mathbb{C}_+$, converges uniformly to f as $\Im v \downarrow 0$, and note that the integral can be approximated by finite Riemann sums.

The proof of Theorem 4.1 parallels that of Molchanov's theorem 1.3. First, we need a generalization of Proposition 3.1 to the case of mixed Cauchy distributions.

Proposition 4.2. *Mixed Cauchy distributions have the following properties:*

(1) *If* $\nu_1, \nu_1 \in \mathcal{M}$, *then*
$$k_{\nu_1} \star k_{\nu_2} = k_{\nu_1 \star \nu_2}.$$
(2) *If* $\tau(z) = -\overline{z}$ *or if* $\tau(z) = (az+b)/(cz+d)$, *where* $a, b, c, d \in \mathbb{R}$, $ad - bc > 0$, *then*
$$\tau k_\nu = k_{\tau\nu},$$
for all $\nu \in \mathcal{M}$.

(3) *For all* $\nu \in \mathcal{M}$ *and integers* $n \geq 2$,
$$A_n k_\nu = k_{A_n \nu}.$$

The proof is a straight forward consequence of the definitions together with Propostion 3.1 and Fubini's theorem.

Proof of Theorem 1.3. Let E be the set of full Lebesgue measure as in the beginning of the proof of Theorem 1.3, and let $e \in E$ be fixed. It follows from (3.2) and from Proposition 4.2 that $\{\gamma_\omega(Q_0(x); e) : x \in X\}$ are i.i.d. mixed Cauchy random variables with distribution k_{ν_0}, where $\nu_0 = \tau_0\nu$, $\tau_0(z) = -(e - \overline{z})^{-1}$. Since the distance from the support of ν to $\mathbb{R}$ is strictly positive, there is a closed disk $D_0 \subset \mathbb{C}_+$ such that $\mathrm{supp}(\nu_0) \subset \mathrm{D}_0$. The relation (3.3) and Proposition 4.2 yield that for $r \geq 1$, $\{\gamma_\omega(Q_r; e) : Q_r \in P_r\}$ are i.i.d. mixed Cauchy random variables with distribution k_{ν_r}, where

$$\nu_r = \tau_{p_r} A_{n_r} \nu_{r-1}.$$

For $r \geq 0$, let $D_r := \tau_{\lambda_r} D_0$. Since τ_{λ_r} is a fractional linear transformation taking $\mathbb{C}_+$ onto itself, D_r is again a closed disk in $\mathbb{C}_+$. We claim that for all $r \geq 0$,

$$(*) \qquad \mathrm{supp}(\nu_\mathrm{r}) \subset \mathrm{D_r}.$$

The proof is by induction. By construction of D_0, $(*)$ holds for $r = 0$. Let $r \geq 1$ and suppose $(*)$ holds for $r - 1$. Since $\mathrm{supp}(\nu_{\mathrm{r}-1}) \subset \mathrm{D}_{\mathrm{r}-1}$, and D_{r-1} is convex, we have that $\mathrm{supp}(\mathrm{A_{n_r}}\nu_{\mathrm{r}-1}) \subset \mathrm{D}_{\mathrm{r}-1}$. Hence $\mathit{support}(\tau_{p_r} A_{n_r} \nu_{r-1}) \subset \tau_{p_r} D_{r-1} = D_r$, and $(*)$ holds for r, proving the claim.

Let D be the closure of $\cup_{r\geq 0} D_r$. D is a compact subset of $\mathbb{C}_+$ and for all $r \geq 0$, $\mathrm{supp}(\nu_\mathrm{r}) \subset \mathrm{D}$. It follows that there is a constant $K(e) < \infty$, such that

$$\mathbb{P}\left(|\gamma_\omega(Q_r; e)| > u\right) \leq \frac{K(e)}{u},$$

for all real $u > 0$, integer $r \geq 0$ and $Q_r \in P_r$. The rest of the proof is the same as in Theorem 1.3, with the more general sequence u_r instead of $r^{1+\varepsilon}$. $\square$

5. Appendix

Proof of Lemma 1.2. Let $e \in S$. Given a cluster Q_r or rank r and given $\varepsilon > 0$, consider the event

$$B(Q_r, \varepsilon) := \left\{ \omega : \max_{x \in Q_r} |\omega(x) - e| < \varepsilon \right\}.$$

Since $\nu((e-\varepsilon, e+\varepsilon)) > 0$ and the random variables $\{\omega(x) : x \in Q_r\}$ are independent, we have that

$$\mathbb{P}\left(B(Q_r, \varepsilon)\right) = \left(\nu((e-\varepsilon, e+\varepsilon))\right)^{N_r} > 0.$$

According to the hierarchical structure, X is a countable union of disjoint clusters of rank r: $X = \bigcup_{j=1}^{\infty} Q_{r,j}$. Let $B_{r,\varepsilon} := \limsup_j B(Q_{r,j}, \varepsilon)$. Since the events $\{B(Q_{r,j}, \varepsilon), j = 1, \ldots, \infty\}$ are independent and $\sum_{j=1}^{\infty} \mathbb{P}\left(B(Q_{r,j}, \varepsilon)\right) = \infty$, the second Borel-Cantelli Lemma yields that $\mathbb{P}(B_{r,\varepsilon}) = 1$. Letting $B := \cap_{r \geq 0} B_{r,1/(r+1)}$, we also have $\mathbb{P}(B) = 1$. Hence, for $\omega \in B$, there is a sequence $(Q_{r,j_r} : r \geq 0)$ of clusters of increasing rank with the property that

$$(*) \qquad \max_{x \in Q_{r,j_r}} |\omega(x) - e| < \frac{1}{r+1},$$

for all $r \geq 0$. Let $\lambda \in \sigma(\Delta)$. Then $\lambda = \lambda_s$ for some $s \in \{1, \ldots, \infty\}$. For $r \geq 0$, we let $q_r = \min(r, s)$. It is easy to see that for each $r \geq 0$, there exists a unit eigenvector ψ_r of Δ corresponding to the eigenvalue λ_{q_r}, and such that $\psi_r(x) = 0$ for $x \notin Q_{r,j_r}$. Indeed, take a unit vector ψ_r constant on every rank q_r subcluster of Q_{r,j_r}, zero outside Q_{r,j_r} and such that $E_{q_r+1}\psi_r = 0$. Then

$$\lim_{r \to \infty} \|\Delta\psi_r - \lambda_s \psi_r\| = 0,$$

and, by (*),

$$\lim_{r \to \infty} \|H_\omega \psi_r - (\lambda_s + e)\psi_r\| = 0,$$

for $\omega \in B$. Hence, ψ_r is a Weil sequence for the operator H_ω and $\lambda_s + e \in \sigma(H_\omega)$. We have shown that for $\mathbb{P}$-a.e. $\omega \in \Omega$, $\sigma(\Delta) + S \subset \sigma(H_\omega)$. The inclusion $\sigma(H_\omega) \subset (\sigma(\Delta) + ch(S)) \cap ([0,1] + S)$ follows from general facts about the spectrum of a sum of two self-adjoint operators. $\square$.

Lemma 5.1. *Let A be a hermitian $N \times N$ matrix and $v \in \mathbb{C}^N$. Then for all $M > 0$*

$$m\left(\left\{e : \left\|(A-e)^{-1}v\right\|_2^2 \geq M\right\}\right) \leq 4\sqrt{\frac{N}{M}}\,\|v\|_2$$

Proof. (Following [M2]) Let $\lambda_1, \ldots, \lambda_N$ be the eigenvalues of A and $\psi_1, \ldots, \psi_N$ the corresponding orthonormal basis of eigenvectors. In this basis, we have $v = \sum_{i=1}^N v_i \psi_i$, $\|v\|_2^2 = \sum_{i=1}^N |v_i|^2$ and $(A-e)^{-1}v = \sum_{i=1}^N (\lambda_i - e)^{-1} v_i \psi_i$. For each $i = 1, \ldots, N$, consider the open interval $\Delta_i = (\lambda_i - v_i/\sqrt{M}, \lambda_i + v_i/\sqrt{M})$. The total length of these N intervals is

$$m\left(\bigcup_{i=1}^N \Delta_i\right) \leq \sum_{i=1}^N m\left(\Delta_i\right) = \frac{2}{\sqrt{M}} \sum_{i=1}^N |v_i| \leq \frac{2}{\sqrt{M}} \sqrt{N}\,\|v\|_2\,.$$

Chebyshev inequality yields

$$
\begin{aligned}
m\left(\left\{e \in \mathbb{R}\backslash \bigcup_{i=1}^{N} \Delta_i : \left\|(A-e)^{-1}v\right\|_2^2 \geq M\right\}\right) &\leq \frac{1}{M}\int_{\mathbb{R}\backslash \bigcup_{i=1}^{N}\Delta_i} \left\|(A-e)^{-1}v\right\|_2^2 de \\
&= \frac{1}{M}\int_{\mathbb{R}\backslash \bigcup_{i=1}^{N}\Delta_i} \sum_{i=1}^{N}\left|\frac{v_i}{\lambda_i - e}\right|^2 de \\
&\leq \frac{1}{M}\sum_{i=1}^{N}\int_{\mathbb{R}\backslash\Delta_i}\left|\frac{v_i}{\lambda_i - e}\right|^2 de \\
&= \frac{1}{M}\sum_{i=1}^{N} 2\int_{|v_i|/\sqrt{M}}^{\infty}\frac{|v_i|^2}{t^2}dt \\
&= \frac{2}{M}\sqrt{M}\sum_{i=1}^{N}|v_i| \\
&\leq \frac{2}{\sqrt{M}}\sqrt{N}\left\|v\right\|_2,
\end{aligned}
$$

and the result follows. □

Acknowledgements. We are grateful to Vojkan Jaksic for suggesting this research project and for giving us constructive feedback. We would also like to thank Stanislav Molchanov for very helpful discussions and enthusiastic encouragements.

References

[Bo] Bovier, A.: The density of states in the Anderson model at weak disorder: a renormalization group analysis of the hierarchical model. J. Statist. Phys. **59** (1990), no. 3-4, 745–779.

[BS] Bleher, P. M., Sinai, Ya. G.: Investigation of the Critical Point in Models of the Type of Dyson's Hierarchical Models. Commun. Math. Phys. **33**, (1973).

[CL] Carmona R., Lacroix J.: Spectral Theory of Random Schrodinger Operators. Birkhauser, 1990.

[D] Dyson, F.J.: Existence of a phase-transition in a one dimensional Ising Ferromagnet. Comm. Math. Phys. **12**, **91** (1969).

[K1] Kritchevski, E.: Spectral Localization in the Hierarchical Anderson Model. To appear in the Proceedings of AMS.

[K2] Kritchevski, E.: PhD. thesis, McGill University. In preparation.

[M1] Molchanov, S.: Lectures on random media. Lectures on probability theory (Saint-Flour, 1992), 242–411, Lecture Notes in Math., **1581**, Springer, Berlin, 1994.

[M2] Molchanov S.: Hierarchical random matrices and operators. Application to Anderson model. Multidimensional statistical analysis and theory of random matrices (Bowling Green, OH, 1996), 179–194, VSP, Utrecht, 1996.

[SW] Simon B., Wolff T.: Singular continuous spectrum under rank one perturbations and localization for random Hamiltonians. Communications in Pure and Applied Mathematics **49** (1986), 75.

Department of Mathematics and Statistics. McGill University, 805 Sherbrooke Street West Montreal, QC, H3A 2K6 Canada

E-mail address: `ekritc@math.mcgill.ca`