

# GENERALIZED EIGENVALUE-COUNTING ESTIMATES FOR THE ANDERSON MODEL

JEAN-MICHEL COMBES, FRANÇOIS GERMINET, AND ABEL KLEIN

ABSTRACT. We show how spectral averaging for rank one perturbations, combined with a consequence of the min-max principle, can be used to derive generalized eigenvalue-counting estimates for the Anderson model. In particular, we present a simple and transparent proof of Minami's estimate for the probability of two or more eigenvalues in an interval. Our estimates generalize Minami's estimate and its extensions to  $n$  eigenvalues in two ways: we allow for singular measures and for  $n$  arbitrary intervals. As an application, we derive new results about the multiplicity of eigenvalues and Mott's formula for the ac-conductivity when the single site probability distribution is Hölder continuous.

## 1. INTRODUCTION

Consider the random self-adjoint operator

$$H_\omega = H_0 + \omega \Pi_\varphi \quad \text{on } \mathcal{H}, \quad (1.1)$$

where  $H_0$  is a self-adjoint operator on the Hilbert space  $\mathcal{H}$ ,  $\varphi \in \mathcal{H}$  with  $\|\varphi\| = 1$ , and  $\omega$  is a random variable with a non-degenerate probability distribution  $\mu$  with compact support. By  $\Pi_\varphi$  we denote the orthogonal projection onto  $\mathbb{C}\varphi$ , the one-dimensional subspace spanned by  $\varphi$ . Let  $P_\omega(J) = \chi_J(H_\omega)$  for a Borel set  $J \subset \mathbb{R}$ . There is a fundamental spectral averaging estimate: for all bounded intervals  $I \subset \mathbb{R}$  we have

$$\mathbb{E}_\omega \{ \langle \varphi, P_\omega(I)\varphi \rangle \} := \int d\mu(\omega) \langle \varphi, P_\omega(I)\varphi \rangle \leq Q_\mu(|I|), \quad (1.2)$$

where  $Q_\mu(s) = 8S_\mu(s)$  for  $s \geq 0$ , with  $S_\mu(s) := \sup_{a \in \mathbb{R}} \mu([a, a+s])$  being the concentration function of  $\mu$ . In this generality, i.e.,  $\mu$  arbitrary and  $Q_\mu(s) = 8S_\mu(s)$ , this is a recent result of Combes, Hislop and Klopp [CoHK2, Eq. (3.16)]. (We present a proof in Appendix A for completeness.) The estimate (1.2) is useful when the measure  $\mu$  has no atoms, i.e.,  $\lim_{s \downarrow 0} Q_\mu(s) = 0$ , which we assume from now on. If  $\mu$  has a bounded density  $\rho$ , (1.2) was known to hold with  $Q_\mu(s) = \|\rho\|_\infty s$  (e.g. [W, FS, CKM, CoHK, Ki]; a simple proof is given in Appendix A). If  $\mu$  is Hölder continuous, i.e.,  $S_\mu(s) \leq Cs^\alpha$  with  $\alpha \in ]0, 1[$ , (1.2) was known with  $Q_\mu(s) = C(1-\alpha)^{-1}s^\alpha$  [CKM, Theorem 6.2]. All we will require of  $Q_\mu$  is the validity of (1.2).

In this article we show how the fundamental spectral averaging estimate (1.2), combined with a consequence of the min-max principle applied to rank one perturbations, can be used to derive generalized eigenvalue-counting estimates for the

---

*Date:* Version of April 13, 2008.

*2000 Mathematics Subject Classification.* Primary 82B44; Secondary 47B80, 60H25.

Anderson model. As an application, we derive new results about the multiplicity of eigenvalues and Mott's formula for the ac-conductivity when the single site probability distribution is Hölder continuous.

We consider the generalized Anderson model given by the random Hamiltonian

$$H_\omega = H_0 + V_\omega \quad \text{on} \quad \ell^2(\mathbb{Z}^d), \quad (1.3)$$

where  $H_0$  is a bounded self-adjoint operator and  $V_\omega$  is the random potential given by  $V_\omega(j) = \omega_j$ , i.e.,  $V_\omega = \sum_{j \in \mathbb{Z}^d} \omega_j \Pi_j$  with  $\Pi_j = \Pi_{\delta_j}$ . Here  $\omega = \{\omega_j\}_{j \in \mathbb{Z}^d}$  is a family of independent random variables, such that for each  $j \in \mathbb{Z}^d$  the random variable  $\omega_j$  has a probability distribution  $\mu_j$  with no atoms and compact support. We set  $Q_j = Q_{\mu_j}$ . We write  $\mathbb{E}_{\omega_j}$  for the expectation with respect to the random variable  $\omega_j$ , and write  $\mathbb{E} = \mathbb{E}_\omega$  for the joint expectation. We also set  $\omega_k^\perp = \{\omega_j\}_{j \in \mathbb{Z}^d \setminus \{k\}}$  and let  $\mathbb{E}_{\omega_k^\perp}$  denote the corresponding expectation.

Restrictions of  $H_\omega$  to finite volumes  $\Lambda \subset \mathbb{Z}^d$  are denoted by  $H_{\omega, \Lambda}$ , a self-adjoint operator of the form

$$H_{\omega, \Lambda} = H_{0, \Lambda} + \sum_{j \in \Lambda} \omega_j \Pi_j \quad \text{on} \quad \ell^2(\Lambda), \quad (1.4)$$

with  $H_{0, \Lambda}$  a self-adjoint restriction of  $H_0$  to the finite-dimensional Hilbert space  $\ell^2(\Lambda)$ . (Our results are not sensitive to the choice of  $H_{0, \Lambda}$ , they hold for any boundary condition.) Given a Borel set  $J \subset \mathbb{R}$ , we write  $P_\omega^{(\Lambda)}(J) = P_{H_\omega}^{(\Lambda)}(J) = \chi_J(H_{\omega, \Lambda})$  for the associated spectral projection. We set  $Q_\Lambda(s) := \max_{j \in \Lambda} Q_j(s)$ .

The Wegner estimate [W] measures the probability that  $H_{\omega, \Lambda}$  has an eigenvalue in an interval  $I$ :

$$\mathbb{P} \left\{ \text{tr} P_{H_\omega}^{(\Lambda)}(I) \geq 1 \right\} \leq \mathbb{E} \left\{ \text{tr} P_{H_\omega}^{(\Lambda)}(I) \right\} \leq Q_\Lambda(|I|) |\Lambda|. \quad (1.5)$$

The Wegner estimate holds for the generalized Anderson model. It is an immediate consequence of (1.2):

$$\mathbb{E} \left\{ \text{tr} P_{H_\omega}^{(\Lambda)}(I) \right\} = \sum_{j \in \Lambda} \mathbb{E}_{\omega_j^\perp} \left\{ \mathbb{E}_{\omega_j} \left\{ \langle \delta_j, P_{H_\omega}^{(\Lambda)}(I) \delta_j \rangle \right\} \right\} \leq Q_\Lambda(|I|) |\Lambda|. \quad (1.6)$$

Minami [M] estimated the probability that  $H_{\omega, \Lambda}$  has at least two eigenvalues in an interval  $I$ . Assuming that all  $\mu_j$  have bounded densities  $\rho_j$ , Minami proved that

$$2 \mathbb{P} \left\{ \text{tr} P_{H_\omega}^{(\Lambda)}(I) \geq 2 \right\} \leq \mathbb{E} \left\{ \left( \text{tr} P_{H_\omega}^{(\Lambda)}(I) \right)^2 - \text{tr} P_{H_\omega}^{(\Lambda)}(I) \right\} \leq \left( \pi \rho_\infty^{(\Lambda)} |I| |\Lambda| \right)^2, \quad (1.7)$$

where  $\rho_\infty^{(\Lambda)} := \max_{j \in \Lambda} \|\rho_j\|_\infty$ . Minami's proof required  $H_0$  to have real matrix elements, i.e.,  $\langle \delta_j, H_0 \delta_k \rangle \in \mathbb{R}$  for all  $j, k$ . This restriction was recently removed by Bellissard, Hislop and Stolz [BHS] and by Graf and Vaghi [GrV]. They also estimated the probability that  $H_{\omega, \Lambda}$  has at least  $n$  eigenvalues in  $I$  for all  $n \in \mathbb{N}$ , assuming, as Minami, that all  $\mu_j$  have bounded densities  $\rho_j$ ,

Minami's estimate has important consequences for the physical behavior of the Anderson model in the localized (insulator) regime. It is the crucial ingredient in Minami's proof of the absence of eigenvalue repulsion, showing that the properly rescaled eigenvalues behave according to a Poisson process [M]. (See [N, KN] for further developments.) It was shown to imply simplicity of eigenvalues by Klein and Molchanov [KLM]. It is an important ingredient in the derivation of a rigorous form of Mott's formula for the ac-conductivity by Klein, Lenoble and Müller [KILM].

In this paper we present a new approach to eigenvalue-counting inequalities, obtaining a simple and transparent proof of Minami's estimate, based on (1.2) and a consequence of the min-max principle applied to rank one perturbations (see Lemma 4.1). Our proof also generalizes Minami's estimate and its extensions to  $n$  eigenvalues in two ways: we allow for singular measures and for  $n$  arbitrary intervals.

The crucial step in Minami's proof, namely [M, Lemma 2], estimates the average of a determinant whose entries are matrix elements of the imaginary part of the resolvent; the proofs in [BHS, GrV] have similar steps. In contrast, our proof only averages spectral projections. Although in this paper we use results specific to rank one perturbations, we are hopeful that our approach can be used to finally provide a proof of a Minami-like estimate to the continuum Anderson Hamiltonian. This would lead to a proof of Poisson eigenvalue statistics and simplicity of eigenvalues for the continuum Anderson Hamiltonian.

The paper is organized as follows. In Section 2 we state our main results, namely our generalized Minami estimate together with its application to eigenvalue-counting estimates. In Section 3, we consider the Anderson model with a Hölder continuous probability distribution, for which we extend previous results on the multiplicity of the spectrum and Mott's formula for energies in the region of Anderson localization. In Section 4 we prove the statements of Section 2; we give first a simple proof of the original Minami estimate, allowing arbitrary probability distributions, and then turn to the proof of the general case. In Appendix A we provide proofs for the fundamental spectral averaging estimate (1.2).

## 2. THE MAIN RESULTS

We will now state our main results. The proofs will be given in Section 4. We start with our extension of Minami's estimate, i.e., (1.7). Recall we may take

$$Q_\Lambda(s) := \max_{j \in \Lambda} Q_j(s), \text{ with } Q_j(s) = \begin{cases} \|\rho_j\|_\infty s & \text{if } \mu_j \text{ has a bounded density } \rho_j \\ 8S_{\mu_j}(s) & \text{otherwise} \end{cases}.$$

**Theorem 2.1.** *Fix a finite volume  $\Lambda \subset \mathbb{Z}^d$ . For any two bounded intervals  $I_1, I_2$  we have*

$$\mathbb{E} \left\{ \left( \text{tr } P_\omega^{(\Lambda)}(I_1) \right) \left( \text{tr } P_\omega^{(\Lambda)}(I_2) \right) - \min \left\{ \text{tr } P_\omega^{(\Lambda)}(I_1), \text{tr } P_\omega^{(\Lambda)}(I_2) \right\} \right\} \leq 2 Q_\Lambda(|I_1|) Q_\Lambda(|I_2|) |\Lambda|^2. \quad (2.1)$$

If  $I_1 \subset I_2$ , we have

$$\mathbb{E} \left\{ \left( \text{tr } P_\omega^{(\Lambda)}(I_1) \right) \left( \text{tr } P_\omega^{(\Lambda)}(I_2) - 1 \right) \right\} \leq Q_\Lambda(|I_1|) Q_\Lambda(|I_2|) |\Lambda|^2. \quad (2.2)$$

In particular, for all bounded intervals  $I$  we have

$$\mathbb{E} \left\{ \left( \text{tr } P_\omega^{(\Lambda)}(I) \right) \left( \text{tr } P_\omega^{(\Lambda)}(I) - 1 \right) \right\} \leq (Q_\Lambda(|I|) |\Lambda|)^2. \quad (2.3)$$

**Remark 2.2.** *Note that*

$$\left( \text{tr } P_\omega^{(\Lambda)}(I_1) \right) \left( \text{tr } P_\omega^{(\Lambda)}(I_2) \right) - \min \left\{ \text{tr } P_\omega^{(\Lambda)}(I_1), \text{tr } P_\omega^{(\Lambda)}(I_2) \right\} \geq 0. \quad (2.4)$$

Note also that the intervals  $I_1$  and  $I_2$  in (2.1) may be disjoint. In this case the usual Minami's estimate (1.7) would yield  $\pi^2 \left(\rho_\infty^{(\Lambda)}\right)^2 |I|^2 |\Lambda|^2$ , with an interval  $I \supset I_1 \cup I_2$ , while (2.1) gives  $2 \left(\rho_\infty^{(\Lambda)}\right)^2 |I_1| |I_2| |\Lambda|^2$ .

We now turn to the general case of  $n$  arbitrary intervals, extending the results of [BHS, GrV]. Given  $n \in \mathbb{N}$ , we let  $\mathcal{S}_n$  denote the group of all permutations of  $\{1, 2, \dots, n\}$ , and recall that  $|\mathcal{S}_n| = n!$ . Given a finite volume  $\Lambda \subset \mathbb{Z}^d$  and bounded intervals  $I_1, \dots, I_n$  (not necessarily distinct), we pick  $\sigma_\omega = \sigma_\omega^{(\Lambda)}(I_1, \dots, I_n) \in \mathcal{S}_n$  such that

$$\mathrm{tr} P_\omega^{(\Lambda)}(I_{\sigma_\omega(1)}) \leq \mathrm{tr} P_\omega^{(\Lambda)}(I_{\sigma_\omega(2)}) \leq \dots \leq \mathrm{tr} P_\omega^{(\Lambda)}(I_{\sigma_\omega(n)}), \quad (2.5)$$

in which case we have

$$\left(\mathrm{tr} P_\omega^{(\Lambda)}(I_{\sigma_\omega(1)})\right) \left(\mathrm{tr} P_\omega^{(\Lambda)}(I_{\sigma_\omega(2)}) - 1\right) \cdots \left(\mathrm{tr} P_\omega^{(\Lambda)}(I_{\sigma_\omega(n)}) - (n-1)\right) \geq 0. \quad (2.6)$$

To avoid ambiguity, we select  $\sigma_\omega$  uniquely by requiring  $\sigma_\omega(i) < \sigma_\omega(j)$  if  $i < j$  and  $\mathrm{tr} P_\omega^{(\Lambda)}(I_{\sigma_\omega(i)}) = \mathrm{tr} P_\omega^{(\Lambda)}(I_{\sigma_\omega(j)})$ . (Note that the product in the left hand side of (2.6) is independent of the choice of  $\sigma_\omega \in \mathcal{S}_n$  satisfying (2.5).) We let  $\mathcal{S}_n(I_1, \dots, I_n)$  be the collection permutations  $\sigma \in \mathcal{S}_n$  such that  $\sigma = \sigma_\omega$  for some  $\omega$ , and let  $M(I_1, \dots, I_n)$  denote the cardinality of  $\mathcal{S}_n(I_1, \dots, I_n)$ . Note that  $1 \leq M(I_1, \dots, I_n) \leq n!$ . We have  $M(I_1, \dots, I_n) = n!$  if the  $n$  intervals are incompatible, i.e.,  $I_j \subset I_k$  implies  $j = k$ , and  $M(I_1, \dots, I_n) = 1$  if  $I_1 \subset I_2 \subset \dots \subset I_n$ .

**Theorem 2.3.** *Fix a finite volume  $\Lambda \subset \mathbb{Z}^d$ , let  $n \in \mathbb{N}$ , and consider  $n$  bounded intervals  $I_1, \dots, I_n$  (not necessarily distinct). Then, setting  $\sigma_\omega = \sigma_\omega^{(\Lambda)}(I_1, \dots, I_n)$ , we have*

$$\begin{aligned} \mathbb{E} \left\{ \left(\mathrm{tr} P_\omega^{(\Lambda)}(I_{\sigma_\omega(1)})\right) \left(\mathrm{tr} P_\omega^{(\Lambda)}(I_{\sigma_\omega(2)}) - 1\right) \cdots \left(\mathrm{tr} P_\omega^{(\Lambda)}(I_{\sigma_\omega(n)}) - (n-1)\right) \right\} & (2.7) \\ & \leq M(I_1, \dots, I_n) \left( \prod_{k=1}^n Q^{(\Lambda)}(|I_k|) \right) |\Lambda|^n. \end{aligned}$$

In the special case when  $I_1 \subset I_2 \subset \dots \subset I_n$ , we have

$$\begin{aligned} \mathbb{E} \left\{ \left(\mathrm{tr} P_\omega^{(\Lambda)}(I_1)\right) \left(\mathrm{tr} P_\omega^{(\Lambda)}(I_2) - 1\right) \cdots \left(\mathrm{tr} P_\omega^{(\Lambda)}(I_n) - (n-1)\right) \right\} & (2.8) \\ & \leq \left( \prod_{k=1}^n Q^{(\Lambda)}(|I_k|) \right) |\Lambda|^n. \end{aligned}$$

In particular, for any bounded interval  $I$  we have

$$\mathbb{E} \left\{ \left(\mathrm{tr} P_\omega^{(\Lambda)}(I)\right) \left(\mathrm{tr} P_\omega^{(\Lambda)}(I) - 1\right) \cdots \left(\mathrm{tr} P_\omega^{(\Lambda)}(I) - (n-1)\right) \right\} \leq \left(Q^{(\Lambda)}(|I|) |\Lambda|\right)^n. \quad (2.9)$$

As a corollary, we get probabilistic estimates on the number of eigenvalues of  $H_{\omega, \Lambda}$  in intervals.

**Corollary 2.4.** *Fix a finite volume  $\Lambda \subset \mathbb{Z}^d$ . For all  $n \in \mathbb{N}$  and  $I$  a bounded interval, we have*

$$\mathbb{P} \left\{ \mathrm{tr} P_\omega^{(\Lambda)}(I) \geq n \right\} \leq \frac{1}{n!} \left(Q^{(\Lambda)}(|I|) |\Lambda|\right)^n. \quad (2.10)$$

Furthermore, for all bounded intervals  $I_1, \dots, I_n$  we get

$$\begin{aligned} \mathbb{P} \left\{ \operatorname{tr} P_{\omega}^{(\Lambda)}(I_{\sigma_{\omega}(1)}) \geq 1, \operatorname{tr} P_{\omega}^{(\Lambda)}(I_{\sigma_{\omega}(2)}) \geq 2, \dots, \operatorname{tr} P_{\omega}^{(\Lambda)}(I_{\sigma_{\omega}(n)}) \geq n \right\} \\ \leq M(I_1, \dots, I_n) \left( \prod_{k=1}^n Q^{(\Lambda)}(|I_k|) \right) |\Lambda|^n, \end{aligned} \quad (2.11)$$

and, in the special case when  $I_1 \subset I_2 \subset \dots \subset I_n$ , we have

$$\mathbb{P} \left\{ \operatorname{tr} P_{\omega}^{(\Lambda)}(I_1) \geq 1, \operatorname{tr} P_{\omega}^{(\Lambda)}(I_2) \geq 2, \dots, \operatorname{tr} P_{\omega}^{(\Lambda)}(I_n) \geq n \right\} \leq \left( \prod_{k=1}^n Q^{(\Lambda)}(|I_k|) \right) |\Lambda|^n. \quad (2.12)$$

**Remark 2.5.** Given bounded intervals  $I_1$  and  $I_2$ , let  $d(I_1, I_2)$  denote the distance between the two intervals. It follows from (2.12) that

$$\begin{aligned} \mathbb{P} \left\{ \operatorname{tr} P_{\omega}^{(\Lambda)}(I_1) \geq 1 \text{ and } \operatorname{tr} P_{\omega}^{(\Lambda)}(I_2) \geq 1 \right\} \\ \leq \left( \min \left\{ Q^{(\Lambda)}(|I_1|), Q^{(\Lambda)}(|I_2|) \right\} Q^{(\Lambda)}(d(I_1, I_2) + |I_1| + |I_2|) \right) |\Lambda|^2. \end{aligned} \quad (2.13)$$

### 3. APPLICATIONS TO HÖLDER CONTINUOUS DISTRIBUTIONS

The (standard) Anderson model is given by  $H_{\omega}$  as in (1.3), with  $H_0 = -\Delta$ , the centered discrete Laplacian, and  $\omega = \{\omega_j\}_{j \in \mathbb{Z}^d}$  a family of independent identically distributed random variables with joint probability distribution  $\mu$ , which we assume to have no atoms and compact support. Localization for the Anderson model has been well studied, mostly for  $\mu$  with a bounded density  $\rho$ , cf. [FS, FMSS, DLS, SW, DrK, AM, A] and many others, as well as for probability distributions  $\mu$  that are Hölder continuous [CKM, DrK, H, ASFH, GK1], i.e.,  $Q_{\mu}(s) \leq U s^{\alpha}$  for  $s$  small, for some constants  $U$  and  $\alpha \in ]0, 1[$ . If the probability distribution  $\mu$  has a bounded density, Minami's estimate (1.7) was a crucial ingredient in Klein and Molchanov's proof of simplicity of eigenvalues [KIM] and in Klein, Lenoble and Müller derivation of a rigorous form of Mott's formula for the ac-conductivity [KILM]. In this section we show that with (2.3) these proofs extend to the case when  $\mu$  is only Hölder continuous.

**3.1. Multiplicity of the spectrum.** Let  $H_{\omega}$  be a generalized Anderson model as in (1.3), let  $\alpha \in ]0, 1[$ , and assume that the probability distributions  $\mu_j$  are uniformly  $\alpha$ -Hölder continuous, i.e., there is a constant  $U$  and  $s_0 > 0$  such that

$$\sup_{j \in \mathbb{Z}^d} Q_j(s) \leq U s^{\alpha} \quad \text{for all } s \in [0, s_0]. \quad (3.1)$$

In this case we say that  $H_{\omega}$  is an  $\alpha$ -Hölder continuous generalized Anderson model.

We say that the generalized Anderson model  $H_{\omega}$  exhibits Anderson localization in some interval  $I$  if, with probability one,  $H_{\omega}$  has pure point spectrum in  $I$  and the corresponding eigenfunctions decay exponentially. Given  $x \geq 0$ , we let  $[x]$  denote the integer part of  $x$ . Following Klein and Molchanov [KIM], we prove the following result.

**Theorem 3.1.** *Let  $H_{\omega}$  be an  $\alpha$ -Hölder continuous generalized Anderson model. Suppose  $H_{\omega}$  exhibits Anderson localization in some interval  $I$ . Then, with probability one, every eigenvalue of  $H_{\omega}$  in  $I$  has multiplicity  $\leq [\alpha^{-1}]$ . In particular, if  $\alpha > \frac{1}{2}$ , with probability one every eigenvalue of  $H_{\omega}$  in  $I$  is simple.*

**Remark 3.2.** For the standard Anderson model, where  $\mu_j = \mu$  for all  $j \in \mathbb{Z}^d$ , this theorem was originally proved by Simon [S] when  $\mu$  has a bounded density (i.e.,  $\alpha = 1$ ). Our proof is based on a simple proof later provided by Klein and Molchanov [KLM], based on Minami's estimate (1.7). For singular measures (i.e.,  $\alpha < 1$ ), the best previously known result for the standard Anderson model is the finite multiplicity of the eigenvalues [CoH, GK3]; uniform boundedness of the multiplicity was not previously known. Thus Theorem 3.1 improves on both results.

*Proof of Theorem 3.1.* We proceed as in [KLM]. We call  $\varphi \in \ell^2(\mathbb{Z}^d)$   $\alpha$ -fast decaying if it has  $\beta$ -decay, that is,  $|\varphi(x)| \leq C_\varphi (1 + |x|)^{-\beta}$  for some  $C_\varphi < \infty$ , with

$$\beta > \left( \frac{1}{2} + \left( \alpha - ([\alpha^{-1}] + 1)^{-1} \right)^{-1} \right) d. \quad (3.2)$$

To prove the theorem, we will show that, with probability one, an  $\alpha$ -Hölder continuous generalized Anderson model  $H_\omega$  cannot have an eigenvalue with  $[\alpha^{-1}] + 1$  linearly independent  $\alpha$ -fast decaying eigenfunctions.

We set  $N = [\alpha^{-1}] + 1$ , so that  $N\alpha > 1$ . For a given open interval  $I$  we pick

$$q > \frac{Nd}{N\alpha - 1} = \left( \alpha - ([\alpha^{-1}] + 1)^{-1} \right)^{-1} d. \quad (3.3)$$

Given a scale  $L > 0$ , we let  $\Lambda_L$  denote the cube of side  $L$  centered at 0, and cover  $I$  by  $2 \left( \lceil \frac{L^q}{2} |I| \rceil + 1 \right) \leq L^q |I| + 2$  intervals of length  $2L^{-q}$ , in such a way that any subinterval  $J \subset I$  with length  $|J| \leq L^{-q}$  will be contained in one of these intervals. We consider the event  $\mathcal{B}_{L,I,q}$ , which occurs if there exists an interval  $J \subset I$  with  $|J| \leq L^{-q}$  such that  $\text{tr } P_\omega^{(\Lambda_L)}(J) \geq N$ . Its probability can be estimated, using (2.10) and (3.1), by

$$\mathbb{P}\{\mathcal{B}_{L,I,q}\} \leq \frac{1}{N!} (L^q |I| + 2) (U (2L^{-q})^\alpha L^d)^N \leq (|I| + 1) \frac{(2^\alpha U)^N}{N!} L^{-(N\alpha - 1)q + Nd}. \quad (3.4)$$

In view of (3.3), taking scales  $L_k = 2^k$ , it follows from the Borel-Cantelli Lemma that, with probability one, the event  $\mathcal{B}_{L_k,I,q}$  eventually does not occur.

Now, suppose that for some  $\omega$  there exists  $E \in I$  which is an eigenvalue of  $H_\omega$  with  $N$  linearly independent  $\alpha$ -fast decaying eigenfunctions, so they all have  $\beta$ -decay for some  $\beta$  as in (3.2). It follows, as in [KLM, Lemma 1], that for  $L$  large enough the finite volume operator  $H_{\omega,\Lambda_L}$  has at least  $N$  eigenvalues in the interval  $J_{E,L} = [E - \varepsilon_L, E + \varepsilon_L]$ , where  $\varepsilon_L = CL^{-\beta + \frac{d}{2}}$  for an appropriate constant  $C$  independent of  $L$ . In view of (3.2), we can pick  $q$ , satisfying (3.3), such that  $\beta - \frac{d}{2} > q$ , and hence  $\varepsilon_L < L^{-q}$  for all large  $L$ . But with probability one this is impossible since the event  $\mathcal{B}_{L_k,I,q}$  does not occur for large  $L_k$ .  $\square$

**3.2. Generalized Mott's formula.** Let  $\alpha \in ]0, 1[$ , and consider the  $\alpha$ -Hölder continuous Anderson model, that is, the Anderson model  $H_\omega$  with the single-site probability distribution  $\mu$  an  $\alpha$ -Hölder continuous measure:

$$Q_\mu(s) \leq Us^\alpha \quad \text{for all } s \in [0, s_0], \quad (3.5)$$

where  $U$  and  $s_0 > 0$  are constants. The fractional moment method can be applied to such measures, leading to exponential decay of the expectation of some fractional power of the Green's function [H, ASFH]. We may then define the region of complete localization  $\Xi^{\text{CL}}$ , introduced in [GK2, GK3], as in [KILM, Definition 2.1]. However, [KILM, Eqs. (4.1), (4.3) and (4.4)] have not been derived from the fractional moment method for  $\mu$  with compact support satisfying only the condition

(3.5). ([H, ASFH, Appendix A] assumes that  $\mu$  has a bounded density in the derivation of such estimates.) But in this region of complete localization we can always perform a multiscale analysis as in [GK1] with only hypothesis (3.5), and get the estimates [KILM, Eqs. (4.1), (4.3)] with sub-exponential decay [GK1, GK3], and hence conclude that [KILM, Assumption 3.1 and Eq. (4.4)] are satisfied. Thus, given a Fermi energy  $E_F \in \Xi^{\text{CL}}$ , the analysis in [KILM] applies and we may define the *average in phase conductivity*  $\bar{\sigma}_{E_F}^{\text{in}}(\nu)$  as in [KILM, Eq. (2.17)]. We have the following extension of [KILM, Theorem 2.3]. Note that we get  $\bar{\sigma}_{E_F}^{\text{in}}(\nu) \leq C\nu^{2\alpha} (\log \frac{1}{\nu})^{d+2}$  for small  $\nu$ , consistent with  $C\nu^2 (\log \frac{1}{\nu})^{d+2}$  for  $\alpha = 1$  as in [KILM].

**Theorem 3.3.** *Given  $\alpha \in ]0, 1[$ , let  $H_\omega$  be a  $\alpha$ -Hölder continuous Anderson model, with  $\mu$  as in (3.5). Consider a Fermi energy in its region of complete localization:  $E_F \in \Xi^{\text{CL}}$ . Then*

$$\limsup_{\nu \downarrow 0} \frac{\bar{\sigma}_{E_F}^{\text{in}}(\nu)}{\nu^{2\alpha} (\log \frac{1}{\nu})^{d+2}} \leq B \ell_{E_F}^{d+2}, \quad (3.6)$$

where  $\ell_{E_F}$  is given in [KILM, Eq. (2.3)], and the constant  $B$  depends only on  $d, U$  and  $\alpha$ .

*Proof.* The proof of [KILM, Theorem 2.3] applies, with modifications due to the use of (1.5) and (2.3) with  $Q(s) = Q_\mu(s)$  as in (3.5). The modifications are as follows (we use the notation of [KILM]):

- (1) We systematically use (1.2) instead of [KILM, Eq. (4.5)]. In particular, [KILM, Eq. (4.10)] becomes

$$\langle \langle Y_{E_F}, \chi_B(\mathcal{H}_L) Y_{E_F} \rangle \rangle \leq W_\beta Q(|B|)^\beta \quad \text{for all Borel sets } B \subset \mathbb{R}. \quad (3.7)$$

To derive this estimate, we use the sub-exponential decay of the Fermi projection given in [GK3, Theorem 3], i.e., we use [KILM, Eq. (4.1)] but with sub-exponential decay. This can be done because we only use summability of this decay. As a consequence, [KILM, Eq. (4.6)] becomes

$$\Psi_{E_F}(B_+ \times B_-) \leq W_\beta (\min \{Q(|B_+|), Q(|B_-|)\})^\beta. \quad (3.8)$$

- (2) Using (2.3), [KILM, Eq. (4.51)] becomes

$$\mathbb{E} \{ \langle \delta_0, F_{-,L} X_1 F_{+,L} X_1 F_{-,L} \delta_0 \rangle \} \leq \frac{1}{4} Q(|J|)^2 L^{d+2}. \quad (3.9)$$

- (3) In [KILM, Lemma 4.6], we cannot use the estimate [KILM, Eq. (4.28)]. But proceeding as in the proof of [KILM, Eq. (4.30)], we can replace it by

$$\mathbb{E} \{ |\langle \delta_x, F_\pm \delta_y \rangle|^p \} \leq C_I \{ \{ f_\pm \} \}_2 e^{-\frac{1}{2}|x-y|} \quad \text{for all } p \in [1, \infty[ \text{ and } x, y \in \mathbb{Z}^d. \quad (3.10)$$

As a consequence, the right hand side of [KILM, Eq. (4.27)] becomes

$$C \left( \left( \{ \{ f_+ \} \}_2 \{ \{ f_- \} \}_3 \{ \{ f_- \} \}_4 \right)^{\frac{1}{3}} + \left( \{ \{ f_- \} \}_3^2 \{ \{ f_+ \} \}_4 \right)^{\frac{1}{3}} \right) L^{\frac{4}{3}d} e^{-\frac{1}{12\ell}L}. \quad (3.11)$$

- (4) Using (2.3) instead of [KILM, Eq. (4.47)], and taking into account the above modifications, [KILM, Eq. (4.61)] becomes

$$\begin{aligned} \Psi_{E_F}(I_+ \times I_-) &\leq \frac{1}{4} Q(2\nu)^2 L^{d+2} + C' \nu^{-15} L^{\frac{4}{3}d} e^{-\frac{1}{12\ell}L} + 4W_{\frac{1}{2}} Q(\nu^4)^{\frac{1}{2}} \\ &\leq 4^{\alpha-1} U^2 \nu^{2\alpha} L^{d+2} + C' \nu^{-15} L^{\frac{4}{3}d} e^{-\frac{1}{12\ell}L} + 4U^{\frac{1}{2}} W_{\frac{1}{2}} \nu^{2\alpha}, \end{aligned} \quad (3.12)$$

where we used (3.5).

- (5) As in [KILM, Eq. (4.62)], we choose  $L = A\ell \log \frac{1}{\nu}$ , where  $A$  is some suitable constant, depending on  $d$ ,  $\alpha$  and  $U$ , such that, similarly to [KILM, Eq. (4.63)], we get

$$\Psi_{E_F}(I_+ \times I_-) \leq B\ell^{d+2}\nu^{2\alpha} \left(\log \frac{1}{\nu}\right)^{d+2} + C''\nu^{2\alpha}, \quad (3.13)$$

where  $B$  and  $C''$  are constants, with  $B$  depending only on  $d$ ,  $\alpha$  and  $U$ , from which (3.6) follows. □

#### 4. PROOFS OF THE MAIN RESULTS

Our proofs are based on the fundamental spectral averaging estimate (1.2) and the following lemma.

**Lemma 4.1.** *Consider the self-adjoint operator  $H_s = H_0 + s\Pi_\varphi$  on the Hilbert space  $\mathcal{H}$ , where  $H_0$  is a self-adjoint operator on  $\mathcal{H}$ ,  $\varphi \in \mathcal{H}$  with  $\|\varphi\| = 1$ , and  $s \in \mathbb{R}$ . Let  $P_s(J) = \chi_J(H_s)$  for an interval  $J$ , and suppose  $\text{tr } P_0(]-\infty, c]) < \infty$  for all  $c \in \mathbb{R}$ . Then, given  $a, b \in \mathbb{R}$  with  $a < b$ , we have*

$$\text{tr } P_s(]a, b]) \leq 1 + \text{tr } P_t(]a, b]) \quad \text{for all } 0 \leq s \leq t. \quad (4.1)$$

*Proof.* Let  $0 \leq s \leq t$ . We recall that for any  $c \in \mathbb{R}$  we always have

$$0 \leq \text{tr } P_s(]-\infty, c]) - \text{tr } P_t(]-\infty, c]) \leq 1, \quad (4.2)$$

the last inequality being a consequence of the min-max principle applied to rank one perturbations, e.g. [Ki, Lemma 5.22]. Thus

$$\begin{aligned} \text{tr } P_s(]a, b]) &= \text{tr } P_s(]-\infty, b]) - \text{tr } P_s(]-\infty, a]) \\ &\leq \text{tr } P_s(]-\infty, b]) - \text{tr } P_t(]-\infty, a]) \\ &= \text{tr } P_s(]-\infty, b]) - \text{tr } P_t(]-\infty, b]) + \text{tr } P_t(]a, b]) \\ &\leq 1 + \text{tr } P_t(]a, b]). \end{aligned} \quad (4.3)$$

□

When we fix  $\Lambda$ , we may drop it from the notation. Given  $\omega \in \mathbb{R}^d$ , we write  $P(I) = P_\omega^{(\Lambda)}(I) = \chi_I(H_{\omega, \Lambda})$ . Given  $j \in \Lambda$ , we write  $\omega = (\omega_j^\perp, \omega_j)$  and  $P_{\omega_j=s}(I) = P_{(\omega_j^\perp, s)}(I)$  when we want to make explicit the value of  $\omega_j$ . We also write  $P_{\omega_j \rightarrow s}(I)$  to denote that  $\omega_j$  was replaced by  $s$ .

Since we assumed that the measures  $\mu_j$  have no atoms, it follows from (1.5) that  $\mathbb{E}_\omega \left\{ \text{tr } P_\omega^{(\Lambda)}(\{c\}) \right\} = 0$  for any  $c \in \mathbb{R}$ . Thus it does not matter if the intervals are open or closed at the endpoints, so in the proofs we may take all intervals to be of the form  $]a, b]$ , which allows the use of Lemma 4.1.

Theorem 2.1 is a particular case of Theorem 2.3, but in order to illustrate the simplicity of our approach we first give a proof of Theorem 2.1 and then prove the general case.



*Proof of Theorem 2.1.* Fix  $\Lambda \subset \mathbb{Z}^d$  and let  $I_1, I_2$  be bounded intervals. Using Lemma 4.1, for  $\tau_j \geq \omega_j$  we always have

$$\begin{aligned} (\operatorname{tr} P_{\omega}(I_1)) (\operatorname{tr} P_{\omega}(I_2) - 1) &= \sum_{j \in \Lambda} \{ \langle \delta_j, P_{\omega}(I_1) \delta_j \rangle (\operatorname{tr} P_{\omega}(I_2) - 1) \} \\ &\leq \sum_{j \in \Lambda} \left\{ \left\langle \delta_j, P_{(\omega_j^{\pm}, \omega_j)}(I_1) \delta_j \right\rangle \left( \operatorname{tr} P_{(\omega_j^{\pm}, \tau_j)}(I_2) \right) \right\}. \end{aligned} \quad (4.4)$$

We now take  $\tau_j \geq \max \operatorname{supp} \mu_j$  for all  $j \in \Lambda$ , and average over the random variables  $\omega = \{\omega_j\}_{j \in \mathbb{Z}^d}$ , where each  $\omega_j$  has the probability distribution  $\mu_j$ . Using (1.2), we get

$$\begin{aligned} \mathbb{E}_{\omega} \{ (\operatorname{tr} P_{\omega}(I_1)) (\operatorname{tr} P_{\omega}(I_2) - 1) \} & \\ &\leq \sum_{j \in \Lambda} \mathbb{E}_{\omega_j^{\pm}} \left\{ \left( \operatorname{tr} P_{(\omega_j^{\pm}, \tau_j)}(I_2) \right) \left( \mathbb{E}_{\omega_j} \left\{ \left\langle \delta_j, P_{(\omega_j^{\pm}, \omega_j)}(I_1) \delta_j \right\rangle \right\} \right) \right\} \\ &\leq Q_{\Lambda} (|I_1|) \sum_{j \in \Lambda} \mathbb{E}_{\omega_j^{\pm}} \left\{ \operatorname{tr} P_{(\omega_j^{\pm}, \tau_j)}(I_2) \right\}. \end{aligned} \quad (4.5)$$

This holds for all  $\tau_j \geq \max \operatorname{supp} \mu_j$ ,  $j \in \Lambda$ , so we now take  $\tau_j = \max \operatorname{supp} \mu_j + \tilde{\omega}_j$ , where  $\tilde{\omega} = \{\tilde{\omega}_j\}_{j \in \mathbb{Z}^d}$  and  $\omega = \{\omega_j\}_{j \in \mathbb{Z}^d}$  are two independent, identically distributed collections of random variables, and average over these random variables. We get

$$\begin{aligned} \mathbb{E}_{\omega} \{ (\operatorname{tr} P_{\omega}(I_1)) (\operatorname{tr} P_{\omega}(I_2) - 1) \} &= \mathbb{E}_{\tilde{\omega}} \{ \mathbb{E}_{\omega} \{ (\operatorname{tr} P_{\omega}(I_1)) (\operatorname{tr} P_{\omega}(I_2) - 1) \} \} \\ &\leq Q_{\Lambda} (|I_1|) \sum_{j \in \Lambda} \mathbb{E}_{(\omega_j^{\pm}, \tilde{\omega}_j)} (\operatorname{tr} P_{(\omega_j^{\pm}, \tau_j)}(I_2)) \leq Q_{\Lambda} (|I_1|) Q_{\Lambda} (|I_2|) |\Lambda|^2, \end{aligned} \quad (4.6)$$

where we used the Wegner estimate (1.5).

The estimates (2.2) and (2.3) follow immediately from (4.6). To get (2.1), we use (4.6) and the obvious estimate

$$\begin{aligned} (\operatorname{tr} P_{\omega}(I_1)) (\operatorname{tr} P_{\omega}(I_2)) - \min \{ \operatorname{tr}_{\omega} P(I_1), \operatorname{tr} P_{\omega}(I_2) \} & \\ &\leq (\operatorname{tr} P_{\omega}(I_1)) (\operatorname{tr} P_{\omega}(I_2) - 1) + (\operatorname{tr} P_{\omega}(I_2)) (\operatorname{tr} P_{\omega}(I_1) - 1). \end{aligned} \quad (4.7)$$

□

We now turn to the general case.

*Proof of Theorem 2.3.* Although (2.9) is a particular case of (2.7), it has a simpler proof, so will prove it first. We fix the bounded interval  $I$  and proceed by induction on  $n$ . The case  $n = 1$ , is just Wegner's inequality (1.5). Let us assume that (2.9) holds for  $n$ , for all possible probability distributions  $\mu_j$  with compact support. Then, given  $j \in \Lambda$  and  $\tau_j \geq \max \operatorname{supp} \mu_j$ , we have, using (4.1), that for all  $k = 1, 2, \dots, n$ ,

$$\operatorname{tr} P_{\omega}(I) - k \leq 1 + \operatorname{tr} P_{(\omega_j^{\pm}, \tau_j)}(I) - k = \operatorname{tr} P_{(\omega_j^{\pm}, \tau_j)}(I) - (k - 1). \quad (4.8)$$

Note that  $\left( \operatorname{tr} P_{(\omega_j^{\pm}, \tau_j)}(I) \right) (\operatorname{tr} P_{(\omega_j^{\pm}, \tau_j)}(I) - 1) \cdots (\operatorname{tr} P_{(\omega_j^{\pm}, \tau_j)}(I) - (n - 1)) \geq 0$ . Since either  $\operatorname{tr} P_{\omega}(I) (\operatorname{tr} P_{\omega}(I) - 1) \cdots (\operatorname{tr} P_{\omega}(I) - n) = 0$  or  $\operatorname{tr} P_{\omega}(I) - k > 0$  for  $k = 0, 1, \dots, n$ , it follows that we always have

$$\begin{aligned} (\operatorname{tr} P_{\omega}(I)) (\operatorname{tr} P_{\omega}(I) - 1) \cdots (\operatorname{tr} P_{\omega}(I) - n) & \\ &\leq \sum_{j \in \Lambda} \langle \delta_j, P_{\omega}(I) \delta_j \rangle \left( \operatorname{tr} P_{(\omega_j^{\pm}, \tau_j)}(I) \right) (\operatorname{tr} P_{(\omega_j^{\pm}, \tau_j)}(I) - 1) \cdots (\operatorname{tr} P_{(\omega_j^{\pm}, \tau_j)}(I) - (n - 1)). \end{aligned} \quad (4.9)$$

Using (1.2), we have

$$\begin{aligned} & \mathbb{E} \left\{ \langle \delta_j, P_\omega(I) \delta_j \rangle \operatorname{tr} P_{(\omega_j^\perp, \tau_j)}(I) (\operatorname{tr} P_{(\omega_j^\perp, \tau_j)}(I) - 1) \cdots (\operatorname{tr} P_{(\omega_j^\perp, \tau_j)}(I) - (n-1)) \right\} \\ &= \mathbb{E}_{\omega_j^\perp} \left\{ \mathbb{E}_{\omega_j} \left\{ \langle \delta_j, P_\omega(I) \delta_j \rangle \right\} (\operatorname{tr} P_{(\omega_j^\perp, \tau_j)}(I)) (\operatorname{tr} P_{(\omega_j^\perp, \tau_j)}(I) - 1) \times \right. \\ & \quad \left. \cdots (\operatorname{tr} P_{(\omega_j^\perp, \tau_j)}(I) - (n-1)) \right\} \\ & \leq Q_\Lambda(|I|) \mathbb{E}_{\omega_j^\perp} \left\{ (\operatorname{tr} P_{(\omega_j^\perp, \tau_j)}(I)) (\operatorname{tr} P_{(\omega_j^\perp, \tau_j)}(I) - 1) \cdots (\operatorname{tr} P_{(\omega_j^\perp, \tau_j)}(I) - (n-1)) \right\}. \end{aligned} \quad (4.10)$$

We now take  $\tau = \{\tau_j = a_j + \tilde{\omega}_j\}_{j \in \Lambda}$ , where  $\tilde{\omega} = \{\tilde{\omega}_j\}_{j \in \Lambda}$  are independent random variables, independent of  $\omega$ , such that  $\tilde{\omega}_j$  has  $\mu_j$  for probability distribution, and  $a_j = \max \operatorname{supp} \mu_j$ . Using (4.9) and (4.10), plus the induction hypothesis, we get

$$\begin{aligned} & \mathbb{E}_\omega \left\{ \operatorname{tr} P_\omega(I) (\operatorname{tr} P_\omega(I) - 1) \cdots (\operatorname{tr} P_\omega(I) - n) \right\} \\ &= \mathbb{E}_{(\omega, \tau)} \left\{ \operatorname{tr} P_\omega(I) (\operatorname{tr} P_\omega(I) - 1) \cdots (\operatorname{tr} P_\omega(I) - n) \right\} \\ & \leq Q_\Lambda(|I|) \sum_{j \in \Lambda} \mathbb{E}_{(\omega_j^\perp, \tau_j)} \left\{ (\operatorname{tr} P_{(\omega_j^\perp, \tau_j)}(I)) (\operatorname{tr} P_{(\omega_j^\perp, \tau_j)}(I) - 1) \times \right. \\ & \quad \left. \cdots (\operatorname{tr} P_{(\omega_j^\perp, \tau_j)}(I) - (n-1)) \right\} \\ & \leq Q_\Lambda(|I|) \sum_j (Q_\Lambda(|I|) |\Lambda|)^n = (Q_\Lambda(|I|) |\Lambda|)^{n+1}. \end{aligned} \quad (4.11)$$

We now turn to the proof of (2.7). The case  $n = 1$  is just (1.5), and  $n = 2$  is (2.1), so we assume  $n \geq 3$ . Let  $I_1, I_2, \dots, I_n$  be bounded intervals. For a fixed  $\omega$ , we have (2.5) and (2.6). Let us suppose

$$\Xi_\omega := \left( \operatorname{tr} P_\omega^{(\Lambda)}(I_{\sigma_\omega(1)}) \right) \left( \operatorname{tr} P_\omega^{(\Lambda)}(I_{\sigma_\omega(2)}) - 1 \right) \cdots \left( \operatorname{tr} P_\omega^{(\Lambda)}(I_{\sigma_\omega(n)}) - (n-1) \right) > 0, \quad (4.12)$$

and note that in this case we must have

$$\operatorname{tr} P_\omega(I_{\sigma_\omega(k)}) - k + 1 \geq 1, \quad \text{i.e.,} \quad \operatorname{tr} P_\omega(I_{\sigma_\omega(k)}) \geq k, \quad \text{for all } k = 1, \dots, n. \quad (4.13)$$

Then, using Lemma 4.1 repeatedly, we get

$$\begin{aligned} \Xi_\omega & \leq \sum_{j_1 \in \Lambda} \left\{ \langle \delta_{j_1}, P_\omega(I_{\sigma_\omega(1)}) \delta_{j_1} \rangle (\operatorname{tr} P_{\omega^{(j_1)}}(I_{\sigma_\omega(2)})) \cdots (\operatorname{tr} P_{\omega^{(j_1)}}(I_{\sigma_\omega(n)}) - n) \right\} \\ & \leq \dots \dots \dots \\ & \leq \sum_{j_1, j_2, \dots, j_{n-1} \in \Lambda} \left\{ \langle \delta_{j_1}, P_\omega(I_{\sigma_\omega(1)}) \delta_{j_1} \rangle \langle \delta_{j_2}, P_{\omega^{(j_1)}}(I_{\sigma_\omega(2)}) \delta_{j_2} \rangle \cdots \times \right. \\ & \quad \left. \langle \delta_{j_{n-1}}, P_{\omega^{(j_1, j_2, \dots, j_{n-2})}}(I_{\sigma_\omega(n-1)}) \delta_{j_{n-1}} \rangle (\operatorname{tr} P_{\omega^{(j_1, j_2, \dots, j_{n-1})}}(I_{\sigma_\omega(n)})) \right\}, \end{aligned} \quad (4.14)$$

where  $\omega^{(j_1)}$  is  $\omega$  with  $\omega_{j_1} \rightarrow \tau_{j_1}^{(1)}$ ,  $\omega^{(j_1, j_2)}$  is  $\omega^{(j_1)}$  with  $\omega_{j_2} \rightarrow \tau_{j_2}^{(2)}$ . . . . ,  $\omega^{(j_1, j_2, \dots, j_{n-1})}$  is  $\omega^{(j_1, j_2, \dots, j_{n-2})}$  with  $\omega_{j_{n-1}} \rightarrow \tau_{j_{n-1}}^{(n-1)}$ . To be able to apply Lemma 4.1 we must have

$$\omega_{j_1} \leq \tau_{j_1}^{(1)}, \omega_{(j_2)} \leq \tau_{j_2}^{(2)}, \dots, \omega_{(j_{n-1})} \leq \tau_{j_{n-1}}^{(n-1)}, \quad j_k \in \Lambda, \quad k = 1, 2, \dots, n-1. \quad (4.15)$$

We then take

$$\tau = \left\{ \tau_{j_k}^{(k)} = a_{j_k}^{(k)} + \omega_{j_k}^{(k)}; \quad j_k \in \Lambda, \quad k = 1, 2, \dots, n-1 \right\},$$

where  $\widehat{\omega} = \{\omega_{j_k}^{(k)}; j_k \in \Lambda, k = 1, 2, \dots, n-1\}$  are independent random variables, independent of  $\omega$ , such that  $\omega_{j_k}^{(k)}$  has  $\mu_{j_k}$  for probability distribution, and the real numbers  $a_{j_k}^{(k)}$  are chosen such that (4.15) holds  $\mathbb{P}_{(\omega, \widehat{\omega})}$ -almost surely.

Since the last expression in (4.14) is obviously nonnegative, it follows that (4.14) holds also when  $\Xi_\omega = 0$ , and hence it holds  $\mathbb{P}_{(\omega, \widehat{\omega})}$ -almost surely.

Given  $\sigma \in \mathcal{S}_n$ , let

$$\begin{aligned} \Xi_{\omega, \tau, \sigma} := & \sum_{j_1, j_2, \dots, j_{n-1} \in \Lambda} \left\{ \langle \delta_{j_1}, P_\omega(I_{\sigma(1)}) \delta_{j_1} \rangle \langle \delta_{j_2}, P_{\omega^{(j_1)}}(I_{\sigma(2)}) \delta_{j_2} \rangle \cdots \times \right. \\ & \left. \langle \delta_{j_{n-1}}, P_{\omega^{(j_1, j_2, \dots, j_{n-2})}}(I_{\sigma_{\omega(n-1)}}) \delta_{j_{n-1}} \rangle \left( \text{tr } P_{\omega^{(j_1, j_2, \dots, j_{n-1})}}(I_{\sigma_{\omega(n)}}) \right) \right\}. \end{aligned} \quad (4.16)$$

It follows that  $\mathbb{P}_{(\omega, \widehat{\omega})}$ -almost surely we have

$$\Xi_\omega \leq \sum_{\sigma \in \mathcal{S}_n(I_1, \dots, I_n)} \Xi_{\omega, \tau, \sigma}, \quad (4.17)$$

and hence

$$\mathbb{E}_\omega \{\Xi_\omega\} = \mathbb{E}_{(\omega, \widehat{\omega})} \{\Xi_\omega\} \leq \sum_{\sigma \in \mathcal{S}_n(I_1, \dots, I_n)} \mathbb{E}_{(\omega, \widehat{\omega})} \{\Xi_{\omega, \tau, \sigma}\}. \quad (4.18)$$

By performing the integrations in the right order, using (1.2)  $n-1$  times, and then using the Wegner estimate (1.5), we get

$$\mathbb{E}_{(\omega, \widehat{\omega})} \{\Xi_{\omega, \tau, \sigma}\} \leq \left( \prod_{k=1}^n Q^{(\Lambda)}(|I_k|) \right) |\Lambda|^n. \quad (4.19)$$

Since  $M(I_1, \dots, I_n) = |\mathcal{S}_n(I_1, \dots, I_n)|$ , the estimate (2.7) follows.  $\square$

*Proof of Corollary 2.4.* The estimate (2.10) follows from (2.9) and the inequality

$$\begin{aligned} \mathbb{P} \{ \text{tr } P_\omega(I) \geq n \} & \leq \mathbb{P} \{ (\text{tr } P_\omega(I)) (\text{tr } P_\omega(I) - 1) \cdots (\text{tr } P_\omega(I) - (n-1)) \geq n! \} \\ & \leq \frac{1}{n!} \mathbb{E} \{ (\text{tr } P_\omega(I)) (\text{tr } P_\omega(I) - 1) \cdots (\text{tr } P_\omega(I) - (n-1)) \}. \end{aligned} \quad (4.20)$$

To obtain (2.11), we use (2.7) with

$$\begin{aligned} \mathbb{P} \left\{ \text{tr } P_\omega^{(\Lambda)}(I_{\sigma_\omega(1)}) \geq 1, \text{tr } P_\omega^{(\Lambda)}(I_{\sigma_\omega(2)}) \geq 2, \dots, \text{tr } P_\omega^{(\Lambda)}(I_{\sigma_\omega(n)}) \geq n \right\} \\ \leq \mathbb{E} \left\{ \left( \text{tr } P_\omega^{(\Lambda)}(I_{\sigma_\omega(1)}) \right) \left( \text{tr } P_\omega^{(\Lambda)}(I_{\sigma_\omega(2)}) - 1 \right) \cdots \left( \text{tr } P_\omega^{(\Lambda)}(I_{\sigma_\omega(n)}) - (n-1) \right) \right\}. \end{aligned} \quad (4.21)$$

Similarly, (2.12) follows from (2.8).  $\square$

## APPENDIX A. THE FUNDAMENTAL SPECTRAL AVERAGING ESTIMATE

For the reader convenience we present a proof of the fundamental spectral averaging result (1.2) Recall  $H_\omega = H_0 + \omega \Pi_\varphi$ , where  $\omega$  is a random variable with probability distribution  $\mu$ . Given  $z$  with  $\Im z > 0$ , we have, as in [CKM, Proof of Lemma 6.1], that

$$\langle \varphi, (H_\omega - z)^{-1} \varphi \rangle = \left( \langle \varphi, (H_0 - z)^{-1} \varphi \rangle^{-1} + \omega \right)^{-1}. \quad (A.1)$$

**A.1. The probability distribution  $\mu$  has a bounded density  $\rho$ .** In this case, we use

$$\int_{\mathbb{R}} d\omega \Im \langle \varphi, (H_\omega - z)^{-1} \varphi \rangle = \pi, \quad (\text{A.2})$$

a consequence of (A.1) (cf. [CKM, Proof of Lemma 6.1]). It then follows from Stone's formula (cf. [RS, Theorem VII.13]) that

$$\int_{\mathbb{R}} d\omega \langle \varphi, \frac{1}{2} \{P_\omega([a, b]) + P_\omega(]a, b[)\} \varphi \rangle \leq (b - a). \quad (\text{A.3})$$

In particular,  $\int_{\mathbb{R}} d\omega \langle \varphi, P_\omega(\{c\}) \varphi \rangle = 0$ , and hence for any bounded interval  $I$  we get

$$\int_{\mathbb{R}} d\omega \langle \varphi, P_\omega(I) \varphi \rangle \leq |I|. \quad (\text{A.4})$$

Since  $\mu$  has a bounded density  $\rho$ , we get

$$\int_{\mathbb{R}} d\mu(\omega) \langle \varphi, P_\omega(I) \varphi \rangle = \int_{\mathbb{R}} d\omega \rho(\omega) \langle \varphi, P_\omega(I) \varphi \rangle \leq \|\rho\|_\infty |I|. \quad (\text{A.5})$$

**Remark A.1.** *The reader may notice that [CKM] has an extra factor of  $\pi$  in the right hand side of (A.5); the difference comes from using Stone's formula instead of the simple estimate (A.10). Since in Theorem 2.1 we obtain (2.3) as a consequence of (A.5), and (1.7) is a particular case of (2.3), we do not have the factor of  $\pi^2$  in the right hand side (1.7): the estimate is just  $(\rho_\infty^{(\Lambda)} |I| |\Lambda|)^2$ .*

**A.2. Arbitrary probability distribution  $\mu$ .** We consider an interval  $I = [E - \varepsilon, E + \varepsilon]$ ,  $\varepsilon > 0$ , and set  $z = E + i\varepsilon$  and  $R_0(z) = (H_0 - z)^{-1}$ . Given  $\kappa > 0$ , we define real numbers  $a$  and  $b$  by

$$a - ib = \frac{\kappa}{2\varepsilon} \langle \varphi, R_0(z) \varphi \rangle^{-1}, \quad (\text{A.6})$$

and note that we always have

$$\frac{2}{\kappa} b = \frac{\frac{1}{\varepsilon} \Im \langle \varphi, R_0(z) \varphi \rangle}{|\langle \varphi, R_0(z) \varphi \rangle|^2} = \frac{\|R_0(z) \varphi\|^2}{|\langle \varphi, R_0(z) \varphi \rangle|^2} \geq 1. \quad (\text{A.7})$$

From (A.6) and (A.1) we get,

$$\varepsilon \Im \langle \varphi, (H_\omega - z)^{-1} \varphi \rangle = \frac{\kappa}{2} \frac{b}{(a + \frac{\kappa}{2\varepsilon} \omega)^2 + b^2}. \quad (\text{A.8})$$

Proceeding as in [CoHK2], and using (A.8), [CoHK2, Lemma 3.1] and (A.7), we get

$$\begin{aligned} \varepsilon \int d\mu(\omega) \Im \langle \varphi, (H_\omega - z)^{-1} \varphi \rangle &= \frac{\kappa}{2} \sum_{n \in \mathbb{Z}} \int_{[n \frac{2\varepsilon}{\kappa}, (n+1) \frac{2\varepsilon}{\kappa}[} d\mu(\omega) \frac{b}{(a + \frac{\kappa}{2\varepsilon} \omega)^2 + b^2} \\ &\leq \frac{\kappa}{2} S_\mu \left( \frac{2\varepsilon}{\kappa} \right) \sum_{n \in \mathbb{Z}} \sup_{y \in [0, 1[} \frac{b}{(a + n + y)^2 + b^2} \leq \frac{\kappa}{2} \pi \left(1 + \frac{1}{b}\right) S_\mu \left( \frac{2\varepsilon}{\kappa} \right) \leq \pi \left(1 + \frac{\kappa}{2}\right) S_\mu \left( \frac{2\varepsilon}{\kappa} \right). \end{aligned} \quad (\text{A.9})$$

We may now get (1.2) in two ways. Using the simple inequality

$$P_\omega(I) \leq 2\varepsilon \Im (H_\omega - z)^{-1}, \quad (\text{A.10})$$

(1.2) follows immediately from (A.9) with

$$Q_\mu(|I|) = \inf_{\kappa > 0} \left( \pi \left(2 + \kappa\right) S_\mu \left( \frac{1}{\kappa} |I| \right) \right) \leq 3\pi S_\mu(|I|). \quad (\text{A.11})$$

We can improve the constant slightly by using the more sophisticated inequality given in [CoHK2, Eq. (3.1)], that is,

$$P_\omega(I) \leq \frac{4}{\pi} \int_I d\lambda \Im(H_\omega - \lambda - i|I|)^{-1}, \quad (\text{A.12})$$

together with (A.9) with  $\kappa = 2$ , getting (1.2) with

$$Q_\mu(|I|) = \inf_{\kappa > 0} (4(1 + \kappa)S_\mu(\frac{1}{\kappa}|I|)) \leq 8S_\mu(|I|). \quad (\text{A.13})$$

*Acknowledgement.* The authors thank E. Kritchovski for pointing to them the use of Stone's formula in (A.3)

## REFERENCES

- [A] Aizenman, M.: Localization at weak disorder: some elementary bounds. *Rev. Math. Phys.* **6**, 1163-1182 (1994)
- [AM] Aizenman, M., Molchanov, S.: Localization at large disorder and extreme energies: an elementary derivation. *Commun. Math. Phys.* **157**, 245-278 (1993)
- [ASFH] Aizenman, M., Schenker, J., Friedrich, R., Hundertmark, D.: Finite volume fractional-moment criteria for Anderson localization. *Commun. Math. Phys.* **224**, 219-253 (2001)
- [BHS] Bellissard, J., Hislop, P., Stolz, G.: Correlation estimates in the Anderson model. *J. Stat. Phys.* **129**, 649-662 (2007)
- [CKM] Carmona, R., Klein, A., Martinelli, F.: Anderson localization for Bernoulli and other singular potentials. *Commun. Math. Phys.* **108**, 41-66 (1987)
- [CoH] Combes, J.M., Hislop, P.D.: Localization for some continuous, random Hamiltonians in d-dimension. *J. Funct. Anal.* **124**, 149-180 (1994)
- [CoHK] Combes, J.-M., Hislop, P. D., Klopp, F.: Regularity properties for the density of states of random Schrödinger operators. *Waves in periodic and random media* (South Hadley, MA, 2002), 15-24, *Contemp. Math.*, 339, 2003.
- [CoHK2] Combes, J.M., Hislop, P.D., Klopp, F.: Optimal Wegner estimate and its application to the global continuity of the integrated density of states for random Schrödinger operators. *Duke Math. J.* **140**, 469-498 (2007)
- [DLS] Delyon, F., Lévy, Y., Souillard, B.: Anderson localization for multidimensional systems at large disorder or large energy. *Comm. Math. Phys.* 100, no. 4, 463-470 (1985)
- [DrK] von Dreifus, H., Klein, A.: A new proof of localization in the Anderson tight binding model. *Commun. Math. Phys.* **124**, 285-299 (1989)
- [FMSS] Fröhlich, J., Martinelli, F., Scoppola, E., Spencer, T.: Constructive proof of localization in the Anderson tight binding model. *Commun. Math. Phys.* **101**, 21-46 (1985)
- [FS] Fröhlich, J., Spencer, T.: Absence of diffusion with Anderson tight binding model for large disorder or low energy. *Commun. Math. Phys.* **88**, 151-184 (1983)
- [GK1] Germinet, F., Klein, A.: Bootstrap multiscale analysis and localization in random media. *Commun. Math. Phys.* **222**, 415-448 (2001)
- [GK2] Germinet, F., Klein, A.: A characterization of the Anderson metal-insulator transport transition. *Duke Math. J.* **124**, 309-351 (2004).
- [GK3] Germinet, F., Klein, A.: New characterizations of the region of complete localization for random Schrödinger operators. *J. Stat. Phys.* **122**, 73-94 (2006)
- [GrV] Graf, G.-M., Vaghi, A.: A remark on an estimate by Minami, *Lett. Math. Phys.* **79**, 17-22 (2007)
- [H] Hundertmark, D.: On the time-dependent approach to Anderson localization. *Math. Nachr.* **214**, 25-38 (2000)
- [KN] Killip, R., Nakano, F.: Eigenfunction statistics in the localized Anderson model. *Ann. Henri Poincaré* **8**, 27-36 (2007)
- [Ki] Kirsch, W.: An invitation to random Schrödinger operators, *Panorama et Synthèses* 2008.
- [KILM] Klein, A., Lenoble, O., Müller, P. : On Mott's formula for the ac-conductivity in the Anderson model. *Annals of Math.* **166**, 549-577 (2007)
- [KIM] Klein, A., Molchanov, S.: Simplicity of eigenvalues in the Anderson model. *J. Stat. Phys.* **122**, 95-99 (2006)

- [M] Minami N.: Local fluctuation of the spectrum of a multidimensional Anderson tight binding model. *Commun. Math. Phys.* **177**, 709-725 (1996)
- [N] Nakano, F.: The repulsion between localization centers in the Anderson model. *J. Stat. Phys.* **123**, 803-810 (2006)
- [RS] Reed, M., Simon, B.: *Methods of Modern Mathematical Physics I: Functional Analysis*, revised and enlarged edition. Academic Press, 1980
- [S] Simon, B.: Cyclic vectors in the Anderson model. Special issue dedicated to Elliott H. Lieb. *Rev. Math. Phys.* **6**, 1183-1185 (1994)
- [SW] Simon, B., Wolff, T.: Singular continuum spectrum under rank one perturbations and localization for random Hamiltonians. *Commun. Pure. Appl. Math.* **39**, 75-90 (1986)
- [W] Wegner, F.: Bounds on the density of states in disordered systems, *Z. Phys.* **B44** 9-15 (1981)

(Combes) UNIVERSITÉ DU SUD: TOULON ET LE VAR, DÉPARTEMENT DE MATHÉMATIQUES, F-83130 LA GARDE, FRANCE

*E-mail address:* `combes@cpt.univ-mrs.fr`

(Germinet) UNIV CERGY-PONTOISE, CNRS UMR 8088, IUF, DÉPARTEMENT DE MATHÉMATIQUES, F-95000 CERGY-PONTOISE, FRANCE

*E-mail address:* `germinet@math.u-cergy.fr`

(Klein) UNIVERSITY OF CALIFORNIA, IRVINE, DEPARTMENT OF MATHEMATICS, IRVINE, CA 92697-3875, USA

*E-mail address:* `aklein@uci.edu`