

Localization for heavy-tailed Anderson models

Omar Hurtado

Abstract. Using recent results on uniform large deviation estimates for random matrix products obtained by Omar Hurtado and Sidhanth Raman (2025), we prove localization for one-dimensional Anderson models with heavy tails.

1. Introduction

1.1. Main results

This article is concerned with the Anderson model with heavy tailed distributions. This is a fundamental model in the study of disordered materials, whose spectral properties have been extensively studied. Mathematically it is given by a (random) operator acting on $\ell^2(\mathbb{Z})$ of the form

$$[H\psi](n) = \psi(n+1) + \psi(n-1) + V_n\psi(n), \quad (1.1)$$

where V_n is an i.i.d. random potential. Such random operators allow one to study the effect on electronic transport of disorder in a material, and said operators are expected (and to a large degree, now known) to display a phenomenon known as Anderson localization. For any probability measure on μ , we call the operator H defined by (1.1) with V_i having law μ the *Anderson model with single site distribution μ* . There are stronger and weaker notions of localization, the stronger notions formulated in terms of moments of the position operator applied to the time evolution associated to H ; this article focuses on purely spectral notions of localization.

Definition 1.1. An operator H is spectrally localized if it has no continuous spectrum and its eigenvectors decay exponentially.

The current state of the art concerning spectral localization for general operators of the form (1.1) is the following, from [5] by Carmona, Klein, and Martinelli.

Mathematics Subject Classification 2020: 81Q10 (primary); 47B80 (secondary).

Keywords: Anderson localization, spectral theory, Schrödinger operators.

Theorem 1.2 ([5]). *If μ satisfies*

$$\int |x|^\alpha d\mu(x) < \infty$$

for some $\alpha > 0$ and is not concentrated on a single point, with single site distribution μ is almost surely spectrally localized.

Importantly, Carmona, Klein, and Martinelli make no regularity assumptions in their work, and we will not either. In this work we treat models which satisfy the weaker moment condition:

$$\int (\log^+(|x|))^p d\mu(x) < \infty. \quad (1.2)$$

(Here $\log^+(x) = \max\{\log x, 0\}$.) We obtain certain non-trivial estimates as soon as $p > 1$, but our main result concerns the case where $p > 11$.

By applying results obtained in [23] and combining them with variations on arguments from [5] and a result from [33] which we use as a black box, we obtain the following theorem.

Theorem 1.3. *If μ satisfies (1.2) for $p > 11$ and is not supported on a single point, then the Anderson model associated to μ is almost surely spectrally localized.*

At least in terms of relaxing the moment assumptions on μ , this is the first improvement on the work of Carmona, Klein, and Martinelli which does not require any regularity assumption, and partially answers a conjecture posed in [28].

1.2. Background on random Schrödinger type operators

These models have been quite well studied, and are fundamental models in the study of disordered materials going back to the groundbreaking (and Nobel Prize winning) work of P. W. Anderson near the middle of the twentieth century. We do not try and give a full account of the history, but we mention some useful reference works, as well as works concerning localization for similar models, with special attention paid to those which treat singular distributions in one dimension.

Indeed, good accounts for the general, possibly multidimensional, case appear in [2, 6]. The one-dimensional theory is markedly different because of the transfer matrix method. While there are other methods which are unique to one dimension, e.g., the spectral averaging of [25], the modern study of random one-dimensional models (and closely related one-dimensional ergodic models of various other flavors) is dominated by the use of this method, which studies ergodic operators by studying associated linear cocycles. For an account of the specifically one-dimensional theory, discussing also work for more general ergodic models, see e.g., [7].

While there were many results capable of treating the case where the random noise is sufficiently regular (see [12, 13, 19, 25]), the first result capable of treating the case with singular potentials (e.g., V_n Bernoulli variables) was the groundbreaking work of Carmona, Klein, Martinelli in [5]. Already, some aspects of the theory of random matrix products were exploited here to facilitate the proof. This proof combined the Multiscale Analysis (MSA) method developed in [13] with certain facts from the theory of random matrix products.

Since then, various “single-scale” proofs of localization in one dimension which eschew MSA and leverage even further properties of random matrix products have appeared, see e.g., [4, 20, 24]. See also [22, 28, 29] for more works which used methods introduced in [24] to prove localization for wider classes of random operators, and [8, 21] which built on methods from [20]. (It is worth mentioning also the important paper [31] which provided a proof based on harmonic analysis via results in harmonic analysis concerning a variant of the uncertainty principle.) In the “single-scale” proofs, uniform large deviation estimates are of crucial importance. It would be interesting to see if a “single-scale” proof of the present results is possible; all of the current single-scale approaches leverage exponential large deviation estimates, which are not available in the heavy-tail regime.

As this history makes clear, lack of regularity in μ complicates proofs of localization. (This is also true in higher dimensions, and in fact localization for e.g., the Bernoulli–Anderson model was only recently solved in dimension 2 [9] and 3 [27], and remains open in dimensions 4 and higher; this work used ideas developed for related continuum models in the landmark work of Bourgain and Kenig [3].) Except for the highly technical version of MSA developed in [3, 18] and various works which built upon these, applications of MSA generally require some moment condition but can treat Hölder regular potentials in higher dimensions. As soon as one has e.g., μ given by a bounded density with respect to Lebesgue measure, one does not require any moment condition. Indeed, the celebrated fractional moment method introduced by Aizenman and Molchanov in [1] gives the following.

Theorem 1.4 ([1]). *Let μ be a distribution generated by a bounded density on \mathbb{R} . Then the Anderson model generated by μ is almost surely spectrally localized.*

In particular, heavy tailed models with regular potentials can be treated via FMM; our localization results are novel in the context of potentials with heavy tails and a lack of regularity.

1.3. Approach of the paper

The main new tool is the use of certain uniform large deviation estimates which were proven by Raman and the author in [23]. The results of said paper are reasonably

general, requiring only irreducibility and proximality assumptions. Specialized to the context of the current paper, we specifically obtain polynomial large deviation estimates for the associated transfer matrices when (1.2) is satisfied for $p > 1$; this is Theorem 2.4. This more or less automatically yields one of the necessary assumptions for a localization proof via MSA for p sufficiently large; the so-called “initial scale estimate”, Theorem 3.4.

These polynomial large deviation estimates also allow us to prove the second necessary assumption; the Wegner estimate which controls the probability of an eigenvalue falling with a small interval, in this case Theorem 4.1. However, here it is not quite automatic, and we prove it by adapting an argument appearing in [5] to prove a similar bound, though it is worth pointing out that we obtain a weaker bound under weaker assumptions. We prove a polynomial bound, and though the outline is the same as in [5] there are a few technical difficulties only arising in our case, and at the same time we simplify certain parts of the argument using our bounds which are uniform in energy. (Such uniform bounds were not yet proven when [5] was published, though the pointwise versions were and figured crucially in the proof.) The polynomial bounds we obtain suffice to prove localization via a theorem of von Dreifus and Klein in [33], which makes use of MSA; we use it as a black box and do not carry out an MSA argument ourselves.

2. Preliminaries

2.1. Notation

Throughout, we always use \mathbb{P} to denote the probability of an event. Generally, we use calligraphic letters to denote events, and \mathcal{A}^C is the complement of the event \mathcal{A} . For an operator H on a Hilbert space, $\sigma(H)$ denotes the spectrum of H , i.e., the set of those $E \in \mathbb{C}$ such that $H - E$ does not have a bounded inverse. All operators considered are self-adjoint (by our assumption that V is real-valued) and so in particular the spectrum is actually a subset of \mathbb{R} . Whenever we deal with random phenomena we use $\mathbb{P}[\mathcal{A}]$ to denote probability of an event \mathcal{A} ; for a random variable, we denote the probability of an event depending on its value by e.g., $\mathbb{P}[X \leq \varepsilon]$.

2.2. Generalized eigenfunctions and transfer matrices

Much of this section recalls facts which are well known in the study of random or more generally ergodic Schrödinger operators; the fundamentally new inputs are Theorem 2.4 (which is proven in [23]) and Proposition 2.7. More comprehensive accounts of many of these identities appear in e.g., [22, 24, 29].

Fundamental objects in the study of spectral questions concerning the Anderson model are generalized eigenfunctions.

Definition 2.1. We say E and $\psi \in \mathbb{R}^{\mathbb{Z}}$ are a *generalized eigenvalue* and *generalized eigenfunction*, respectively, if

$$H\psi = E\psi \tag{2.1}$$

and ψ is of strictly slower than exponential growth, i.e.,

$$\limsup_{|n| \rightarrow \infty} \frac{1}{|n|} \log |\psi(n)| \leq 0.$$

Clearly, all bona fide ℓ^2 eigenfunctions are generalized eigenfunctions, but formal solutions with e.g., very slow decay or even polynomial growth are also permitted. It is a general fact that “intermediate” growth which is superpolynomial but subexponential is impossible for formal solutions, so one can restrict to the class of polynomially bounded formal solutions, but this fact is not important for our purposes.

The importance of these objects in spectral theory is a consequence of a result known as Sch’*n*ol’s theorem, which roughly says that the spectrum of H is supported on the set of the generalized eigenvalues of H . We will not make this precise since we do not need to make use of Sch’*n*ol’s theorem directly, but it is implicit via our use of Theorem 3.2 below (which is a special case of a result of von Dreifus and Klein). Because of this relationship, in localization arguments asymptotics of formal solutions end up being central. In the one-dimensional context, these can be studied fruitfully via the transfer matrix method. Specifically, we introduce first the one step transfer matrices:

$$T_n^E = \begin{pmatrix} E - V_n & -1 \\ 1 & 0 \end{pmatrix}.$$

For ψ solving $H\psi = E\psi$, we have

$$\begin{pmatrix} \psi(n + 1) \\ \psi(n) \end{pmatrix} = T_n^E \begin{pmatrix} \psi(n) \\ \psi(n - 1) \end{pmatrix}. \tag{2.2}$$

Naturally, one can multiply one step transfer matrices to get a matrix which gives one $\begin{pmatrix} \psi(n+1) \\ \psi(n) \end{pmatrix}$ from any $\begin{pmatrix} \psi(m+1) \\ \psi(m) \end{pmatrix}$; we define

$$T_{n,m}^E := \begin{cases} T_n^E T_{n-1}^E \cdots T_{m+1}^E, & n > m, \\ I, & n = m, \\ (T_{n+1}^E)^{-1} (T_n^E)^{-1} \cdots (T_m^E)^{-1}, & n < m. \end{cases}$$

By induction, (2.2) gives the following:

$$\begin{pmatrix} \psi(n + 1) \\ \psi(n) \end{pmatrix} = T_{n,m}^E \begin{pmatrix} \psi(m + 1) \\ \psi(m) \end{pmatrix}.$$

Note that $T_{n,m}^E$ is a product of $|n - m|$ i.i.d. random matrices, with distribution of each term depending on the sign of $n - m$; we will focus on the case $n > m$ specifically. We let μ^E be the distribution of T_0^E on $SL(2, \mathbb{R})$. Under certain assumptions on the distribution, some of them “geometric” and some in terms of moments, we can obtain exponential behavior of random products with high probability. By exponential behavior, we refer to linear growth of $\log \|T_{[a,b]}\|$ and related quantities in $|b - a|$; in particular the following well-known fact is a consequence of work of Furstenberg and collaborators.

Proposition 2.2 ([14–16]). *Under the assumptions of Theorem 1.3, there is a function $\lambda: \mathbb{R} \rightarrow \mathbb{R}$ which is strictly positive and continuous such that*

$$\lim_{b-a \rightarrow +\infty} \frac{1}{b-a} \log \|T_{[a,b]}^E\| = \lambda(E)$$

almost surely.

A similar statement holds for the quantities $\log \|T_{[a,b]}^E x\|$ and $\log |\langle x, T_{[a,b]}^E y \rangle|$, where x, y are non-zero elements of \mathbb{R}^2 . (There is also $\lambda^-(E)$ corresponding to the limit $b - a \rightarrow -\infty$; in many cases of interest, including the particular case of Schrödinger cocycles it turns out that this “backwards” Lyapunov exponent coincides with the “forwards” one.)

Remark 2.3. To be precise all the works cited treat products of random matrices, and make various assumptions, said assumptions varying from paper to paper. Furstenberg and Kesten demonstrated the existence of the limit under very weak assumptions; later Furstenberg showed positivity under assumptions which hold in the present context. Continuity was demonstrated by Furstenberg and Kifer under assumptions which also hold in the present context. That the various hypotheses necessary to apply these results are satisfied is a standard argument, see e.g., [5].

Under strong moment conditions (like e.g., those assumed in the paper of Carmona, Klein, and Martinelli), strong quantitative bounds on this convergence are known; these were proven (in a pointwise form) by Le Page in [26] and exploited in the work of Carmona, Klein, and Martinelli. Versions uniform in energy have since been proven in e.g., [4, 10, 11, 32]; note that except for the work of Tsay all of these results assume boundedness, which is stronger than the assumptions made by Carmona, Klein, and Martinelli.

In [23], Raman and the author showed (among other things) uniform estimates under weaker moment assumptions than those previously considered, in [23, Theorem 1.20].

Theorem 2.4 ([23]). *We let H be a random operator as in Theorem 1.3, and $T_{[a,b]}^E$ the associated transfer matrices. Then for any $I \subset \mathbb{R}$ compact and $\varepsilon > 0$, there is*

$C > 0$ (depending on μ, I, ε) such that all of the following estimates hold:

$$\begin{aligned} \mathbb{P}[|\log \|T_{[a,b]}^E\| - n\lambda(E)| > n\varepsilon] &\leq CL^{-p}, \\ \mathbb{P}[|\log \|T_{[a,b]}^E x\| - n\lambda(E)| > n\varepsilon] &\leq CL^{-p} \quad \text{for } \|x\| = 1, \\ \mathbb{P}[|\log |\langle x, T_{[a,b]}^E y \rangle| - n\lambda(E)| > n\varepsilon] &\leq CL^{-p} \quad \text{for } \|x\| = \|y\| = 1. \end{aligned}$$

Remark 2.5. In fact, these bounds hold for operators of the form (1.1) satisfying (1.2) with $p \geq 3$. Even for those satisfying it for some $p \in (1, 3)$ a similar bound holds but with CL^{-p} replaced by $CL^{1.5-1.5p}$.

These bounds on the transfer matrices are the main new input enabling the localization result. They enable a Wegner estimate, though the proof is somewhat technical and deferred to the last section. They also provide nice bounds on what are called the *Green’s functions* of the operator, which is a sort of finite volume resolvent.

In order to introduce the Green’s function, we must first define a few other quantities. First, we let $H_{[a,b]}$ denote the truncation of H to the interval $[a, b]$. (More precisely, $H_{[a,b]}$ is the corner $P_{[a,b]}HP_{[a,b]}$, where $P_{[a,b]}$ is the appropriate projection.) For any integers $a < b$ and $x, y \in [a, b]$ (here and throughout $[a, b]$ denotes the interval in \mathbb{Z}), we define the Green’s function as follows:

$$G_{[a,b]}^E(x, y) = \langle \delta_x, (H_{[a,b]} - E)^{-1} \delta_y \rangle$$

and formally set $G_{[a,b]}^E = \infty$ if $E \in \sigma(H_{[a,b]})$. (That is, we will use e.g., $\mathbb{P}[G_{[a,b]}^E > C]$ as a shorthand for $\mathbb{P}[E \notin \sigma(H_{[a,b]}) \text{ and } G_{[a,b]}^E > C]$.)

Remark 2.6. $H_{[a,b]}$ can be quite naturally identified with an operator on $\ell^2([a, b])$, i.e., a $(b - a + 1)^2$ matrix in the standard basis. In particular, $G_{[a,b]}^E$ is defined in terms of this identification; by $(H_{[a,b]} - E)^{-1}$ we mean the $\ell^2([a, b])$ inverse, not the inverse with respect to the whole space.

The Green’s function is a key object because it is closely tied to the behavior of formal solutions to (2.1), but it satisfies certain useful identities which make it more amenable to analysis than the formal solutions themselves. Specifically, we have the following well-known identity relating the Green’s function and any formal solution of (2.1):

$$\psi(x) = -G_{[a,b]}^E(x, a)\psi(a - 1) - G_{[a,b]}^E(x, b)\psi(b + 1) \quad \text{for } x \leq y. \tag{2.3}$$

At the same time, the Green’s function satisfies many well-known identities which allow one to relate its asymptotics to those of the transfer matrices, as one would suspect from the relation to the formal solutions via (2.3).

We let

$$P_{[a,b]}^E = \begin{cases} 1 & \text{for } b \leq a, \\ \det(E - H_{[a,b]}) & \text{for } a < b. \end{cases}$$

Then we have the following:

$$|G_{[a,b]}^E(x, y)| = \frac{|P_{[a,x-1]}^E| \cdot |P_{[y+1,b]}^E|}{|P_{[a,b]}^E|} \quad \text{for } a \leq x \leq y \leq b.$$

In particular,

$$\begin{aligned} |G_{[x-L/2,x+L/2]}^E(x, x + L/2)| &= \frac{|P_{[x-L/2,x-1]}^E|}{|P_{[x-L/2,x+L/2]}^E|}, \\ |G_{[x-L/2,x+L/2]}^E(x, x - L/2)| &= \frac{|P_{[x+1,x+L/2]}^E|}{|P_{[x-L/2,x+L/2]}^E|}. \end{aligned} \tag{2.4}$$

(Note that we use self-adjointness of $(H_{[a,b]} - E)^{-1}$ for the second calculation.) Finally, we recall the following well-known fact:

$$P_{[a,b]}^E = \left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix}, T_{[a,b]}^E \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\rangle \quad \text{for } a \leq b. \tag{2.5}$$

From these identities, we can extract one of the crucial ingredients towards a localization result: probabilistic estimates on the Green’s function. While the large deviation results quoted above are a new input, the basic strategy goes back to work of Fröhlich and Spencer.

Proposition 2.7. *Fix $I \subset \mathbb{R}$ a compact interval, and H an Anderson model with single site distribution μ not supported on a single point and satisfying (1.2) for some $p > 3$. Then there is $C > 0$ (depending on I) such that*

$$\mathbb{P}[|G_{[x-L/2,x+L/2]}^E(x, x \pm L/2)| \geq e^{-\lambda(E)L/4}] \leq CL^{-p}$$

for all $L \in \mathbb{N}$ and $E \in I$.

Before commencing the proof in earnest, we define

$$\lambda_{\min} = \min_{E \in I} \lambda(E).$$

We will use this shorthand throughout.

Proof. By Theorem 2.4, we have, for any $\varepsilon > 0$ and compact $I \subset \mathbb{R}$, the existence of $C, c > 0$ (depending on I and μ) such that with probability at least $1 - L^{1-p''}$, we

have the following bounds:

$$\begin{aligned} |P_{[x-L/2, x+L/2]}^E| &\geq e^{(L+1)(\lambda(E)-\varepsilon)}, \\ |P_{[x-L/2, x+1]}^E| &\leq e^{(L/2+1)(\lambda(E)-\varepsilon)}, \\ |P_{[x-1, x+L/2]}^E| &\leq e^{(L/2+1)(\lambda(E)-\varepsilon)}, \end{aligned}$$

for any $p'' \in (p', p)$ and L sufficiently large. Indeed, (2.5) gives precisely this. One thus obtains, via (2.4), that with probability at least $1 - L^{1-p'}$, we have

$$|G_{[x-L/2, x+L/2]}^E(x, x \pm L/2)| \leq e^{(3L/2+2)\varepsilon - \lambda(E)L/2}.$$

Taking $\varepsilon < \lambda_{\min}/6$ gives the result. ■

3. Multiscale analysis

As has been discussed in the introduction, multiscale analysis (MSA) is a central technique in the study of random Schrödinger operators. In some sense, the strongest version of MSA known to work with unbounded operators is the bootstrap MSA developed by Germinet and Klein in [17]; stronger versions have since appeared (see e.g., [3, 9, 18, 30]), but to our knowledge the arguments bridging from the very weak Green’s functions estimates this new MSA can obtain to localization have only been done in the bounded context.

In general, a multiscale analysis is an inductive method by which one can establish bounds on what is called the *Green’s function*. While we have already defined one type of “interval”, in this section we will also consider another type of interval in order to align with usual MSA arguments. For L even, we let $\Lambda_L(x)$ denote the interval $[x - L/2, x + L/2]$; note that unless we are discussing multiple possibly overlapping boxes, the choice of center does not matter for any probabilistic questions; this is a consequence of stationarity and independence of the potentials at different sites; in particular our theorem statements will discuss $\Lambda_L := \Lambda_L(0)$ explicitly, but it should be understood that such results hold for any choice of center. MSA requires the notion of regular “boxes” (intervals in the 1D setting). There are various notions appearing across different papers, but we use (a slight variation of) the notion appearing in [33].

Definition 3.1. We say $\Lambda_L(x)$ is (m, E) regular if

$$|G_{[x-L/2, x+L/2]}^E(x, x \pm L/2)| \leq e^{-mL/2}.$$

MSA is a method whereby we can obtain bounds on $\mathbb{P}[\Lambda_L \text{ is } (m_0, E)\text{-regular}]$; the essential insight is that if most boxes at a smaller scale are regular, and we can

avoid a “resonance”, then we will have regularity (with a slight loss to exponential rate m) at larger scales.

In particular, von Dreifus and Klein proved the following in [33] specifically in the context of Schrödinger operators on \mathbb{Z} .

Theorem 3.2 ([33]). *Let H be an Anderson model of the form (1.1). Let $E_0 \in \mathbb{R}$. Suppose that there is $L_0 > 0$ such that*

- (P1) $\mathbb{P}[\Lambda_{L_0}$ is (m_0, E_0) -regular] $\geq 1 - L_0^{-q_1}$ for some $q_1 > 1$;
- (P2) *there exist $\beta \in (0, 1)$, $q_2 > 4q_1 + 6$ and $\varepsilon > 0$ such that $\mathbb{P}[\text{dist}(\sigma(H_{\Lambda_L}), E) \leq e^{-L^\beta}] \leq 1/L^{q_2}$ for all $L \geq L_0$ and E satisfying $|E - E_0| \leq \varepsilon$*

Then given any $m \in (0, m_0)$, there exists $B < \infty$ (depending on the various parameters at play excluding L_0) such that if $L_0 > B$, we can find δ (depending on the various parameters at play) such that with probability one, H is exponentially localized in $(E_0 - \delta, E_0 + \delta)$, i.e., the portion of the spectrum contained in $(E_0 - \delta, E_0 + \delta)$ is pure point and all eigenfunctions associated to E in this range decay exponentially with rate m .

Remark 3.3. We have formulated a specifically one-dimensional version of the result from [33]; the result is true in arbitrary dimension with some modifications. Moreover, there have been significant advances since said work. Notably a bootstrap MSA was introduced by Germinet and Klein in [17], which can work with a much weaker estimate than (P1). However, said variant also requires a much stronger version of (P2) than we can prove. As has been mentioned before, a variant of MSA capable of dealing with even weaker estimates was introduced and developed in [3, 18], but currently it seems this variant requires boundedness.

While the theorem of von Dreifus and Klein, on the face of it, only proves localization in small intervals, we will use basic union bounds to extend the result to the whole real line. In particular, the proof of Theorem 1.3 is more or less reduced to verifying the two conditions condition (P1) and condition (P2) for all energies. This is the content of Theorems 3.4 and 4.1 below.

The first result, verifying condition (P1), we have more or less already proven; it follows from Proposition 2.7, and the derivation is very brief. The second condition, often called a *Wegner estimate*, we prove using roughly the argument of Carmona, Klein, and Martinelli in [5] to obtain the same under their stronger moment assumption. There are some technical differences in the argument owing to the estimates being uniform in energy but weaker than those exploited in [5].

Theorem 3.4. *Given any compact interval $I \subset \mathbb{R}$, and H with single site distribution μ satisfying (1.2) for $p > 2$, there exists $m > 0$ such that*

$$\mathbb{P}[\Lambda_L \text{ is } (E, m)\text{-regular}] \geq 1 - L^{1-p}$$

for L sufficiently large.

Proof. For

$$m < \frac{1}{4} \lambda_{\min},$$

the desired regularity follows by Proposition 2.7 with probability $1 - CL^{-p}$; for sufficiently large scales $CL^{-p} \leq L^{1-p}$. ■

4. Wegner estimate

Here we prove the Wegner estimate, verifying that condition (P2) holds.

Theorem 4.1. *Given H with single site distribution μ non-trivial and satisfying (1.2) for $p > 2$, and fixed compact $I \subset \mathbb{R}$ we have, for any $E \in I$ and $\beta \in (0, 1)$:*

$$\mathbb{P}[\text{dist}(\sigma(H_{\Lambda_L}), E) \leq e^{-L^\beta}] \leq L^{1-\eta p'}$$

for any $0 < \eta < \beta$ and $p' \in (1, p)$ and L sufficiently large, with the requisite largeness of L depending on the interval I , β , p' , η and μ .

Our proof is based on the strategy of [5, Theorem 4.1], and the following lemma is a technical lemma appearing in their paper.

Lemma 4.2 ([5]). *Let H be as in Theorem 4.1. Then given any $I \subset \mathbb{R}$ compact, there are ρ and C positive, depending on μ and I , such that*

$$\mathbb{P}[\mathcal{E}_{E,L,\varepsilon}] \leq CL\varepsilon^\rho,$$

where $\mathcal{E}_{E,L,\varepsilon}$ denotes the event that there is an eigenvalue $E' \in [E - \varepsilon, E + \varepsilon]$ of $H_{[-L,L]}$ with a normalized eigenfunction φ' such that $|\varphi'(-L)|^2 + |\varphi'(L)|^2 \leq \varepsilon^2$.

It was proven in [5] for operators generated by potentials with fractional moments, but the proof only required Hölder continuity of the density of states, which in particular was shown (in the same paper, see [5, Theorem A.1]) for non-trivial μ satisfying (1.2) with $p \geq 1$.

The last necessary notion before we prove Theorem 4.1 is a certain function which agrees with a power of the logarithmic function at large values, but is linear at small

values, introduced (to our knowledge at least) in [23]. Specifically, we define a function for $x \geq 0$:

$$\log^{p^*}(x) = \begin{cases} (\log x)^p & \text{for } x \geq e^p, \\ \left(\frac{p}{e}\right)^p x & \text{for } x < e^p. \end{cases}$$

These functions have the following important properties for $p \geq 1$:

- (1) \log^{p^*} is concave;
- (2) \log^{p^*} is subadditive;
- (3) $\log^{p^*}(xy) \leq C_p \log^{p^*}(x) \log^{p^*}(y)$ for some $C_p \geq 1$.

In fact, if $p \geq e$, one can take $C_p = 1$; in general one can achieve $C_p = 1$ by rescaling. Since we are principally interested in the case $p > 11$, we will proceed under the assumption $C_p = 1$ without rescaling and omit it to keep notation lighter.

Crucially, if a measure μ satisfies (1.2) for some $p \geq 1$, then for the same p it satisfies:

$$\int \log^{p^*}(|x|) d\mu(x).$$

However, moments with respect to \log^{p^*} are more tractable due to its various nice properties.

Now, we prove Theorem 4.1, adapting the argument from [5].

Proof of Theorem 4.1. As we have discussed, if μ satisfies (1.2), then

$$\mathbb{E}[\log^{p^*}(\|T_0^E\|)] < \infty$$

for any given E , and moreover for any compact $I \subset \mathbb{R}$,

$$M := \sup_{E \in I} \mathbb{E}[\log^{p^*}(\|T_0^E\|)] < \infty.$$

Throughout, to keep computations legible, we omit floor symbols; when L^η appears somewhere an integer argument is necessary we mean $\lfloor L^\eta \rfloor$. The $O(1)$ difference is negligible in all calculations and estimates. We consider two families of events:

$$\begin{aligned} \mathcal{A}_{\theta,L}^E &:= \left\{ \left\| T_{[-L, -L+L^\eta]}^E \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\| \geq \exp(\beta L^\eta) \right\}, \\ \mathcal{B}_{\theta,L}^E &:= \left\{ \left\| T_{[L-L^\eta, L]}^E \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\| \geq \exp(\beta L^\eta) \right\}. \end{aligned}$$

Letting E' range over eigenvalues of $H_{[-L,L]}$, we can bound $\mathbb{P}[\text{dist}(\sigma(H_{[-L,L]}), E) < e^{-L^\beta}]$ by the sum of probabilities for the following four events, for some $\kappa > 0$ to be

determined:

$$\begin{aligned} \mathcal{C}_1 &:= \{\text{dist}(\sigma(H_{[-L,L]}, E) < e^{-L^\beta})\} \cap \left(\bigcap_{|E'-E| < e^{L^\beta}} (\mathcal{A}_{\kappa/2,L}^{E'} \cap \mathcal{B}_{\kappa/2,L}^{E'}) \right); \\ \mathcal{C}_2 &:= \mathcal{A}_{\kappa,L}^E \cap \mathcal{B}_{\kappa,L}^E \cap \left(\bigcup_{|E'-E| < e^{-L^\beta}} (\mathcal{A}_{\kappa/2,L}^{E'})^C \right); \\ \mathcal{C}_3 &:= \mathcal{A}_{\kappa,L}^E \cap \mathcal{B}_{\kappa,L}^E \cap \left(\bigcup_{|E'-E| < e^{-L^\beta}} (\mathcal{B}_{\kappa/2,L}^{E'})^C \right); \\ \mathcal{C}_4 &:= (\mathcal{A}_{\kappa,L}^E)^C \cup (\mathcal{B}_{\kappa,L}^E)^C. \end{aligned}$$

This same cover of the event was carried out in the proof of [5, Theorem 4.1]. We estimate $\mathbb{P}[\mathcal{C}_1]$ using Lemma 4.2. By a standard argument using the Poisson formula (see the proof of [5, Theorem 4.1]), the event \mathcal{C}_1 implies, for every E' satisfying $|E' - E| < e^{-L^\beta}$, the existence of a corresponding normalized eigenfunction $\varphi' \in \ell^2([-L, L])$ with

$$\max\{|\varphi'(L)|^2, |\varphi'(-L)|^2\} \leq 2 \exp(-\kappa L^\eta/2).$$

Note that for large L the intervals $[-L + L^\eta, L]$ and $[-L, L - L^\eta]$ are of size at least $L/2$; we obtain

$$\mathbb{P}[\mathcal{C}_1] \leq CL \exp(-\rho \kappa L^\eta/2)$$

by Lemma 4.2. (Note that in particular we take $\varepsilon = \exp(-\rho \min\{\kappa L^\eta/2, L^\beta\})$, which for sufficiently large L , is just $e^{-\rho \kappa L^\eta/2}$).

The arguments for bounding $\mathbb{P}[\mathcal{C}_2]$ and $\mathbb{P}[\mathcal{C}_3]$ are almost identical; we only treat the former explicitly. Note that one step transfer operators $\begin{pmatrix} E - V_k & -1 \\ 1 & 0 \end{pmatrix}$ and $\begin{pmatrix} E' - V_k & -1 \\ 1 & 0 \end{pmatrix}$ are separated by a rank one operator of norm equal to $|E' - E|$. In particular, if \mathcal{C}_2 holds, then we have

$$\begin{aligned} \exp(\kappa L^\beta/2) &\geq \left\| T_{[-L, -L+L^\eta]}^{E'} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\| \\ &\geq \left\| T_{[-L, -L+L^\eta]}^E \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\| - \|T_{[-L, -L+L^\eta]}^{E'} - T_{[-L, -L+L^\eta]}^E\| \\ &\geq \exp(\kappa L^\beta) - \|T_{[-L, -L+L^\eta]}^{E'} - T_{[-L, -L+L^\eta]}^E\|. \end{aligned}$$

In particular, one obtains for large L the estimate $\|T_{[-L, -L+L^\eta]}^{E'} - T_{[-L, -L+L^\eta]}^E\| \geq \exp(-\kappa L^\beta)/2$. Letting $\Gamma := \|T_{[-L, -L+L^\eta]}^{E'} - T_{[-L, -L+L^\eta]}^E\|$, by Chebyshev,

$$\begin{aligned} \mathbb{P}[\mathcal{C}_2] &\leq \mathbb{P}[\Gamma > \exp(\kappa L^\beta)/2] \\ &\leq \mathbb{P}[\log^{p^*}(\Gamma) > \log^{p^*}(\exp(\kappa L^\beta)/2)] \\ &\leq C \kappa^p L^{-\beta p} \mathbb{E}[\log^{p^*}(\Gamma)]. \end{aligned}$$

We can bound Γ as follows:

$$\Gamma \leq \sum_{k=1}^{L^n+1} \sum_{i_1 < \dots < i_k} \exp(-kL^\beta) \left\| \begin{pmatrix} E - V_{i_k} & -1 \\ 1 & 0 \end{pmatrix} \right\| \dots \left\| \begin{pmatrix} E - V_{i_1} & -1 \\ 1 & 0 \end{pmatrix} \right\|.$$

(This is actually a dramatic overestimate; it is not necessary to count over all possible choices of k indices. However, by overcounting here we can make use of nice combinatorial identities later.) Using concavity, subadditivity, the fact that it is submultiplicative up to a constant, and the fact that \log^{p^*} is linear for small values, we obtain for some $C > 1$,

$$\begin{aligned} \mathbb{E}[\log^{p^*}(\Gamma)] &\leq \mathbb{E} \left[\log^{p^*} \left(\sum_{k=1}^{L^n+1} \sum_{i_1 < \dots < i_k} |E - E'|^k \|T_{L^n+1-k}^E\| \right) \right] \\ &\leq \mathbb{E} \left[\sum_{k=1}^{L^n+1} \sum_{i_1 < \dots < i_k} \log^{p^*}(|E - E'|)^k \prod_{\ell \notin \{i_1, \dots, i_k\}} \log^{p^*}(\|T_\ell^E\|) \right] \\ &\leq \sum_{k=1}^{L^n} \binom{L^n+1}{k} \left(\frac{p}{e}\right)^{pk} \exp(-kL^\beta) M^{L^n+1-k} \\ &= \left(CM + \left(\frac{p}{e}\right)^p \exp(-L^\beta) \right)^{L^n+1} - M^{L^n+1} \\ &\leq C(L^n+1)M^{L^n+1} \left(\frac{p}{e}\right)^p \exp(-L^\beta). \end{aligned}$$

(Recall that $M = \sup_{E \in I} \mathbb{E}[\log^{p^*}(\|T_0^E\|)]$.) We have exploited that $\exp(-L^\beta)$ is small and linearized $x \mapsto x^{L^n+1}$ in the last step.

Because $2(L^n+1)(M+1)^{L^n} \exp(-L^\beta)$ is small one obtains finally

$$\mathbb{E}[\log^{p^*}(\Gamma)] \leq CL^n(M+1)^{L^n} \exp(-L^\beta)$$

for some $C > 0$. This yields

$$\mathbb{P}[\mathcal{C}_2] \leq C\kappa^p L^{\eta-\beta p} (M+1)^{L^n} \exp(-L^\beta),$$

which is, for large L , dominated by $\exp(-L^\beta)/2$. A similar argument gives

$$\mathbb{P}[\mathcal{C}_3] \leq \exp(-L^\beta)/2.$$

Finally, to bound \mathcal{C}_4 , we note that it is a union of at most $2L$ large deviation events. We make this explicit in the case of $\mathcal{A}_{\kappa,L}^E$. Clearly, said event holds if and only if

$$\log \left\| T_{[-L, -L+L^n]}^E \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\| / L^\eta \geq \kappa.$$

By Theorem 2.4, if $\kappa < \lambda_{\min}$, we obtain a bound $\mathbb{P}[\mathcal{C}_4] \leq 2L^{1-\eta p''}$ for any $p'' \in (p', p)$. Hence, we obtain

$$\mathbb{P}[\mathcal{C}_1 \cup \mathcal{C}_2 \cup \mathcal{C}_3 \cup \mathcal{C}_4] \leq L^{1-\eta p'},$$

for L sufficiently large by combining the bounds on all terms. ■

Finally, we prove Theorem 1.3.

Proof of Theorem 1.3. If $p > 11$, one can pick $q_1 > 1$, $q_2 > 4q_1 + 6$, $\beta \in (0, 1)$, $\eta \in (0, \beta)$ and $p' < p$ such that particular $1 - \eta p' < -q_2$. Indeed, let $\kappa = \min\{p - 11, .01\}$, and take $q_1 = 1 + \kappa/16$, $q_2 = 10 + \kappa/4$, $\beta = \left(\frac{11+\kappa/4}{11+\kappa/2}\right)^{1/2}$, $\eta = \frac{11+\kappa/4}{11+\kappa/2}$, and $p' = 11 + \kappa/2$. By Theorem 3.4, we have condition (P1), and by Theorem 4.1, we have condition (P2). Thus, the hypotheses of Theorem 3.2 hold, and so we have almost sure localization on a small interval around any fixed energy E . We take a countable collection of intervals $(E_i - \delta_i, E_i + \delta_i)$ covering \mathbb{R} for which almost sure localization holds via said theorem; taking an intersection over all the events gives almost sure localization for all energies. Such a countable collection exists by the fact that \mathbb{R} is Lindelöf, i.e., any open cover has a countable subcover, and the theorem follows. ■

Acknowledgments. The author thanks Lana Jitomirskaya for suggesting the question of localization for heavy tailed distributions.

Funding. The author was supported in part by NSF Grants Nos. DMS-2052899 and DMS-2155211, and Simons Grant No. 896624.

References

- [1] M. Aizenman and S. Molchanov, [Localization at large disorder and at extreme energies: an elementary derivation](#). *Comm. Math. Phys.* **157** (1993), no. 2, 245–278
Zbl 0782.60044 MR 1244867
- [2] M. Aizenman and S. Warzel, [Random operators](#). Grad. Stud. Math. 168, American Mathematical Society, Providence, RI, 2015 Zbl 1333.82001 MR 3364516
- [3] J. Bourgain and C. E. Kenig, [On localization in the continuous Anderson–Bernoulli model in higher dimension](#). *Invent. Math.* **161** (2005), no. 2, 389–426 Zbl 1084.82005
MR 2180453
- [4] V. Bucaj, D. Damanik, J. Fillman, V. Gerbuz, T. VandenBoom, F. Wang, and Z. Zhang, [Localization for the one-dimensional Anderson model via positivity and large deviations for the Lyapunov exponent](#). *Trans. Amer. Math. Soc.* **372** (2019), no. 5, 3619–3667
Zbl 1422.35022 MR 3988621

- [5] R. Carmona, A. Klein, and F. Martinelli, [Anderson localization for Bernoulli and other singular potentials](#). *Comm. Math. Phys.* **108** (1987), no. 1, 41–66 Zbl 0615.60098 MR 0872140
- [6] H. L. Cycon, R. G. Froese, W. Kirsch, and B. Simon, *Schrödinger operators with application to quantum mechanics and global geometry*. Texts Monogr. Phys., Springer Study Ed., Springer, Berlin, 1987 Zbl 0619.47005 MR 0883643
- [7] D. Damanik and J. Fillman, *One-dimensional ergodic Schrödinger operators. II. Specific classes*. Grad. Stud. Math. 249, American Mathematical Society, Providence, RI, 2024 Zbl 1562.58001 MR 4840232
- [8] D. Damanik, A. Gorodetski, and V. Kleptsyn, Localization for random schrödinger operators defined by block factors. 2025, arXiv:2504.08153v1
- [9] J. Ding and C. K. Smart, [Localization near the edge for the Anderson Bernoulli model on the two-dimensional lattice](#). *Invent. Math.* **219** (2020), no. 2, 467–506 Zbl 1448.60148 MR 4054258
- [10] P. Duarte and S. Klein, *Lyapunov exponents of linear cocycles*. Atlantis Stud. Dyn. Syst. 3, Atlantis Press, Paris, 2016 Zbl 1364.37001 MR 3468528
- [11] P. Duarte and S. Klein, [Large deviations for products of random two-dimensional matrices](#). *Comm. Math. Phys.* **375** (2020), no. 3, 2191–2257 Zbl 1443.37039 MR 4091498
- [12] J. Fröhlich, F. Martinelli, E. Scoppola, and T. Spencer, [Constructive proof of localization in the Anderson tight binding model](#). *Comm. Math. Phys.* **101** (1985), no. 1, 21–46 Zbl 0573.60096 MR 0814541
- [13] J. Fröhlich and T. Spencer, [Absence of diffusion in the Anderson tight binding model for large disorder or low energy](#). *Comm. Math. Phys.* **88** (1983), no. 2, 151–184 Zbl 0519.60066 MR 0696803
- [14] H. Furstenberg, [Noncommuting random products](#). *Trans. Amer. Math. Soc.* **108** (1963), 377–428 Zbl 0203.19102 MR 0163345
- [15] H. Furstenberg and H. Kesten, [Products of random matrices](#). *Ann. Math. Statist.* **31** (1960), 457–469 Zbl 0137.35501 MR 0121828
- [16] H. Furstenberg and Y. Kifer, [Random matrix products and measures on projective spaces](#). *Israel J. Math.* **46** (1983), no. 1-2, 12–32 Zbl 0528.60028 MR 0727020
- [17] F. Germinet and A. Klein, [Bootstrap multiscale analysis and localization in random media](#). *Comm. Math. Phys.* **222** (2001), no. 2, 415–448 Zbl 0982.82030 MR 1859605
- [18] F. Germinet and A. Klein, [A comprehensive proof of localization for continuous Anderson models with singular random potentials](#). *J. Eur. Math. Soc. (JEMS)* **15** (2013), no. 1, 53–143 Zbl 1267.82066 MR 2998830
- [19] I. J. Gol'dšėid, S. A. Molčanov, and L. A. Pastur, A pure point spectrum of the stochastic one-dimensional Schrödinger operator. *Funkcional. Anal. i Priložen.* **11** (1977), no. 1, 1–10, 96. [English traslation](#). *Funct. Anal. Appl.* **11** (1977), 1–8 Zbl 0368.34015 MR 0470515
- [20] A. Gorodetski and V. Kleptsyn, [Parametric Furstenberg theorem on random products of \$SL\(2, \mathbb{R}\)\$ matrices](#). *Adv. Math.* **378** (2021), article no. 107522 Zbl 1458.37060 MR 4184298

- [21] A. Gorodetski and V. Kleptsyn, [Non-stationary Anderson localization](#). *Commun. Am. Math. Soc.* **5** (2025), 81–143 Zbl 08037087 MR 4894553
- [22] O. Hurtado, [A “lifting” method for exponential large deviation estimates and an application to certain non-stationary 1D lattice Anderson models](#). *J. Math. Phys.* **64** (2023), no. 6, article no. 061902 Zbl 1516.82042 MR 4604656
- [23] O. Hurtado and S. Raman, [Uniform estimates for random matrix products and applications](#). 2025, arXiv:2504.08226v2
- [24] S. Jitomirskaya and X. Zhu, [Large deviations of the Lyapunov exponent and localization for the 1D Anderson model](#). *Comm. Math. Phys.* **370** (2019), no. 1, 311–324 Zbl 07083976 MR 3982697
- [25] H. Kunz and B. Souillard, [Sur le spectre des opérateurs aux différences finies aléatoires](#). *Comm. Math. Phys.* **78** (1980/81), no. 2, 201–246 Zbl 0449.60048 MR 0597748
- [26] É. Le Page, [Théorèmes limites pour les produits de matrices aléatoires](#). In *Probability measures on groups (Oberwolfach, 1981)*, pp. 258–303, Lecture Notes in Math. 928, Springer, Berlin etc., 1982 Zbl 0506.60019 MR 0669072
- [27] L. Li and L. Zhang, [Anderson–Bernoulli localization on the three-dimensional lattice and discrete unique continuation principle](#). *Duke Math. J.* **171** (2022), no. 2, 327–415 Zbl 1489.82044 MR 4375618
- [28] D. Macera and S. Sodin, [Anderson localisation for quasi-one-dimensional random operators](#). *Ann. Henri Poincaré* **23** (2022), no. 12, 4227–4247 Zbl 07622222 MR 4512235
- [29] N. Rangamani, [Singular-unbounded random Jacobi matrices](#). *J. Math. Phys.* **60** (2019), no. 8, article no. 081904 Zbl 1426.82032 MR 3997122
- [30] N. Rangamani and X. Zhu, [Dynamical localization for the singular Anderson model in \$\mathbb{Z}^d\$](#) . *J. Math. Phys.* **66** (2025), no. 2, article no. 023505 Zbl 1559.82002 MR 4859979
- [31] C. Shubin, R. Vakilian, and T. Wolff, [Some harmonic analysis questions suggested by Anderson–Bernoulli models](#). *Geom. Funct. Anal.* **8** (1998), no. 5, 932–964 Zbl 0920.42005 MR 1650106
- [32] J. Tsay, [Some uniform estimates in products of random matrices](#). *Taiwanese J. Math.* **3** (1999), no. 3, 291–302 Zbl 0941.60017 MR 1705982
- [33] H. von Dreifus and A. Klein, [A new proof of localization in the Anderson tight binding model](#). *Comm. Math. Phys.* **124** (1989), no. 2, 285–299 Zbl 0698.60051 MR 1012868

Received 14 May 2025; revised 11 January 2026.

Omar Hurtado

Department of Mathematics, University of California, Irvine, 340 Rowland Hall, Irvine, CA 92697, USA; ohurtad1@uci.edu

Author IDs: zbMATH [hurtado.omar](#) MR 1565148 ORCID 0009-0005-3268-7649