

LYAPUNOV EXPONENTS AND LARGE DEVIATIONS ANALYSIS OF EIGENFUNCTIONS IN ANDERSON MODELS ON GRAPHS

VICTOR CHULAEVSKY

DÉPARTEMENT DE MATHÉMATIQUES UNIVERSITÉ DE REIMS, MOULIN DE LA HOUSSE, B.P. 1039 51687 REIMS
CEDEX 2, FRANCE

*CORRESPONDING AUTHOR: VICTOR.TCHOULAEVSKI@UNIV-REIMS.FR

COPYRIGHT ©2013 HORIZON RESEARCH PUBLISHING ALL RIGHTS RESERVED.

Abstract We propose a new probabilistic approach to the analysis of decay of the Green's functions and the eigenfunctions of the Anderson Hamiltonians on countable graphs. Our method is close in spirit to the Fractional Moment Method, but we show how the use of the *fractional moments* can be avoided, so that exponential decay of the Green's functions can be established in some models where the fractional moments diverge, due to low regularity of the random potential. We elucidate the exceptional role of the Hölder continuity condition, usual in the FMM, in terms of Cramer's condition in the large deviations problem for a suitably constructed rigorous path expansion.

Keywords Anderson Localization, Schur Complement, Large Deviations Estimates

1. INTRODUCTION. DECOUPLING TECHNIQUES IN THE LOCALIZATION THEORY

We consider random Anderson Hamiltonians on locally finite countable graphs \mathcal{G} , endowed with the canonical graph distance $d(\cdot, \cdot)$, of the form¹

$$\begin{aligned} (H(\omega)\psi)(x) &= (\epsilon\Delta_{\mathcal{G}}\psi)(x) + V(x;\omega)\psi(x) \\ &= \sum_{y: d(x,y)=1} \epsilon(\psi(y) - \psi(x)) + V(x;\omega)\psi(x) \end{aligned}$$

where $\Delta_{\mathcal{G}}$ is the canonical graph Laplacian on \mathcal{G} , $V : \mathcal{G} \times \Omega \rightarrow \mathbb{R}$ is an IID random field on \mathcal{G} relative to some probability space $(\Omega, \mathfrak{F}, \mathbb{P})$, and $\epsilon > 0$ is a parameter; in fact, ϵ^{-1} measures the amplitude of the disorder.

Starting from the very first mathematical works (cf. [17, 18, 29, 13]) on Anderson localization in multidimensional random environment, the analysis of the decay properties of the Green's functions (= matrix elements of the resolvent in a suitable basis) was carried out with the help of one or another decoupling technique, allowing to effectively decompose a large (ultimately, infinite) system into smaller subsystems. This was less pronounced in [17, 18], but even there the multiscale approach referred to the local Hamiltonians in the subsystems of finite size. In the reformulation of the Multi-Scale Analysis (MSA) performed in [29, 13] and later works, the

entire inductive procedure was based on the analysis of the resolvents of the finite-dimensional subsystems. The main tool of such analysis is the well-known second resolvent equation. The drawback of this technique is the necessity to keep track of the "resonances", or "small denominators", occurring in the inductively treated finite (but growing) subsystems. The MSA has become by now a powerful method – or, rather, family of methods – successfully applied to a very large class of Anderson models in discrete media (such as periodic lattices or, more generally, countable graphs with tempered rate of growth of balls) and continuous media (including those in Euclidean spaces and in the so-called "quantum graphs"). However, the above mentioned drawback is still prohibiting the application of the MSA to the graphs with exponential growth of balls, including the non-degenerate Cayley trees.

A important step forward has been made in 1993 by Aizenman and Molchanov [2] who proposed a different approach, based on the complete decoupling of the fractional moments of the Green's functions in small subsystems of a larger system. Initially applied to the strongly disordered systems, and having to rely on the Simon–Wolff criterion of localization [31] for the proof of spectral localization (a.s. pure point spectrum), this technique, known today as the Fractional Moment Method (FMM) has been improved and generalized in a series of works by Aizenman et al. (cf. e.g., [3, 5, 4]).

Compared to the MSA, the FMM is much less sensitive to the combinatorial complexity of the underlying graph² \mathcal{Z} in which the Anderson localization problem is considered; most notably, FMM applies with no difficulty to the graphs with exponential growth of the volume of balls, including the regular trees. The reverse of the medal is the FMM's greater sensitivity to the regularity of the probability distribution of the random potential. Specifically, the key objects of the FMM, the expectations of some fractional power $s \in (0, 1)$ of the Green's functions, are well-defined under the assumption of Hölder-continuity of the marginal probability distribution of the random potential. In the framework of the FMM, this comes as a matter-of-fact constataion; since the method simply stops working for less regular random potentials, the analysis of the role of the Hölder continuity also stops there.

¹We write the kinetic energy as $\epsilon\Delta$ only to avoid more cumbersome factors $(-\epsilon)$ in our formulae.

²In the present paper, we focus on the discrete models.

On the other hand, the MSA excels in the situations where the random potential has an extremely poor marginal distribution (most notably, the Anderson–Bernoulli model in \mathbb{R}^d ; cf. [9]). Joint efforts by Germinet–Klein and Aizenman–Warzel resulted recently in a remarkable improvement of the celebrated Wegner estimate [6] and, ultimately, in the proof of Anderson localization in Euclidean space with alloy-type potential with *any nontrivial* probability distribution of the scatterers’ amplitudes (cf. [22]). (The lattice counterpart of this model remains a challenging open problem).

In the present work, we elucidate the deep reasons for the exceptional role of the Hölder continuity in the moment analysis of the Anderson Hamiltonians: it can be interpreted as Cramer’s condition in a related large deviations analysis, appearing in a special, rigorous path expansion for the Green’s functions. In the theory of the Large Deviations Estimates³ (LDE), the crucial role of Cramer’s condition is well-known. It is also well-understood how weaker LDE can be obtained under weaker assumptions on the tail probabilities of the random (e.g., IID = independent and identically distributed) summands. In our approach, the role of these summands is played by the *logarithms* of the Green’s functions. An advantage of this language is that the logarithms can have finite moments of any given order under a very weak assumption of log-Hölder (and not necessarily Hölder) continuity of the marginal probability distribution of the potential.

Although the path expansions used in our paper, and originating in earlier works [5, 23, 30, 32], do not solve all problems appearing in the moment analysis when only a relatively low regularity is assumed, they clearly evidence that there is much more place for the ergodic theory in the localization analysis than there used to be, even beyond the particular class of trees, starting with the one-dimensional lattice \mathbb{Z}^1 , where one or another variant of the transfer-operator can be employed. Basically, the two most important implications of the underlying ergodicity in general multi-dimensional media used so far were

- the a.s.-non-randomness of the spectrum and of its components (a.c., s.c. and p.p. spectra);
- the large deviations estimates in the framework of the “Lifshitz tails”, resulting in the non-perturbative (holding for an arbitrarily small amplitude of the disorder) result on localization near the spectral edges.

From the analytic perspective, the properly constructed path expansions allow to achieve a rigorous decoupling of a large system into smaller subsystems *before* any probabilistic estimates are made; by comparison, the AM/FMM techniques achieve the efficient decoupling *only* in the expectations (fractional moments), and the latter may or may not exist.

Aiming at the decay bounds on the eigenfunction correlators, we also have to turn, at some point, to the expectations, but an important distinction from the general FMM strategy is that we calculate the expectations

of the eigenfunction correlators, e.g.,

$$\langle \delta_x | P_I(H(\omega)) | \delta_y \rangle,$$

where P_I is the spectral projection on some interval $I \subset \mathbb{R}$, and these quantities are bounded (by 1), while the Green’s functions are not. (This important fact has been pointed out in numerous works on the FMM techniques.) Therefore, one can achieve a satisfactory upper bound on the eigenfunction correlators, even if the related estimates on the Green’s functions are obtained only with high probability. In the FMM approach, the convergence of the fractional moments is a *sine qua non* condition.

In a way, our approach compared to the FMM reminds the Cheshire cat’s smile: it can be considered as the Fractional Moment Method ... possibly *without* fractional moments; the latter may appear in a disguised form – in the Cramer’s condition, if and when it is fulfilled.

I would like to summarize the above discussion by saying the following:

- First, the main point of the new approach, building on and further developing the moment analysis and the path expansions used earlier (cf. [2, 5, 23]), is that the decoupling is to be performed as early as possible in the analysis of the resolvents, while the use of expectations, on the contrary, has to be postponed to the latest possible stage.
- Secondly, in the main body of the decay analysis of the resolvents, the powerful techniques of the ergodic theory should (and, as we show, can, at least in some classes of models) be used more systematically, to obtain the key estimates in probability, when their counterparts in expectation are not available.

The initial motivation for this work was an attempt to find a method capable to “interpolate” between the MSA and the FMM. The technique presented here ostensibly gravitates toward the AM/FMM approach, but the remaining difficulties might require a more substantial use of the general MSA’s philosophy.

In this paper, we focus mainly on the positivity of the “Lyapunov exponents” (= decay exponents of the Green’s functions), in the perspective of its application to the analysis of eigenfunction correlators (cf. Section 4). As to spectral localization, it is discussed only in passing (cf. subsection 4.2).

The derivation of the localization bounds from the fixed-energy analysis can be done (and has been done in the past) in different ways. Recall that Martinelli and Scoppola [26] proved the absence of the a.c. spectrum on a lattice under the assumption of fast decay of the Green’s functions. Their argument, based on the Chebyshev inequality in the energy-disorder product space, has been further developed and used, e.g., by Bourgain and Kenig [9], where it is a part of an elaborate, all-in-one scaling procedure, and, in a more distinctly encapsulated form, by Elgart et al. [15]. This allows to easily transform the fixed-energy estimates into their energy-interval counterparts. As to the spectral localization, it was derived from the energy-interval estimates for the

³Note that Aizenman and Warzel (cf., e.g., [8] and references therein) used the large deviations estimates in their deep analysis of the delocalization phenomena on trees, under the assumption of Lipschitz continuity of the IID random potential.

Green’s functions by Fröhlich et al. [18] and, in a modified way, by von Dreifus and Klein [13]. Dynamical localization was inferred from the energy-interval estimates by Germinet–De Bièvre [20] and by Damanik–Stollmann [12] (in a stronger form). A particularly short and transparent derivation was developed by Germinet and Klein [21], and we use their approach (with minor adaptations) in Section 4.

The passage from the Lyapunov exponents to the localization is not automatic, however. The two most spectacular counterexamples are provided by the random Anderson model on a non-degenerate (not one-dimensional) Cayley tree, and by the quasi-periodic Almost Mathieu (a.k.a. Harper’s) operator in \mathbb{Z}^1 in the strong disorder regime. In the former model, the exponential decay of the Green’s functions can be insufficient to overcome the exponential growth of the spheres of radius $L \rightarrow \infty$, while in the latter the spectrum may be purely singular continuous, due to a result by Gordon [19], if the basic frequency is abnormally fast approximated by rational numbers. Nevertheless, we prefer to address in this paper mainly the Lyapunov exponents.

The structure of the paper is as follows:

- (1) The main path (“slim wormhole”) expansion is introduced in Section 2 and further developed in Section 3. It appeared earlier in [23] (and less explicitly, in [5]). See also recent works [30], [32]. Unlike [23], we derive it from the Schur complement formula, without the loop elimination in the formal random walk expansion.
- (2) Some other expansions are discussed in Section 5. In particular, we propose the “fat” version of the wormhole expansion, not considered in [23], and a more promising “sandwiched” expansion.
- (3) In Section 3, we employ the “slim wormhole” expansions and standard methods of the large deviations analysis to prove exponential decay of Green’s functions for Anderson Hamiltonians on various types of locally finite graphs.
- (4) The derivation of spectral and dynamical localization from the positivity of “Lyapunov exponents” is discussed in Section 4, where we follow essentially the works by Elgart et al. [15] and Germinet–Klein [21].
- (5) The large deviations estimates used in the main text are proven in Appendix.

2. SCHUR FORMULA AND WORMHOLE EXPANSIONS

2.1. Schur formula for block-inversion. The algorithm for inverting a finite-dimensional matrix represented in a block form has many names; an enlightening discussion of its amazing story can be found, e.g., in the book [34]. In the physical literature, particularly in quantum chemistry and *ab initio* numerical methods thereof, one usually refers to the Feshbach (–Fano) method. Since we are going to apply this method only to finite-dimensional operators, we refer to it as the Schur method (cf. [28]).

The popularity of the Feshbach method in physics, makes it – *today!* – quite surprising that it had not been

used in the early physical papers on the Anderson localization. A more problematic, perturbative path expansion, employed already in the seminal paper [1] by P. W. Anderson, results in a much more perilous journey towards the proof of localization, even for the fixed-energy Green’s functions. There seems to be no obstacle for using a regularized, well-defined self-avoiding path expansion stemming from the Schur–Feshbach approach, which can even be finite yet provide all information a physicist may need, while satisfying the most rigorous mathematician.

For a block matrix, which need not be real-symmetric or Hermitian, one has

$$\begin{aligned} & \begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1} \\ &= \begin{pmatrix} M_A^{-1} & -M_A^{-1}BD^{-1} \\ -D^{-1}CM_A^{-1} & D^{-1} + D^{-1}CM_A^{-1}BD^{-1} \end{pmatrix} \end{aligned} \tag{2.1}$$

where M is the so-called Schur complement of D :

$$M = A - BD^{-1}C,$$

provided that the matrices D and $A - BD^{-1}C$ are invertible. Naturally, in the case of Hermitian matrices, $A^* = A$, $D^* = D$ and $B^* = C$.

2.2. The case rank $A = 1$. Consider a finite connected graph \mathcal{G} , the graph Laplacian $\Delta = \Delta_{\mathcal{G}}$ associated with it, and a function $V : \mathcal{G} \rightarrow \mathbb{R}$. Explicitly,

$$\begin{aligned} (\Delta f)(x) &= \sum_{y \in S_1(x)} (f(x) - f(y)) \\ &= n_{\mathcal{G}}(x) f(x) - \sum_{y \in S_1(x)} f(y), \end{aligned}$$

where $n_{\mathcal{G}}(x) = |S_1(x)|$ is called the coordination number (the number of the nearest neighbors) of the point x in the graph \mathcal{G} . Introduce the discrete Schrödinger operator $H = \epsilon\Delta + gV$, and its matrix in the standard delta-basis in $\ell^2(\mathcal{G})$. The parameter $\epsilon > 0$ is often referred to as the “hopping” amplitude, while $|g|$ measures the amplitude of the disorder. Certainly, by rescaling of the energy one can eliminate one of the parameters ϵ and g . For the purposes of path expansions, it will be convenient to use essentially ϵ , and only occasionally g . Clearly, small values of $\epsilon > 0$ correspond to the strong disorder regime.

Next, consider the decomposition of the graph $\mathcal{G} = \{x_0\} \sqcup (\mathcal{G} \setminus \{x_0\})$ and the corresponding block-decomposition of H , with $A = |\mathbf{1}_{x_0}\rangle\langle\mathbf{1}_{x_0}|$ (we use the Dirac’s “bra-ket” notations throughout the paper). Respectively,

$$\begin{aligned} B &= \sum_{y \in S_1(x_0)} |\mathbf{1}_{x_0}\rangle\langle\mathbf{1}_y|, & C &= \sum_{y \in S_1(x_0)} |\mathbf{1}_y\rangle\langle\mathbf{1}_{x_0}|, \\ D &= \Pi_{\mathcal{G} \setminus \{x_0\}} H \Pi_{\mathcal{G} \setminus \{x_0\}}, \end{aligned}$$

where $S_1(x_0)$ is the sphere of radius 1 with center x_0 : $S_1(x_0) = \{y : d(x_0, y) = 1\}$, and $\Pi_{\mathcal{G} \setminus \{x_0\}}$ is the orthogonal projection onto $\mathcal{G} \setminus \{x_0\}$.

Denote by $G_{\mathcal{G} \setminus \{x_0\}}(x, y; E)$ the matrix elements (in the delta-basis) of the resolvent $(D - E)^{-1}$ (where the latter is well-defined). Let

$$\mathcal{D}(x_0; \mathcal{G}; E) := \sum_{y \in S_1(x_0)} \sum_{z \in S_1(x_0)} G_{\mathcal{G} \setminus \{x_0\}}(y, z; E),$$

then

$$\begin{aligned} & B(D - E)^{-1}C \\ &= \sum_{y \in S_1(x_0)} \sum_{z \in S_1(x_0)} |\mathbf{1}_{x_0}\rangle \langle \mathbf{1}_y | G_{\mathcal{G} \setminus \{x_0\}}(E) |\mathbf{1}_z\rangle \langle \mathbf{1}_{x_0}| \\ &= \mathcal{D}(x_0; \mathcal{G}; E) \cdot \Pi_{x_0}. \end{aligned}$$

For notational brevity, set

$$\tilde{E}(x_0; \mathcal{G}; E) := E - n_{\mathcal{G}}(x) - \mathcal{D}(x_0; \mathcal{G}; E), \quad (2.2)$$

where $n_{\mathcal{G}}(x) := |S_1(x)|$ is the coordination number of $x \in \mathcal{G}$, then

$$(A - E - BG_{\mathcal{G} \setminus \{x_0\}}(E)C)^{-1} = (V(x_0) - \tilde{E}(x_0; E))^{-1} \cdot \Pi_{x_0}. \quad (2.3)$$

and for $y_0 \neq x_0$, we have

$$\begin{aligned} & G_{\mathcal{G}}(x_0, y_0; E) \\ &= \langle \mathbf{1}_{x_0} | (A - E - BG_{\mathcal{G} \setminus \{x_0\}}(E)C)^{-1} BG_{\mathcal{G} \setminus \{x_0\}}(E) | \mathbf{1}_{y_0} \rangle. \end{aligned}$$

The energy shift in $\tilde{E}(\cdot)$ is usually referred to as the "self-energy".

2.3. The "wormhole" expansion. We associate with each subgraph $\mathcal{G}' \subseteq \mathcal{G}$ the resolvent $G_{\mathcal{G}'}(E)$ of the respective discrete Schrödinger operator $H_{\mathcal{G}'} = \epsilon \Delta_{\mathcal{G}'} + V$. At least for $E \in \mathbb{C}_+$, we have

$$\begin{aligned} & G_{\mathcal{G}}(x_0, y_0; E) \\ &= (V(x_0) - \tilde{E}(x_0; \mathcal{G}; E))^{-1} \sum_{u \in S_1(x_0)} G_{\mathcal{G} \setminus \{x_0\}}(u, y_0; E). \end{aligned} \quad (2.4)$$

Note that for $\Im E > 0$, $\Im \mathcal{D}(x_0; \mathcal{G}; E) > 0$, therefore, $\Im \tilde{E}(x_0; \mathcal{G}; E) > \Im E > 0$.

Now assume we have an IID random field $V : \mathcal{G} \times \Omega \rightarrow \mathbb{R}$, relative to a probability space $(\Omega, \mathfrak{B}, \mathbb{P})$. Let \mathfrak{F}_{x_0} be the sigma-algebra generated by $V(x_0; \omega)$ and, respectively, $\mathfrak{F}_{\neq x_0}$ the sigma-algebra generated by $\{V(x; \cdot), x \neq x_0\}$. Then $\tilde{E}(x_0; \mathcal{G}; E; \omega)$ and $G_{\mathcal{G} \setminus \{x_0\}}(u, y_0; E; \omega)$ are $\mathfrak{F}_{\neq x_0}$ -measurable.

The set $\mathcal{G} \setminus \{x_0\}$ may be disconnected, i.e., decomposed into a union of disjoint connected subgraphs,

$$\mathcal{G} \setminus \{x_0\} = \sqcup_{i=1}^n \mathcal{G}_i,$$

among which exactly one, say, $\mathcal{G}_{i(y_0)}$, contains y_0 . More generally, given a subset of vertices $\mathcal{X} \subset \mathcal{G} \setminus \{y_0\}$, there is a disjoint union of connected subgraphs

$$\mathcal{G} \setminus \mathcal{X} = \sqcup_{i=1}^n \mathcal{G}_i,$$

and a unique index $i(y_0)$ such that $\mathcal{G}_{i(y_0)} \ni y_0$. In this case, denote $\mathcal{G} \ominus \mathcal{X} := \mathcal{G}_{i(y_0)}$. Clearly, for any $x_1 \in \mathcal{G} \setminus \mathcal{X}$, one has

$$G_{\mathcal{G} \setminus \mathcal{X}}(x_1, y_0; E) = G_{\mathcal{G} \ominus \mathcal{X}}(x_1, y_0; E) = G_{i(y_0)}(x_1, y_0; E).$$

We see that the first step expands the initial Green's function in a sum of the "reduced" ones relative to $\mathcal{G} \setminus \{x_0\}$, with x_1 ranging over the entire set $\mathcal{G} \setminus \{x_0\}$.

By recursion, we derive from (2.4)

$$\begin{aligned} & G_{\mathcal{G}}(x_0, y_0; E) \\ &= \sum_{x_1 \in S_1(x_0)} \sum_{x_2 \in S_1(x_1) \setminus \{x_0\}} G_{\mathcal{G} \ominus \{x_0, x_1\}}(x_2, y_0; E) \times \\ & \quad \times \prod_{j=0}^1 \frac{1}{V(x_j) - \tilde{E}(x_j; E)} \end{aligned} \quad (2.5)$$

where some terms may be zero, due to the possible disconnectedness of the set $\mathcal{G} \setminus \{x_0\}$ (but not all, since \mathcal{G} is connected).

Call a self-avoiding path (SAP) any finite sequence of pairwise distinct points going over the edges:

$$\begin{aligned} & \gamma = (x_0, \dots, x_k), \quad d(x_j, x_{j+1}) = 1, \\ & |\gamma| := k = \#\{x_0, \dots, x_k\} - 1 \end{aligned}$$

($|\gamma|$ is the number of steps made by γ), and denote

$$\gamma(t) := x_t, \quad t = 0, \dots, |\gamma| + 1.$$

It is also useful to have a notation for the *subpaths* of γ ; for $j \geq 0$, we set

$$\gamma_j := \{\gamma(t), t = 0, \dots, j\}, \quad j = 0, \dots, |\gamma| + 1,$$

while $\gamma_{-1} := \emptyset$. If γ' is a subpath of γ (denoted as $\gamma' \subset \gamma$), we shall say that γ' is an *ancestor* of γ , and γ is a *descendant* of γ' .

Unless otherwise specified, below **all paths will be assumed to be self-avoiding** (SAP).

Denote:

- by $\Gamma_n(x)$ the set of all paths γ in \mathcal{G} with $\gamma(0) = x$ and $|\gamma| = n$;
- by $\Gamma(x, y)$ the set of all paths γ starting at x and ending at y ($\gamma : x \rightsquigarrow y$);
- by $\Gamma_n(x, y)$ the set of all paths $\gamma : x \rightsquigarrow y$ of length n . We call such paths "*n*-bridges" (between x and y).

Now (2.5) becomes a particular case of a more general formula, easily obtained by recursion. Fix two distinct points $x_0, y_0 \in \mathcal{G}$. Next, pick any $0 \leq n \leq d(x_0, y_0)$. Then we have the following identity: for any $E \in \mathbb{C}_+$

$$\begin{aligned} G_{\mathcal{G}}(x_0, y_0; E) &= \sum_{\gamma \in \Gamma_n(x_0)} G_{\mathcal{G} \ominus \gamma}(x_n, y_0; E) \times \\ & \quad \times \prod_{j=0}^{|\gamma|} \frac{1}{V(x_j) - \tilde{E}(x_j; \mathcal{G} \ominus \gamma_{j-1}; E)}. \end{aligned} \quad (2.6)$$

The above identity also holds, of course, for some real energies, viz. for those E for which the denominators $V(x_j) - \tilde{E}(x_j; E)$ do not vanish and $G_{\mathcal{G} \ominus \gamma}(x_n, y_0; E)$ is well-defined.

The role of the paths is two-fold:

- the point x_{j+1} may (but not necessarily does) come one step closer than x_j to the point y_0 ;
- the step $x_j \rightsquigarrow x_{j+1}$ eliminates at least one point (possible more) from the domain where the matrix element $G_{\mathcal{G} \ominus \{x_0, \dots, x_{j+1}\}}(x_{j+1}, y_0; E)$ is evaluated.

The resolvents evolve along the excluded "tunnel" $\gamma(j+1) = \{x_0, \dots, x_{j+1}\}$ across the space between x_0 and y_0 , so we call the resulting sum a *wormhole* expansion. In the simplest version that we introduced, the tunnel is 1-point wide, so we call it *slim*. More general ("fat") wormholes are discussed in Section 5.

2.4. Assessing the remainder factor $G_{\mathcal{G} \ominus \gamma}$. Aizenman and Molchanov [2] and Aizenman et al. [5] used the Krein formula to reduce a similar estimation to a two-dimensional, random matrix problem, regardless of the size of the ambient graph \mathcal{G} .

Consider the paths $\gamma \in \Gamma(x_0, y_0)$; recall that $|\gamma|$ is the number of jumps (x_i, x_{i+1}) made by γ ; each jump is associated with the "hopping" factor ϵ from $\epsilon\Delta$. Let

$$\mathcal{X}_j(\gamma) := \mathcal{G} \ominus \gamma_{j-1}, \quad j = 0, \dots, |\gamma| + 1, \quad (2.7)$$

so $\mathcal{X}_{|\gamma|+1} = \mathcal{G} \ominus \gamma$, and, with notations of (2.6), set for brevity

$$\lambda_j(\gamma) = \lambda_j(\gamma; E) := \tilde{E}(x_j; \mathcal{X}_j(\gamma); E) \equiv \tilde{E}(x_j; \mathcal{G} \ominus \gamma_{j-1}; E).$$

An important observation is that λ_j is independent of $\{V(x_i; \cdot), 0 \leq i \leq j\}$. Then we obtain the expansion (cf. (2.6))

$$\begin{aligned} G_{\mathcal{G}}(x_0, y_0; E) &= \sum_{\gamma \in \Gamma(x_0; y_0)} G_{\mathcal{X}_{|\gamma|+1}(\gamma)}(y_0, y_0; E) \prod_{j=0}^{|\gamma|} \frac{\epsilon}{V(x_j) - \lambda_j(\gamma; E)}. \end{aligned} \quad (2.8)$$

The factor $G_{\mathcal{G} \ominus \gamma}(y_0, y_0; E)$ can be re-written in the form similar to the other factors, again with the help of the Schur formula, applied to the operator $H_{\mathcal{X}_{|\gamma|+1}(\gamma)}$, where the block operator A is the restriction $\mathbf{1}_{y_0} H_{\mathcal{X}_{|\gamma|+1}(\gamma)} \mathbf{1}_{y_0}$ to $\mathcal{X}_{|\gamma|+1}(\gamma)$. By the Schur formula (2.1) for the diagonal entry,

$$\begin{aligned} G_{\mathcal{X}_{|\gamma|+1}(\gamma)}(y_0, y_0; E) &= \langle \mathbf{1}_{y_0} | (A - BD^{-1}C)^{-1} | \mathbf{1}_{y_0} \rangle \\ &= \frac{1}{V(y_0; \omega) - \lambda_{|\gamma|+1}(\gamma; E)}. \end{aligned}$$

Now (2.8) takes a more homogeneous form: with $x_j = \gamma(j)$,

$$G_{\mathcal{G}}(x_0, y_0; E) = \sum_{\gamma \in \Gamma(x_0; y_0)} \epsilon^{|\gamma|} \prod_{j=0}^{|\gamma|+1} \frac{1}{V(x_j) - \lambda_j(\gamma; E)}. \quad (2.9)$$

This is a rigorous identity; if $|\mathcal{G}| < \infty$, then the sum is actually finite. Note that we could stop expanding the resolvent after any given number of steps $n < \infty$.

The identity (2.9) implies the following inequality, which plays the crucial role in our analysis:

$$\begin{aligned} |G_{\mathcal{G}}(x_0, y_0; E)| &\leq \sum_{\gamma \in \Gamma(x_0; y_0)} \epsilon^{|\gamma|} \exp \left[\sum_{j=0}^{|\gamma|} \ln |V(x_j; \omega) - \lambda_j(\gamma; E; \omega)|^{-1} \right]. \end{aligned} \quad (2.10)$$

2.5. From paths to ergodic theory. Given a self-avoiding path $\gamma = (x_0, \dots, x_n)$, introduce the sequence of random variables

$$X_j = X_j(\gamma; E; \omega) := \ln |V(z_j; \omega) - \lambda_j(\gamma; E; \omega)|^{-1},$$

$j = 0, 1, \dots, n$. Given a subpath $\gamma_j = (x_0, \dots, x_j)$ of γ , introduce the sigma-algebra \mathfrak{F}_j generated by $\{V(u; \cdot), u \in \mathcal{G} \ominus \gamma_{j-1}\}$. Clearly, $\mathfrak{F}_j \supseteq \mathfrak{F}_{j+1}$. It is readily seen that $\lambda_j(\gamma)$ is \mathfrak{F}_{j+1} -measurable, while $V(x_j; \cdot)$ (hence, $X_j(\gamma)$) is \mathfrak{F}_j -measurable.

Although the sequence (X_0, \dots, X_n) is not IID (not even independent), it will not be difficult to adapt standard methods of the large deviations theory to the sums

$$S_n = S_n(\gamma; E; \omega) := \sum_{j=0}^n X_j(\gamma; E; \omega).$$

In terms of $X_j(\gamma)$ and $S_n(\gamma)$, the inequality (2.10) takes the form

$$|G_{\mathcal{G}}(x_0, y_0; E)| \leq \sum_{\gamma \in \Gamma(x_0; y_0)} \epsilon^{|\gamma|} e^{S_n(\gamma; E; \omega)}. \quad (2.11)$$

Suppose that $\sup_{x,y} |\Gamma_n(x, y)| \leq e^{C_{BR}n}$ (here "BR" stands for "bridges"). Setting

$$m = -\ln \epsilon, \quad \tilde{m} = m - C_{BR},$$

and assuming $\tilde{m} > 0$, we obtain an expansion suitable for finite and infinite graphs with $|\Gamma_n(x, y)| \leq e^{C_{BR}n}$: allowing the value $+\infty$ for positive series, we have

$$\begin{aligned} |G_{\mathcal{G}}(x_0, y_0; E)| &\leq \sum_{n=0}^{\infty} e^{-mn} \sum_{\gamma \in \Gamma(x_0; y_0; n)} e^{S_n(\gamma; E; \omega)} \\ &\leq \sum_{n=0}^{\infty} e^{-\tilde{m}n} \max_{\gamma \in \Gamma_n(x_0; y_0)} e^{S_n(\gamma; E; \omega)} \end{aligned}$$

with the usual convention that the sum over the empty set of indices is zero. This occurs, for example, when $n < d(x_0, y_0)$; if the graph \mathcal{G} is finite, then the length of any SAP γ is bounded by $|\mathcal{G}| < \infty$, rendering finite the above sum over n .

2.6. Large Deviations Estimates. Some natural limitations. Fix an integer $n \geq 0$, two distinct points $x_0, y_0 \in \mathcal{G}$, and set for brevity $\Gamma_n = \Gamma_n(x_0; y_0)$. We have, for any $a > 0$,

$$\begin{aligned} &\mathbb{P} \left\{ e^{-\tilde{m}n} \max_{\gamma \in \Gamma_n} e^{S_n(\gamma; E; \omega)} > e^{an} \right\} \\ &= \mathbb{P} \left\{ \max_{\gamma \in \Gamma_n} S_n(\gamma; E; \omega) > (a + \tilde{m})n \right\}. \end{aligned}$$

Let us first discuss the problem at hand informally.

- If the ergodic theorem were applicable to the sums $\sum_{j=0}^n X_j$, then the probability in the above RHS would tend to 0 as $n \rightarrow \infty$, for any $a > \max_j \mathbb{E}[X_j]$.
- The cardinality of Γ_n may grow exponentially in n , even for d -dimensional lattices with $d > 1$. If the r.v. X_j had some finite exponential moment, this would not be a problem, at least in the strong disorder regime, for one could apply the standard LDE method based on Cramer's condition (finiteness of an exponential moment).
- It is not difficult to see (cf. the discussion in subsection 2.7) that Cramer's condition on X_i amounts to the Hölder continuity of the marginal PDF F_V of the IID random potential. However, our goal is to find a method which can also afford probabilistic bounds much weaker than exponential, at least under some additional hypotheses upon the combinatorial properties of the underlying graph \mathcal{G} .
- The (unwanted) events in the family

$$\left\{ \omega : \sum_{j=0}^n X_j > an \right\}, \quad \gamma \in \Gamma_n,$$

are highly correlated; at least, a large number of them may be. Yet, it is not quite obvious how to turn to our advantage such a high correlation.

Some of the above mentioned difficulties may be the price to pay for limiting the analysis to the simplest path expansions ("slim wormholes"), where at each step a subspace of dimension 1 is split off with the help of the Schur formula. Quite possibly, some more elaborate variants of this procedure (see Section 5) could bypass these difficulties.

Summarizing these observations, it seems difficult to assess directly the sums of the original variables X_j , unless they obey Cramer's condition. The general experience accumulated in the LDE theory suggests a different approach: replacing X_j by their truncated counterparts,

$$\tilde{X}_j(\omega) := X_j(\omega) \cdot \mathbf{1}_{\{X_j \leq b_n\}}(\omega),$$

with a judiciously chosen threshold $b_n < +\infty$. See the details in Appendix C.

In our case, we cannot restrict the truncation procedure to the sites occupied by one given path γ , since the random variables to be truncated depend upon the path through $\lambda_j(\gamma)$. What makes things even worse, and considerably, is that the number of relevant paths grows exponentially already in the periodic lattices \mathbb{Z}^d (even in the strip $\{0, 1\} \times \mathbb{Z}^1$!).

There is a particular situation where the above mentioned difficulty does not occur: the combinatorial explosion concerns the number of *bridges* between two distant points x_0, y_0 with $d(x_0, y_0) = n$, and *not the volume* of the ball $B_n(x_0)$. From this point of view, the trees (starting with \mathbb{Z}^1) represent the simplest case. A moderate number of loops can also be tolerated, if they result only in a tempered growth of the number of n -bridges. The slower grows the number of n -bridges as $n \rightarrow \infty$, and the lower regularity of the random potential can be tolerated.

Again, we see that the model where the Aizenman–Molchanov method excels in the localization analysis – the regular Anderson model on a Cayley tree – appears, with no surprise, in both categories which we have discussed above: in the case where F_V is Lipschitz- or Hölder-continuous (hence the Cramer's condition for X_j), even if the balls grow exponentially and the loops are present, and in the category of trees.

Remark 2.1. There is no point trying to hide the fact that we make use of the condition of tempered growth of n -bridges only out of necessity. Nevertheless, progress in nanophysics of the "quantum graphs" might lead to the study of some micro- or mesoscopic structures with moderate number of loops, where such a condition may be fulfilled. In relatively small (or moderately large) graphs, the distinction between an exponential and sub-exponential bound is merely the matter of choosing appropriate constants.

2.7. Right vs. left tails of X_j . Observe that the conditional right-tail probabilities $\mathbb{P}\{X_j > t \mid \mathfrak{F}_{j-1}\}$, albeit they are themselves random (dependent upon the condition), admit a common a.s. upper bound in terms of the continuity modulus \mathfrak{s}_V of the IID r.v. $V(\cdot; \omega)$: owing

to \mathfrak{F}_{j-1} -measurability of λ_j , we have

$$\begin{aligned} & \text{ess sup } \mathbb{P}\{X_j > t \mid \mathfrak{F}_{j-1}\} \\ &= \text{ess sup } \mathbb{P}\left\{\ln |V(z_j; \omega) - \lambda_j(\gamma; E; \omega)|^{-1} > t \mid \mathfrak{F}_{j-1}\right\} \\ &\leq \sup_{\lambda \in \mathbb{R}} \mathbb{P}\left\{|V(z_j; \omega) - \lambda|^{-1} < e^{-t}\right\} \\ &\leq \mathfrak{s}_V(2e^{-t}). \end{aligned}$$

Unlike the quantities $|V(z_j; \omega) - \lambda_j(\gamma; E; \omega)|^{-1}$ which cannot have finite moments of order $r \geq 1$ even for Lipschitz-continuous random potentials V , the new variables X_j may have finite moments of any fixed order, under a very weak assumption of log-Hölder continuity of the PDF F_V . The Hölder-continuity of F_V results, obviously, in finiteness of some *exponential* moment of X_j .

Specifically, if $\mathfrak{s}_V(t) \leq C |\ln |t||^{-c}$, $c > 0$, then for any $0 < r < c$, we have (cf. [16], Section V.6, Eqn (6.3))

$$\begin{aligned} & \int_0^\infty t^r dF_{X_j}(t) \\ &= r \int_0^\infty t^{r-1} (1 - F_{X_j}(t)) dt \\ &\leq r \int_0^\infty t^{r-1} \mathfrak{s}_V(2e^{-t}) dt \leq C' r + r \int_1^\infty t^{r-1} \mathfrak{s}_V(2e^{-t}) dt \\ &\leq C' r + C'' r \int_1^\infty \frac{t^{r-1}}{t^c} dt \leq C''' \frac{r}{c-r} < +\infty. \end{aligned}$$

The role of the *left* tails of the probability distribution of X_j is quite different: in the context of the *localization* analysis, the heavy *left* tails are more than welcome.

Note that

$$\mathbb{P}\{S_n > t\} \leq \mathbb{P}\{S_n^{(+)} > t\},$$

where

$$S_n^{(+)} = \sum_{j=1}^n X_j^{(+)},$$

$$X_j^{(+)} = \mathbf{1}_{\{X_j \geq 0\}} X_j = \ln_+ |V(z_j; \omega) - \lambda_j|^{-1}.$$

Clearly, this is *not* an efficient way to assess the large deviations, for one consciously loses the benefit of the lower (possibly negative!) value of the expectation $\mathbb{E}[X_j]$, replacing X_j by $X_j^{(+)}$. However, this simplifies a number of technicalities in the subsequent analysis. After this reduction, the behavior of the tail probabilities of the random potential V becomes irrelevant, and the finiteness (resp., divergence) of the moments of $S_n^{(+)}$ is determined only by the continuity modulus \mathfrak{s}_V .

3. SLIM WORMHOLES AND POSITIVITY OF LYAPUNOV EXPONENTS

Formally speaking, the Lyapunov exponents for the eigenfunctions of random operators can be defined in the context of one-dimensional, or quasi-one-dimensional, media, such as \mathbb{Z}^1 or strips of finite width. Nevertheless, one often considers the decay exponent of the Green's functions on lattices \mathbb{Z}^d and more general graphs, as a direct analog of the Lyapunov exponent(s). In this section, we will employ the slim wormhole expansions to obtain some simple, single-point criteria of positivity of such exponents.

We investigate some specific classes of models; the most important parameters distinguishing these classes are:

- the combinatorial characteristics of the underlying graph \mathcal{G} ;
- the local regularity properties of the marginal distribution of the IID random potential $V : \mathcal{G} \times \Omega \rightarrow \mathbb{R}$.

3.1. Graphs with exponential growth of balls and bridges. We call a bridge between two distinct points $x, y \in \mathcal{G}$ any SAP γ connecting these points. (Self-crossing bridges are indeed rare in real life.) Recall that the set of bridges of length n (n -bridges) between x and y was denoted by $\Gamma_n(x, y)$.

The case where $|\Gamma_n(x, y)|$ may grow exponentially in n is the most challenging for our method; it has *de facto* been addressed by Aizenman and Molchanov [2] and, with the help of a different version of the fractional-moment analysis, by Aizenman et al. [5]. Actually, for the methods of [2, 5] only the fact of exponentially bounded growth of balls is relevant (and represents, of course, the main difficulty that the MSA cannot overcome). We have to further distinguish between subclasses of this exponential class of graphs, depending on the asymptotic behavior of their number of n -bridges. This sets apart the trees, where specific recursive methods apply, as they do in the one-dimensional case, but also the “moderately looped” graphs, where one does not have the benefit of simple algebraic recursions, while the exponential explosion remains a combinatorial challenge.

Here we are forced to apply the classical assumption of the FMM: the Hölder regularity of the random potential. Note that our approach clarifies the true role of the Hölder regularity in the moment analysis. What might seem to be a lucky accident in the elegant Aizenman–Molchanov decoupling lemma, is actually interpreted as Cramer’s condition in the LDE component of our analysis. In the LDE theory, the special role of Cramer’s condition is well-understood: it is basically the only tool allowing to obtain *exponential* large deviations estimates.

Remark 3.1. While it does not follow from our analysis, it seems to be a reasonable **conjecture** that, under weaker regularity assumptions than Hölder-continuity of the marginal PDF F_V , the sub-exponential estimates provided by the MSA (or any reasonable improvements thereof) cannot be made exponential. I would be glad to be proven wrong on this point, if some more advanced variant of the MSA could establish exponential decay of eigenfunction correlators under weak regularity assumptions on $V(x; \cdot)$.

As to the FMM itself, it simply does not apply to the probability distributions not satisfying the Hölder-continuity condition.

In the next subsection, where the number of bridges will be required to grow sub-exponentially (or not at all, as on the tree graphs), we will show that the AM/FMM assumption of the Hölder continuity can be relaxed, depending upon the growth asymptotic of the number of bridges.

We stress that in this section we are concerned only with the proof of decay properties of the *Green’s functions*. The derivation of the almost sure p.p. spectrum will be the subject of Section 4.

The main statement of this subsection is the following Theorem 3.1. The result in itself is not new, since the model in question has been studied by Aizenman and Molchanov [2] and subsequent works on the FMM. However, we establish here an important link between the localization analysis and the classical large deviations theory.

Theorem 3.1. *Assume the following conditions:*

- (A) *The number of n -bridges in the graph \mathcal{G} satisfies the upper bound $|\Gamma_n| \leq e^{C_{BR} n}$, with $C_{BR} < +\infty$, (“BR” stands for “bridges”).*
- (B) *The common PDF F_V is Hölder-continuous of some order $\delta > 0$.*

Set

$$a := \inf_{\lambda \in \mathbb{R}} \int_{\mathbb{R}} \ln |v - \lambda|^{-1} dF_V(v). \tag{3.1}$$

Then there exists $c = c(F_V) \in \mathbb{R}$ such that for any $0 < |\epsilon| < e^{-a - C_{BR} - c}$, i.e., for $|\epsilon| = e^{-m}$ with $m > a + C_{BR} + c$,

$$\begin{aligned} \max_{y: d(x,y)=L} \mathbb{P} \left\{ |G_{B_L(x)}(x, y; E; \omega)| > \frac{e^{-(m-a)L}}{1 - e^{-(m-a)}} \right\} \\ \leq \frac{e^{-(m-c-C_{BR})L}}{1 - e^{-(m-c-C_{BR})}}. \end{aligned}$$

Proof. Step 1. Setting up the wormhole expansion in the ball $\mathcal{G} = B_L(x)$.

$$\begin{aligned} |G_{\mathcal{G}}(x, y; E; \omega)| \\ \leq \sum_{n \geq d(x,y)} \epsilon^n \sum_{\gamma \in \Gamma(x,y;n)} e^{S_n(\gamma)} \\ = \sum_{n \geq L} e^{-mn} \sum_{\gamma \in \Gamma_n} e^{S_n(\gamma)}. \end{aligned} \tag{3.2}$$

Let $m = \ln \epsilon^{-1}$, $\tilde{m} = m - C_{BR}$. Then for any $a \in (0, \tilde{m})$,

$$\begin{aligned} \mathbb{P} \left\{ |G_{\mathcal{G}}(x, y; E; \omega)| > \frac{e^{-(\tilde{m}-a)L}}{1 - e^{-(\tilde{m}-a)}} \right\} \\ \leq \sum_{n \geq L} \mathbb{P} \left\{ \sum_{\gamma \in \Gamma_n} e^{S_n(\gamma)} > e^{(a+\tilde{m})n} \right\}. \end{aligned} \tag{3.3}$$

Indeed, if neither of the events figuring in the RHS of (3.3) occurs, then $|G_{\mathcal{G}}(x, y; E; \omega)|$ is upper-bounded by the sum of the convergent geometric series.

Step 2. Paths and probabilities. By hypothesis (B), $|\Gamma_n| \leq e^{C_{BR} n}$, so for any n ,

$$\begin{aligned} \mathbb{P} \left\{ \sum_{\gamma \in \Gamma_n} e^{S_n(\gamma)} > e^{(a+\tilde{m})n} \right\} \\ \leq e^{C_{BR} n} \max_{\gamma \in \Gamma_n} \mathbb{P} \left\{ S_n(\gamma) > (a + \tilde{m})n \right\}. \end{aligned} \tag{3.4}$$

Therefore, the bound

$$\mathbb{P} \left\{ |G_{\mathcal{G}}(x, y; E; \omega)| > \frac{e^{-(\tilde{m}-a)L}}{1 - e^{-(\tilde{m}-a)}} \right\} \leq \frac{e^{-(m-c-C_{BR})L}}{1 - e^{-(m-c-C_{BR})}}$$

with some $c \in \mathbb{R}$ and $m > c + C_{BR}$, would follow, if we could prove, for all $\gamma \in \Gamma_n$ and the given value of c ,

$$\mathbb{P} \left\{ S_n(\gamma) > (a + \tilde{m})n \right\} \leq e^{-(m-c)n}. \tag{3.5}$$

Step 3. Large deviations along a fixed path.

We postpone the large deviations analysis until Appendix B; here we only state the final result, which is a reformulation of Proposition B.1 proven in Appendix B.

Proposition 3.2 (Cf. Proposition B.1). *Assume that the marginal PDF F_V of the IID random field $V : \mathcal{Z} \times \Omega \rightarrow \mathbb{R}$ is Hölder continuous of some order $\delta > 0$, and set*

$$a := \inf_{\lambda \in \mathbb{R}} \int_{\mathbb{R}} \ln |gv - \lambda|^{-1} dF_V(v). \tag{3.6}$$

Then there exist positive numbers $I(a) > 0$, $t^ > 0$ such that, for any $\tilde{m} > 0$, any $n \geq 1$ and any SAP γ of length n , with notations of this section, one has*

$$\mathbb{P} \{ S_n(\gamma) > (a + \tilde{m})n \} \leq e^{-n(I(a)+t^*\tilde{m})}. \tag{3.7}$$

Conclusion. The claim follows from the results of Steps 1–3. \square

Remark 3.2. It is worth mentioning that the exponential decay of the Green’s functions established by Theorem 3.1 does not reveal yet another lower bound on the ”mass” $m > 0$ required for the spectral localization. We prove that the maximum value of the Green’s functions is exponentially small with high probability, but localization requires that it remain small after multiplication by the surface of the L -sphere.

3.2. Tempered growth of bridges. Now we will allow for a lower regularity of the probability distribution of the random potential, loosing the benefit of Cramer’s condition in the large deviations analysis. The price to pay will be a stronger restriction on the growth rate of the number of n -bridges.

Perhaps, it would be appropriate to cite here the opening lines of the well-known paper by Dobrushin and Shlosman [14] (also devoted to the finite-volume criteria, in statistical mechanics), who in turn cite the opening lines of Lev N. Tolstoy’s roman ”Anna Karenina”: ”All happy families are alike, each unhappy family is unhappy in its own fashion”. Put simply, the universal happiness ends where Cramer’s condition ends (and the application of the FMM becomes impossible). To render things less dramatic, we abandon the very idea of presenting a reasonably complete set of results of the LDE theory and their applications to the path expansions in localization theory. Instead, in this subsection we make the weakest possible regularity assumption under which the MSA can still work in discrete, non-one-dimensional models.

The main result of this subsection is Theorem 3.3. Observe that the guaranteed rate of decay of the Green’s functions remains exponential, but without Hölder continuity of F_V , we have to be content with sub-exponential probabilistic bounds. In Theorem 3.3, we consider a particular case, where the balls may grow polynomially and the PDF F_V is log-Hölder continuous (of sufficiently high order). While the combinatorial part of adaptation to the graphs with the rate of growth of balls, intermediate between power-law and exponential, is straightforward, it is more difficult (actually, impractical) to fit the large deviations estimates in one, universal formula interpolating between these two extreme cases. This has been well-known in probability theory since the early results

by Linnik [25], who introduced several classes of probability laws, each requiring a specific adaptation of his general method. We refer to the review by Nagaev [27] and the bibliography given there; there are many more reviews and books available by now on this subject.

Theorem 3.3 does not establish⁴ Anderson localization; we only prove exponential decay of the Green’s functions. The novelty of this result, apart from the use of the large deviations theory, is that it applies, in particular, to a class of Anderson models on graphs of exponential growth which are ”moderately looped”; the latter property is expressed in terms of the number $|\Gamma_n(x, y)|$ of bridges of length n connecting two given points. For example, in a tree, $|\Gamma_n(x, y)| = 1$, while the balls may grow exponentially, e.g., in the Cayley tree \mathcal{G}_K with the constant coordination number $n_{\mathcal{G}_K} \equiv K + 1$, $K \geq 2$. This renders inapplicable the existing MSA techniques; on the other hand, the FMM does not apply either, if the common marginal distribution of the IID random potential is only log-Hölder (but not Hölder) continuous, so that the fractional moments diverge.

Theorem 3.3. *Assume that the number of n -bridges in the graph \mathcal{Z} admits a power-law bound:*

$$|\Gamma_n| \leq C_B n^d, \quad C_B, d < \infty, \tag{3.8}$$

and the continuity modulus of the PDF F_V of an IID random field $V : \mathcal{Z} \times \Omega \rightarrow \mathbb{R}$ on \mathcal{Z} , relative to some probability space $(\Omega, \mathfrak{B}, \mathbb{P})$, satisfies the following upper bound: for some $A > d + 1$ and $C_H < \infty$,

$$\forall s \in (0, 1/2) \quad \mathfrak{s}_V(s) \leq C_H |\ln s|^{-A}. \tag{3.9}$$

Then for any $x \in \mathcal{G}$ and L large enough, with some $\delta > 0$, $a \in \mathbb{R}$ (cf. (3.1)), and all $\epsilon = e^{-m} \in (0, e^{-a})$

$$\begin{aligned} & \max_{y: d(x,y)=L} \mathbb{P} \left\{ |G_{B_L(x)}(x, y; E; \omega)| > \frac{e^{-(m-a)L}}{1 - e^{-(m-a)}} \right\} \\ & \leq \frac{1}{L^{A-d-1-2\delta}} \xrightarrow{L \rightarrow \infty} 0. \end{aligned} \tag{3.10}$$

In view of the general setting used in the proof of Theorem 3.1, we only have to modify the large deviations estimate (3.7) used in **Step 3**, so the the main ingredient of the proof of Theorem 3.3 becomes the following

Theorem 3.4. *Assume that the number of n -bridges in the graph \mathcal{Z} admits a power-law bound:*

$$|\Gamma_n| \leq C_B n^d, \quad C_B, d < \infty, \tag{3.11}$$

and the continuity modulus of the PDF F_V of an IID random field $V : \mathcal{Z} \times \Omega \rightarrow \mathbb{R}$ on \mathcal{Z} , relative to some probability space $(\Omega, \mathfrak{B}, \mathbb{P})$ satisfies the following upper bound: for some $A > d + 1$ and $C_H < \infty$,

$$\forall s \in (0, 1/2) \quad \mathfrak{s}_V(s) \leq C_H |\ln s|^{-A}. \tag{3.12}$$

⁴Anderson localization actually follows from Theorem 3.3, if the growth balls in \mathcal{G} is also polynomially bounded. In this case, Steps 1 and 2 of the proof of Theorem 3.1, combined with the fairly standard large deviations analysis, results in a very short proof of the fixed-energy MSA-type bounds, under the assumption $|\Gamma_n(x, y)| \leq C_B n^d$.

Then, with notations of Section 2 and subsection 3.1, one has, for some $a \in \mathbb{R}$, $\delta > 0$ and all $L \geq L_0(\delta)$,

$$\sum_{n \geq L} \mathbb{P} \left\{ \max_{\gamma \in \Gamma_n} S_n(\gamma) > an \right\} \leq \frac{1}{L^{A-d-1-2\delta}} \xrightarrow{L \rightarrow \infty} 0. \tag{3.13}$$

Proof. We will see (cf. (3.16)) that the implicit requirement $L \geq L_0$ can be replaced by a more explicit, but also more cumbersome probability bound valid for any $L \geq 1$.

Fix $A > d + 1$ and pick some $\delta > 0$ such that $A > d + 1 + 2\delta$.

It follows from (3.12) that, for any $t > 0$, and uniformly in $\lambda \in \mathbb{R}$,

$$\begin{aligned} & \mathbb{P} \{ \ln |V(x; \omega) - \lambda|^{-1} > t \} \\ &= \mathbb{P} \{ |V(x; \omega) - \lambda| < e^{-t} \} \leq C_H t^{-A}. \end{aligned}$$

Denoting $X_+(\omega) = X_+(\omega; x) = \ln_+ |V(x; \omega) - \lambda| \geq 0$, we have for any $\alpha \in (0, A)$, uniformly in $x \in \mathcal{Z}$ and $\lambda \in \mathbb{R}$ (cf., e.g., [16], Section V.6, Eqn (6.3))

$$\begin{aligned} \mathbb{E} [X_+^\alpha] &= \int_0^{+\infty} t^\alpha dF_{X_+}(t) = \alpha \int_0^{+\infty} t^{\alpha-1} F_{X_+} dt \\ &\leq Const \int_0^{+\infty} t^{-(1+A-\alpha)} dt < +\infty. \end{aligned}$$

Now we set $\alpha = A - \delta > d + 1 + \delta$, so that $A - \alpha > 0$ and the above expectation converges.

Next, fix any path $\gamma \in \Gamma_n$ and set

$$X_j(\gamma) = \ln |V(x_i; \omega) - \lambda_i|^{-1}, \quad i = 0, \dots, n - 1 = |\gamma|,$$

and $S_n(\gamma) = \sum_j X_j(\gamma)$. Further let

$$X_i^+(\omega) = \ln_+ |V(x; \omega) - \lambda_i|^{-1}$$

and $S_n^+(\gamma) = \sum_{i=0}^{|\gamma|} X_i^+(\gamma)$. Since $S_n(\gamma) \leq S_n^+(\gamma)$, we have for any $a \in \mathbb{R}$

$$\mathbb{P} \{ S_n(\gamma) > an \} \leq \mathbb{P} \{ S_n^+(\gamma) > an \}.$$

This makes unnecessary any restrictions on the decay rate of the left-tail probabilities $\mathbb{P} \{ X_j < -t \}$, i.e., the tail probabilities $\mathbb{P} \{ |V(x; \omega)| > s \}$, as $s \rightarrow +\infty$.

From the hypothesis (3.11), we infer, therefore, that

$$\mathbb{P} \left\{ \max_{\gamma \in \Gamma_n} S_n(\gamma) > an \right\} \leq Cn^d \max_{\gamma \in \Gamma_n} \mathbb{P} \{ S_n^+(\gamma) > an \}. \tag{3.14}$$

Recall that we associated with each SAP γ the sequence of decreasing sigma-algebras \mathfrak{F}_i so that X_i (hence, X_i^+) is \mathfrak{F}_i -measurable, while λ_i is \mathfrak{F}_{i+1} -measurable. Note that

$$\begin{aligned} \mathbb{E} [|X_i^+|^\alpha | \mathfrak{F}_{i+1}] &= \mathbb{E} [\ln_+ |V(x_i; \omega) - \lambda_i(\omega)|^{-1} | \mathfrak{F}_{i+1}] \\ &\leq \sup_{\lambda \in \mathbb{R}} \mathbb{E} [\ln_+ |V(x_i; \omega) - \lambda|^{-1} | \mathfrak{F}_{i+1}] < +\infty. \end{aligned} \tag{3.15}$$

Now we need the following large deviation estimate proven in Appendix C (formulated there in a more general fashion, making no reference to graphs and paths).

Proposition 3.5 (Cf. Proposition C.1). *Let be given $d \in (0, +\infty)$. Assume that (3.15) is fulfilled with $\alpha > \max\{d + 1, 2\}$, and let*

$$\mu := \max_{1 \leq i \leq n} \sup_{\lambda \in \mathbb{R}} \mathbb{E} [\ln_+ |V(x_i; \omega) - \lambda|^{-1} | \mathfrak{F}_{i+1}].$$

Then for any $a > \mu$ there exists $\tilde{c} > 1$ such that for any path $\gamma \in \Gamma_n$

$$\mathbb{P} \{ S_n^+(\gamma) > an \} \leq \frac{\ln \tilde{c} n}{n^{\alpha-1}} \leq \frac{C'}{n^{\alpha-1-o(1)}}, \quad \text{as } n \rightarrow \infty. \tag{3.16}$$

Given $d > 0$, let $\alpha > d + 1$. For any $0 < \delta < \alpha - d - 1$ and some $n_0(\delta) < \infty$, we obtain for all $n \geq n_0(\delta)$, applying (3.16),

$$\mathbb{P} \{ S_n^+(\gamma) > an \} \leq \frac{C''}{n^{\alpha-1-\delta}}. \tag{3.17}$$

Collecting (3.14) and (3.17), the assertion follows:

$$\begin{aligned} & \sum_{n \geq L} \mathbb{P} \left\{ \sum_{\gamma \in \Gamma_n} e^{S_n(\gamma)} > e^{an} \right\} \\ & \leq C'' \sum_{n \geq L} \frac{n^d}{n^{\alpha-1-\delta}} \leq \frac{1}{L^{A-d-1-2\delta}} \xrightarrow{L \rightarrow \infty} 0. \end{aligned}$$

(the last inequality is valid for L large enough). \square

Replacing X_j by X_j^+ is the simplest, but not the most efficient way to bound $\mathbb{P} \{ S_n > an \}$, since this gives rise to a larger value of the expectation μ than for the original variables X_j . It is worth mentioning also that the requirement of finiteness of the expectation

$$\mathbb{E} [\ln_- |V(x; \omega) - \lambda|^{-1}] = \mathbb{E} [\ln_+ |V(x; \omega) - \lambda|]$$

is actually a very mild condition on the decay of the tail probabilities for the random potential V , since for any $t > 0$,

$$\mathbb{P} \{ \ln |V(x; \omega) - \lambda| > t \} = \mathbb{P} \{ |V(x; \omega) - \lambda| > e^t \}.$$

4. FROM WORMHOLE EXPANSIONS TO EXPECTATIONS

Most of the facts discussed in this section are well-known; we state them in order to render our presentation reasonably self-contained. In essence, the following meta-theorem holds true:

Fast decay of Green's functions implies strong dynamical and spectral localization.

Certainly, the above statement is too vague and too general to be true. However, presenting a large variety of particular cases where its formal, precise instance holds true, is beyond the scope of this work, focusing, as was pointed out in the Introduction, on the positivity of Lyapunov exponents. To find a compromise between brevity and logical completeness of the text, we describe below the principal steps, leading to the proof of the above statement in specific models. The necessary additional assumptions are introduced along the way, on the as-needed basis.

4.1. Eigenfunction correlators. Step 1. From fixed energy to an energy interval.

We start with the following statement, essentially going back to [15] and reformulated in [11, 10]. It is the core of one of the forms of the spectral reduction, inferring probabilistic localization bounds in an entire energy interval $I \subset \mathbb{R}$ from the pointwise estimates valid for each $E \in I$. Note that the parameter $a > 0$ below does not have the same meaning as in Section 3.

Lemma 4.1. [Cf. Theorem 2.5.1 in [10]] Let $L \geq 0$, $x \in \mathcal{G}$, $N := |\mathbb{B}_L(x)|$. Let $\{\lambda_j, j = 1, \dots, N\}$ be the eigenvalues of $H_{\mathbb{B}_L(x)}$ counting multiplicities, and $I \subset \mathbb{R}$ an interval of length $|I| < \infty$. Let be given positive numbers a, b, c, q_L such that

$$b \leq \min\{N^{-1}ac^2, c\} \quad (4.1)$$

and

$$\sup_{E \in I} \mathbb{P}\{M_x(E) > a\} \leq q_L. \quad (4.2)$$

There is an event $\mathcal{B}_x(b) \subset \Omega$ with $\mathbb{P}\{\mathcal{B}_x(b)\} \leq b^{-1}|I|q_L$ such that $\forall \omega \notin \mathcal{B}_x(b)$ the set

$$\mathcal{E}_x(2a; \omega) = \{E \in I : M_x(E) \geq 2a\} \subset \mathbb{R} \quad (4.3)$$

has Lebesgue measure $\text{mes}(\mathcal{E}_x(2a; \omega)) \leq b$ and is covered by a union of intervals $\mathcal{I}(\omega) = \cup_{j=1}^N I_j$, $I_j := \{E : |E - \lambda_j| \leq c\}$, $\lambda_j \in I$.

Moreover, since the function M_x is continuous outside its poles, the intersection $\mathcal{I}' := \mathcal{I} \cap \mathcal{E}_x(2a; \omega)$ actually is a finite union of intervals, $\mathcal{I}' = \cup_{i=1}^{N'} I'_i$, $N' < \infty$, of Lebesgue measure bounded by b .

Corollary 4.1 (Cf. Theorem 2.5.2 in [10]). Assume that $V : \mathcal{Z} \times \Omega \rightarrow \mathbb{R}$ is IID, and its marginal probability distribution admits continuity modulus $\mathfrak{s}_V(\cdot)$. Let $I \subset \mathbb{R}$ be a bounded interval, and consider a ball $\mathbb{B}_L(x)$ of some radius $L > 0$. Let $a(L), b(L), c(L), q(L)$ be some numbers satisfying (4.1) and such that

$$\sup_{E \in I} \mathbb{P}\{M_x(E; \omega) > a(L)\} \leq q(L).$$

Then

$$\begin{aligned} & \mathbb{P}\{\exists E \in I : \min(M_x, M_y) > a\} \\ & \leq |\mathbb{B}_L(x)|^2 \mathfrak{s}_V(2c(L)) + \frac{2|I|q(L)}{b(L)} \end{aligned} \quad (4.4)$$

Example 1. Assume that $|\mathbb{B}_L| \leq C_d L^d$, for some $d, C_d \in (0, +\infty)$. Then (4.1) is satisfied with

$$\begin{aligned} a(L) &= L^{-\frac{p}{5}}, \quad b(L) = L^{-\frac{4p}{5}}, \\ c(L) &= (C_d)^{\frac{d}{2}} L^{-\frac{p}{5}}, \quad q(L) = L^{-p}, \end{aligned}$$

provided that $p > 5d$.

Example 2. In the case where $|\mathbb{B}_L| \leq e^{\alpha L}$, $\alpha \leq \frac{1}{2}\mu$, for some $\mu > 0$, inequality (4.1) is satisfied with

$$\begin{aligned} a(L) &= e^{-\frac{1}{2}\mu L}, \quad b(L) = e^{-\frac{1}{2}\mu L}, \\ c(L) &= e^{-\frac{1}{4}\mu L}, \quad q(L) = e^{-\mu L}. \end{aligned} \quad (4.5)$$

If we assume in addition that $\mathfrak{s}_V(s) \leq C_H s^\delta$, $\delta \in (0, 1]$, then

$$\begin{aligned} & \mathbb{P}\{\exists E \in I : \min(M_x, M_y) > a\} \\ & \leq e^{-(\frac{\delta\mu}{4} - 2\alpha)L} + 2|I|e^{-\mu L + \frac{1}{2}\mu L} \\ & \leq \max(2, 4|I|)e^{-(\frac{\delta\mu}{4} - 2\alpha)L}. \end{aligned} \quad (4.6)$$

Of course, the obtained bound is efficient when $\mu > 8\delta^{-1}\alpha$. With the Hölder exponent δ fixed, this means that either μ is to be sufficiently large (given the rate of growth of balls $\alpha < \infty$), or the the rate of growth of balls α is small enough (with given Lyapunov exponent $\mu > 0$). With $\alpha = 0$ (sub-exponential graphs), any $\mu > 0$ suffices, for L large enough.

Step 2. Using the Bessel inequality (cf. [21].)

The following statement is an adaptation of the Germinet–Klein argument [21] to finite subgraphs of a countable graph. Its assertion is logically independent of Theorem 4.1, so we are free to allow here any interval $I \subset \mathbb{R}$. However, to avoid confusing differences between the assumptions of various statements given in this section, we prefer to address here only the eigenfunction correlators in an arbitrarily large, but bounded (and fixed) interval I . Naturally, this setting suffices for the localization analysis of a.s.-uniformly bounded random operators, including discrete Schrödinger operators with a.s.-uniformly bounded random potential.

One possible extension to the case where both the random potential and the interval I are allowed to be unbounded, is discussed in [10], where the tail probabilities are required to satisfy the power-law bound

$$\mathbb{P}\{|V(x; \omega)| > t\} \leq Ct^{-A'}, \quad A' > 0$$

The main idea is that the maximum value $|V(x, \omega)|$ for $x \in \mathbb{B}_L(u)$ (hence, the largest eigenvalue in $\mathbb{B}_L(u)$) can be easily upper-bounded in probability, and then the upper bound on the eigenfunction correlators can be obtained with the help of Lemma 4.1 in a suitable chosen interval $I = I_L = [-R(L), R(L)]$.

Below we denote by $\mathcal{B}_1(I)$ the set of all complex-valued Borel functions ϕ with $\text{supp } \phi \subset I$ and $\|\phi\|_\infty \leq 1$.

Theorem 4.1 (Cf. Theorem 2.5.3 in [10]). Assume that, in the graph \mathcal{Z} , the rate of growth of spheres $S_L(x) = \{y \in \mathcal{Z} : d(x, y) = L\}$ satisfies some uniform bound,

$$\sup_{x \in \mathcal{Z}} \text{card } S_L(x) \leq \mathbf{S}(L).$$

Further, fix an energy interval $I \subset \mathbb{R}$ and assume that for given balls $\mathbb{B}_L(x), \mathbb{B}_L(y)$ with $d(x, y) \geq 2L + 1$ and some $h(L) \in [0, 1]$ one has

$$\mathbb{P}\{\exists E \in I : \min(M_x, M_y) > a(L)\} \leq h(L).$$

Then for any finite subgraph $\mathcal{Z} \supset \mathcal{G} \supset (\mathbb{B}_L(x) \cup \mathbb{B}_L(y))$ one has

$$\mathbb{E}\left[|\langle \delta_x | \phi(H_{\mathcal{G}}) | \delta_y \rangle|\right] \leq 4\mathbf{S}(L)a(L) + h(L). \quad (4.7)$$

Corollary 4.2. Assume that the growth of balls in a graph \mathcal{Z} is exponentially bounded and that the marginal PDF F_V of an IID random field $V : \mathcal{Z} \times \Omega \rightarrow \mathbb{R}$ is Hölder continuous. Then for $\epsilon > 0$ small enough, the eigenfunction correlators for the operator $H = -\epsilon\Delta + V(\omega)$ decay exponentially fast: for some $m(\epsilon) > 0$,

$$\mathbb{E}\left[\sup_{\phi \in \mathcal{B}_1(I)} |\langle \delta_x | \phi(H) | \delta_y \rangle|\right] \leq Ce^{-m(\epsilon)d(x, y)}.$$

Proof. It suffices to combine (4.7) and Lemma 4.1 with parameters $a(L), b(L), c(L), q(L)$ chosen as in (4.5). Since $\mathbf{S}(L) \leq e^{\alpha L}$ for some $\alpha < \infty$ and the decay exponent for the Green's functions is lower-bounded by $|\ln \epsilon| - \text{Const}$, taking $\epsilon > 0$ small enough, we can have $a(L) = e^{-\frac{1}{2}\mu L}$ with arbitrarily large $\mu > 2\alpha$. At the same time, $h(L)$ in the RHS of (4.7) also admits an exponentially small upper bound, if ϵ is small enough. \square

Corollary 4.3. Assume that the growth of balls in the graph \mathcal{Z} is polynomially bounded,

$$|\mathbb{B}_L(x)| \leq C_d L^d, \quad L \geq 1,$$

and that the marginal PDF F_V of an IID random field $V : \mathcal{Z} \times \Omega \rightarrow \mathbb{R}$ is log-Hölder continuous:

$$\mathfrak{s}_V(s) \leq C |\ln s|^{-A}.$$

Then for any $B > 0$, with $A > 0$ large enough and ϵ small enough, the eigenfunction correlators for the operator $H = -\epsilon\Delta + V(\omega)$ admit a power-law decay bound:

$$\mathbb{E} [|\langle \delta_x | \phi(H) | \delta_y \rangle|] \leq C (d(x, y))^{-B}.$$

Proof. The claim follows from the estimate (3.13), where the key probabilities decay at least as $O(L^{-A+d+2})$, and $A - d - 2 \rightarrow +\infty$ as $A \rightarrow \infty$. \square

Naturally, the regularity interpolating between the log-Hölder continuity and Hölder continuity results (e.g., in the graphs with a power-law bound on the volume of balls) in the decay bound on eigenfunction correlators intermediate between the power-law and exponential one. For example, assuming that the continuity modulus satisfies

$$\mathfrak{s}(s) \leq \frac{C}{|\ln s|^{c|\ln s|}}, \quad c > 0,$$

one can prove strong dynamical localization with rate

$$\mathbb{E} [|\langle \delta_x | e^{-itH} | \delta_y \rangle|] \leq e^{-\ln^{1+a} d(x, y)}, \quad a > 0,$$

hence, strong dynamical localization of all orders $b > 0$:

$$\forall b > 0 \quad \mathbb{E} [|\langle \delta_x | e^{-itH} | \delta_y \rangle|] \leq C(b) (1 + d(x, y))^{-b}.$$

On the other hand, one can also trade higher regularity of the random potential for higher rate of growth of balls. It is clear from Eqn (4.7) that, in the graphs with exponential growth of balls, one needs exponential bounds on the probabilities of large deviations to overcome the exponentially large factor $\mathbf{S}(L)$, so our technique, just as the FMM, requires Hölder continuity of the marginal distribution for the proof of dynamical localization in this class of models.

Step 3. From finite subgraphs to the entire configuration space.

Only finite (but large) configuration spaces are relevant for the application to the physical models of localization/quantum transport (except, perhaps, for general cosmology ...). The actually infinite models may present only a mathematical challenge. Fortunately enough, the transition from large (but finite) to infinite models is quite easy, once some uniform bounds on the eigenfunction correlators are established in arbitrarily large finite configuration spaces.

Fix a graph \mathcal{G} and an interval $I \subseteq \mathbb{R}$. Assume that for any pair $x, y \in \mathcal{Z}$ and for all sufficiently large subgraphs $\mathcal{G} \subset \mathcal{Z}$ containing x and y , for some function $f : \mathbb{N} \rightarrow \mathbb{R}$ and any $f \in \mathcal{B}_1(I)$ we have

$$\mathbb{E} [|\langle \delta_x | \phi(H_{\mathcal{G}}) | \delta_y \rangle|] \leq f(d(x, y)). \quad (4.8)$$

The derivation of the analog of (4.8) for $\mathcal{G} = \mathcal{Z}$ is well-known by now; see [5, 4]; it is based on the Fatou lemma, applied to the (real-valued) spectral measures $\nu_{\mathcal{G}}^{(x, y)}$, associated with the eigenfunction correlators and defined for compactly supported continuous (or bounded Borel) functions ϕ by

$$\int \phi(\lambda) d\nu_{\mathcal{G}}^{(x, y)}(\lambda) = \langle \delta_x | \phi(H_{\mathcal{G}}) | \delta_y \rangle;$$

here, it is assumed that $x, y \in \mathcal{G}$. The operators $H_{\mathcal{G}}$ converge in the strong resolvent sense as $\mathcal{G} \uparrow \mathcal{Z}$ (cf. [24]), so the measures $\nu_{\mathcal{G}}^{(x, y)}$ converge vaguely to the respective spectral measure $\nu_{\mathcal{Z}}^{(x, y)}$ ($= \nu_{\mathcal{Z}}^{(x, y)}$) for the operator $H = H_{\mathcal{Z}}$. By the Fatou lemma combined with standard approximation arguments, one can conclude that

$$\mathbb{E} [|\langle \delta_x | \phi(H) | \delta_y \rangle|] \leq f(d(x, y)).$$

See the details in [5, 4].

A more direct approach to strong dynamical localization in an infinite configuration space is also known. In fact, as was pointed out, Theorem 4.1 is itself an adaptation of the result by Germinet–Klein [21] which applies directly to the infinite systems. This requires an a priori upper bound on the growth of generalized spectral projections for H on \mathcal{Z} . See details in [21]. Recall that the techniques by Germinet–De Bièvre [20] and Damanik–Stollmann [12] also work in the infinite configuration space (and also require some a priori functional-analytic bounds).

4.2. Towards the spectral localization. By the well-known RAGE (Ruelle–Amrein–Georgescu–Enss) theorems, a sufficiently rapid⁵ decay rate of the eigenfunction correlators implies a.s. point spectrum (spectral localization). However, an efficient, quantitative estimate on the decay rate of the eigenfunctions requires additional arguments, first of all a priori bounds of Shnol–Simon type on the *growth rate* of the generalized eigenfunctions. The usual procedure goes back to the seminal work by Fröhlich et al. [18], and in a reformulated form, to the paper by von Dreifus and Klein [13].

Lemma 4.1 suffices for assessing the eigenfunction correlators (which is indeed the main axis of the present paper. However, it has to be emphasized that the above approach does not provide the optimal bound on the eigenfunction decay. One of the alternatives to Lemma 4.1, based on a different analytic idea, was proposed in our work [10], but we are not going to discuss further this issue.

5. OTHER TYPES OF EXPANSIONS

5.1. More general stopping rules. It is clear that, in the recursive procedure described in Sections 2–3, one can set up more general rules to decide, which paths γ starting at $x_0 \in \mathcal{G}$ are to be expanded at a given length $n \geq 0$. Naturally, these rules have to be consistent with the recursion: given a path γ , either *all* or *none* of its descendants are to be retained in a more elaborate expansion, dictated by the stopping rule. (The latter term is inspired by a similar notion of stopping rule, or Markov moment, used in the theory of Markov processes and martingales.)

Definition 5.1. Let be given a graph \mathcal{G} and a function $V : \mathcal{G} \rightarrow \mathbb{R}$. Consider the Hamiltonians $H_{\mathcal{G}'} = -\Delta_{\mathcal{G}'} + V_{\mathcal{G}'}$, $V_{\mathcal{G}'} := V|_{\mathcal{G}'}$. We call a stopping rule any function $\tau : \Gamma_{\mathcal{G}} \rightarrow \mathbb{N}$ such that for any $\gamma \in \Gamma_{\mathcal{G}}$,

- (i) either $\tau(\gamma') = |\gamma|$ for all descendants $\gamma'(\supset \gamma)$ of γ ,
- (ii) or $\tau(\gamma') > |\gamma|$ for all descendants $\gamma' \supset \gamma$.

⁵The required rate of decay also depends on the rate of growth of balls.

In the case (i), the path γ is stopped at the moment $\tau(\gamma) = |\gamma|$, and all its descendants are suppressed by the stopping rule τ . Respectively, in the case (ii) the stopping rule allows *all descendants* $\gamma' \supset \gamma$ with $|\gamma'| = |\gamma| + 1$, i.e., *at least one more step* is allowed by τ for the descendants of γ , in *all* possible directions.

A simple example of a stopping rule is given by the function

$$\tau : \gamma \mapsto |\gamma| \wedge n \equiv \max(|\gamma|, n),$$

with $n \geq 0$. This stopping rule had already appeared implicitly in the previous subsection: it stops all paths after exactly n steps, regardless of the shape of the path.

Now we can generalize the wormhole expansions: given any stopping rule τ , we have the identity for all $E \in \mathbb{C}_+ \cup \mathcal{R}(\mathcal{G}; V)$ (i.e., for complex or allowed real energies)

$$\begin{aligned} G_{\mathcal{G}}(x_0, y_0; E) &= \sum_{\gamma \in \Gamma(x_0)} G_{\mathcal{G} \ominus \gamma}(x, y_0; E) \times \\ &\times \prod_{j=0}^{\tau(\gamma)} \frac{\epsilon}{V(x_j) - \tilde{E}(x_j; \mathcal{G} \ominus \gamma_{j-1}; E)}. \end{aligned} \quad (5.1)$$

The summation in the RHS is formally carried over all self-avoiding paths, even infinite if and when \mathcal{G} is infinite, but only the finite subpath

$$\gamma_{\tau} := \{\gamma(t), t = 0, \dots, \tau(\gamma)\} \subset \gamma$$

contributes to the RHS. (Recall: by definition, a stopping rule takes integer – hence, finite – values.)

In the present paper, we do not develop an elaborate diagrammatic technique, so introducing general stopping rules may appear superfluous and not quite natural. However, it seems equally unnatural to limit the wormhole expansions to those corresponding to stopping rules of the form $\tau(\gamma) = f(|\gamma|)$. We merely point out that rigorous – and quite simple – identities can be easily obtained, in a large variety of forms, for the Green's functions of Anderson-type Hamiltonians on arbitrary (locally finite) countable graphs. The choice of a particular stopping rule may be dictated by a specific model.

5.2. Fat wormholes. One obvious generalization of the slim wormhole expansion is obtained by splitting off at each step a subspace in $\ell^2(\mathcal{G})$ of dimension large than 1, i.e., taking the decomposition of the form $\mathcal{G} = \mathcal{G}_A \sqcup \mathcal{G}_D$ with $|\mathcal{G}_A| > 1$. For example, in the procedure presented in Sections 2–3, one can replace $\{x_0\} = B_0(x_0)$ by $B_{\ell}(x_0) \cap \mathcal{G}$ with $\ell > 0$. The scalar factors $|V(x_j) - \lambda_j|^{-1}$ are to be replaced by the respective local resolvents, making the estimates less explicit. However, this can be used in the framework of numerical analysis and provide a better description of the zone in the parameter space where localization occurs.

The “fat” version of the wormhole expansions were not discussed in [23], where the starting point is the perturbative, formal random walk expansion which is to be regularized with the help of the loop elimination. The Schur complement formula seems to be the easiest way to obtain such an expansion.

5.3. The Schur sandwiches. Consider now the case where \mathcal{G} is partitioned into three subgraphs,

$$\mathcal{G} = \mathcal{G}_{D_1} \sqcup \mathcal{G}_{A_1} \sqcup \mathcal{G}_{D'_2}$$

such that the complement of \mathcal{G}_{A_1} is itself disjoint:

$$\mathcal{G} \setminus \mathcal{G}_{A_1} = \mathcal{G}_{D_1} \sqcup \mathcal{G}_{D'_2}.$$

In other words, the subgraph \mathcal{G}_A is sandwiched between the components \mathcal{G}_{D_1} and $\mathcal{G}_{D'_2}$ which it separates completely. Applying the Schur decomposition formula (2.1) with $A = H_{\mathcal{G}_{D_1}}$, $D = H_{\mathcal{G}_{D_1} \cup \mathcal{G}_{D'_2}}$, we obtain for any $x \in \mathcal{G}_{D_1}$ and $y \in \mathcal{G}_{D'_2}$:

$$\begin{aligned} G_{\mathcal{G}}(x, y; E) &= G_D(x, y; E) + \left[G_D C M_{A_1}^{-1} B G_D \right](x, y) \\ &= 0 + \left[G_{D_1} C M_{A_1}^{-1} B G_{D'_2} \right](x, y) \end{aligned} \quad (5.2)$$

where

$$M_{A_1} = A - B(D - E)^{-1}C = gV|_{\mathcal{G}_{A_1}} - E - K(E)$$

and $K(E) = K(E; \omega)$ is measurable with respect to $\mathfrak{F}_{\mathcal{G}_{D_1} \cup \mathcal{G}_{D'_2}}$. This procedure can be iterated. For example, after two steps, using the decomposition

$$\mathcal{G} = \mathcal{G}_{D_1} \sqcup \mathcal{G}_{A_1} \sqcup \left(\mathcal{G}_{D_2} \sqcup \mathcal{G}_{A_2} \sqcup \mathcal{G}_{D'_3} \right), \quad (5.3)$$

we obtain the representation for $x \in \mathcal{G}_{D_1}$, $y \in \mathcal{G}_{D'_3}$

$$G_{\mathcal{G}}(x, y; E) = \left[G_{D_1} C M_{A_1}^{-1} B G_{D_2} C M_{A_2}^{-1} B G_{D'_3} \right](x, y), \quad (5.4)$$

and so on.

The “sandwiched” decomposition (5.2) is similar in spirit to the one introduced by Aizenman et al. [5], but, owing to the use of the Schur formula instead of the second resolvent equation, the middle component $G_{\mathcal{G}_{A_1}}$ is not the full resolvent $G_{\mathcal{G}}$ as it was in [5]. In (5.2), it is a modified resolvent of a small⁶ subsystem \mathcal{G}_{A_1} . It is, of course, rather implicit, but for the purposes of the resonance analysis (proximity of the spectrum of $gV|_{\mathcal{G}_{A_1}} - E - K(E)$ to E), it is basically as good as the explicit multiplication operator $gV|_{\mathcal{G}_{A_1}} - E$, conditional on the random potential outside \mathcal{G}_{A_1} . Indeed, conditional on $\mathfrak{F}_{\mathcal{G}_{D_1} \cup \mathcal{G}_{D'_2}}$, which renders $K(E; \omega)$ nonrandom, the distribution of the eigenvalues of M_{A_1} is at least as regular as for $K(E) = 0$, for only the regularity of the diagonal elements of M counts, with $K(E)$ fixed.

Note also that in (5.4), the resolvents \mathcal{G}_{D_1} , \mathcal{G}_{D_2} and $\mathcal{G}_{D'_3}$ are independent and explicit, particularly when $|\mathcal{G}_{D_1}| = |\mathcal{G}_{D_2}| = 1$ ($\mathcal{G}_{D'_3}$ is to be decomposed further), while \mathcal{G}_{A_1} and \mathcal{G}_{A_2} become conditionally independent (albeit not necessarily identically distributed), given the sigma-algebra $\mathfrak{F}_{\mathcal{G}_{D_1} \cup \mathcal{G}_{D_2} \cup \mathcal{G}_{D'_3}}$.

We plan to address in a forthcoming paper the case where $|\mathcal{G}_{D_j}| = 1$, $j = 1, 2, \dots, n$ (“slim sandwiched expansion”).

6. CONCLUSION

The two most popular methods of the mathematical theory of Anderson localization, the Multi-Scale Analysis and the Fractional Moment Method, have a large common domain of application, but also exclusive areas where each of them is the indisputable champion.

⁶Ultimately, the middle component of the sandwiched resolvent is reduced in [5] to a two-point subsystem, with the help of the Krein formula. Cf. also [2].

It is a legitimate question, what are the natural boundaries for the FMM techniques (or rather general philosophy), in terms of the regularity of the underlying probability distribution. In the present paper, we give only some partial answers in this direction, but hope that a more systematic use of the methods of ergodic theory can be beneficial for extending the general FMM's strategy to a larger class of models where the fractional moments *per se* do not exist.

It is also intriguing to find out, if the MSA's strategy can be improved⁷ so as to apply to the exponential graphs, or to achieve exponential bounds on the eigenfunction correlators. In the author's opinion, which may of course be incorrect, in the case where the random potential is not Hölder continuous, there may be some natural limits to the decay rate of eigenfunction correlators, independent of the techniques employed, due to the limitations following from the accurate large deviations analysis of rare, but inevitable resonances.

It seems likely that more general results could be obtained by the synthesis of the "mono-scale" (FMM-type) and the multi-scale approaches.

APPENDIX A. PROGRESSIVE MEASURABILITY ISSUES

Given an SAP $\gamma = (x_0, \dots, x_n)$, $n = |\gamma| \geq 1$, we have a decreasing sequence of subgraphs (cf. (2.7))

$$\mathcal{Y}_i = \mathcal{G} \ominus \{x_0, \dots, x_i\}, \quad 0 \leq i \leq n,$$

where, for any given subset $\mathcal{X} \subset \mathcal{G}$, $\mathcal{G} \ominus \mathcal{X}$ is defined as the connected subgraph of $\mathcal{G} \setminus \mathcal{X}$ containing x_n . Clearly, $\mathcal{Y}_0 \supseteq \mathcal{Y}_1 \cdots \supseteq \mathcal{Y}_n$. Further, we have a decreasing sequence of sigma-algebras $\mathfrak{F}_i = \sigma[V(x; \cdot), x \in \mathcal{Y}_i]$ and the random variables

$$\lambda_i = \lambda_i(\gamma) := \tilde{E}(x_i; \mathcal{Y}_i; E)$$

where the "self-energy" $\tilde{E}(x_i; \mathcal{G} \ominus \gamma_i; E)$ is defined as in (2.6). We will never use any property of λ_j other than "regressive" measurability with respect to the filtration \mathfrak{F}_\bullet :

$$\forall i = 0, \dots, n, \quad \lambda_i(\gamma; E; \cdot) \text{ is } \mathfrak{F}_{i+1}\text{-measurable.}$$

(Usually, one deals in probability theory with "progressive" measurability with respect to a growing family – "filtration" – of sigma-algebras.) At the same time, $V(x_i; \omega)$ are \mathfrak{F}_i -measurable. We will say that $(V(x_i; \cdot))_{i \geq 0}$ are adapted to $(\mathfrak{F}_i)_{i \geq 0}$. Clearly, the random variables

$$X_i(\gamma; \omega) = \ln |V(x_i; \omega) - \lambda_i(\gamma; \omega)|^{-1}$$

are also adapted to $(\mathfrak{F}_i)_{i \geq 0}$.

The above notations and facts will be used below without further notice.

APPENDIX B. LARGE DEVIATIONS UNDER CRAMER'S CONDITION

In Proposition B.1, we adapt a classical argument, historically applied first to independent random variables, to the sequence $\{X_j\}$ adapted to the sigma-algebras $\{\mathfrak{F}_i\}$. There is a vast literature devoted to the large deviations for the martingales, but in our situation the variables at hand do not necessarily form a martingale.

It is certainly possible to introduce a martingale, using the so-called compensator process, but a straightforward adaptation of Cramer's method is short enough and more transparent. The final result, of course, is not new, although it would be difficult to make a direct, formal reference to a published paper or monograph, without any adaptation.

In Section 5, we use Proposition B.1 in the particular case where X_1, \dots, X_n are independent (this is obviously allowed by the hypotheses).

Proposition B.1. *Let X_1, \dots, X_n , $n \geq 1$, be random variables adapted to a decreasing family of sigma-algebras \mathfrak{F}_i , $i = 0, \dots, n$. Assume the following properties:*

- (1) *There exists some $A_1 < \infty$ such that for all $i \geq 1$, $\mathbb{E}[|X_i| | \mathfrak{F}_{i-1}] \leq A_1$, thus*

$$\exists \mu \in \mathbb{R} : \mu_i := \mathbb{E}[X_i | \mathfrak{F}_{i-1}] \stackrel{\text{a.s.}}{\leq} \mu.$$

- (2) *There exists some $t_0 > 0$ and $A_2 < \infty$ such that for all $i \geq 1$, $\mathbb{E}[e^{t_0 X_i} | \mathfrak{F}_{i-1}] \leq A_2$, and, consequently, the random variables*

$$\psi_i : t \mapsto \mathbb{E}[e^{t X_i} | \mathfrak{F}_{i-1}] \tag{B.1}$$

are a.s. well-defined and uniformly bounded for $t \in [0, t_0]$.

Then for any $a > 0$ there exists $I(a) > 0$ such that

$$\mathbb{P}\left\{ \sum_{i=1}^n X_i \geq (\mu + a)n \right\} \leq e^{-nI(a)}.$$

Proof. Fix any $t_1 \in (0, t_0)$ and consider on $[0, t_1]$ random functions $\varphi_i = (\ln \psi_i)'$, i.e.,

$$\begin{aligned} \varphi_i : t &\mapsto \left(\ln \mathbb{E}[e^{t X_i} | \mathfrak{F}_{i-1}] - t \mathbb{E}[X_i | \mathfrak{F}_{i-1}] \right), \quad t \in [0, t_1], \\ \varphi : t &\mapsto \max_{1 \leq i \leq n} \varphi_i(t). \end{aligned}$$

The conditional probability distribution of X_i , given \mathfrak{F}_{i+1} , is a (random) probability measure on \mathbb{R} , depending upon the condition; for brevity, we denote it by \mathbb{P}_i . The respective conditional expectation is denoted by $\mathbb{E}_{i+1}[\cdot]$.

Clearly, for each i , $\varphi_i(0) = 0$, and its derivative

$$\varphi'_i : t \mapsto \frac{\mathbb{E}_{i+1}[X_i e^{t X_i}]}{\mathbb{E}_{i+1}[e^{t X_i}]} - 1$$

is well-defined and continuous in $[0, t_1]$, and we have

$$\varphi'_i(0; \omega) \stackrel{\text{a.s.}}{=} 0.$$

Therefore,

$$\forall a > 0 \quad I_i(a; \omega) := \sup_{t \in [0, t_1]} (at - \varphi_i(t; \omega)) \stackrel{\text{a.s.}}{>} 0,$$

since the function $t \mapsto at - \varphi_i(t)$ is continuously differentiable in $[0, t_1]$, vanishes at $t = 0$ along with its first derivative, and has positive second derivative at this point, thus grows in some interval $(0, t_1)$, $t_1 > 0$.

Due to the continuity of the function $t \mapsto at - \varphi_i(t)$ in $[0, t_1]$, one can safely replace $\sup_{[0, t_1]}$ by $\max_{[0, t_1]}$, and we have

$$t_i^*(a) := \arg \max_{t \in [0, t_1]} (at - \varphi_i(t)) > 0. \tag{B.2}$$

Furthermore, due to the a.s.-uniform boundedness of the exponential moments $\mathbb{E}_{i+1}[e^{t_0 X_i}]$, the second

⁷I thank Abel Klein for fruitful exchanges on this subject.

derivatives φ_i'' also admit an a.s.-uniform upper bound, hence t_i^* admit a uniform lower bound:

$$\exists t^*(a) > 0 : \forall i \geq 1 \quad t_i^*(a) \stackrel{\text{a.s.}}{\geq} t^*(a) > 0.$$

In turn, this results in the following a.s.-uniform lower bound:

$$\forall a > 0 \quad \exists I(a) > 0 : \forall i \geq 1 \quad I_i(a) \stackrel{\text{a.s.}}{\geq} I(a).$$

Set $R_i = X_i + \dots + X_n$, $1 \leq i \leq n$, so that $S_{i-1} + R_i = S_n$. (We always use the standard convention that the sum over the empty set of indices is zero.) Now assess the Laplace transform of $S_n(\omega)$ as follows:

$$\begin{aligned} \mathbb{E} [e^{tS_n}] &= \mathbb{E} [e^{tX_1 + tR_2}] = \mathbb{E} [\mathbb{E}_{i+1} [e^{tR_2} e^{tX_1}]] \\ &= \mathbb{E} [e^{tR_2} \mathbb{E}_{i+1} [e^{tX_1}]] \\ &= \mathbb{E} [e^{tR_2} \mathbb{E}_{i+1} [e^{t\mu_1} e^{t(X_1 - \mu_1)}]] \end{aligned}$$

(recall: $\mu_i := \mathbb{E}[X_i | \mathfrak{F}_{i-1}]$)

$$\leq \mathbb{E} [e^{tR_2}] \cdot \text{ess sup} (e^{t\mu_1} \mathbb{E}_{i+1} [e^{tX_1}])$$

(by the assumption (B.1))

$$\leq e^{\varphi_1(t) + t\mu} \mathbb{E} [e^{tR_2}].$$

Similarly, we obtain by recursion

$$\mathbb{E} [e^{tS_n}] \leq e^{\sum_{i=1}^n (\varphi_i(t) + t\mu)}$$

and by Chebyshev's inequality, we conclude that

$$\begin{aligned} \mathbb{P} \{ S_n \geq (\mu + a)n \} &\leq \exp \left\{ - \sum_{i=1}^n (at - \varphi_i(t)) \right\} \\ &\leq e^{-(n+1)I(a)}, \quad I(a) > 0. \end{aligned}$$

□

The most important point in the above assertion is that the decay exponent $I(a)$ of the probability of large (linear in n) deviations is *strictly positive* for an *arbitrarily small* excess $a > 0$ of the expectation. As to the large values of $a > 0$, taking $t^* = t^*(a_0)$ for some $a_0 > 0$, as in (B.2), we see that, owing to the strict positivity of t^* , one has

$$\begin{aligned} I(a_0 + a) &:= \sup_{t \in [0, t_1]} ((a_0 + a)t - \varphi(t)) \\ &\geq (a_0 + a)t^* - \varphi(t^*) \\ &= I(a_0) + at^* \xrightarrow{a \rightarrow +\infty} +\infty. \end{aligned}$$

APPENDIX C. LARGE DEVIATIONS FOR HEAVY-TAILED SUMS

Here we use essentially, and adapt where necessary, the technique from [27], where the sums of *independent* random variables were considered. Again, it is to be emphasized that the result is not genuinely new and may constitute, at best, an exercise in a standard course of probability theory, covering the large deviations estimates.

Proposition C.1. *Suppose that the random variables X_i are adapted to a decreasing family $\{\mathfrak{F}_i\}$ and have uniformly bounded absolute moments of order $\alpha \geq 2$:*

$$\begin{aligned} \max_i \mathbb{E} [|X_i|^\alpha | \mathfrak{F}_{i-1}] &\leq C < \infty, \\ \mu &:= \max_i \text{ess sup} \mathbb{E} [X_i | \mathfrak{F}_{i+1}] \in \mathbb{R}. \end{aligned}$$

Then for any $a > \mu$ and some $c > 0$, the following estimates hold true:

$$\begin{aligned} \mathbb{P} \left\{ \sum_{i=1}^n X_i \geq an \right\} &\leq C' \frac{\ln^{2c\alpha}}{n^{\alpha-1}} + n^{-\ln^c n (\ln^{c+1} n - Ca^{-1})} \\ &= O \left(n^{-(\alpha-1-o(1))} \right), \quad n \rightarrow \infty. \end{aligned} \tag{C.1}$$

Proof. As usual in the estimation of heavy-tailed r.v., introduce the truncated r.v.

$$\tilde{X}_i(\omega) = X_i(\omega) \mathbf{1}_{\{X_i \leq b\}}(\omega),$$

$$\tilde{S}_n = \sum_{i=1}^n \tilde{X}_i, \quad \tilde{R}_i = \tilde{S}_n - \tilde{S}_{i-1}.$$

It is readily seen that for any $x \in \mathbb{R}$

$$\mathbb{P} \{ S_n \geq x \} \leq \mathbb{P} \{ \tilde{S}_n \geq x \} + n \max_{1 \leq i \leq n} \mathbb{P} \{ X_i > b \}. \tag{C.2}$$

Since all \tilde{X}_i are bounded, \tilde{S}_n is also bounded and has finite exponential moments of any order, but we will have to choose judiciously a value $h_n > 0$, taking into account the cut-off threshold b , and then use the Chebyshev inequality

$$\mathbb{P} \{ \tilde{S}_n \geq x \} \leq e^{-h_n x} \mathbb{E} [e^{h_n \tilde{S}_n}].$$

Step 1. Since X_1 (hence \tilde{X}_1) is \mathfrak{F}_0 -measurable, and $\mathfrak{F}_0 \subset \mathfrak{F}_i$ for each i , all X_i with $i \geq 2$ are \mathfrak{F}_0 -measurable, thus

$$\begin{aligned} \mathbb{E} [e^{h_n \tilde{S}_n}] &= \mathbb{E} [e^{h_n \tilde{R}_2} \mathbb{E} [e^{h_n \tilde{X}_1} | \mathfrak{F}_0]] \\ &\leq \mathbb{E} [e^{h_n \tilde{R}_2}] \cdot \text{ess sup} \mathbb{E} [e^{h_n \tilde{X}_1} | \mathfrak{F}_0]. \end{aligned}$$

Using the recursion on $i = 1, \dots, n$, we obtain the upper bound

$$\begin{aligned} \ln \mathbb{E} [e^{h_n \tilde{S}_n}] &\leq \sum_{i=1}^n \ln \left(\text{ess sup} \mathbb{E} [e^{h_n \tilde{X}_i} | \mathfrak{F}_{i-1}] \right) \\ &\leq n \cdot \max_{1 \leq i \leq n} \ln \left(\text{ess sup} \mathbb{E} [e^{h_n \tilde{X}_i} | \mathfrak{F}_{i-1}] \right). \end{aligned}$$

Step 2. Fix the following parameters:

$$\begin{aligned} x &= x_n = an, \quad a > 0, \\ b &= b_n = \frac{an}{\ln^{2c} n}, \quad c > 0, \\ h &= h_n = \frac{\ln^{c+1} n}{an} \equiv \frac{\ln^{c+1} n}{x_n}, \end{aligned}$$

so that

$$\begin{aligned} h_n x_n &= \ln^{2c+1} n \gg 1 \quad \text{for } n \gg 1, \\ h_n b_n &= \frac{1}{\ln^{c-1} n} \ll 1 \quad \text{for } n \gg 1. \end{aligned}$$

Observe that since $\alpha \geq 2$, we have

$$\sigma^2 := \max_i \mathbb{E} [X_i^2 | \mathfrak{F}_{i-1}] < +\infty. \tag{C.3}$$

Step 3. Assess the truncated exponential moment of order h_n :

$$\begin{aligned} \mathbb{E} [e^{h_n \tilde{X}_i}] &\leq \int_0^{b_n} e^{h_n u} dF_i(u) \\ &= \int_0^{b_n} (1 + h_n u + (e^{h_n u} - 1 - h_n u)) dF_i(u) \\ &\leq 1 + h_n \mu + \int_0^{b_n} (e^{h_n u} - 1 - h_n u) dF_i(u). \end{aligned}$$

For all $u \in [0, b_n]$, we have

$$h_n u \leq h_n b_n = \frac{\ln^{c+1} n}{an} \ll 1,$$

$$0 \leq e^{h_n u} - 1 - h_n u \leq \frac{1}{2} e^{h_n b_n} u^2.$$

Therefore, it follows from (C.3) that

$$\int_0^{b_n} (e^{h_n u} - 1 - h_n u) dF_i(u)$$

$$\leq \frac{e^{h_n b_n}}{2} \int_0^{+\infty} h_n^2 u^2 dF_i(u) \leq \frac{e^{h_n b_n}}{2} \sigma^2 h_n^2$$

thus

$$\mathbb{E} \left[e^{h_n \tilde{X}_i} \right] \leq 1 + h_n \mu + \frac{e^{h_n b_n}}{2} \sigma^2 h_n^2,$$

yielding, by $e^z \leq z - 1$,

$$\ln \mathbb{E} \left[e^{h_n \tilde{X}_i} \right] \leq h_n \left(\mu + \frac{e^{h_n b_n}}{2} \sigma^2 h_n \right) \leq h_n \mu + Ch_n^2.$$

Thus

$$n \ln \mathbb{E} \left[e^{h_n \tilde{X}_i} \right] \leq \frac{Cn \ln^{c+1} n}{an} = Ca^{-1} \ln^{c+1},$$

so

$$h_n x_n - n \ln \mathbb{E} \left[e^{h_n \tilde{X}_i} \right] \geq h_n (x_n - \mu n) - Ch_n^2$$

$$= \frac{(a - \mu)n \ln^{c+1} n}{an} + O(nh_n^2)$$

$$\geq (a - \mu) \ln^{2c+1} n - Ca^{-1} \ln^{c+1}$$

$$= (a - \mu) \ln^{c+1} n (\ln^{c+1} n - Ca^{-1}).$$

Finally,

$$e^{-h_n x_n + n \ln \mathbb{E} \left[e^{h_n \tilde{X}_i} \right]}$$

$$\leq n^{-\ln^c n (\ln^{c+1} n - Ca^{-1})} = O(n^{-\infty}). \tag{C.4}$$

Step 4. By Chebychev's inequality,

$$n \max_{1 \leq i \leq n} \mathbb{P} \{ X_i > b_n \} \leq \frac{n}{(b_n)^\alpha} \mathbb{E} [|X_i|^\alpha]$$

$$\leq \frac{Cn \ln^{2c\alpha} n}{(an)^\alpha} = O \left(\frac{1}{n^{\alpha-1-o(1)}} \right). \tag{C.5}$$

Step 5: Conclusion. We see that the leading term in the RHS of (C.2) is

$$n \max_{1 \leq i \leq n} \mathbb{P} \{ X_i \geq b_n \} \sim n^{-(\alpha-1)},$$

for the other one is of order of $O(n^{-\infty})$, by (C.4), thus

$$\mathbb{P} \{ S_n > b_n \} \leq \frac{C'}{n^{\alpha-1-o(1)}} (1 + o(1)), \quad n \rightarrow \infty. \tag{C.6}$$

Thus the *asymptotical* exponent is $\alpha - 1 > 0$. □

ACKNOWLEDGEMENTS

It is a pleasure to thank Tom Spencer for fruitful discussions and the Institute for Advanced Study, Princeton, for their warm hospitality during my stay at the Institute in the Spring 2012; Sasha Sodin for fruitful discussions at the IAS; Werner Kirsch and the FernUniversität Hagen for their warm hospitality and simulating atmosphere at the PasturFest (May 13–17, 2013); Abel Klein, Nariyuki Minami, Stanislav A. Molchanov, Günter Stolz, Ivan Veselić for fruitful discussions.

REFERENCES

- [1] Anderson, P. W. Absence of diffusion in certain random lattices. Phys. Rev., Vol.109, 1492–1505, 1958.
- [2] Aizenman, M., Molchanov, S. Localization at large disorder and at extreme energies: An elementary derivation. Commun. Math. Phys., Vol.157, 245–278, 1993.
- [3] Aizenman, M. Localization at weak disorder: some elementary bounds. Rev. Math. Phys. (special issue), 1163–1182, 1994.
- [4] Aizenman, M., Elgart, A., Naboko, S., Shenker, J. H., Stoltz, G. Moment analysis for localization in random Schrödinger operators. Invent. Math., Vol. 163, 343–413, 2006.
- [5] Aizenman, M., Shenker, J. H., Fridrich, R. M., Hundertmark, D. Finite-volume fractional-moment criteria for Anderson localization. Commun. Math. Phys., Vol.224, 219–253, 2001.
- [6] Aizenman, M., Germinet, F., Klein, A., Warzel, S. On Bernoulli decompositions for random variables, concentration bounds and spectral localization. Probab. Theory Related Fields, Vol.143, 219–238, 2009.
- [7] Aizenman, M., Warzel, S. Localization bounds for multi-particle systems. Commun. Math. Phys., Vol.290, 903–934, 2009.
- [8] Aizenman, M., Warzel, S. Resonant delocalization for random Schrödinger operators on tree graphs. J. Eur. Math. Soc., Vol.15, 1167–1222, 2013.
- [9] Bourgain, J., Kenig, C. E. On localization in the continuous Anderson-Bernoulli model in higher dimension. Invent. Math., Vol.161, 389–426, 2005.
- [10] Chulaevsky, V., Suhov, Y. Multi-Scale Analysis for Random Quantum Systems with Interaction. Birkhäuser Inc., Boston, 2013.
- [11] Chulaevsky, V. From fixed-energy MSA to dynamical localization: A continuing quest for elementary proofs. Online available from <http://arxiv.org/pdf/1205.5763>, 2012.
- [12] Damanik, D., Stollmann, P. Multi-scale analysis implies strong dynamical localization. Geom. Funct. Anal., Vol.11, 11–29, 2001.
- [13] von Dreifus, H., Klein, A. A new proof of localization in the Anderson tight binding model. Commun. Math. Phys., Vol.124, 285–299, 1989.
- [14] Dobrushin, R.L., Shlosman, S.B. Completely analytical interactions: constructive description. J. Stat. Phys., Vol.46, 983–1014, 1987.
- [15] Elgart, A., Tautenhahn, M., Veselić, I. Anderson localization for a class of models with a sign-indefinite single-site potential via Fractional Moment Method. Annales Henri Poincaré, Vol.12, No.8, 1571–1599, 2010.
- [16] Feller, W. An Introduction to Probability Theory and Its Applications, Vol.2. John Wiley and Sons, Inc., New York London Sidney, 1966.
- [17] Fröhlich, J., Spencer, T. Absence of diffusion in the Anderson tight binding model for large disorder or low energy, Commun. Math. Phys., Vol.88, 151–184, 1983.
- [18] Fröhlich, J., Martinelli, F., Spencer, T., Scoppola, E. Constructive proof of localization in the Anderson tight-binding model. Commun. Math. Phys. Vol.101, 21–46, 1985.
- [19] Gordon, A. Ya. On the point spectrum of the one-dimensional Schrödinger operator. Uspekhi Matem. Nauk (in Russian), Vol.31, 257–258, 1976.
- [20] Germinet, F., De Bièvre, S. Dynamical localization for discrete and continuous random Schrödinger operators. Commun. Math. Phys., Vol.194, 323–341, 1998.

- [21] Germinet, F., Klein, A. Bootstrap multi-scale analysis and localization in random media. *Commun. Math. Phys.*, Vol.222, 415-448, 2001.
- [22] Germinet, F., Klein, A. A comprehensive proof of localization for continuous Anderson models with singular random potentials. *J. Eur. Math. Soc. (JEMS)*, Vol.15, 55-143, 2013.
- [23] Hundertmark, D. A short introduction to Anderson localization. In: *Analysis and Stochastics of Growth Processes and Interface Models*, Oxford University Press, 194-218, 2008.
- [24] Kato, T. *Perturbation theory for linear operators*. Springer-Verlag, New York, 1976.
- [25] Linnik, Yu. V. Limit theorems for sums of independent variables taking into account large deviations. *Theor. Probab. Appl.*, Vol.6-7, 131-148, 345-360 (v.6), 115-129 (v.7), 1961.
- [26] Martinelli, F., Scoppola, E. Remark on the absence of absolutely continuous spectrum for d -dimensional Schrödinger operators with random potential for large disorder or low energy. *Commun. Math. Phys.*, Vol.97, 465-471, 1985.
- [27] Nagaev, S. V. Large deviations of sums of independent random variables. *Annals of Probability*, Vol.7, No.5, 745-789, 1979.
- [28] Schur, I. Über Potenzreihen, die im Innern des Einheitskreises Beschränkt sind [I]. *J. Reine Angew. Math.*, Vol.147, 205-232, 1917.
- [29] Spencer, T. Localization for random and quasi-periodic potentials. *J. Stat. Phys.*, Vol.51, 1009-1019, 1988.
- [30] Suzuki, F. Anderson localization with self-avoiding walk representation. *J. Phys.: Conf. ser.*, Vol.410, 012010, 2010.
- [31] Simon, B., Wolff, T. Singular continuous spectrum under rank-one perturbations and localization for random Hamiltonians. *Commun. Pure App. Math.*, Vol.39, 75-90, 1986.
- [32] Tautenhahn, M. Localization criteria for Anderson models on locally finite graphs. Online available from <http://arxiv.org/pdf/1008.4503>, 2011.
- [33] Wegner, F. Bounds on the Density of States in Disordered Systems. *Z. Phys. B, Condensed Matter*, Vol.44, 9-15, 1981.
- [34] Zhang, Fuzhen (ed.). *The Schur Complement and Its Applications. Numerical Methods and Algorithms*, Vol.4, Springer Science Business Media, Inc., 2005.