Victor Chulaevsky

Exponential scaling limit of the single-particle Anderson model via adaptive feedback scaling

Abstract We propose a reformulation of the bootstrap version of the Multi-Scale Analysis (BMSA), developed by Germinet and Klein, to make explicit the fact that BMSA implies asymptotically exponential decay of eigenfunctions (EFs) and of EF correlators (EFCs), in the lattice Anderson models with diagonal disorder, viz. with an IID random potential. We also show that the exponential scaling limit of EFs and EFCs holds true for a class of marginal distributions of the random potential with regularity lower than Hölder continuity of any positive order.

1 Introduction

We consider Anderson models with diagonal disorder in a periodic lattice $\mathbb{Z}^d$, $d \geq 1$. Such models have been extensively studied over the last thirty years; the two principal tools of the modern Anderson localization theory are the Multi-Scale Analysis (MSA) and the Fractional Moment Method (FMM). In the framework of lattice systems (and more generally, systems on graphs with sub-exponential growth of balls) the MSA proved to be more flexible; in particular, it is less exigent to the regularity properties of the probability distribution generating the local disorder – in the simplest case, the single-site marginal distribution of the IID (in-

V. Chulaevsky
Université de Reims, Département de mathématiques, Moulin de la Housse, B.P. 1039, 51687 Reims Cedex, France.
E-mail: victor.tchoulaevski@univ-reims.fr

dependent and identically distributed) values of the external random potential. On the other hand, a considerable advantage of the FMM in the same class of models is to provide exponential decay bounds for the (averaged) eigenfunction correlators (EFCs), under the condition of Hölder continuity of the single-site marginal distribution. By comparison, the original MSA scheme by Fröhlich et al. [23], reformulated by von Dreifus and Klein [19], proved only a power-law decay of the key probabilistic estimates in finite volumes. When the MSA was adapted to the proofs of strong dynamical localization (cf. [24,20,25]), this resulted in power-law decay of EFCs.

Germinet and Klein [25] significantly narrowed the gap between the EFC decay bounds provided by the MSA and FMM. Specifically, using the bootstrap MSA, involving several interconnected scaling analyses, they proved sub-exponential decay bounds with rate $L \mapsto \mathrm{e}^{-L^{\delta}}$ for any $\delta \in (0,1)$. Recently Klein and Nguyen [26,27] have adapted the BMSA to the multi-particle Anderson Hamiltonians.

In theoretical physics, the celebrated scaling theory, put forward by the "Gang of Four" (Abrahams, Anderson, Licciardello and Ramakrishnan, [2]) and further developing the Anderson localization theory [1], predicted – under certain assumptions including also those sufficient for the MSA or FMM to apply – that the functionals F_L related to the quantum transport, first of all conductance, for systems of large size L, should admit a limiting behaviour in the double logarithmic coordinate system, with the independent variable to be $\ln L$ rather than L. While the existence of a.c. spectrum for systems on a periodic lattice or in a Euclidean space remains an intriguing challenge for the mathematicians, we show that in the parameter zone(s) where various forms of localization can be established with the help of existing techniques, the rate of decay $F(L)$ of eigenfunction correlators (EFCs) at large distances L admits the limit

$$\lim_{L\to\infty} \frac{\ln\ln F_L}{\ln L} = 1.$$

Below we will call such a behavior *exponential scaling limit* (ESL). Formally speaking, we obtain, as usual, only upper bounds, but the example of one-dimensional systems shows that decay faster than exponential should not be expected.

The main goal of the present paper is a transformation of the Germinet–Klein multi-stage bootstrap MSA procedure from [25] into a single scaling algorithm, replacing several interconnected scaling analyses in the bootstrap method and establishing the ESL in the traditional Anderson model.

The motivation for the present work came from an observation, made in Refs. [12] (cf. [12, Theorem 6]), [14] (see Theorem 8 in [14] and discussion after its proof), and some earlier works, that already in the von Dreifus–Klein method from [19] there were some unexploited resources, giving rise to "self-improving" estimates in the course of the induction on the length scales L_k, $k \geq 0$, following the recursion $L_k = \lfloor L_{k-1}^{\alpha} \rfloor \sim L_0^{\alpha^k}$, $\alpha > 1$. Specifically, it was observed that the k-th induction step actually produces more decay of the GFs than required for merely reproducing the desired decay rate at the step $k+1$, and that this excess can be put in a feedback loop, improving the master parameters of the scaling scheme. The net result is the decay of the GFs (and ultimately, EFCs) faster than any power law[1], viz. $L \mapsto \mathrm{e}^{-a \ln^{1+c} L}$, with $a, c > 0$.

The benefits of such a feedback-based self-enhancement of the master scaling parameters become much greater, when the scales grow multiplicatively, as in the first stage of the bootstrap MSA (BMSA): $L_k = Y L_{k-1} = Y^k L_0$, with $Y \geq 2$. A fairly simple calculation shows that essentially the same feedback loop as the one used in [12, 16, 14] for the scales $L_k \sim L_0^{\alpha^k}$, $k \geq 0$, gives rise in this case to a fractional-exponential decay $L \mapsto \mathrm{e}^{-L^{\delta}}$, with some $\delta > 0$.

Acting in the spirit of the bootstrap MSA, we implement a technically more involved scaling procedure than the above mentioned "simple feedback scaling", aiming to render more explicit and constructive the statement of the BMSA (cf. [25]) that any (viz. arbitrarily close to 1) value of the exponent δ in the above formulae can be achieved for L large enough. To this end, we replace the first two stages of the BMSA (with fixed parameters) by an adaptive scaling algorithm. The latter makes the multiplicative growth factor Y, figuring in the scaling relation $L_k = Y L_{k-1}$, scale-dependent: $Y_k = \mathcal{Y}(k, L_k)$. In fact, the BMSA scheme includes another geometrical parameter – an integer $S_k \in [1, Y_k)$; see Section 3.

[1] This result holds true under a very weak regularity of the random potential, just barely stronger than the conventional log-Hölder continuity of the marginal distribution. See Assumption (W3) (Eqn. (2.15)) in [14].

However, the "simple feedback scaling" – with Y_k and S_k fixed – is still required during an initial "boost" stage, where the effects of localization are almost imperceptible, particularly in the probabilistic estimates. Since the scales grow with k (viz. $L_k = Y^k L_0$), writing formally $Y_k = L_k^{\tau_k}$ results in a finite, initial sub-sequence $\{\tau_1, \ldots, \tau_{\mathfrak{K}-1}\}$, with some $\mathfrak{K}$ depending upon the model parameters, which is actually decreasing. (As such, the values $\{\tau_1, \ldots, \tau_{\mathfrak{K}-1}\}$ are simply unused.) It is only later, for $k \geq \mathfrak{K}$, that we fix $\tau_k = \tau^* > 0$, thus effectively switching to the super-exponential growth $L_k \sim CL_{\mathfrak{K}}^{(1+\tau^*)^k}$. Of course, depending on the reader's personal point of view, the presence of this switching point may be considered as another form of the Germinet–Klein multi-stage technique.

Taking account of abundance of various scaling parameters in our scheme, we keep τ_k fixed for the rest of the scaling procedure. However, the algorithm's efficiency can be further improved by making τ_k also k-dependent (and growing). This may prove useful in a numerical implementation of the adaptive scaling algorithm, as well as in specific models (including the multi-particle models with slowly decaying interaction). We show that the "gap" between the genuine exponential decay (viz. the value $\delta = 1$) and the exponent δ_k achieved at the k-th step, decays at least exponentially fast in k. In a way, it provides a rigorous complement to the predictions of the physical scaling theory on the convergence to the ESL, at least in the parameter zone(s) where localization can be proved with the existing scaling methods.

Speaking of the consecutive phases (analyses) in the Germinet–Klein BMSA, it is to be pointed out that we do not perform the last stage where a genuine exponential decay of the Green functions is established in cubes of size L_k with probability $\sim \mathrm{e}^{-L_k^{\delta_k}}$, where $\delta_k = \delta$ is made arbitrarily close to 1 by the results of [25]; one would expect $\delta_k \nearrow 1$ in the framework of the present paper. We do not analyze the behaviour of such probabilities related to the *exponential* decay of the GFs in finite volume. As was already said, this paper focuses on the exponential scaling *limit* – for the Green functions, eigenfunctions and eigenfunction correlators. The actual road map is as follows: $GFs \rightsquigarrow EFCs \rightsquigarrow EFs$, so the decay rate of the EFs is shaped by that of the EFCs. Naturally, one can switch at any moment from the analysis of the "almost exponential" decay to that of the exponential one, by simply following the Germinet-Klein approach, but our main goal is the construction of a single algorithm which takes care of all exponents δ close to 1. In the author's opinion, there can be various further developments of the BMSA technoology from [25].

Finally, we show that the proposed adaptive scaling technique allows for a lower regularity of the marginal distribution of the IID random potential than Hölder continuity of any positive order. In the realm of the FMM proofs of localization, is is known that the absolute continuity of the marginal distribution can be safely and easily relaxed to Hölder continuity of any positive order β (cf. [3]); a similar observation was made in the works following [25]; yet, the MSA in general is renowned for its tolerance to a lower regularity of the probability distribution of the disorder. So, while the question on the lowest regularity compatible with the FMM approach to the *exponential* strong dynamical localization remains open, our results evidence that Hölder continuity is not required for the exponential scaling *limit* of the EF correlators.

As was said, strong dynamical localization at *some* fractional-exponential rate $\delta \in (0,1)$ actually follows from the initial, weak hypotheses through a simpler scaling procedure, under the assumption of Hölder continuity of the marginal PDF of the random potential.

1.1 The model

We focus on the case where $\mathcal{Z} = \mathbb{Z}^d$ and consider the random Hamiltonian $H(\omega)$ of the form

$$\big(H\psi\big)(x) = \sum_{|y-x|=1} \big(\psi(x) - \psi(y)\big) + V(x;\omega)\psi(x),$$

where $V : \mathbb{Z}^d \times \Omega \to \mathbb{R}$ is an IID random field relative to some probability space $(\Omega, \mathfrak{F}, \mathbb{P})$. Until Section 7, we assume that marginal probability distribution function (PDF) F_V, of the random field V,

$$F_V(t) := \mathbb{P}\left\{V(0;\omega) \le t\right\}, \quad t \in \mathbb{R},$$

is Hölder-continuous of some order $\beta \in (0,1)$. In Section 7 we show that the assumption of Hölder-continuity can be slightly relaxed (cf. Eqn. (7.1)).

The second-order lattice Laplacian can be easily replaced by any (self-adjoint) finite-difference Hamiltonian of finite order, without any significant modification of our algorithm. Indeed, we replace the form of the Geometric Resolvent Inequality most often employed in the MSA of lattice models, with its variant traditional for the MSA in continuous systems (in $\mathbb{R}^d$). It is based on a simple commutator relation, so that the range (order) of a finite-difference kinetic energy operator becomes irrelevant, provided the initial length scale L_0 is large enough. For clarity, we work only with the standard lattice Laplacian.

1.2 Structure of the paper

- The principal objects and notations are introduced in Section 2.
- In Section 3, we present the main analytic tool of the scaling analysis – the Geometric Resolvent Equation (GRE) and Inequality (GRI) stemming from it. The exposition is closer to the form of the GRE/GRI used in the continuous systems than to the one traditionally used in the lattice models, starting from the pioneering papers [22,23,18,31,19]. This is required for the geometrical optimizations à la Germinet–Klein [25] and the proofs of the scale-independent, percolation-type probabilistic bounds.
- The core of the paper is Section 4.
- The derivation of the exponential scaling limit from the results of Section 4 is given in Section 5.
- Section 6 is devoted to a "soft" derivation of strong dynamical localization from the fixed-energy analysis carried out in Section 4.
- In Section 7, we relax the Hölder-continuity assumption on the marginal probability distribution of the random potential.

In theoretical physics, a sufficiently fast decay of the Green functions away from the diagonal is usually considered as one of equivalent "signatures" of Anderson localization. Speaking mathematically, this is a higher-dimensional analog of positivity of Lyapunov exponents in one-dimensional (or quasi-one-dimensional) systems. While it is known that this analog does not imply in general spectral localization, first, it has been shown long ago by Martinelli and Scoppola [28] that it rules out a.c. spectrum with probability one, and, secondly, it has been observed that the s.c. spectrum occurs in systems with some strong “degeneracies” in the probability distribution of the ergodic (not necessarily IID or weakly correlated) potential. Under reasonable assumptions on regularity of the ergodic potential, fast decay of the GFs implies indeed spectral and strong dynamical localization, and the role of Section 6 to summarize the progress achieved in this direction and to show, in a fairly simple way, that the fixed-energy analysis is the heart of the localization analysis of the conventional lattice Anderson model.

The continuous systems are not considered in the present paper, since working with unbounded differential operators would certainly require an additional technical discussion pertaining to the domains, self-adjointness, etc. But as was already said, we focus mainly on the scaling algorithm that could be applied, essentially in the same way, both to the discrete and continuous systems.

2 Basic geometric objects and notations

Following essentially Ref. [25] (where the Anderson-type models in a continuous space $\mathbb{R}^d$ were considered), we work with a hierarchical collection of lattice cubes, with specific centers and positive integer side lengths L_k. For our purposes, it is more convenient to start with the cardinalities of the cubes and those of their one-dimensional projections: we fix odd positive integers Y, ℓ_0 and set

$$L_k = Y^k \cdot 3\ell_0 = 3 \cdot Y^k \ell_0 =: 3\ell_k.$$

Next, we consider the lattice cubes with coordinate projections of cardinality L_k:

$$\mathrm{B}_{L_k}(x) := \left\{ y \in \mathbb{Z}^d : \; |y - x| \le \frac{L_k}{2} \right\}.$$

Since $L_k = 3Y^k\ell_0$ is odd, the upper bound in the above definition of the cube $\mathrm{B}_{L_k}(x)$ could have been replaced with $(L_k - 1)/2$, resulting in the same lattice subset. However, having in mind the canonical embedding $\mathbb{Z}^d \hookrightarrow \mathbb{R}^d$, the above definition looks more natural when transformed as follows: with $y \in \mathbb{Z}^d \hookrightarrow \mathbb{R}^d$,

$$\mathrm{B}_{L_k}(x) := \left\{ y \in \mathbb{Z}^d \hookrightarrow \mathbb{R}^d : \; |y - x| \le \frac{L_k}{2} \right\},$$

so that the ”fictitious” radius of the ball is precisely $L_k/2$.

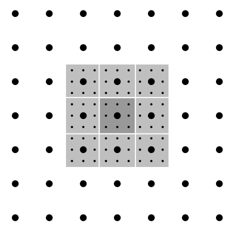

Fig. 1 *Cubes and cells*. Here $d = 2$, $L_0 = 9$. 3^2 cells (gray), including the core (dark gray).

Sometimes it is more convenient to refer to the spherical layers and balls relative to the max-distance, with a clearly identified integer *radius*:

$$\mathcal{L}_r(u) = \left\{ x \in \mathbb{Z}^d : |x - u|_\infty = r \right\},$$
$$\Lambda_r(u) = \left\{ x \in \mathbb{Z}^d : |x - u|_\infty \le r \right\} \equiv \mathrm{B}_{2r+1}(u). \tag{2.1}$$

Notice that one has $\mathrm{B}_{L_k}(u) = \Lambda_{\frac{L_k-1}{2}}(u)$.

The cube $\mathrm{B}_{L_k}(u)$ is partitioned into 3^d adjacent cubes called k-cells,

$$\mathbf{C}_k(c) := \mathrm{B}_{\ell_k}(c) = \Lambda_{\frac{\ell_k-1}{2}}(c)$$

with centers in the sub-lattice $(3\mathbb{Z})^d$.

– The central cell $\mathbf{C}_k(u)$ of a cube $\mathrm{B}_{L_k}(u)$ will be called the *core* of $\mathrm{B}_{L_k}(u)$;
– the complementary annulus, formed by the remaining 3^d-1 cells of $\mathrm{B}_{L_k}(u)$, will be called the *shell* of $\mathrm{B}_{L_k}(u)$.

Given any length scale $L_k = Y^k L_0$, we shall always work with the family of L_k-cubes whose cells form the uniquely defined partition of $\mathbb{Z}^d$ including the cube centered at the origin, $\mathrm{B}_{\ell_k}(0)$; these cores, as well as their centers, will be called *admissible* at the scale L_k. The centers of the admissible ℓ_k-cores form a sublattice of $\mathbb{Z}^d$ which we will denote by $\mathcal{C}^k$. Sometimes we use notation $\langle c, c' \rangle$, meaning that $c, c' \in \mathcal{C}^k$ are two nearest neighbors (in $\mathcal{C}^k$) relative to the max-distance: $|c - c'| = \ell_k$. By a slight abuse of notations, we will write, e.g., $\sum_{\langle c,c' \rangle \in \mathcal{C}^k}$ instead of $\sum_{\langle c,c' \rangle \in \mathcal{C}^k \times \mathcal{C}^k}$ Each point $c \in \mathcal{C}^k$ has $3^d - 1$ nearest neighbors.

See Fig. 1 where

– an admissible square of size $L = 9$ (thus with 9 vertices along each side) is shown in gray color; it is partitioned into $3^d = 3^2$ congruent cells separated visually by thin white lines; the admissibility means that the periodic sublattice of the cell centers (large black dots) includes the origin $0 \in \mathbb{Z}^d$;

- the *core*, i.e., the central cell, is shown in a darker shade of gray;
- each cell is composed – in this example – of 3^2 points.

The larger dots on Fig. 1 represent the centers of the cells of size $\ell_0 = L_0/3 = 3$ admissible in the geometrical constructions referring to the cubes of such size. In this case, the minimal spacing between the centers of admissible cores $= 3$. Considering $L_0 = 9$, we have the spacing $\ell_0 = L_0/3$. The admissible cells of a given size form a partition of $\mathbb{Z}^d$, and we denote by $\mathbf{C}_k(x)$ the unique admissible cell of size $\ell_k = L_k/3$, containing a given point x.

It will be convenient to endow the set of the admissible cell centers in $\mathrm{B}_{L_{k+1}}(u)$ with the natural graph structure, with edges formed by the pairs of nearest neighbors c, c' with respect to the max-distance, i.e., those with $|c-c'|_\infty = \ell_k = L_k/3$. Such a graph $\mathcal{B}_{k+1}$ will be called the *skeleton graph* of $\mathrm{B}_{L_{k+1}}$.

The main tool for the analysis of the Green functions in such balls is the Geometric Resolvent Inequality (GRI). In its basic form, used in [18,31,19] and in numerous subsequent works, a single application of the GRI moves one from the center of a given ball $\mathrm{B}_L(x)$ to (any) point y of the exterior boundary $\partial^+\mathrm{B}_L(x) := \big\{z:\, \mathrm{d}\big(z, \mathrm{B}_L(x)\big) = 1)\big\}$. Here $\mathrm{d}(\cdot\,,\cdot)$ stands for the graph-distance in the lattice $\mathbb{Z}^d$, with edges formed by the nearest neighbors in the Euclidean norm $|\cdot|_2$. The notion of the exterior boundary is relative to an ambient set $\Lambda \supset \mathrm{B}_L(x)$ (a subgraph of $\mathbb{Z}^d$), when the analysis is carried out in a proper subset Λ of the lattice.

Given a finite subset $\Lambda \subset \mathbb{Z}^d$, we introduce the local Hamiltonian $H_\Lambda := \mathbf{1}_\Lambda H \mathbf{1}_\Lambda \restriction \ell^2(\Lambda)$, acting in the finite-dimensional space $\ell^2(\Lambda)$ canonically injected into $\ell^2(\mathbb{Z}^d)$. H_Λ is self-adjoint; it is often considered as the restriction of H to the subset Λ with Dirichlet boundary conditions outside Λ, but the terminology here varies from one source to another.

Further, given any point $x \in \mathbb{Z}^d$ and the nearest k-admissible center, denoted by c_x, we denote by Γ^k_x the boundary annulus of width 2 of $\mathrm{B}_{L_k}(c_x)$: with $R_k := (L_k - 1)/2$,

$$\Gamma^k_x \equiv \Gamma^k_{c_x} := \Lambda_{R_k}(c_x) \setminus \Lambda_{R_k - 2}(c_x), \tag{2.2}$$

and by $\mathbf{\Gamma}^k_x$ the operator of multiplication by $\mathbf{1}_{\Gamma^k_x}$.

Let ϕ be a compactly supported function ϕ on $\mathbb{Z}^d$, where $\operatorname{supp}\phi \subset \Lambda \subset \Lambda'$, with $\mathrm{d}(\partial^+\Lambda, \partial^-\Lambda') \geq 2$, and Φ the operator of multiplication by ϕ. For any $u \in \Lambda$,

$H_\Lambda \mathbf{1}_u = H_{\Lambda'} \mathbf{1}_u$, since H is a finite-difference operator of order 2, so one has the operator identity $\Phi H_\Lambda = \Phi H_{\Lambda'}$, thus for any $E \in \mathbb{R}$,

$$\Phi(H_\Lambda - E) = (H_{\Lambda'} - E)\Phi + [\Phi, (H_{\Lambda'} - E)] \tag{2.3}$$

Below E will be fixed and omitted from notation in the resolvents $G_{\Lambda'} = (H_{\Lambda'} - E)^{-1}$, $G_\Lambda = (H_\Lambda - E)^{-1}$. Denoting $W = [\Phi, (H_{\Lambda'} - E)]$ and multiplying (2.3) by $G_{\Lambda'}$ on the left and by G_Λ on the right, we obtain the identity

$$G_{\Lambda'}\Phi = \Phi G_\Lambda + G_{\Lambda'} W G_\Lambda \,.$$

Now let $k \ge 0$, $R_k = \frac{L_k - 1}{2}$, $\Lambda = \mathrm{B}_{L_k}(u) = \Lambda_{R_k}(u)$, $\phi = \mathbf{1}_{\Lambda_{R_k - 1}(u)}$, so Φ is the projection onto $\Lambda_{R_k-1}(u)$, and the commutator

$$W = \mathbf{1}_{\Lambda_{R_k-1}(u)} \circ (H_{\Lambda'} - E) - (H_{\Lambda'} - E) \circ \mathbf{1}_{\Lambda_{R_k-1}(u)}$$

satisfies the operator identity $W = \mathbf{\Gamma}^k W \mathbf{\Gamma}^k$, where

$$\mathbf{\Gamma}^k := \sum_{c \in \Gamma^k_x} \chi^k_c \mathbf{1}_{\Gamma^k_x}, \; \Gamma^k_x := \big\{ c : \, |c - c_x|_\infty \in \{\ell_k - 1, \ell_k\} \, \big\},$$

and χ^k_c the indicator function of the admissible cell of size $L_k/3$ centered at c. Thus for any subset $A \subset \Lambda' \setminus \Lambda_{R_k}(u)$, one has

$$\begin{aligned}
\mathbf{1}_A G_{\Lambda'} \chi^k_u &= \mathbf{1}_A \Phi G_\Lambda \chi^k_u + \mathbf{1}_A G_{\Lambda'} W G_\Lambda \chi^k_u \\
&= \big(\mathbf{1}_A G_{\Lambda'} \mathbf{\Gamma}^k W\big) \big(\mathbf{\Gamma}^k G_\Lambda \chi^k_u\big),
\end{aligned}$$

so we come to the following form of the Geometric Resolvent Inequality:

$$\big\|\mathbf{1}_A G_{\Lambda'} \chi^k_u\big\| \le \big\|W\big\| \cdot \big\|\mathbf{1}_A G_{\Lambda'} \mathbf{\Gamma}^k\big\| \cdot \big\|\mathbf{\Gamma}^k G_\Lambda \chi^k_u\big\|.$$

Introduce a slightly abusive but convenient notation

$$\big\|G_{\mathrm{B}_{L_k}(x)}\big\|^{\curlywedge} := 3^d C_W \big\|\mathbf{1}_{\mathbf{\Gamma}_{L_k}(u)} G_{\mathrm{B}_{L_k}(x)}(E) \mathbf{1}_{\overline{\mathbf{C}}_k(u)}\big\|, \;\; C_W := \|W\|.$$

Here $\curlywedge$ symbolizes the decay from the center to the boundary of a ball. A more accurate, but also more cumbersome notation would include the dependence of the symbol $\curlywedge$ upon the ball B. Let $\mathrm{B}' = \mathrm{B}_{L_{k+1}}(u)$, $\mathrm{B} = \mathrm{B}_{L_k}(x)$. With $A = \Gamma^{k+1}$, we have

$$\begin{aligned}
\big\|\mathbf{\Gamma}^{k+1}_u G_{\mathrm{B}'} \chi^k_x\big\| &\le C_W \cdot \big\|\mathbf{\Gamma}^{k+1}_u G_{\mathrm{B}'} \mathbf{\Gamma}^k_x\big\| \cdot \big\|\mathbf{\Gamma}^k_x G_{\mathrm{B}} \chi^k_x\big\| \\
&\le C_W \big\|G_{\mathrm{B}}\big\|^{\curlywedge} \sum_{c:\chi^k_c \cap \Gamma^k_x \neq \varnothing} \big\|\mathbf{\Gamma}^{k+1}_u G_{\mathrm{B}'} \chi^k_c\big\| \\
&\le \big\|G_{\mathrm{B}}\big\|^{\curlywedge} \max_{\langle c, x\rangle \in \mathcal{C}^k} \big\|\mathbf{\Gamma}^{k+1}_u G_{\mathrm{B}'} \chi^k_c\big\|.
\end{aligned} \tag{2.4}$$

More generally, in the case where $x \in \mathrm{B}$ is not necessarily the center of the cube $\mathrm{B} \subset \mathrm{B}'$, with $\operatorname{diam} \mathrm{B}' \leq Y \operatorname{diam} \mathrm{B}$, we obtain

$$\begin{aligned} \left\|\mathbf{\Gamma}(\mathrm{B}')G_{\mathrm{B}'}\chi_x^k\right\| &\leq \left\|\mathbf{\Gamma}(\mathrm{B}')G_{\mathrm{B}'}\mathbf{\Gamma}(\mathrm{B})\right\|\left\|\mathbf{\Gamma}(\mathrm{B})G_{\mathrm{B}}\chi_x^k\right\| \\ &\leq \left\|\mathbf{\Gamma}(\mathrm{B})G_{\mathrm{B}}\chi_x^k\right\| \sum_{c\in\mathbf{\Gamma}(\mathrm{B})\cap\mathcal{C}^k} \left\|\mathbf{\Gamma}(\mathrm{B}')G_{\mathrm{B}'}\chi_c^k\right\| \\ &\leq Y^d \left\|\mathbf{\Gamma}(\mathrm{B})G_{\mathrm{B}}\chi_x^k\right\| \max_{c\in\mathbf{\Gamma}(\mathrm{B})\cap\mathcal{C}^k} \left\|\mathbf{\Gamma}(\mathrm{B}')G_{\mathrm{B}'}\chi_c^k\right\| \end{aligned}$$

By self-adjointness of the Hamiltonians at hand, we also have an upper bound

$$\begin{aligned} \left\|\mathbf{\Gamma}(\mathrm{B})G_{\mathrm{B}}\chi_x^k\right\| &\leq \left\|\mathbf{\Gamma}(\mathrm{B})\right\| \left\|\chi_x^k\right\| \ (\operatorname{dist}(E, \Sigma(H'_{\mathrm{B}}))^{-1} \\ &\leq (\operatorname{dist}(E, \Sigma(H'_{\mathrm{B}}))^{-1}, \end{aligned} \tag{2.5}$$

hence

$$\left\|\mathbf{\Gamma}(\mathrm{B}')G_{\mathrm{B}'}\chi_x^k\right\| \leq (\operatorname{dist}(E, \Sigma(H'_{\mathrm{B}}))^{-1} \max_{c\in\mathbf{\Gamma}(\mathrm{B})\cap\mathcal{C}^k} \left\|\mathbf{\Gamma}(\mathrm{B}')G_{\mathrm{B}'}\chi_c^k\right\|. \tag{2.6}$$

3 Dominated decay and EVC bounds

Consider a cube $\mathrm{B} = \mathrm{B}_{L_{k+1}}(u)$ along with its skeleton graph $\mathcal{B}$, and introduce the function $f : \mathcal{B} \to \mathbb{R}_+$ given by

$$f : x \mapsto \left\|\mathbf{\Gamma}_u^{k+1} G_{\mathrm{B}_{L_{k+1}}(u)}\chi_x^k\right\|.$$

Then by GRI (2.4),

$$f(x) \leq \left\|G_{\mathrm{B}_{L_k}(x)}\right\|^{\curlywedge} \max_{\langle c,x\rangle\in\mathcal{C}^k} f(c). \tag{3.1}$$

An inequality of the form (3.1) is most useful when $\left\|G_{\mathrm{B}_{L_k}(x)}\right\|^{\curlywedge} \leq q < 1$; in this case, using an iterated application of the GRI, it is not difficult to prove the bound $f(u) \leq q^{Y-1}$. Below we prove an analog of this simple bound in a more general situation where there are at most $S \geq 1$ vertices $\mathrm{c} \in \mathcal{B}$ where $\left\|\mathrm{B}_{L_k}(\mathrm{c})\right\|^{\curlywedge}$ fails to be smaller than q.

Definition 1 Let be given an integer $k \geq 0$ and real numbers $\epsilon > 0$ and E.

- A cube $\mathrm{B}_L(u)$ is called (E, ϵ)-NR, iff $\operatorname{dist}(\Sigma_{u,L}, E) \geq \epsilon$;
- A cube $\mathrm{B}_{L_{k+1}}(u)$ is called (E, ϵ)-CNR, iff for all $j = 1, \dots, Y_k$ the cube $\mathrm{B}_{jL_k/3}(u)$ is (E, ϵ)-NR.

Recall that, according to the discussion in Section 2, $L_k/3$ is an integer – the size of the cells of order k. The role of the cubes concentric with $\mathrm{B}_{L_{k+1}}(u)$ and composed of entire adjacent $L_k/3$-cells is explained in Appendix A where Lemma 1 is proved. Notice that with $j = Y_{k+1}$, we have $\mathrm{B}_{jL_k/3}(u) = \mathrm{B}_{L_{k+1}}(u)$.

Definition 2 Let be given an integer $k \geq 0$ and real numbers $\epsilon > 0$ and E. A cube $\mathrm{B}_{L_k}(u)$ is called (E, ϵ)-NS, if $E \notin \Sigma_{u,L}$, and $\left\| G_{\mathrm{B}_{L_k}(u)} \right\|^{\curlywedge} \leq \epsilon$.

Below we choose the sizes L of cubes $\mathrm{B}_L(u)$ and the parameter $\epsilon > 0$ figuring in Definitions 1 and 2 in a specific way. First, we take $L \in \{L_k,\, k \geq 0\}$, with L_k defined in (4.9); $\epsilon = L_k^{-b_k}$ in the context of Definition 1, while in the property (E, ϵ)-CNR we set $\epsilon = L_k^{-s_k}$, with recursively constructed sequences b_k and s_k, $k \geq 0$.

Lemma 1 *Suppose that for some $S \geq 0$, a cube $\mathrm{B} = \mathrm{B}_{L_{k+1}}(u)$*

(i) *is $(E, L_k^{-s_k})$-CNR, $s_k \leq b_k$;*

(ii) *contains no collection of $(S_{k+1} + 1)$ disjoint $(E, L_k^{-b_k})$-S cubes of size L_k with admissible centers $c \in \mathcal{C}^k$.*

Then one has

$$\left\| G_{\mathrm{B}}(E) \right\|^{\curlywedge} \leq Y_{k+1}^d \, L_{k+1}^{s_k} \, L_k^{-b_k(Y - 6S_{k+1} - 1)}. \tag{3.2}$$

See the proof in Appendix A.

As usual in the MSA, we also need an eigenvalue concentration (EVC) estimate to bound the norm of the resolvent near the spectrum.

Lemma 2 *Assume that the marginal probability distribution of an IID random potential V is Hölder-continuous of order $\beta \in (0, 1)$. Then for any cube of size L one has*

$$\mathbb{P}\left\{ \mathrm{B}_L(u) \textit{ is not } (E, L^{-s}\text{-CNR} \right\} \leq \mathrm{Const}\, L^{-\beta s}. \tag{3.3}$$

In the case where V admits a bounded probability density, hence $\beta = 1$, this is the classical result by Wegner [32]; cf. also a short proof in [15]. A simple adaptation to Hölder-continuous (and more general) marginal distributions, sufficient for our purposes, can be found in [13], where it is shown that an EVC bound for the potentials with Lipschitz-continuous marginal PDF F_V can be automatically transformed into its counterparts for PDF with an arbitrary continuity modulus. Optimal Wegner bounds have been proved earlier for various types of operators; cf., e.g., [10,11,27].

4 Adaptive feedback scaling

4.1 Technical assumptions and some useful inequalities

In the recursive construction of the sequences $(b_k)_{k\geq 0}$ and $(s_k)_{k\geq 0}$, mentioned in the previous section, the crucial parameter is b_0. Given the marginal distribution F_V of the random potential $V : \mathbb{Z}^d \times \Omega \to \mathbb{R}$, which we assume Hölder-continuous of order $\beta \in (0,1]$ until Section 7, we always assume that $b_0 > d/\beta$ and introduce the scaling parameters

$$\eta := \frac{1}{2}(\beta b_0 - d) > 0, \tag{4.1}$$

$$s_0 := \frac{d}{\beta} + \frac{\eta}{\beta} \equiv b_0 - \frac{\eta}{\beta}. \tag{4.2}$$

Next, denote

$$\kappa_d = d\ln 3 + \ln 2 \lessapprox 1.1d + 0.7 < 2d. \tag{4.3}$$

Given an integer $L_0 \geq 1$, set

$$Y_1 = 9, \quad S_1 = 1, L_1 = Y_1 L_0, \quad a_1 := (3Y_1 - 4)^d. \tag{4.4}$$

Under the crucial assumption,

$$p_0 := \mathbb{P}\left\{ \mathrm{B}_{L_0}(0) \text{ is not } (E, L_k^{-b_0}\text{-NS} \right\} < a_1^{-2d} = \frac{1}{529^d}, \tag{4.5}$$

we introduce the parameters $\theta_0 \in (0, 1/3)$ and $\sigma > 0$ by setting

$$1 - \frac{\ln a_1}{\ln p_0^{-1}} = \frac{1 + 3\theta_0}{2}$$

and

$$\sigma_0 = \min\left[\frac{\ln a_1}{\ln p_0^{-1}}, \frac{1}{16}\right]. \tag{4.6}$$

The scale-free probability threshold in the RHS of (4.5) is slightly larger (hence better) than 841^{-d} given in [25]. This marginal modification is due to a geometrical strategy of the proof of Lemma 1 which deviates from that of the analogous Germinet–Klein argument in [25]. It is clear, however, that the importance of the scale-free probability bounds from [25] goes far beyond the explicit numerical estimates for specific periodic lattices.

Further, introduce an integer $\mathfrak{K} = \mathfrak{K}(p_0)$ (one might want to add Y_1 to the list of arguments of the function $\mathfrak{K}(\cdot)$, but with $Y_1 = 9$ fixed, this becomes unnecessary):

$$\mathfrak{K} = \min\{k \geq 1 : \ (1+\theta_0)^k \geq 2d/\sigma_0\}, \tag{4.7}$$

and define the integer sequences $(Y_k)_{k\geq 1}$, $(S_k)_{k\geq 1}$, and $(L_k)_{k\geq 1}$ as follows:

$$Y_k = \begin{cases} Y_1 = 9, & k \leq \mathfrak{K}, \\ \left\lfloor L_{k-1}^{1/8} \right\rfloor, & k > \mathfrak{K}, \end{cases} \tag{4.8}$$

$$L_k := Y_k L_{k+1}, \tag{4.9}$$

$$S_k := \begin{cases} S_1 = 1, & k \leq \mathfrak{K}, \\ \left\lfloor \frac{1}{9} Y_k \right\rfloor, & k > \mathfrak{K}. \end{cases} \tag{4.10}$$

For further use, notice that we can have $Y_k \leq L_k^{\tau_k}$, for all $k \geq 1$, taking

$$\tau_k := \begin{cases} \tau_1 := \min\left[\frac{\ln Y_1}{\ln L_0}, 1/8\right], & k \leq \mathfrak{K}, \\ 1/8, & k > \mathfrak{K}. \end{cases}$$

Here τ_k, $1 \leq k \leq \mathfrak{K}$, depend upon L_0 and can actually be very small, so in most difficult cases the constraint $\tau_1 \leq 1/8$ is a pure formality, but this allows us to avoid the case-by-case study of different parameter zones.

In the course of the scale induction, we will also assume that, for some integer $M \geq 18$,

$$L_0 \geq L_0(\eta, \tau_1) := \max\left[(M+1)^{1/\tau_1^2}, \ \exp\left(\frac{4d}{\eta}\right)\right]. \tag{4.11}$$

4.2 Unbounded growth of the geometric scaling parameters

The following statement is an important ingredient of the proof of exponential scaling limit in the scheme with varying scaling parameters Y_k, S_k (cf. Sect. 5).

Lemma 3 *Let be given an integer $L_0 \geq L_0(\eta)$, with $L_0(\eta)$ given by (4.11). Then the sequences $(L_k)_{k\geq 0}$, $(Y_k)_{k\geq \mathfrak{K}}$ and $(S_k)_{k\geq \mathfrak{K}}$, given by (4.8)–(4.10), are strictly monotone increasing, and for all $k \geq 1$ one has*

$$\frac{1}{10} Y_k \leq S_k \leq \frac{1}{9} Y_k. \tag{4.12}$$

Proof Pick an integer $M \geq 18$, and let $L_0 \geq (M+1)^{1/\tau^2}$. Then we have

$$\begin{aligned} Y_{k+1} &= \lfloor L_k^\tau \rfloor = \lfloor Y_k^\tau L_{k-1}^\tau \rfloor = \lfloor (\lfloor L_{k-1}^\tau \rfloor)^\tau L_{k-1}^\tau \rfloor \geq \lfloor (\lfloor L_0^\tau \rfloor)^\tau L_{k-1}^\tau \rfloor \\ &\geq \left\lfloor \left(\left\lfloor (M+1)^{\tau/\tau^2} \right\rfloor \right)^\tau L_{k-1}^\tau \right\rfloor \geq \left\lfloor \left(\left\lfloor M^{\tau/\tau^2} + 1 \right\rfloor \right)^\tau L_{k-1}^\tau \right\rfloor \\ &\geq \left\lfloor \left(M^{\tau/\tau^2} \right)^\tau L_{k-1}^\tau \right\rfloor \geq M \lfloor L_{k-1}^\tau \rfloor = MY_k > Y_k. \end{aligned}$$

Furthermore,

$$S_{k+1} = \left\lfloor \frac{1}{9} Y_{k+1} \right\rfloor \geq \left\lfloor M \cdot \frac{1}{9} Y_k \right\rfloor \geq M \left\lfloor \frac{1}{9} Y_k \right\rfloor > S_k.$$

Therefore, the sequences $(Y_k)_{k \geq \mathfrak{K}}$ and $(S_k)_{k \geq \mathfrak{K}}$ are strictly increasing.

To prove the RHS inequality in (4.12), notice that $Y_{\mathfrak{K}+1} \geq MY_{\mathfrak{K}} \geq 18 \cdot 9$, and for any real $y \geq 90$ one has $\lfloor \frac{y}{9} \rfloor \geq \frac{y}{9} - 1 \geq \frac{y}{10}$, hence $\forall\, k \geq \mathfrak{K} + 1$

$$\frac{1}{10} Y_k \leq S_k = \left\lfloor \frac{1}{9} Y_k \right\rfloor \leq \frac{1}{9} Y_k, \tag{4.13}$$

as asserted. □

4.3 Scaling of the GFs

Lemma 4 *Let be given the integers $k \geq 0$, $1 \leq S_{k+1} \leq Y_{k+1}$ and real numbers $b_k \geq s_k > d/\beta$. Assume that*

(i) *the ball $\mathrm{B}_{L_{k+1}}(u)$ is $(E, L_{k+1}^{-s_k})$-CNR and contains no collection of $S_{k+1} + 1$ pairwise disjoint $(E, L_k^{-b_k})$-S balls of radius L_k with admissible centers;*
(ii) *$Y_{k+1} \leq L_k^{\tau_k}$ for some $0 < \tau_k \leq 1/8$;*
(iii) $N_{k+1} := Y_{k+1} - 5S_{k+1} - 1 \geq 3$.

Set

$$b_{k+1} := A_{k+1} b_k, \quad \text{with } A_{k+1} := \frac{7}{12} N_{k+1} \geq \frac{7}{4}. \tag{4.14}$$

Then $\mathrm{B}_{L_{k+1}}(u)$ is $(E, L^{-b_{k+1}})$-NS.

Proof Denote

$$f(u, L_{k+1}) = \left\| G_{\mathrm{B}_{L_{k+1}}(u)} \right\|^\curlywedge,$$

then by Lemma 1, we have

$$f(u, L_{k+1}) \leq Y^d L_{k+1}^{s_k} L_k^{-b_k N_{k+1}} \leq Y_{k+1}^{d+s_k} L_k^{-b_k N_{k+1} + s_k}.$$

Therefore, recalling that $b_k \ge s_k > d/\beta$, $N_{k+1} \ge 3$, we obtain

$$
\begin{aligned}
-\frac{\ln f(u, L_{k+1})}{\ln L_k} &\ge b_k N_{k+1} - s_k - \frac{(d+s_k)\cdot \ln Y_{k+1}}{\ln L_k} \\
&\ge b_k\, N_{k+1}\left(1 - \frac{s_k}{N_{k+1}b_k} - \frac{d+s_k}{N_{k+1}b_k}\cdot \frac{\ln L_k^{\tau_{k+1}}}{\ln L_k}\right) \\
&\ge b_k\, N_{k+1}\left(\left(1 - \frac{1}{N_{k+1}}\right) - \frac{2}{N_{k+1}}\tau_{k+1}\right) \\
&\ge b_k\, N_{k+1}\left(\frac{2}{3} - \frac{2}{3}\cdot\frac{1}{8}\right) \ge \frac{7N_{k+1}}{12} b_k \ge \frac{7}{4} b_k\,.
\end{aligned}
\tag{4.15}
$$

With $L_{k+1} = L_k Y_{k+1}$, $\ln Y_{k+1}/\ln L_k \le 1/8$, we have (cf. (4.14))

$$
\begin{aligned}
-\frac{\ln f(u, L_{k+1})}{\ln L_{k+1}} = b_{k+1} \ge b_k \frac{7}{12} N_{k+1}\cdot \frac{1}{1+\frac{1}{8}} = A_{k+1}\, b_k, \\
A_{k+1} = \frac{14}{27} N_{k+1} \ge \frac{14}{9}.
\end{aligned}
\tag{4.16}
$$

□

4.4 Scaling of the probabilities

Denote, as before,

$$
\begin{aligned}
p_k &:= \mathbb{P}\left\{ \mathrm{B}_{L_k}(u) \text{ is } (E, L_k^{-b_k})\text{-S} \right\}, \\
\mathfrak{w}_{k+1} &:= \mathbb{P}\left\{ \mathrm{B}_{L_k}(u) \text{ is not } (E, L_{k+1}^{-s_k})\text{-CNR} \right\}
\end{aligned}
\tag{4.17}
$$

Further, let

$$
\varrho_{k+1} = \beta s_k \gamma_k, \quad \gamma_k := \begin{cases} \frac{1}{2}\left(1 - \frac{d}{\beta s_0}\right) > 0, & k = 0, \\ \frac{1}{8}, & k \ge 1. \end{cases}
\tag{4.18}
$$

In the next statement, we establish an important technical ingredient of the proof of the key Lemma 6. Specifically, we assess the probability of "admissible resonances" by using a Wegner estimate. As usual in the MSA, such upper bounds essentially shape those on probability of "insufficient decay" of the Green functions and, ultimately, of the eigenfunction correlators. Pictorially, one cannot get bounds better than those stemming from a Wegner-type analysis, so we have to make sure the latter is compatible with the exponential scaling limit.

Lemma 5 *Consider the sequences of positive integers* $(L_k)_{k\ge 0}$, $(Y_k)_{k\ge 1}$, $(S_k)_{k\ge 1}$ *defined as in* (4.8)–(4.10)*, and let* $\mathfrak{w}_k$, $k \ge 0$*, be given by* (4.17)*. Assume that* $L_0 \ge \exp(4d/\eta)$ *(cf.* (4.11)*). Then for all* $k \ge 1$*, the following bound holds:*

$$\mathfrak{w}_k \le L_k^{-\varrho_k}. \tag{4.19}$$

Proof By Wegner estimate (cf. (3.3)),

$$\begin{aligned} \frac{-\ln\left(\frac{1}{2}\mathfrak{w}_{k+1}\right)}{\ln L_{k+1}} &\ge \frac{-\ln\left((S_{k+1}+1)(3L_{k+1})^d\, L_{k+1}^{-\beta s_k}\right)}{\ln L_{k+1}} \\ &= \beta s_k\left(1 - \frac{d}{\beta s_k} - \frac{d\ln 3}{\beta s_k\, \ln L_{k+1}} - \frac{\ln(S_{k+1}+1)}{\beta s_k\, \ln L_{k+1}}\right) \end{aligned} \tag{4.20}$$

Recall $\kappa_d = d\ln 3 + \ln 2 < 2d$, $L_0 \ge \mathrm{e}^{4d/\eta} > \mathrm{e}^{2\kappa_d/\eta}$, so for $k=0$, (4.20) becomes

$$\frac{-\ln\left(\frac{1}{2}\mathfrak{w}_1\right)}{\ln L_1} \ge \beta s_0 - d - \frac{\kappa_d}{\ln L_1} \ge \eta - \frac{2d}{\ln L_0} \ge \frac{1}{2}\eta = \varrho_1\ . \tag{4.21}$$

Hence $\frac{1}{2}\mathfrak{w}_1 \le L_1^{-\varrho_1} =: \frac{1}{2}q_1$.

Now let $k \ge 1$. With $\beta s_1 = \frac{3}{2}\beta b_0 > \frac{3d}{2}$, we have

$$\frac{-\ln\left(\frac{1}{2}\mathfrak{w}_{k+1}\right)}{\ln L_{k+1}} \ge \beta s_k\left(1 - \frac{d}{\beta s_k} - \frac{d\ln 3}{\beta s_k\, \ln L_{k+1}} - \frac{\ln(S_{k+1}+1)}{\beta s_k\, \ln L_{k+1}}\right) =: \beta s_k \widetilde{\gamma}_{k+1},$$

where we denoted temporarily

$$\begin{aligned} \widetilde{\gamma}_{k+1} &:= 1 - \frac{d}{\beta s_k} - \frac{d\ln 3}{\beta s_k\, \ln L_{k+1}} - \frac{\ln(S_{k+1}+1)}{\beta s_k\, \ln L_{k+1}} \\ &\ge 1 - \frac{2}{3} - \frac{2\ln 3}{3\, \ln L_{k+1}} - \frac{\ln Y_{k+1}}{\beta s_k\, \ln L_{k+1}} \\ &\ge \frac{1}{3} - \frac{2\ln 3}{3\, \ln L_{k+1}} - \frac{2\ln L_k^{1/8}}{3\, \ln L_k} \ge \left(\frac{1}{3} - \frac{1}{12}\right) - \frac{2\ln 3}{3\, \ln L_{k+1}} \\ &\ge \frac{1}{4} - \frac{2\ln 3}{3\, \ln L_0} \ge \frac{1}{8}, \end{aligned}$$

provided $L_0 \ge 27 > 3^{8/3}$. □

Recall that we have defined in (4.7) an integer $\mathfrak{K} = \min\{k \ge 1 :\ (1+\theta_0)^k \ge 2d/\sigma_0\}$, with σ_0 defined in (4.6).

Lemma 6 *Consider the sequences of positive integers* $(L_k)_{k\geq 0}$, $(Y_k)_{k\geq 1}$, $(S_k)_{k\geq 1}$ *defined as in* (4.8)–(4.10)*, and let* $\{p_k,\ k\geq 0\}$ *be given by* (4.17)*. Assume that* $L_0 \geq \exp(4d/\eta)$ *(cf.* (4.11)*), and we have*

$$p_0 < (3Y_1 - 4)^{-2d}. \tag{4.22}$$

Define recursively a sequence of positive numbers $(\sigma_k)_{k\geq 0}$ *:*

$$\sigma_k = \mathbf{B}_k \sigma_0, \quad \mathbf{B}_k = B_1 \cdots B_k, \tag{4.23}$$

$$B_j = \begin{cases} 1+\theta_0 \in (1,2), & j = 1, \dots \mathfrak{K} \\ \frac{7}{9}(S_j+1) > \frac{3}{2}, & j \geq \mathfrak{K}+1. \end{cases} \tag{4.24}$$

Then for all $k \geq 1$*, the following bound holds:*

$$p_k \leq L_k^{-\sigma_k}. \tag{4.25}$$

Proof By Lemma 4, if $\mathrm{B}_{L_{n+1}}(u)$ is singular, then it must be either $(S_{n+1}+1)$-bad or not $(E, L_{n+1}^{-s_n})$-CNR. Therefore, with $a_{n+1} = (3Y_{n+1}-4)^d$,

$$p_{n+1} \leq \frac{1}{2}\big(a_{n+1}p_n\big)^{S_{n+1}+1} + \frac{1}{2}\mathfrak{w}_{n+1}. \tag{4.26}$$

By Lemma 5, we have $\mathfrak{w}_{k+1} \leq q_{k+1} = L_{k+1}^{\varrho_{k+1}}$, thus

$$p_{k+1} \leq \frac{1}{2}(a_{k+1}p_k)^{S_{k+1}+1} + \frac{1}{2}q_{k+1}.$$

If $p_{k+1} \leq q_{k+1}$, we simply keep this bound and proceed to the conclusion of the scaling step; otherwise, we argue as follows.

► First, let $k < \mathfrak{K}(\theta_0)$, so that $Y_j = Y_1$, $a_j \equiv (3Y_j - 4)^d = a_1$ for all $1 \leq j \leq k+1$. By induction in $j = 1, \dots, k$, we know that σ_j is monotone increasing, thus

$$1 - \frac{d \ln a_{k+1}}{\sigma_k \ln L_k} \geq 1 - \frac{d \ln a_1}{\sigma_0 \ln L_0} = 1 - \frac{d \ln a_1}{\ln p_0^{-1}} = \frac{1+3\theta_0}{2}.$$

Therefore, with $S_{k+1} = S_k = \cdots = S_1 = 1$,

$$\frac{\ln p_{k+1}^{-1}}{\ln L_{k+1}} = \widetilde{\sigma}_{k+1} \geq \sigma_k \cdot (S_{k+1}+1)\left(1 - \frac{d \ln a_{k+1}}{\sigma_k \ln L_k}\right)\frac{\ln L_k}{\ln L_{k+1}} \tag{4.27}$$

$$\geq \sigma_k \cdot (1+3\theta_0) \cdot \left(1 - \frac{\ln Y_1}{\ln L_n + \ln Y_1}\right) \tag{4.28}$$

$$\geq \sigma_k \cdot (1+3\theta_0) \cdot (1-\tau_0) \tag{4.29}$$

$$\geq \sigma_k \cdot (1+\theta_0) \overset{\text{by induction}}{=} \sigma_0 \cdot (1+\theta_0)^k. \tag{4.30}$$

► Now let $k \geq \mathfrak{K}$, so $\mathbf{B}_k \geq 2d/\sigma_0$. Then we have, with $Y_{k+1} = \lfloor L_k^{\tau_{k+1}} \rfloor \leq L_k^{\tau_{k+1}}$, $\tau_{k+1} = 1/8$,

$$\frac{2d \ln Y_{k+1}}{\sigma_k \ln L_k} \leq \frac{2d\tau_{k+1} \ln L_k}{\sigma_0 \mathbf{B}_k \ln L_k} \leq \frac{1}{8}.$$

Therefore,

$$\frac{\ln p_{k+1}^{-1}}{\ln L_{k+1}} = \widetilde{\sigma}_{k+1} \geq \sigma_k \cdot (S_{k+1}+1)\left(1 - \frac{d \ln a_{k+1}}{\sigma_k \ln L_k}\right) \frac{\ln L_k}{\ln L_{k+1}} \tag{4.31}$$

$$\geq \sigma_k \cdot (S_{k+1}+1) \cdot \left(1 - \frac{1}{8}\right) \frac{\ln L_k}{\ln L_{k+1}(1+\tau_{k+1})} \tag{4.32}$$

$$\geq \sigma_k \cdot (S_{k+1}+1) \cdot \frac{7}{8} \cdot \frac{8}{9} \tag{4.33}$$

$$\geq \sigma_k \cdot \frac{7}{9}(S_{k+1}+1) \geq \frac{14}{9}\sigma_k > \frac{3}{2}\sigma_k. \tag{4.34}$$

by definition of σ_k.

Further, $A_j = \frac{7}{12} N_j$, with $N_j = Y_j - 5S_j - 1 \geq 9S_j - 6S_j = 3S_j$, thus $A_j \geq \frac{7}{12} \cdot 3S_j = \frac{7}{4} S_j$.

$$A_j \geq \frac{7}{4} S_j > \frac{14}{9} S_j \geq \max\left[\frac{7}{9}(S_j+1),\, 1+\theta_0\right] \geq B_j.$$

Therefore, using $b_0 \geq d/\beta \geq 1$, $\sigma_0 \leq 1/16$ (cf. (4.6)),

$$\varrho_k = \frac{b_0}{16}\mathbf{A}_k \geq \sigma_0 \mathbf{B}_k = \sigma_k.$$

The asserted inductive bound (4.25) is proved. □

5 Exponential scaling limit

We have shown that

$$\mathbb{P}\left\{ \mathrm{B}_{L_k} \text{ is } (E, L_k^{-b_k}\text{-S} \right\} \leq L_k^{-\sigma_k}.$$

Our aim now is to show that the above bound can be re-written as follows:

$$\mathbb{P}\left\{ \mathrm{B}_{L_k} \text{ is } (E, \mathrm{e}^{-(L_k)^{\delta_k}}\text{-S} \right\} \leq \mathrm{e}^{-(L_k)^{\kappa_k}},$$

where $\delta_k, \kappa_k \nearrow 1$ as $k \to +\infty$. This is a matter of simple calculations.

Indeed, by induction, $b_k = \mathbf{A}_k b_0$. Since $S_j \leq Y_j/9$, we have

$$A_j = \frac{7}{12} N_j \geq \frac{7}{12}(Y_j - 5S_j - 1) \geq \frac{7}{12} \cdot \frac{3}{9} Y_j > \frac{1}{4} Y_j. \tag{5.1}$$

Therefore,

$$\begin{aligned} b_k &> b_0 4^{-k} \prod_{j=1}^{k} Y_j = \frac{b_0}{L_0} 4^{-k} L_k, \\ &= L_k^{1-\frac{1}{\ln L_k}\left(\ln \frac{L_0}{b_0} + k \ln 4\right)} > L_k^{1-\frac{\ln L_0 + 2k}{\ln L_k}} = L_k^{1-\mathrm{o}(1)} \end{aligned} \tag{5.2}$$

since $Y_j \nearrow +\infty$, thus $k/\ln L_k \to 0$. Consequently,

$$L_k^{-b_k} \leq \mathrm{e}^{-\ln L_k \cdot L_k^{1-o(1)}} = \mathrm{e}^{-c_k\, L_k^{1-o(1)}}, \quad c_k \underset{k\to+\infty}{\longrightarrow} +\infty.$$

More precisely, $L_k \approx L_{k-1}^{9/8}$ for $k > \mathfrak{K}$, so for some $1 < q \approx 9/8$,

$$\ln L_k \geq C + C' q^{k-\mathfrak{K}} \geq C'' q^k.$$

Thus

$$\frac{\ln\ln L_k^{-b_k}}{\ln L_k} \geq 1 - \frac{C'''}{(1+\epsilon)^k}, \quad \epsilon \approx 1/8.$$

Similarly, for the probabilities $p_k \leq L_k^{-\sigma_k}$ we have

$$\ln p_k^{-1} \geq \sigma_k \ln L_k \geq \sigma_0 B_1 \dots B_k,$$

where $B_j \geq CS_j \geq C'Y_j$, $C, C' > 0$, for all $j \geq \mathfrak{K}$. By taking a sufficiently small constant $C'' > 0$, one can extend this lower bound to $B_1, \dots, B_{\mathfrak{K}}$:

$$\ln p_k^{-1} \geq C'' c^k Y_1 \cdots Y_k \geq C''' c^k L_k \geq L_k^{1-\alpha(k)},$$

with $\alpha(k) \leq h^k$, $h \in (0,1)$.

6 ESL for the eigenfunctions and their correlators

It is well-known by now that a sufficiently fast decay of the Green functions, proved with sufficiently high probability at each energy E in a given interval $I \subseteq \mathbb{R}$, implies both spectral localization (a.s. pure point spectrum in I with rapidly decaying eigenfunctions) and strong dynamical localization, with rapidly decaying averaged EF correlators. Such implications can be established with the help of different methods. For example, in the bootstrap method presented in Ref. [25], the fixed-energy estimates in probability, proved at a given energy E_0, are extended to an interval $I_0 = [E_0 - \epsilon, E_0 + \epsilon]$ with sufficiently small $\epsilon > 0$, by means of the energy-interval (a.k.a. variable-energy) MSA induction; the core procedure goes back to earlier works [18, 31, 19].

In our work [14] (cf. also the book [16]), we proposed an alternative approach based on an argument employed by Elgart et al. [21] in the general context of the FMM and encapsulated in a fairly general, abstract spectral reduction (FEMSA ⇒ VEMSA). Similar ideas, in essence going back to the work by Martinelli and Scoppola [28], were used in other papers; cf., e.g., [9].

We formulate the spectral reduction in the following way (cf. [14, 16]). (*Notice that the boldface* $\mathbf{b}_L$ *are unrelated to the sequence of scaling exponents* b_k.)

Theorem 1 *Let be given a bounded interval* $I \subset \mathbb{R}$*, an integer* $L \geq 0$*, two disjoint balls* $\mathrm{B}_L(x)$*,* $\mathrm{B}_L(y)$*, and the positive numbers* $\mathbf{a}_L, \mathbf{b}_L, \mathbf{c}_L, Q_L$ *satisfying*

$$\mathbf{b}_L \leq \min \left[\mathbf{a}_L \mathbf{b}_L^2, \mathbf{c}_L \right] \tag{6.1}$$

and such that

$$\forall\, E \in I \qquad \max_{z \in \{x,y\}} \mathbb{P}\left\{ \mathbf{F}_z > \mathbf{a}_L \right\} \leq Q_L. \tag{6.2}$$

Assume also that, for some function $f : (0,1] \to \mathbb{R}_+$*,*

$$\forall\, \epsilon \in (0,1] \qquad \mathbb{P}\left\{ \mathrm{dist}\left(\Sigma(H_{\mathrm{B}_L(x)}),\, \Sigma(H_{\mathrm{B}_L(y)}) \right) \leq \epsilon \right\} \leq f(\epsilon) \tag{6.3}$$

Then

$$\mathbb{P}\left\{ \sup_{E \in I} \max \left[\mathbf{F}_x(E),\, \mathbf{F}_y(E) \right] > \mathbf{a}_L \right\} \leq \frac{|I|\, Q_L}{\mathbf{b}_L} + f(2\mathbf{c}_L) \tag{6.4}$$

Consequently, taking into account the results of Section 4, for some $\delta_k \nearrow 1$ *as* $k \to +\infty$*, one has*

$$\mathbb{P}\left\{ \exists\, E \in I :\ \ \mathrm{B}_L(x) \text{ and } \mathrm{B}_L(y) \text{ are } (E, L_k^{-b_k})\text{-}\mathbf{S} \right\} \leq \mathrm{e}^{-L^{\delta_k}}.$$

The proof given below is based on the following

Lemma 7 *Let be given positive numbers* $\mathbf{a}_L, \mathbf{b}_L, \mathbf{c}_L, Q_L$ *such that*

$$\mathbf{b}_L \leq \min \left[\mathbf{a}_L \mathbf{b}_L^2, \mathbf{c}_L \right] \tag{6.5}$$

and

$$\forall\, E \in I \quad \mathbb{P}\left\{ \mathbf{F}_x(E) > \mathbf{a}_L \right\} \leq Q_L. \tag{6.6}$$

There is an event $\mathcal{B}_x$ *such that* $\mathbb{P}\left\{ \mathcal{B}_x \right\} \leq b^{-1}Q$ *and for any* $\omega \notin \mathcal{B}_x$*, the set* $\mathscr{E}_x(2a) := \{E : \mathbf{F}_x(E) > 2\mathbf{a}_L\}$ *is contained in a union of intervals* $\cup_{j=1}^{N} I_j :=$ $\{E : |E - E_j| \leq 2\mathbf{c}_L\}$*, centered at the eigenvalues* $E_j \in \Sigma(H(\omega)) \cap I$*.*

Proof Consider the random subsets of the interval I parameterized by $a' > 0$,

$$\mathscr{E}(a';\omega) = \{E : \mathbf{F}_x(E) \geq a'\}$$

and the events parameterized by $b' > 0$,

$$\mathcal{B}(b') = \{\omega \in \Omega : \operatorname{mes}(\mathscr{E}(a) > b'\} = \left\{\omega \in \Omega : \int_I \mathbf{1}_{\mathbf{F}_x(E) \geq \mathbf{a}_L} \, dE \; > b' \right\}.$$

Using the hypotheses (6.5)-(6.6), apply Chebyshev's inequality and the Fubini theorem:

$$\begin{aligned} \mathbb{P}\left\{\mathcal{B}(\mathbf{b}_L)\right\} &\leq \mathbf{b}_L^{-1} \mathbb{E}\left[\operatorname{mes}(\mathscr{E}(\mathbf{a}_L))\right] \\ &= b^{-1} \int_I dE \, \mathbb{E}\left[\mathbf{1}_{\mathbf{F}_x(E) \geq \mathbf{a}_L}\right] \leq \mathbf{b}_L^{-1} \mathbb{P}\left\{\mathbf{F}_x(E) \geq \mathbf{a}_L\right\}. \end{aligned}$$

Fix any $\omega \notin \mathcal{B}(b)$, so $\operatorname{mes}(\mathscr{E}(\mathbf{a}_L;\omega)) \leq \mathbf{b}_L$.

Further, consider the random sets patameterized by $c' > 0$,

$$\mathcal{R}(c') = \{\lambda \in \mathbb{R} : \min_j |\lambda_j(\omega) - \lambda| \geq c'\}.$$

Note that for $\mathbf{a}_L \in (0, \mathbf{c}_L)$, $\mathcal{A}_{\mathbf{b}_L} := \{E : \operatorname{dist}(E, \mathcal{R}(2\mathbf{c}_L)) < \mathbf{b}_L\} \subset \mathcal{R}(\mathbf{c}_L)$, hence the complement $\mathcal{A}^{\mathrm{c}}_{\mathbf{b}_L}$ is a union of sub-intervals at distance at least $\mathbf{c}_L$ from the spectrum.

Let us show by contraposition that, for any $\omega \notin \mathcal{B}(\mathbf{b}_L)$, one has

$$\{E : \mathbf{F}_x(E;\omega) \geq 2\mathbf{a}_L\} \cap \mathcal{R}(2\mathbf{c}_L) = \varnothing.$$

Assume otherwise and pick any point λ^* in the non-empty intersection on the LHS. Let $J := \{E' : |E - \lambda^*| < b\} \subset \mathcal{A}_b \subset \mathcal{R}(c)$. By the first resolvent identity

$$\begin{aligned} \|G(E')\| &\geq \|G(\lambda^*)\| - |E' - \lambda^*| \, \|G(E')\| \|G(\lambda^*)\| \\ &\geq 2\mathbf{a}_L - \mathbf{b}_L \cdot (2\mathbf{c}_L)^{-1} (\mathbf{c}_L)^{-1} \geq \mathbf{a}_L, \end{aligned}$$

owing to the assumption (6.1) on $\mathbf{a}_L, \mathbf{b}_L, \mathbf{c}_L$. We also used here the bounds $\|G(\lambda^*)\| \leq (2\mathbf{c}_L)^{-1}$ and

$$\|G(E')\| \leq (\operatorname{dist}(E', \Sigma))^{-1} \leq (\operatorname{dist}(\lambda^*, \Sigma) - |E' - \lambda^*|)^{-1} \leq (2\mathbf{c}_L - \mathbf{b}_L)^{-1},$$

with $\mathbf{b}_L \leq \mathbf{c}_L$. Consequently, the entire interval $(\lambda^* - \mathbf{b}_L, \lambda^* + \mathbf{b}_L)$ of length $2\mathbf{a}_L > \mathbf{b}_L$ is a subset of $\mathscr{E}(\mathbf{a}_L;\omega)$, which is impossible for any $\omega \notin \mathcal{B}(\mathbf{b}_L)$. This contradiction completes the proof. $\square$

Proof of Theorem 1 Define the events $\mathcal{B}_x, \mathcal{B}_y$ related to the points x, y in the same way as the event $\mathcal{B}_x$ relative to x in the proof of Lemma 7, and let $\mathcal{B} = \mathcal{B}_x \cup \mathcal{B}_y$. Let $\omega \notin \mathcal{B}$. Then for both values of $z \in \{x, y\}$, the set $\mathscr{E}_z(a)$ is contained in the union of at most K intervals $J_{z,i} = [E_i^{(z)} - 2\mathbf{c}_L, E_i^{(z)} + 2\mathbf{c}_L]$. Therefore, the event

$$\mathbb{P}\left\{\omega : \inf_{E\in I} \max\left[\mathbf{F}_x(E), \mathbf{F}_y(E)\right] > \mathbf{a}_L\right\} \leq \mathbb{P}\left\{\omega : \operatorname{dist}(\Sigma_x, \Sigma_y) \leq 4\mathbf{c}_L\right\};$$

the latter probability is bounded with the help of the Wegner-type estimate. □

Now the derivation of strong dynamical localization from the VEMSA estimates can be made in the same way as in Ref. [25], directly in the entire lattice $\mathbb{Z}^d$. This requires an a priori, Shnol–Simon polynomial bound (cf., e.g., [29,30]) on the growth rate of spectrally a.e. generalized eigenfunction; the latter becomes unnecessary in arbitrarily large finite balls (cf. [14,16,17]).

Theorem 2 (Cf. [14, Theorem 7]) *Assume that the following bound holds true for a pair of disjoint cubes* $\mathrm{B}_L(x), \mathrm{B}_L(y)$*:*

$$\mathbb{P}\left\{\exists\, E \in I :\ \mathrm{B}_L(x) \text{ and } \mathrm{B}_L(y) \text{ are } (E,\epsilon)\text{-}S\ \right\} \leq h(L).$$

Then for any cube $\mathrm{B}_{L'}(w) \supset (\mathrm{B}_{L+1}(x) \cup \mathrm{B}_{L+1}(x))$ *one has*

$$\mathbb{E}\left[\,\left|\langle \mathbf{1}_x \,|\, \phi\,(H_{\mathrm{B}})\ |\, \mathbf{1}_y\right|\,\right] \leq 4\epsilon + h(L).$$

The extension of the EFC decay bounds to the entire lattice can be done with the help of the Fatou lemma on convergent measures; such a path was laid down in earlier works by Aizenman et al. [4,5,6].

Summarizing, one can say that the essential equivalence of various forms of Anderson localization (decay of the GFs, EFs, EFCs) is firmly established by now for a large class of random Hamiltonians.

7 Lower regularity

Theorem 3 *The results of Section 4 remain valid for the marginal probability distributions with continuity modulus* $\mathfrak{s}_V(\cdot)$ *satisfying the following condition:*

$$\mathfrak{s}_V(\epsilon) \leq C' \epsilon^{\frac{C}{\ln|\ln\epsilon|}}. \tag{7.1}$$

Proof Consider first the situation where $Y_{k+1} = \lfloor L_k^{\tau} \rfloor$, hence $L_{k+1} \geq CL_k^{1+\tau}$, $\tau > 0$.

The regularity of the marginal distribution of the random potential V must be sufficient for proving a Wegner-type estimate

$$\mathbb{P}\left\{ \| G_{\mathrm{B}_{L_k}}(E) \| > L_k^{s_k} \right\} \leq L_k^{-\beta_k s_k},$$

where $\beta_k s_k$, replacing βs_k used in the previous section, has to be compatible with our main estimates. Denoting $\epsilon_k = L_k^{-s_k}$, we thus should have

$$\mathbb{P}\left\{ \| G_{\mathrm{B}_{L_k}}(E) \| > \epsilon_k \right\} \leq \epsilon^{\beta_k}.$$

Up to some inessential factors (depending on L_k), the above estimate can be inferred in a standard way from the continuity of the marginal PDF F_V with the continuity modulus of the form $\mathfrak{s}_V(\epsilon) \leq C\epsilon^{C'\beta_k}$.

Next, observe that one has $\epsilon_k^{-1} = L_k^{s_k} \leq \mathrm{e}^{c_1 L_k}$: indeed, our estimates by $L_k^{-b_k}$ and $L_k^{\pm s_k}$ are not truly exponential in L_k (although that would be very welcome), so we only have $\epsilon_k \sim \mathrm{e}^{\pm L_k^{1-o(1)} \ln L_k} = \mathrm{e}^{\pm L_k^{1-o(1)}}$. Thus

$$\ln \ln \ln \epsilon_k^{-1} = \ln \ln \ln L_k^{s_k} \leq \ln \ln (c_1 L_k) \leq \ln \ln (c_2 L_0^{q^k}) \leq c_3 k.$$

At the same time, with $\beta_k = \frac{\beta_0}{(1+\kappa)^k}$, we have $\ln \beta_k^{-1} \geq c_4 k$, hence one can proceed with the scaling algorithm even in the case where

$$\ln \beta_k^{-1} \geq c_5 \ln \ln \ln \epsilon_k^{-1} \Longrightarrow \beta_k \leq \frac{c_6}{\ln |\ln \epsilon_k|}.$$

We conclude that the Wegner-type estimates compatible with the adaptive scaling scheme employed in Section 4 can be inferred from the following condition upon the continuity modulus $\mathfrak{s}_V$:

$$\mathfrak{s}_V(\epsilon) \leq C' \epsilon^{\frac{C}{\ln |\ln \epsilon|}},$$

which is – just marginally – weaker than Hölder regularity of any positive order. Pictorially, it can be qualified as Hölder continuity of "almost zero" order.

The proof in the general case can be reduced to the above analysis, since the double-exponential growth $L_k \sim L_0^{q^k}$ takes over the exponential one, $L_k = L_0 Y_1^k$, after a finite number of steps $\mathfrak{K} = \mathfrak{K}(p_0)$. Observe that all intermediate calculations and bounds can be re-written in terms of strict inequalities (for this is the case with the principal hypothesis, $p_0 < a_1^{-2}$), and these strict inequalities can be preserved by replacing $\beta = \mathrm{Const}$ with $\beta_k = Const/(1+\kappa)^k$ during the $\mathfrak{K}$ steps, provided $\kappa > 0$ is small enough – depending of course on $\mathfrak{K}$. The auxiliary constants clearly depend upon the proximity of p_0 to the Germinet-Klein threshold 841^{-d}. After $\mathfrak{K}$ steps, one can start the scaling procedure with $L'_0 := L_{\mathfrak{K}}$. In fact, this would be very close in spirit to the Germinet-Klein first bootstrapping step. $\square$

Appendix A Proof of Lemma 1

Consider the ℓ_k-skeleton graph $\mathcal{B}$ of the cube B. For $r \geq 0$, denote by $\mathcal{B}_r = \mathcal{B}_r(u)$ the balls $\{c \in \mathcal{C}^k : \mathrm{d}_{\mathcal{C}^k}(u,c) \leq r\}$; then $\mathcal{B} = \mathcal{B}_{3K+1}$. To avoid any confusion, recall that the vertices of $\mathcal{B}$ represent the $L_k/3$-cells of the original lattice $\mathbb{Z}^d$.

We will reduce our analysis of the function $\mathcal{B} \ni c \mapsto |G_{\mathrm{B}}(c,y;E)|$ to that of a monotone function of one integer variable

$$F : r \mapsto \max_{c:\, \mathrm{d}_{\mathcal{B}}(u,c) \leq r} |G_{\mathrm{B}}(c,y;E)|;$$

here $r \in I := [0, 3K+1] = \left[0, \frac{Y-1}{2} + Y\right]$. More precisely, we have to assess the decay of F as r runs across the sub-interval $[R-Y-1, R-2]$, $R := 3K+1$.

It is convenient to introduce the spherical layers $\mathcal{L}_r := \{c \in \mathcal{B} : \mathrm{d}_{\mathcal{B}}(u,c) = r\}$ and the function

$$f : r \mapsto \max_{c \in \mathcal{L}_r} |G_{\mathrm{B}}(c,y;E)|,$$

so that $F(r) = \max_{r' \leq r} f(r')$.

Call a vertex $c \in \mathcal{B}$ non-singular if the associated ball $\mathrm{B}_{L_k}(c) \subset \mathbb{Z}^d$ is $(E, L_k^{-b_k})$-NS, and singular, otherwise. Respectively, call $r \in [R-Y-1, R-2]$ non-singular if all vertices c with $\mathrm{dist}_{\mathcal{B}}(u,c)$ are non-singular, and singular, otherwise. The notions of singularity/non-singularity do not apply to $r \in [R-1, R-2]$.

Notice that we have the following inequalities:

(A) for any non-singular $r \in I$,

$$f(r) \leq \max_{r' \in [r-1, r+1]} (3Y_{k+1})^{-d} L_k^{-b_k} f(r');$$

(B) owing to the assumed CNR-property of $\mathrm{B}_{L_{k+1}}(u)$, for any $r \leq r' \leq R-2$ one has, by an application of the GRI,

$$f(r) \leq (3Y_{k+1})^d L_{k+1}^{s_k} f(r'). \tag{A.1}$$

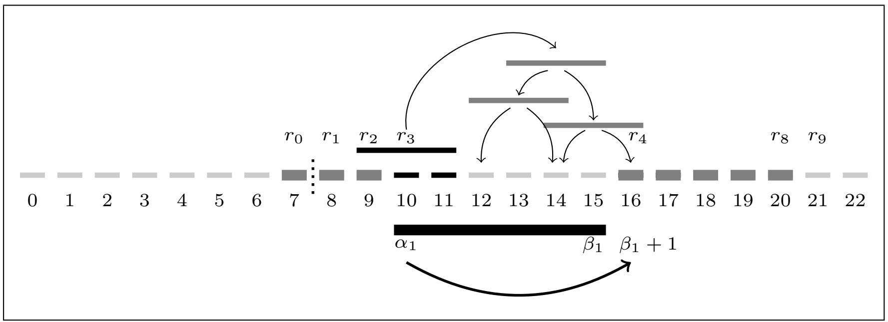

Fig. 2 In this example, $K = 7, Y = 2 \cdot K + 1 = 15, S = 1, N = Y - 5S - 1 = 9$, and one has no pair of disjoint singular intervals of the form $[\rho - 1, \rho + 1]$. r_3 is the smallest integer in $I = [0, 3K + 1]$ which is singular; it is the projection of the center c of a singular ball in the skeleton graph. It is this minimality property which implies that $r_2 = r_3 - 1$ must be non-singular, despite the fact that the intervals $[r_2 - 1, r_2 + 1]$ and $[r_3 - 1, r_3 + 1]$ overlap. On the other hand, due to the overlap of $[r_3 - 1, r_3 + 1]$ with $[r_3, r_3 + 2]$, the point $r_3 + 1$ may (or might) be singular, without producing a disjoint singular pair. Therefore, we still can use the property **(A)** starting off the point r_2 (and aiming at $r_2 + 1 = r_3$), but leaving from r_3, we have to make a longer flight with possible "destinations" (i.e., reference points) ranging in $[(r_3+4)-2, (r_3+4)+2] = [r_3+2, r_3+6]$. The longest flight consumes the distance 6, instead of 1 that we would have for a non-singular departure point; this results in a loss of 5 points. The thick gray intervals indicate the points which provide the factors $q \leq (3Y - 1)^{-d} L^{-b} < 1$ in the "radial descent" induction: $F(r_{i-1}) \leq qF(r_i)$. The point r_9 is used as the last reference point, but we can only bound $F(r_9)$ by the global maximum of F, since the GRI cannot be applied at a center c of the skeleton graph $\mathcal{B}$ with $\mathrm{d}_{\mathcal{B}}(u, c) \geq 3K$. Here we have the guaranteed decay bound $F(r_0) \leq q^9 F(r_9) \leq q^9 F(3K + 1)$.

(A crude bound $(3Y)^d$ can be replaced by $C(d)Y^{d-1}$.) To be more precise, an application of the GRI is required for $r \leq r' - 1$, while for $r = r'$ the inequality (A.1) follows trivially from $(3Y_{k+1})^d L_{k+1}^{s_k} \geq 1$.

Combining **(A)** and **(B)**, we come to the following statement:

(C) Assume that for some $r \leq r' \leq R - 2$, all points $\rho \in [r' + 3, r' + 5]$ are non-singular. Then for all $r \in [0, r' + 5]$

$$F(r) \leq (3Y_{k+1})^{-d} L_k^{-2b_k} L_{k+1}^{s_k} F(r' + 6). \tag{A.2}$$

Notice that for $r = r' + 5$, (A.2) follows immediately from the assumed non-singularity of the point $r' + 5$, so it remains to be established only for $r \leq r' + 4$.

For the proof, we first apply **(B)**:

$$F(r' + 4) = \max_{\rho \leq r' + 4} f(\rho) \leq (3Y)^d L_{k+1}^{s_k} f(r' + 4). \tag{A.3}$$

Next, apply **(A)** to $r' + 4$ (which is non-singular by assumption):

$$f(r' + 4) \leq (3Y_{k+1})^{-d} L_k^{-b_k} \max_{r'' \in [r'+3, r'+5]} f(r''), \tag{A.4}$$

thus

$$F(r'+4) \le (3Y_{k+1})^{-d}\, L_{k+1}^{s_k}\, L_k^{-b_k} \max_{r''\in[r'+3,r'+5]} f(r''). \tag{A.5}$$

Apply **(A)** once again to the three points $r'' \in [r'+3, r'+5]$ (all of which are non-singular by assumption):

$$\begin{aligned}
\max_{r''\in[r'+3,r'+5]} f(r'') &\le (3Y_{k+1})^{-d}\, L_k^{-b_k} \max_{r''\in[r'+3,r'+5]} \max_{r'''\in[r''-1,r''+1]} f(r''') \\
&\le (3Y_{k+1})^{-d}\, L_k^{-b_k} \max_{r'''\in[r'+2,r'+6]} f(r''') \\
&\le (3Y_{k+1})^{-d}\, L_k^{-b_k}\; F(r'+6).
\end{aligned} \tag{A.6}$$

Collecting (A.5) and (A.6), the assertion (C) follows, since $F(r) \le F(r')$ for $r \le r'$.

Pick any maximal collection of disjoint singular cubes $\mathrm{B}_{L_k}(c_i)$, $i = 1, \dots, n \le S$, denote $\rho_i = \mathrm{d}_{\mathcal{B}}(u, c_j)$, and associate with each c_i an interval $[\widetilde{\alpha}_i, \widetilde{\beta}_i] = [\rho_i - 1, \rho_i + 5$. Next, decompose the union of intervals $[\widetilde{\alpha}_i, \widetilde{\beta}_i]$ into a disjoint union of maximal non-overlapping intervals $\mathcal{J}_i = [\alpha_i, \beta_i]$, $1 \le i \le n' \le n$, so that $\beta_i \le \alpha_{i+1} - 1$; the equality $\alpha_{i+1} = \beta_i + 1$ is permitted.

Note that for any i, every points $r \in [\beta_i - 3, \beta_i - 1]$ are non-singular, otherwise we would have to augment $\mathcal{J}_i$ by including the interval $[r-1, r+5]$ overlapping with $\mathcal{J}_i$, which contradicts the maximality of $\mathcal{J}_i$.

Let $I' = I \setminus \cup_i \mathcal{J}_i$ and enumerate the points of $I' = \{r_0, r_1, \dots, r_M\}$ in the natural increasing order; in other words, I' is obtained by collapsing each interval $\mathcal{J}_i$ to a single point, and then we enumerate the new points some of which are images of single points of I and others represent the entire intervals $\mathcal{J}_i$.

If r_i is the image of a non-singular point, then we have $r_{i+1} = r_i + 1$ and

$$F(r_i) \le (3Y_{k+1})^{-d}\, L_k^{-b_k} F(r_{i+1});$$

otherwise, we can apply (C) and obtain

$$\begin{aligned}
F(r_i) &\le (3Y_{k+1})^{-d}\, L_k^{-2b_k} L_{k+1}^{s_k} F(r_{i+1}) \\
&\le \left(L_k^{-(b_k - s_k)} (3Y_{k+1})^{s_k} \right) \cdot (3Y_{k+1})^{-d} L_k^{-b_k} F(r_{i+1}), \\
&\le (3Y_{k+1})^{-d} L_k^{-b_k} F(r_{i+1}),
\end{aligned}$$

since $s_k = b_k/2$ for $k \ge \mathfrak{K}$, while for $k < \mathfrak{K}$ this bound holds true with $Y_k = Y_1 = 9$ and L_0 large enough.

Finally, note that collapsing the intervals $\mathcal{J}_i$ into single points eliminates from I at most $5S_{k+1}$ points; this bound becomes sharp if the radial projections of all singular L_k-balls in the collection (fixed at the beginning) are non-overlapping. Hence $|I'| \geq (Y_{k+1} - 1) - 5S_{k+1}$.

This completes the proof. □

Acknowledgements I thank the Isaac Newton Institute, Cambridge, UK, and the organisers of the program "*Periodic and ergodic spectral problems*" (2015) for the support and opportunity to work for six months in a stimulating atmosphere of the Institute, where the present work was completed.

References

1. Anderson, P. W., Absence of diffusion in certain random lattices. *Phys. Rev.*, **109**, 1492–1505 (1958)
2. Abrahams, E., Anderson, P. W., Licciardello, D.C., Ramakrishnan, T.V., Scaling theory of localization: Absence of quantum diffusion in two dimensions. *Phys. Rev. Lett.*, **42**, 673–676 (1979)
3. Aizenman, M., Molchanov, S., Localization at large disorder and at extreme energies: an elementary derivation. *Commun. Math. Phys.*, **157**, 245–278 (1993)
4. Aizenman, M., Localization at weak disorder: Some elementary bounds. *Rev. Math. Phys.* **6**, 1163–1182 (1994)
5. Aizenman, M., Shenker, J. H., Friedrich, R. M., Hundertmark, D., Finite-volume fractional-moment criteria for Anderson localization. *Commun. Math. Phys.* **224**, 219–253 (2001)
6. Aizenman, M., Elgart, A., Naboko, S., Schenker, J.H., Stolz, G., Moment analysis for localization in random Schrödinger operators. *Invent. Math.* **163**, 343–413 (2006)
7. Damanik, D., Stollmann, P., Multi-scale analysis implies strong dynamical localizaiton. GAFA, Geom. Funct. Anal. **11**, 11–29 (2001)
8. Germinet, F., De Bièvre, S., Dynamical localization for discrete and continuous random Schrödinger operators. *Commun. Math. Physics* **194**, 323–341 (1998)

9. Bourgain, J., Kenig, C.E., On localization in the continuous Anderson-Bernoulli model in higher dimension. *Invent. Math.* **161**, 389–426 (2005)
10. Combes, J.-M., Hislop, P.D., Localization for some continuous, random Hamiltonians in d-dimension. *J. Funct. Anal.* **124**, 149–180 (1994)
11. Combes, J.-M., Hislop, P.D., Klopp, F., An optimal Wegner estimate and its applicaitons to the global continuity of the integrated density of states for random Schrödinger operators. *Duke Math. J.* **140**, 469–498 (2007)
12. Chulaevsky, V., Direct scaling analysis of localization in single-particle quantum systems on graphs with diagonal disorder. *Math. Phys. Anal. Geom.* **15**, 361–399 (2012)
13. Chulaevsky, V., Stochastic regularization and eigenvalue concentration bounds for singular ensembles of random operators. *Int. J. Stat. Mech.* **2013**, 931063 (2013)
14. Chulaevsky, V., From fixed-energy localization analysis to dynamical localization: An elementary path. *J. Stat. Phys.* **154**, 1391–1429 (2014)
15. Carmona, R., Lacroix, J., Spectral Theory of Random Schrödinger Operators. Birkhäuser, Boston (1990)
16. Chulaevsky, V., Suhov, Y., Multi-scale Analysis for Random Quantum Systems with Interaction. Progress in Mathematical Physics. Birkhäuser, Boston (2013)
17. Chulaevsky, V., Suhov, Y., *Efficient Anderson localization bounds for large multi-particle systems*. To appear in: J. Spec. Theory.
18. von Dreifus, H., On effect of randomness in ferromagneic models and Schrödinger operators. PhD diss., New York University, New York (1987)
19. von Dreifus, H., Klein, A., A new proof of localization in the Anderson tight-binding model. *Commun. Math. Phys.* **124**, 285–299 (1989)
20. Damanik, D., Stollmann, P., Multi-scale analysis implies strong dynamical localization. *Geom. Funct. Anal.*, **11**, no. 1, 11–29 (2001)

21. Elgart, A., Tautenhahn, M., Veselić, I., Anderson localization for a class of models with a sign-indefinite single-site potential via fractional moment method. *Ann. Henri Poincaré* **12**, no. 8, 1571–1599 (2010)
22. Fröhlich, J., Spencer, T., Absence of diffusion in the Anderson tight-binding model for large disorder or low energy. *Commun. Math. Phys.* **88**, 151–184, (1983)
23. Fröhlich, J., Martinelli, F., Scoppola, E., Spencer, T., Constructive proof of localization in the Anderson tight-binding model. *Commun. Math. Phys.* **101**, 21–46 (1985)
24. Germinet, F., De Bièvre, S., Dynamical localization for discrete and continuous random Schrödinger operators. *Commun. Math. Physics* **194**, 323–341 (1998)
25. Germinet, F., Klein, A., Bootstrap multi-scale analysis and localization in random media. *Commun. Math. Phys.* **222**, 415–448 (2001)
26. Klein, A., Nguyen, S. T., Bootstrap multiscale analysis for the multi-particle Anderson model. *J. Stat. Phys.* **151**, no. 5, 938–973 (2013)
27. Klein, A., Nguyen, S. T., Bootstrap multiscale analysis for the multi-particle continuous Anderson Hamiltonians. To appear in: J. Spec. Theory.
28. Martinelli, F., Scoppola, E., Remark on the absence of absolutely continuous spectrum for d-dimensional Schrödinger oerators with random potential for large disorder or low energy. *Commun. Math. Phys.* **97**, 465–471 (1985)
29. Shnol, E., On the behaviour of the Schrödinger equation. (in Russian) *Mat. Sbornik* **42**, 273–286 (1957)
30. Simon, B., Schrödinger semigroups. *Bull. Am. Math. Soc.* **7**, 447-526 (1982)
31. Spencer, T.: Localization for random and quasi-periodic potentials. *J. Stat. Phys.* **51**, 1009–1019 (1988)
32. Wegner, F., Bounds on the density of states of disordered systems. *Z. Phys.* **B44**, 9–15 (1981)