Bootstrap Multiscale Analysis and Hilbert-Schmidt Localization in Random Media

François Germinet
UMR 8524 CNRS, UFR de Mathématiques,
Université de Lille 1
F-59655 Villeneuve d'Ascq Cédex, France
e-mail: germinet@agat.univ-lille1.fr

Abel Klein*
Department of Mathematics
University of California, Irvine
Irvine, CA 92697-3875, USA
e-mail: aklein@uci.edu

Revised version, to appear in Commun. Math. Phys

Abstract

We introduce an enhanced multiscale analysis that yields sub-exponentially decaying probabilities for *bad* events. For quantum and classical waves in random media, we obtain exponential decay for the resolvent of the corresponding random operators in boxes of side L with probability higher than $1 - \mathrm{e}^{-L^{\zeta}}$, for any $0 < \zeta < 1$. The starting hypothesis for the enhanced multiscale analysis only requires the verification of polynomial decay of the finite volume resolvent, at some sufficiently large scale, with probability bigger than $1 - \frac{1}{841^d}$ (d is the dimension). Note that from the same starting hypothesis we get conclusions that are valid for any $0 < \zeta < 1$. This is achieved by the repeated use of a bootstrap argument. As an application, we use a generalized eigenfunction expansion to obtain strong dynamical localization of any order in the Hilbert-Schmidt norm, and better estimates on the behavior of the eigenfunctions.

1 Introduction

Quantum and classical waves may be described by first or second order differential equations on a Hilbert space $\mathcal{H} = \mathrm{L}^2(\mathbb{R}^d, \mathrm{d}x; \mathbb{C}^n)$. Quantum waves are described by the Schrödinger equation:

$$i\frac{\partial}{\partial t}\psi_t = H\psi_t \ , \tag{1.1}$$

while classical waves may be described by a second order wave equation with an auxiliary condition:

$$\frac{\partial^2}{\partial t^2}\psi_t = -H\psi_t \ \ , \text{with} \ \ \psi_t = P_H^{\perp}\psi_t \ . \tag{1.2}$$

*Partially supported by NSF Grant DMS-9800883.

In both cases H is a self-adjoint operator on $\mathcal{H}$; for the wave equation we have $H \geq 0$ and $P_H^\perp$ is the orthogonal projection on the orthogonal complement of the kernel of H. Finite energy solutions of the first order equation (1.1) are of the form

$$\psi_t = \mathrm{e}^{-itH}\phi_0 \ , \qquad \phi_0 \in \mathcal{H} \ , \tag{1.3}$$

inasmuch as finite energy solutions of the second order equation (1.2) are given by

$$\psi_t = \cos\left(t\sqrt{H}\right) P_H^\perp \phi_0 + \sin\left(t\sqrt{H}\right) P_H^\perp \eta_0 \ , \qquad \phi_0, \eta_0 \in \mathcal{H} \ . \tag{1.4}$$

In this article we discuss questions concerning localization of waves in random media. A random medium will be modeled by a $\mathbb{Z}^d$-ergodic random self-adjoint operator H_ω, where ω belongs to a probability set Ω with a probability measure $\mathbb{P}$ (e.g., [37, 33, 9, 24, 25, 43, 12]). In this article such a H_ω will be simply called a *random operator*. It follows from ergodicity that there exists a nonrandom set Σ, such that $\sigma(H_\omega) = \Sigma$ with probability one, where $\sigma(A)$ denotes the spectrum of the operator A. In addition, the decomposition of $\sigma(H_\omega)$ into pure point spectrum, absolutely continuous spectrum, and singular continuous spectrum is also independent of the choice of ω with probability one [37, 49, 8, 13].

As an example one can consider the potential

$$V_\omega(x) = \sum_{i \in \mathbb{Z}^d} \lambda_i(\omega) u(x - i), \tag{1.5}$$

where u is a bounded nonnegative function with compact support, and the $\{\lambda_i(\omega)\}_{i \in \mathbb{Z}^d}$ are independent, identically distributed random variables (e.g., [33, 9, 45, 38, 39, 52]). The random Schrödinger operator is given by $H_\omega = -\Delta + \gamma V_\omega(x)$ (plus possibly a bounded periodic potential [45, 38]). The parameter γ measures the amount of disorder.

For classical waves, examples include Maxwell's equations, and the equations of acoustics and elasticity, where the random Schrödinger-like operators in (1.2) are of the form $H_\omega = A_\omega^* A_\omega$ with $A_\omega = \sqrt{R_\omega} D \sqrt{S_\omega}$, where D is a first order partial differential operator with constant coefficients, and R_ω, S_ω are strictly positive matrix valued functions of the form $Y_0(x)(1+\gamma V_\omega(x))^{\pm 1}$, with $Y_0(x)$ a periodic matrix valued function, not necessarily smooth, and $V_\omega(x)$ as in (1.5) [24, 25, 43, 12].

In this paper we restrict ourselves to continuum models. There is a vast literature on the Anderson model and other discrete random operators, e.g., [46, 26, 27, 20, 7, 13, 8, 2, 1, 3, 22, 23, 41, 16, 55]. Our methods and results also apply to discrete random operators, with the obvious modifications.

The main achievement of this paper is an enhanced multiscale analysis, the *bootstrap multiscale analysis*, which is stated in Theorem 3.4. In this context the multiscale analysis (MSA) is a technique, initially developed in [26, 27] and simplified in [19, 20] (see also [21, 40]), for the purpose of proving Anderson localization (pure point spectrum and exponential decay of eigenfunctions). It was later shown to also yield dynamical localization (non spreading of the wave packets) [29], and more recently strong dynamical localization (dynamical localization not only with probability one, but in expectation) up to some order [14]. Our enhancement yields strong dynamical localization up to any order. In fact, it yields more: strong dynamical localization in the Hilbert-Schmidt norm. The usual multiscale analyses, based on von Dreifus and Klein [20], give exponential decay of the resolvent on big boxes with side $L_k \nearrow \infty$, with probability close to 1 up to a polynomially small correction in L_k (i.e., $\geq 1 - L_k^{-p}$ for a fixed $p > 0$). In comparison, the bootstrap multiscale analysis we present here in Theorem 3.4 requires less in the starting hypotheses, and yields far better probability estimates. In fact, it gives any desired sub-exponential decay for the probabilities of *bad* events. An important new feature of the enhanced MSA is that the final probability estimates are independent of the probability estimate in the starting hypothesis. This is achieved by a repeated use of a bootstrap argument. Thus one may look for the weakest possible starting hypothesis without affecting the resulting probability estimates.

An important consequence of this bootstrap MSA is given in Theorem 3.8, which we paraphraze as follows. For a large class of random operators, if the bootstrap MSA starting hypothesis holds at a fixed energy E_0 ($E_0 > 0$ for equation (1.2)), then there exists $\delta_0 > 0$, such that, defining $I(\delta_0) = (E_0 - \delta_0, E_0 + \delta_0)$, one has

$$\mathbb{E}\left(\sup_{\|\|f\|\| \leq 1} \|\chi_x f(H_\omega) E_{H_\omega}(I(\delta_0)) \chi_y\|_2^2 \right) \leq C_\zeta \mathrm{e}^{-|x-y|^\zeta} \tag{1.6}$$

for any $0 < \zeta < 1$, where C_ζ is some finite constant depending only on ζ and on the parameters of the problem (χ_x stands for the characteristic function of a box of side 1 centered at x; the supremum is taken over Borel functions f of a real variable, with $\|\|f\|\| = \sup_{t \in \mathbb{R}} |f(t)|$; $E_{H_\omega}(\ )$ denotes the spectral projection of the operator H_ω; $\|B\|_2$ denotes the Hilbert-Schmidt norm of the operator B). It follows from (1.6), by a fairly straightforward calculation, that for any bounded region Ω and all $q > 0$ we have

$$\mathbb{E}\left(\sup_{\|\|f\|\| \leq 1} \left\| |X|^{\frac{q}{2}} f(H_\omega) E_{H_\omega}(I(\delta_0)) \chi_\Omega \right\|_2^2 \right) < \infty\,, \tag{1.7}$$

in which case we will say that the corresponding wave equation (either (1.1) or (1.2)) exhibits *strong HS-dynamical localization* in the energy interval $I(\delta_0)$. An immediate consequence is *strong dynamical localization*, meaning that for any finite energy solution ψ_t (as in either (1.3) or (1.4)), we have

$$\mathbb{E}\left(\sup_t \left\| \, |X|^{\frac{q}{2}} E_{H_\omega}(I(\delta_0))\, \psi_t \right\|^2 \right) < \infty \tag{1.8}$$

for any $q > 0$ and Cauchy data (either $\phi_0 \in \mathcal{H}$ or $\phi_0, \eta_0 \in \mathcal{H}$) with compact support. (It actually follows from (1.6) that it suffices to have Cauchy data that decays faster than any polynomial in the L^2-sense, i.e., the local L^2-norms decay faster than any polynomial.)

An application of the results of this paper (the bootstrap multiscale analysis and its application to strong HS-dynamical localization) can be found in [32], where we show a discontinuity of the transport properties of the random media at the Anderson metal-insulator transition (if there is one).

Classical waves may be described by first order Schrödinger-like equations of the same form as (1.1), where the self-adjoint operator H is a first order partial differential operator (see [42]); e.g., Maxwell equations. Such an equation yields two second order wave equations of the form (1.2). It turns out that the bootstrap MSA for one of these second order equations implies the estimate (1.6), and hence also (1.7) and (1.8), for the first order classical wave equation, as well as for the other second order wave equation (see [43]).

Dynamical localization is a term commonly used for the almost sure version of (1.8), i.e.,

$$\sup_t \left\| \, |X|^{\frac{q}{2}} E_{H_\omega}(I(\delta_0)) \, \psi_t \right\|^2 < \infty \text{ for a.e. } \omega. \tag{1.9}$$

It was proven in [29] in the context of this paper (the proof is given for Schrödinger operators, but it is also applicable to classical waves). Dynamical localization implies pure point spectrum by the RAGE Theorem (e.g., the argument in [13, Theorem 9.21]), but the converse is not true. Dynamical localization is actually a strictly stronger notion than pure point spectrum, since the latter can take place whereas a quasi-ballistic motion is observed [18]. The question “what is localization?” has been raised in [17] and the last decade has seen many contributions to this subject matter [18, 29, 28, 3, 53].

In the discrete case, the first results on dynamical localization are due to Jona-Lasinio, Martinelli and Scoppola [35] for a hierarchical model, and to Martinelli and Scoppola [47] for the Anderson model. For the latter, with a bounded, absolutely continuous probability distribution for the single site potential, the Aizenman-Molchanov approach [1, 3] gives strong dynamical localization (in fact, it gives exponential decay in (1.6)). But where the Aizenman-Molchanov approach does not apply (e.g., random operators on the continuum, the Anderson model with a Bernoulli potential in one dimension), dynamical localization has been harder to prove.

In the continuum the first results were obtained by by Holden and Martinelli [33], who proved subdiffusive motion for random Schrödinger operators (for the time averaged second moment). Recently, Barbaroux et al. [5] showed the absence of diffusion for the time averaged second moment.

The search for a proof of dynamical localization in the continuum ended when Germinet and De Bièvre [29] proved (almost sure) dynamical localization whenever the MSA applied. More recently, Damanik and Stollman [14] extended the analysis in [29] to prove partial strong dynamical localization. In fact, they proved partial strong operator-dynamical localization. The partial refers to the fact that they obtain (1.8) and (1.10) for all $q < q_0$, for some $q_0 < \infty$ that depends on the parameters of the problem - the disorder, the energy interval where the result takes place, etc. By strong operator-dynamical localization on an interval I we mean (compare with (1.7))

$$\mathbb{E}\left(\sup_{\|f\| \leq 1} \left\| |X|^{\frac{q}{2}} f(H_\omega) E_{H_\omega}(I) \chi_\Omega \right\| \right) < \infty \,. \tag{1.10}$$

We note that the $q < q_0$ limitation comes from the fact that they only had at their disposal the usual MSA. Theorem 3.4 below is sufficient to push their analysis to full strong operator-dynamical localization, i.e., (1.10) for all $q > 0$.

In this article we propose an alternative, and quite natural, way to get strong dynamical localization (see also [30]), which yields strong HS-dynamical localization. As in [28], our method uses a generalized eigenfunction expansion (see Section 2.3) to exploit the fruits of the bootstrap MSA, instead of resorting to centers of localization as in [29, 14]. (See Remark 4.3.)

Let us give an idea of our enhancement of the multiscale analysis. Roughly, the usual MSA works as follows: fix $p > 0$ and a mass $m > 0$. In a way, the parameter p determines how good the final result will be. A box $\Lambda_L(x)$ is said to be regular at an energy E if the resolvent on that box, $R_{\Lambda_L(x)}$, sandwiched between the center of the box and its boundary, is smaller than $e^{-mL/2}$ (see Section 2.1 and Definition 3.2):

$$\|\Gamma_{x,L} R_{x,L}(E) \chi_{x,L/3}\|_{x,L} \leq \mathrm{e}^{-m\frac{L}{2}}. \tag{1.11}$$

The basic result of the usual MSA is to provide an energy interval $I(p)$ and a sequence of scales $L_k \nearrow \infty$, such that the probability of getting regular boxes at the scale L_k at energies $E \in I(p)$ is greater than $1 - L_k^{-p}$ (for precise statements we refer the reader to Sections 3 and 5.1). But this process can only take place once the first step (starting hypothesis) is proven to hold. To achieve the first step, typically one has to work at either the edge of a gap in the spectrum, or at high disorder, or at low energy. The point is that the parameters of the operator (disorder, energy, ...) are fixed depending on p to satisfy this starting hypothesis. As a consequence, in the usual MSA, as $p \to \infty$, then either: (i) at the edge of a gap in the spectrum, the interval $I(p)$ where localization holds will shrink to nothing; (ii) to obtain localization in a specified interval, the required disorder $\gamma = \gamma(p)$ increases to ∞; (iii) at low energy, the energy at which we see localization diverges. If one is interested in the decay of the kernel of the semigroup $e^{-iH_\omega t}$ as in (1.6), then this link between the rate of decay of the probability and the region in the diagram energy $\times$ disorder where the conclusions hold is unfortunate, and limits the scope of the result that can be obtained. More precisely, in that context the usual MSA can only provide results of the type: (i) at the edge of a gap in the spectrum (fixed disorder γ), there exists an interval $I(p)$, shrinking to nothing as $p \to \infty$, and a finite constant C_p, so that

$$\mathbb{E}\left(\sup_{\|f\| \leq 1} \|\chi_x f(H_\omega) E_{H_\omega}(I(p)) \chi_y\| \right) \leq \frac{C_p}{1 + |x-y|^p}; \tag{1.12}$$

(ii) in pre-specified interval I, there exist $\gamma(p) \to \infty$ as $p \to \infty$, so that if $\gamma \geq \gamma(p)$, (1.12) holds on I; (iii) at low energy, there exists $E_p \to \infty$ as $p \to \infty$, such that (1.12) holds on compact intervals $I \subset (-\infty, -E_p)$, and, in the discrete case, also if $I \subset (E_p, \infty)$. These should be compared to (1.6) above, where the desired decay does not affect the starting hypothesis.

In this article we show that once the MSA is performed for one p, then by a bootstrap argument it can be done for *any* p', on the *same* interval $I(p)$ and with the *same* disorder (and of course for the same starting hypothesis). Since this in turn means that the starting hypothesis does not affect the strength of the conclusions, another way to take advantage of this new possibility is to start with the weakest possible starting hypothesis. In a companion paper [31] we shall explore this fact in more detail and, in particular, propose a finite volume criterion that may be implemented numerically. We notice that this type of finite volume criterion is fairly close to results obtained recently in [3], a paper that deals with the discrete setting. Moreover in [3] polynomial decay (of the averaged fractional resolvent) is shown to imply exponential decay. Here, if polynomial decay holds, then any sub-exponential decay follows.

To perform the bootstrap MSA we take advantage of two kinds of multiscale analyses: one where length scales grow by multiplication by a fixed factor: $L_{k+1} = YL_k$, $Y > 1$, and another with exponentially growing length scales: $L_{k+1} = L_k^\alpha$, $\alpha > 1$. Previous proofs yielded only a pre-fixed polynomial decay of the probabilities of bad events (i.e., $\frac{1}{L_k^p}$ for a pre-fixed $p > 0$). In this context, the MSA with exponential growth of length scales is well known, it was put in the present form by von Dreifus [19] and von

Dreifus and Klein [20], simplifying the work of Fröhlich and Spencer [26] and Fröhlich, Martinelli, Scoppola and Spencer [27]. The MSA with multiplicative growth of length scales is less well known and was developed by Figotin and Klein in [24, 25], using ideas of Spencer [51], to improve the starting hypothesis of the MSA, and thereby to weaken the hypotheses of their theorems. We extend both types of multiscale analyses to obtain sub-exponential decay of the probabilities of bad events (i.e., $\mathrm{e}^{-L_k^\zeta}$ for all $0 < \zeta < 1$). A new ingredient in our extension of the MSA with exponentially growing length scales is that we allow the number of bad boxes to grow with the scale. (The bad boxes are controlled by a Wegner estimate in the usual way.) All these multiscale analyses have differing starting hypotheses, the weakest belonging to the Figotin and Klein MSA, which only requires that at some sufficiently large scale we can verify polynomial decay of the finite volume resolvent with some minimal probability, independent of the scale. (In this article we show that a probability bigger than $1 - \frac{1}{841^d}$ suffices.) It is by successively performing the four multiscale analyses, feeding the results of one into the next, thus doing a *bootstrap multiscale analysis*, that we are able to go from the weakest starting hypothesis to the strongest conclusions. Combining the results of this bootstrap multiscale analysis with the generalized eigenfunction expansion leads to (1.6).

The paper is organized as follows. In Section 2 we present, as assumptions, the properties of the random operator H_ω that are required for the multiscale analysis and its applications. In Section 3 we state the main results of this paper, namely the *bootstrap multiscale analysis* (Theorem 3.4), and its application to various manifestations of localization: sub-exponential decay of the kernel of $f(H_\omega)$ (Theorem 3.8), strong HS-dynamical localization (Corollary 3.10), and a SULE property (Theorem 3.11). In Section 4 we assume Theorem 3.4 and prove Theorem 3.8, Corollary 3.10 and Theorem 3.11. In Section 5 we discuss the four multiscale analyses (Theorems 5.1, 5.2, 5.6, and 5.7) that are used in the bootstrap multiscale analysis. In Section 6 we prove Theorem 3.4.

2 Requirements of the multiscale analysis

2.1 Finite volume

Throughout this paper we use the sup norm in $\mathbb{R}^d$:

$$|x| = \max\{|x_i|,\ i = 1, \ldots, d\}\ . \tag{2.1}$$

By $\Lambda_L(x)$ we denote the open box (or cube) of side $L > 0$:

$$\Lambda_L(x) = \{y \in \mathbb{R}^d;\ \ |y - x| < L/2\}, \tag{2.2}$$

and by $\overline{\Lambda}_L(x)$ the closed box. *In this article we will always take boxes centered at sites $x \in \mathbb{Z}^d$ with side $L \in 2\mathbb{N}$.* Very often we will require $L \in 6\mathbb{N}$; given $K \geq 6$, we set

$$[K]_{6\mathbb{N}} = \max\{L \in 6\mathbb{N};\ L \leq K\}. \tag{2.3}$$

The operator $H_{x,L}$ is defined as the restriction of H, either to the open box $\Lambda_L(x)$ with Dirichlet boundary condition, or to the closed box $\overline{\Lambda}_L(x)$ with periodic boundary condition. (We consistently work with either Dirichlet or periodic boundary condition.)

We write $R_{x,L} = (H_{x,L} - z)^{-1}$ for its resolvent. By $\| \|_{x,L}$ we denote the norm or the operator norm on $\mathrm{L}^2(\Lambda_L(x), \mathrm{d}x; \mathbb{C}^n)$.

The characteristic function of a set $\Lambda \subset \mathbb{R}^d$ is denoted by χ_Λ. If $x \in \mathbb{R}^d$ and $\ell > 0$, we let

$$\chi_{x,\ell} = \chi_{\Lambda_\ell(x)}, \quad \chi_x = \chi_{x,1} = \chi_{\Lambda_1(x)}. \tag{2.4}$$

Given a box $\Lambda_L(x)$, we set

$$\Upsilon_L(x) = \left\{ y \in \mathbb{Z}^d;\ |y - x| = \frac{L}{2} - 1 \right\}, \tag{2.5}$$

and define its (boundary) belt by

$$\tilde{\Upsilon}_L(x) = \overline{\Lambda}_{L-1}(x) \backslash \Lambda_{L-3}(x) = \bigcup_{y \in \Upsilon_L(x)} \overline{\Lambda}_1(y) \ ; \tag{2.6}$$

it has the characteristic function

$$\Gamma_{x,L} = \chi_{\tilde{\Upsilon}_L(x)} = \sum_{y \in \Upsilon_L(x)} \chi_y \ \ a.e. \tag{2.7}$$

Note that

$$|\Upsilon_L(x)| = (L-1)^d - (L-2)^d = d \int_{L-2}^{L-1} x^{d-1} dx \le d(L-1)^{d-1}. \tag{2.8}$$

We shall suppress the dependency of a box on its center when not necessary. When using boxes Λ_ℓ contained in bigger boxes Λ_L, we shall need to know that the small box is inside the belt $\tilde{\Upsilon}_L$ of the bigger one. We thus introduce the following definition.

Definition 2.1 *Let $L > \ell + 3$ and $x \in \mathbb{Z}^d$. We say that*

$$\Lambda_\ell \sqsubset \Lambda_L(x) \quad \textit{if} \quad \Lambda_\ell \subset \Lambda_{L-3}(x) \ .$$

2.2 Properties of random operators

We present here, as assumptions, the properties of the random operator H_ω that are required for the multiscale analysis and its applications. These properties are routinely verified for the random operators of interest [26, 27, 20, 33, 9, 24, 25, 42, 43, 52]. The bootstrap multiscale analysis has the same requirements as the usual one.

We fix a compact interval I_0 and an open interval $\tilde{I}_0 \supset I_0$ (we always take $\tilde{I}_0 \subset (0, \infty)$ for equation (1.2)).

2.2.1 Deterministic assumptions

The deterministic assumptions are supposed to hold almost surely, with non-random constants. We omit ω from the notation.

The first assumption is reminiscent of the Simon-Lieb inequality (SLI) in Classical Statistical Mechanics. It relates resolvents in different scales. In the discrete case it is an immediate consequence of the resolvent identity, in this context it was originally used in [26]. In the continuum, its proof requires internal estimates. For Schrödinger operators it was proved in [9]; it was adapted to classical wave operators in [24]. We state it as in [42, Lemma 3.8]. In the continuum, we typically have the constant γ_{I_0} below to be of the form $\gamma_{I_0} = \sup_{E\in I_0} \gamma_E$, with $\gamma_E = const(1 + |E|)^{\frac{1}{2}}$, where the nonrandom constant depends only on nonrandom parameters of the operator (dimension d, bounds on coefficients or potential - see [42, eq. (3.80)] for an explicit expression for classical wave operators, a similar expression holds for Schrödinger operators).

Assumption SLI. *There exists a finite constant γ_{I_0} such that, given $L, \ell', \ell'' \in 2\mathbb{N}$, $x, y, y' \in \mathbb{Z}^d$ with $\Lambda_{\ell''}(y) \sqsubset \Lambda_{\ell'}(y') \sqsubset \Lambda_L(x)$, then for any $E \in I_0$ with $E \notin \sigma(H_{x,L}) \cup \sigma(H_{y',\ell'})$ we have*

$$\|\Gamma_{x,L}R_{x,L}(E)\chi_{y,\ell''}\|_{x,L} \leq \gamma_{I_0}\|\Gamma_{y',\ell'}R_{y',\ell'}(E)\chi_{y,\ell''}\|_{y',\ell'}\|\Gamma_{x,L}R_{x,L}(E)\Gamma_{y',\ell'}\|_{x,L}\,. \tag{2.9}$$

Assumption SLI will be used in the following way: We will take $\ell'' = \ell/3$ with $\ell \in 6\mathbb{N}$, and $\ell' = k\ell/3$ with $3 \leq k \in \mathbb{N}$. By a *cell* we will mean a closed box $\overline{\Lambda}_{\ell/3}(y'')$, with $y'' \in \frac{\ell}{6}\mathbb{Z}^d$. We define $\mathbb{Z}_{\text{even}}$ and $\mathbb{Z}_{\text{odd}}$ to be the sets of even and odd integers. We take $y \in \frac{\ell}{6}\mathbb{Z}^d$, so $\chi_{y,\ell/3}$ is the characteristic function of a cell. We want the closed box $\overline{\Lambda}_{\ell'}(y')$ to be exactly covered by cells (in effect, by k^d cells); thus we specify $y' \in \frac{\ell}{3}\mathbb{Z}^d = \frac{\ell}{6}\mathbb{Z}^d_{\text{even}}$ if k is odd, and $y' \in \frac{\ell}{3}\mathbb{Z}^d + \frac{\ell}{6}(1,1,\ldots,1) = \frac{\ell}{6}\mathbb{Z}^d_{\text{odd}}$ if k is even. We then replace the boundary belt $\tilde{\Upsilon}_{\ell'}(y')$ (of width 1) by a thicker belt $\tilde{\Upsilon}_{\ell',\ell}(y')$ of width $\ell/3$. To do so, we set

$$\Upsilon_{\ell',\ell}(y') = \left\{ y'' \in \frac{\ell}{3}\mathbb{Z}^d;\ |y'' - y'| = \frac{\ell'}{2} - \frac{\ell}{6} \right\}, \tag{2.10}$$

and define the boundary ℓ-belt of $\Lambda_{\ell'}(y')$ by

$$\tilde{\Upsilon}_{\ell',\ell}(y') = \overline{\Lambda}_{\ell'}(y') \backslash \Lambda_{\ell'-2\ell/3}(y') = \bigcup_{y''\in\Upsilon_{\ell',\ell}(y')} \overline{\Lambda}_{\ell/3}(y'')\ , \tag{2.11}$$

with characteristic function

$$\Gamma_{y',\ell',\ell} = \chi_{\tilde{\Upsilon}_{\ell',\ell}(y')} = \sum_{y''\in\Upsilon_{\ell',\ell}(y')} \chi_{y'',\ell/3}\ \ a.e. \tag{2.12}$$

Note that

$$|\Upsilon_{\ell',\ell}(y')| = (k^d - (k-2)^d) \leq k^d\,. \tag{2.13}$$

Since $\Gamma_{y',\ell',\ell}\Gamma_{y',\ell'} = \Gamma_{y',\ell'}$, the projection $\Gamma_{\ell'}$ on the belt of $\Lambda_{\ell'}$ can be replaced by the projection over the thicker belt of width $\ell/3$, which can be decomposed in boxes of side $\ell/3$. Thus (2.9) yields

$$\begin{aligned} &\|\Gamma_{x,L}R_{x,L}(E)\chi_{y,\ell/3}\|_{x,L} \\ &\leq\ \gamma_{I_0}\, k^d\, \|\Gamma_{y',\ell'}R_{y',\ell'}(E)\chi_{y,\ell/3}\|_{y',\ell'}\|\Gamma_{x,L}R_{x,L}(E)\chi_{y'',\ell/3}\|_{x,L}\ , \end{aligned} \tag{2.14}$$

for some $y'' \in \Upsilon_{\ell',\ell}(y')$. *We will say that, after performing the SLI, i.e., using the estimate (2.14), we moved from the cell $\overline{\Lambda}_{\ell/3}(y)$ to the cell $\overline{\Lambda}_{\ell/3}(y'')$.*

Remark 2.2 *While performing a multiscale analysis we will use (2.14) with either $\ell' = \ell$ (for good boxes), or some $\ell' = k\ell/3$, $k > 3$, which will be the side of a bad box. Note that in the first case, $k = 3$, and the geometric factor is $3^d - 1 \leq 3^d$. In that case note also that we must have $y = y'$ and $|y'' - y| = \ell/3$, so after performing the SLI we moved to an adjacent cell, i.e., by $\ell/3$ in the sup norm. (Recall that we are using the sup norm in $\mathbb{R}^d$, so we may move both sidewise and along the diagonals.)*

The second assumption estimates generalized eigenfunctions (see Section 2.3 for a precise definition) in terms of finite volume resolvents. It is not needed for the multiscale analysis, but it plays an important role in obtaining localization from the multiscale analysis [27, 20]. We call it an *eigenfunction decay inequality* (EDI), since it translates decay of finite volume resolvents into decay of generalized eigenfunctions ; we present it as in [42, Lemma 3.9]. It is closely related to the SLI, the proofs being very similar.

Assumption EDI. *There exists a finite constant $\tilde{\gamma}_{I_0}$ such that, given a generalized eigenfunction ψ of H with generalized eigenvalue $E \in I_0$, we have for any $x \in \mathbb{Z}^d$ and $L \in 2\mathbb{N}$ with $E \notin \sigma(H_{x,L})$ that*

$$\|\chi_x \psi\| \leq \tilde{\gamma}_{I_0} \|\Gamma_{x,L} R_{x,L}(E) \chi_x\|_{x,L} \|\Gamma_{x,L} \psi\| \ . \tag{2.15}$$

Typically we have $\tilde{\gamma}_{I_0} = \gamma_{I_0}$, with γ_{I_0} as in (2.9). We will use the following consquence of (2.15):

$$\|\chi_x \psi\| \leq \gamma'_{I_0} L^{d-1} \|\Gamma_{x,L} R_{x,L}(E) \chi_x\|_{x,L} \|\chi_y \psi\| \tag{2.16}$$

for some $y \in \Upsilon_L(x)$, with $\gamma'_{I_0} = d\tilde{\gamma}_{I_0}$.

2.2.2 Probabilistic assumptions

The first probabilistic assumption is *independence at a distance* (IAD). It says that if boxes are far apart, events related to the restrictions of the random operator H_ω to these boxes are independent. We say that an event is based on the box $\Lambda_L(x)$ if it is determined by conditions on the restriction $H_{\omega,x,L}$. Given $\varrho > 0$, we say that two boxes $\Lambda_L(x)$ and $\Lambda_{L'}(x')$ are *ϱ-nonoverlapping* if $|x - x'| > \frac{L+L'}{2} + \varrho$ (i.e., if $\mathrm{d}(\Lambda_L(x), \Lambda_{L'}(x')) > \varrho$).

Assumption IAD. *There exists $\varrho > 0$ such that events based on ϱ-nonoverlapping boxes are independent.*

The second probabilistic assumption is an "a priori" estimate on the average *number of eigenvalues* (NE) of finite volume random operators in a fixed, bounded interval. It is usually proved by a deterministic argument, using the well known bound for the Laplacian [9, 24, 25, 42]. It is, of course, entirely obvious in the discrete case.

Assumption NE *There exists a finite constant C_{I_0} such that*

$$\mathbb{E}\left(\mathrm{tr}_{\mathcal{H}} E_{H_{\omega,x,L}}(\tilde{I}_0)\right) \leq C_{I_0} L^d \tag{2.17}$$

for all $x \in \mathbb{Z}^d$ and $L \in 2\mathbb{N}$.

The final probabilistic assumption is a form of *Wegner's estimate* (W), a probabilistic estimate on the size of the resolvent. It is a crucial ingredient for the MSA, where it is used to control the bad regions.

Assumption W *For some $b \geq 1$ there exists a constant $Q_{I_0} < \infty$, such that*

$$\mathbb{P}\{\mathrm{dist}(\sigma(H_{\omega,x,L}), E) \leq \eta\} \leq Q_{I_0} \eta L^{bd} \ , \tag{2.18}$$

for all $E \in \tilde{I}_0$, $\eta > 0$, $x \in \mathbb{Z}^d$, and $L \in 2\mathbb{N}$.

Remark 2.3 *In the continuum one usually proves the stronger estimate [33, 9, 10, 4, 24, 25, 43, 11]:*

$$\mathbb{E}\left(\operatorname{tr}_{\mathcal{H}} E_{H_{\omega,x,L}}\left([E-\eta, E+\eta]\right)\right) \leq Q_{I_0} \eta L^{bd} \ , \tag{2.19}$$

from which (2.18) follows by Chebychev's inequality. The estimate (2.17) is used as an "a priori" estimate in the proof of (2.19).

Remark 2.4 *In practice we have either $b = 1$ or $b = 2$ in the Wegner estimate (2.18). For random Schrödinger operators with Anderson potential we may have $b = 1$ [9, 45, 4] (including the Landau Hamiltonian). For classical waves in random media, (2.18) has been proven with $b = 2$ [54, 24, 25, 43]. Very recently the correct volume dependency (i.e., $b = 1$) in gaps of the unperturbed operator was obtained in [11], at the price of losing a bit in the η dependency. In this paper, we shall use (2.18) as stated, the modifications in our methods required for the other forms of (2.18) being obvious. Our methods may also accomodate Assumptions NE and W being valid only for large L, and/or Assumption W being valid only for $\eta < \eta_L$ for some appropriate η_L, say $\eta_L = L^{-r}$, some $r > 0$, or $\eta_L = \mathrm{e}^{-L^\beta}$ for some $0 < \beta < 1$. The latter is of importance if one wants to deal with singular probability measures like Bernoulli [7, 36, 16].*

2.3 Generalized eigenfunction expansion

Generalized eigenfunction expansions were originally developed for elliptic partial differential operators with smooth coefficients (we refer to Berezanskii's book [6]). These expansions were extended to Schrödinger operators with singular potentials (Simon [50] and references therein), and to classical wave operators with nonsmooth coefficients by Klein, Koines and Seifert [44].

These expansions construct polynomially bounded generalized eigenfunctions for a set of generalized eigenvalues with full spectral measure. These generalized eigenfunctions were used by Pastur [48] and by Martinelli and Scoppola [47] to prove that certain Schrödinger operators with random potentials have no absolutely continuous spectrum. They played a crucial role in the work by Fröhlich, Martinelli, Spencer and Scoppola [27] and by von Dreifus and Klein [20] on Anderson localization of random Schrödinger operators, providing the crucial link between the multiscale analysis and pure point spectrum: the exponential decay of finite volume Green's functions (obtained by a multiscale analysis) forces polynomially bounded generalized eigenfunctions to be bona fide eigenfunctions, so the spectrum is at most countable and hence pure point.

In this article we go further and, as in [28, 30], use the generalized eigenfunction expansion itself (not just the existence of polynomially bounded generalized eigenfunctions) to provide the link between the multiscale analysis and strong HS-dynamical localization (and hence pure point spectrum). We will state the existence of a generalized eigenfunction expansion as an assumption. Proofs of such an assumption are provided in [50, 44] for Schrödinger operators and classical wave operators. We follow the presentation in [44].

Let T be the operator on $\mathcal{H}$ given by multiplication by the function $(1+|x|^2)^\nu$, where $\nu > d/4$. We define the weighted spaces $\mathcal{H}_\pm$ as follows:

$$\mathcal{H}_\pm = \mathrm{L}^2(\mathbb{R}^d, (1+|x|^2)^{\pm 2\nu} \mathrm{d}x; \mathbb{C}^n) \ . \tag{2.20}$$

$\mathcal{H}_-$ is a space of polynomially L^2-bounded functions. The sesquilinear form

$$\langle \phi_1, \phi_2 \rangle_{\mathcal{H}_+,\mathcal{H}_-} = \int \overline{\phi_1(x)} \cdot \phi_2(x) \mathrm{d}x, \tag{2.21}$$

where $\phi_1 \in \mathcal{H}_+$ and $\phi_2 \in \mathcal{H}_-$, makes $\mathcal{H}_+$ and $\mathcal{H}_-$ conjugate duals to each other. By $O^\dagger$ we will denote the adjoint of an operator O with respect to this duality. By construction, $\mathcal{H}_+ \subset \mathcal{H} \subset \mathcal{H}_-$, the natural injections $\imath_+ : \mathcal{H}_+ \to \mathcal{H}$ and $\imath_- : \mathcal{H} \to \mathcal{H}_-$ being continuous with dense range, with $\imath_+^\dagger = \imath_-$. The operators $T_+ : \mathcal{H}_+ \to \mathcal{H}$ and $T_- : \mathcal{H} \to \mathcal{H}_-$, defined by $T_+ = T\imath_+$, $T_- = \imath_- T$ on $\mathcal{D}(T)$, are unitary with $T_- = T_+^\dagger$. The map $\tau : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H}_+, \mathcal{H}_-)$, with $\tau(C) = T_- C T_+$, is a Banach space isomorphism, as $T_\pm$ are unitary operators. ($\mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$ denotes the Banach space of bounded operators from $\mathcal{H}_1$ to $\mathcal{H}_2$, $\mathcal{B}(\mathcal{H}) = \mathcal{B}(\mathcal{H}, \mathcal{H})$.) If $1 \le q < \infty$, we define $\mathcal{T}_q(\mathcal{H}_+, \mathcal{H}_-) = \tau(\mathcal{T}_q(\mathcal{H}))$, where $\mathcal{T}_q(\mathcal{H})$ denotes the Banach space of bounded operators S on $\mathcal{H}$ with $\|S\|_q = (\mathrm{tr}\,|S|^q)^{\frac{1}{q}} < \infty$. By construction, $\mathcal{T}_q(\mathcal{H}_+, \mathcal{H}_-)$, equipped with the norm $\|B\|_q = \|\tau^{-1}(B)\|_q$, is a Banach space isomorphic to $\mathcal{T}_q(\mathcal{H})$, with $\mathcal{T}_2(\mathcal{H}_+, \mathcal{H}_-)$ being the usual Hilbert space of Hilbert-Schmidt operators from $\mathcal{H}_+$ to $\mathcal{H}_-$.

Note that

$$\|\chi_{x,L}\|_{\mathcal{H},\mathcal{H}_+} = \|\chi_{x,L}\|_{\mathcal{H}_-,\mathcal{H}} \le C_{L,\nu}(1 + |x|^2)^\nu \tag{2.22}$$

for all $x \in \mathbb{R}^d$ and $L > 0$, with $C_{L,\nu}$ a finite constant depending only on L and ν. (Given an operator $B : \mathcal{H}_1 \to \mathcal{H}_2$, $\|B\|_{\mathcal{H}_1,\mathcal{H}_2}$ will denote its operator norm.)

The following assumption guarantees the existence of a generalized eigenfunction expansion (GEE) with the right properties (see [44] for details). The following assumption guarantees the existence of a generalized eigenfunction expansion (GEE) with the right properties (see [44] for details). Recall that $P_H^\perp$ is the orthogonal projection on the orthogonal complement of the kernel of H in the case of classical waves; for convenience we let it be the identity operator in the case of the Schrödinger equation. Note also that we fix $\nu > d/4$ and use the corresponding operator T and weighted spaces $\mathcal{H}_\pm$ as in (2.20).

Assumption GEE. *Fix $\nu > d/4$. The set $\mathcal{D}_+^\omega := \{\phi \in \mathcal{D}(H_\omega) \cap \mathcal{H}_+,\ H_\omega \phi \in \mathcal{H}_+\}$ is dense in $\mathcal{H}_+$ and an operator core for H_ω, with probability one. There exists a bounded, continuous function f, strictly positive on the spectrum of H_ω, such that*

$$tr_{\mathcal{H}} \left(T^{-1} f(H_\omega) P_{H_\omega}^\perp T^{-1} \right) < \infty \tag{2.23}$$

with probability one.

A measurable function $\psi : \mathbb{R}^d \to \mathbb{C}^n$ is said to be a *generalized eigenfunction* of H_ω with generalized eigenvalue λ, if $\psi \in \mathcal{H}_-$ and

$$\langle H_\omega \phi, \psi \rangle_{\mathcal{H}_+,\mathcal{H}_-} = \lambda \langle \phi, \psi \rangle_{\mathcal{H}_+,\mathcal{H}_-} \text{ for all } \phi \in \mathcal{D}_+^\omega.$$

It follows from GEE that if a generalized eigenfunction is in $\mathcal{H}$, then it is a bona fide eigenfunction.

If GEE holds, for almost every ω we have

$$\mathrm{tr}_{\mathcal{H}} \left(T^{-1} E_{H_\omega}(J) P_{H_\omega}^\perp T^{-1} \right) < +\infty \tag{2.24}$$

for all bounded Borel sets J. Thus, with probability one,

$$\mu_\omega(J) = \mathrm{tr}_{\mathcal{H}}\left(T^{-1}E_{H_\omega}(J)P^{\perp}_{H_\omega}T^{-1}\right) \tag{2.25}$$

is a spectral measure for the restriction of H_ω to the Hilbert space $P^{\perp}_{H_\omega}\mathcal{H}$, with

$$\mu_\omega(J) < \infty \quad \text{for} \quad J \quad \text{bounded.} \tag{2.26}$$

In particular, we have a generalized eigenfunction expansion for H_ω: with probability one, there exists a μ_ω-locally integrable function $P_\omega(\lambda)$ from the real line into $\mathcal{T}_1(\mathcal{H}_+,\mathcal{H}_-)$, with

$$P_\omega(\lambda) = P_\omega(\lambda)^\dagger \tag{2.27}$$

and

$$\mathrm{tr}_{\mathcal{H}}\left(T_-^{-1}P_\omega(\lambda)T_+^{-1}\right) = 1 \quad \text{for } \mu_\omega - \text{a.e. } \lambda \ , \tag{2.28}$$

such that

$$\imath_- E_{H_\omega}(J)P^{\perp}_{H_\omega}\imath_+ = \int_J P_\omega(\lambda)\, d\mu_\omega(\lambda) \quad \text{for bounded Borel sets} \quad J \ , \tag{2.29}$$

where the integral is the Bochner integral of $\mathcal{T}_1(\mathcal{H}_+,\mathcal{H}_-)$-valued functions. Moreover, if $\phi \in \mathcal{H}_+$, then $P_\omega(\lambda)\phi \in \mathcal{H}_-$ is a generalized eigenfunction of H_ω with generalized eigenvalue λ, for μ_ω almost every λ.

The following lemma will play an important role in our proof of strong HS-dynamical localization. Note that the constant C in (2.30) is independent of λ, and that$\|\ \|_1$ denotes the trace norm in $\mathcal{H}$.

Lemma 2.5 *Under Assumption GEE, we have, with probability one, that for μ_ω almost every λ,*

$$\|\chi_x P_\omega(\lambda)\chi_y\|_1 \leq C(1+|x|^2)^\nu(1+|y|^2)^\nu \tag{2.30}$$

for all $x, y \in \mathbb{R}^d$, with C a finite constant independent of λ and ω.

Proof: Since

$$\|\chi_x P_\omega(\lambda)\chi_y\|_1 \leq \|\chi_x\|_{\mathcal{H}_-,\mathcal{H}}\|P_\omega(\lambda)\|_{\mathcal{T}_1(\mathcal{H}_+,\mathcal{H}_-)}\|\chi_y\|_{\mathcal{H},\mathcal{H}_+} \ , \tag{2.31}$$

(2.30) follows from (2.22) and (2.28). □

Assumption GEE suffices for proofs of localization [27, 20] and (almost sure) dynamical localization [29, 28]. But for strong HS-dynamical localization we need to strengthen (2.23), as we will use (2.34) below.

Assumption SGEE. *Assumption GEE holds with*

$$\mathbb{E}\left[tr_{\mathcal{H}}\left(T^{-1}f(H_\omega)P^{\perp}_{H_\omega}T^{-1}\right)\right]^2 < \infty \ . \tag{2.32}$$

It follows that

$$\mathbb{E}\left[\mathrm{tr}_{\mathcal{H}}\left(T^{-1}E_{H_\omega}(J)P^{\perp}_{H_\omega}T^{-1}\right)\right]^2 < +\infty \tag{2.33}$$

for all bounded Borel sets J, so we have a stronger version of (2.26):

$$\mathbb{E}\left[\mu_\omega(J)\right]^2 < \infty \quad \text{for} \quad J \quad \text{bounded.} \tag{2.34}$$

Remark 2.6 *Estimate (2.32) is true for the usual random operators. We could have required either the weaker*

$$\mathbb{E}\left[tr_{\mathcal{H}}\left(T^{-1}f(H_\omega)P_{H_\omega}^{\perp}T^{-1}\right)\right] < \infty\,, \tag{2.35}$$

or the stronger

$$\left\|tr_{\mathcal{H}}\left(T^{-1}f(H_\omega)P_{H_\omega}^{\perp}T^{-1}\right)\right\|_\infty < \infty\,. \tag{2.36}$$

If we assume (2.35) instead of (2.32), Theorem 3.8 yields strong operator dynamical localization instead of strong HS-dynamical localization. One usually proves the stronger (2.36) (e.g., [44, 43]), which was one of the assumptions in [14].

3 Statement of the main results

In order to state our results we need first to characterize *good* boxes for random operators. We start with two definitions of good boxes. Note that these are deterministic; we omit ω from the notation when not necessary.

Definition 3.1 *Given $\theta > 0$, $E \in \mathbb{R}$, $x \in \mathbb{Z}^d$, and $L \in 6\mathbb{N}$, we say that the box $\Lambda_L(x)$ is (θ, E)-suitable if $E \notin \sigma(H_{x,L})$ and*

$$\|\Gamma_{x,L}R_{x,L}(E)\chi_{x,L/3}\|_{x,L} \le \frac{1}{L^\theta}\,. \tag{3.1}$$

Definition 3.2 *Given $m > 0$, $E \in \mathbb{R}$, $x \in \mathbb{Z}^d$, and $L \in 6\mathbb{N}$, we say that the box $\Lambda_L(x)$ is (m, E)-regular if $E \notin \sigma(H_{x,L})$ and*

$$\|\Gamma_{x,L}R_{x,L}(E)\chi_{x,L/3}\|_{x,L} \le \mathrm{e}^{-m\frac{L}{2}}. \tag{3.2}$$

Remark 3.3 *Note that a box $\Lambda_L(x)$ is (θ, E)-suitable if and only if it is (m, E)-regular, with $m = 2\theta\frac{\log L}{L}$. The difference between the two definitions is the point of view. In Definition 3.1 we require only polynomial decay in the scale L, while in Definition 3.2 we want exponential decay in L.*

We fix a compact interval I_0 and an open interval $\tilde{I}_0 \supset I_0$ ($\tilde{I}_0 \subset (0,\infty)$ for equation (1.2)). Throughout this paper, by $C = C(a, b, \ldots)$ we mean a positive finite constant C, depending *only* on the parameters $a, b, \ldots$.

The following theorem provides our enhancement of the multiscale analysis.

Theorem 3.4 (Bootstrap Multiscale Analysis) *Let H_ω be a random operator satisfying assumptions SLI, IAD, NE and W in the compact interval I_0. Given $\theta > bd$, there exists a finite scale $\overline{\mathcal{L}} = \overline{\mathcal{L}}(d, \varrho, Q_{I_0}, \gamma_{I_0}, b, \theta)$, such that, if for some $E_0 \in I_0$ we can verify at some finite scale $\mathcal{L} > \overline{\mathcal{L}}$ that*

$$\mathbb{P}\{\Lambda_{\mathcal{L}}(0) \ \text{ is } (\theta, E_0)\text{-suitable}\} > 1 - \frac{1}{841^d}\,, \tag{3.3}$$

then there exists $\delta_0 = \delta_0(d, \varrho, Q_{I_0}, C_{I_0}, \gamma_{I_0}, \theta, \mathcal{L}) > 0$, such that, given any ζ, $0 < \zeta < 1$, and α, $1 < \alpha < \zeta^{-1}$, there is a length scale

$L_0 = L_0(d, \varrho, Q_{I_0}, C_{I_0}, \gamma_{I_0}, \theta, \mathcal{L}, \zeta, \alpha) < \infty$, *and a mass* $m_\zeta = m(\zeta, L_0) > 0$, *so if we set* $L_{k+1} = [L_k^\alpha]_{6\mathbb{N}}$, $k = 0, 1, \ldots$, *we have*

$$\mathbb{P}\,[R\,(m_\zeta, L_k, I(\delta_0), x, y)] \geq 1 - \mathrm{e}^{-L_k^\zeta} \tag{3.4}$$

for all $k = 0, 1, \ldots$, *and* $x, y \in \mathbb{Z}^d$ *with* $|x - y| > L_k + \varrho$, *where* $I(\delta_0) = [E_0 - \delta_0, E_0 + \delta_0] \cap I_0$, *and*

$$\begin{aligned} &R(m, L, I, x, y) = \\ &\qquad \{\textit{for every } E \in I, \ \textit{ either } \Lambda_L(x) \textit{ or } \Lambda_L(y) \textit{ is } (m, E)\textit{-regular}\}\ . \end{aligned} \tag{3.5}$$

Remark 3.5 *If we have the expected volume factor in (2.18), i.e., $b = 1$, we need only $\theta > d$, hence (3.3) is an estimate on the probability that the finite volume resolvent decays faster than the inverse of the volume.*

Remark 3.6 *The initial probability $1 - \frac{1}{841^d}$ in the starting hypothesis (3.3) of Theorem 3.4 does not depend on the initial scale $\mathcal{L}$. It suffices to verify (3.3) for some $\mathcal{L} > \overline{\mathcal{L}}$, with $\overline{\mathcal{L}}$ large enough depending on $d, \varrho, Q_{I_0}, \gamma_{I_0}, b, \theta$. This is not the case in the usual MSA where the required initial probability behaves like $1 - \mathcal{L}^{-p}$ [20]. Estimates on $\overline{\mathcal{L}}$, as well as better numbers for the required initial probability, will be given in [31].*

Remark 3.7 *In some cases one may verify the starting hypothesis (3.3) by proving the stronger condition:*

$$\limsup_{\mathcal{L}\to\infty} \mathbb{P}\{\Lambda_{\mathcal{L}}(0) \ \textit{ is } (\theta, E_0)\textit{-suitable}\} > 1 - \frac{1}{841^d} \quad \textit{for some } \theta > bd. \tag{3.6}$$

In such cases one usually shows that the lim sup *is actually equal to one (e.g., [24, 25]).*

The following result combines Theorem 3.4 and the generalized eigenfunction expansion presented in Section 2.3. Under the hypotheses of Theorem 3.4, we show that one can get any sub-exponential decay of the (averaged) "kernel" of a bounded function of H_ω.

Theorem 3.8 (Decay of the Kernel) *Let H_ω be a random operator satisfying assumptions SLI, IAD, NE and W in the compact interval I_0 as in Theorem 3.4, plus assumptions EDI and SGEE. Suppose (3.3) holds at $E_0 \in I_0$ for some $\theta > bd$, and let δ_0 and $I(\delta_0)$ be as in Theorem 3.4. Then for any $0 < \zeta < 1$ there exists a finite constant $C_\zeta = C(\zeta, d, \varrho, Q_{I_0}, \gamma_{I_0}, \tilde{\gamma}_{I_0}, \theta, \nu)$, such that*

$$\mathbb{E}\left(\sup_{\|f\|\leq 1} \|\chi_x f(H_\omega) E_{H_\omega}(I(\delta_0)) \chi_y\|_2^2\right) \leq C_\zeta\, \mathrm{e}^{-|x-y|^\zeta} \tag{3.7}$$

for all $x, y \in \mathbb{Z}^d$. (The supremum is taken over bounded Borel functions f of a real variable, with $\|f\| = \sup_{t\in\mathbb{R}} |f(t)|$.)

Remark 3.9 *The initial probabilistic estimate (3.3) or (3.6) may be shown to be satisfied either at the edge of a gap in the spectrum, or at low energy, or for sufficiently high disorder in a pre-specified energy interval, e.g., [20, 9] in the Schrödinger case, [24, 25, 43] for classical waves. But, in contrast with the results from the usual MSA, the region where Theorems 3.4 and 3.8 apply (in the diagram energy $\times$ disorder) is not conditioned to the final estimate of the probability of bad events; one always gets* any *sub-exponential decay on a fixed interval $I(\delta_0)$, as shown in (3.7).*

An important application of Theorem 3.8 concerns *strong HS-dynamical localization*, as defined by (1.7).

Corollary 3.10 (Strong HS-Dynamical Localization) *Consider the wave equation (either (1.1) or (1.2)) in a random medium, and assume that the corresponding random operator H_ω satisfies the hypotheses of Theorem 3.8 in the compact interval I_0. Suppose (3.3) holds at $E_0 \in I_0$ for some $\theta > bd$, and let δ_0 and $I(\delta_0)$ be as in Theorem 3.4. Then the wave equation exhibits strong HS-dynamical localization in the energy interval $I(\delta_0)$.*

Related results have been obtained for the almost Mathieu model, a one-dimensional quasi-periodic model: dynamical localization [34, 28], and more recently strong dynamical localization [30] (for the optimal set of coupling constants).

Another measure of localization is "how localized are the eigenfunctions around their center of localization". The criterion SULE [17, 18] deals with this question. Thanks to the sub-exponential decay of the probability in Theorem 3.4, we are able to improve the control on the behavior of the eigenfunctions given in [29].

Theorem 3.11 (SULE) *Let H_ω be a random operator satisfying assumptions SLI, EDI, GEE, IAD, NE and W in the compact interval I_0. Suppose (3.3) holds at $E_0 \in I_0$ for some $\theta > bd$, and let δ_0 and $I(\delta_0)$ be as in Theorem 3.4. Then H_ω exhibits Anderson localization (pure point spectrum) in the interval $I(\delta_0)$. In addition, one gets the following form of SULE: for any $\varepsilon > 0$, there exists a mass $m_\varepsilon > 0$, and for a.e. ω there is a constant $C_{\varepsilon,\omega} < \infty$, such that, if we let $\{\phi_{n,\omega}\}_{n\in\mathbb{N}}$ be the normalized eigenfunctions of H_ω with energy $E_{n,\omega}$ in $I(\delta_0)$, there exist $\{x_{n,\omega}\}_{n\in\mathbb{N}}$, so for any $n \in \mathbb{N}$ and $x \in \mathbb{Z}^d$, we have*

$$\|\chi_x \phi_{n,\omega}\| \leq C_{\varepsilon,\omega}\, \mathrm{e}^{m_\varepsilon (\log |x_{n,\omega}|)^{1+\varepsilon}}\, \mathrm{e}^{-m_\varepsilon |x - x_{n,\omega}|} \ . \tag{3.8}$$

Moreover, the centers of localization $x_{n,\omega}$ can be reordered in such a way that $|x_{n,\omega}|$ increases with n, and

$$|x_{n,\omega}| \geq \tilde{C}_\omega\, n^{\frac{1}{4\nu}} \tag{3.9}$$

for some finite constant $\tilde{C}_\omega > 0$ for a.e. ω, where $\nu > \frac{d}{4}$ is as in GEE.

This improves the result obtained in [29]. First, because the interval $I(\delta_0)$ does not depend anymore on the chosen $\varepsilon > 0$, and second, because the control of the eigenfunctions in terms of the centers of localization $\{x_{n,\omega}\}_{n\in\mathbb{N}}$, given in (3.8), is almost polynomial (we get $\mathrm{e}^{m_\varepsilon (\log |x_{n,\omega}|)^{1+\varepsilon}}$ instead of $\mathrm{e}^{m_\varepsilon |x_{n,\omega}|^\varepsilon}$ as in [29]). Note that exponential decay of the probability of bad events in Theorem 3.4 (i.e., in (3.4)) would provide right away polynomial behavior in $|x_{n,\omega}|$, as expected. In the discrete case, the Aizenman-Molchanov [1, 3] approach supplies that polynomial behavior [18].

If one is interested in proving localization in a specified interval, then sometimes it suffices to take sufficiently large disorder to satisfy the starting hypothesis (3.3) for every energy in the interval. The following corollary re-states Theorem 3.4, Theorem 3.8, Corollary 3.10 and Theorem 3.11 in this case. The proof is a simple compactness argument. Here again, as in Remark 3.9, we improve on former results, since how large the disorder has to be is not anymore conditioned by how good one wants the final probabilistic estimates to be.

Corollary 3.12 *If for some $\theta > bd$ we have (3.3) for every energy E in the compact interval I_0, then Theorem 3.4, Theorem 3.8, Corollary 3.10, and Theorem 3.11 are valid with the whole interval I_0 substituted for $I(\delta_0)$ in the conclusions.*

Remark 3.13 *We note that our results apply to the one-dimensional, discrete Anderson model with a singular potential, like a Bernoulli or alloy potential [7, 36] (for the one-dimensional continuous case see the very recent work [15]). The Wegner estimate proved in [7, 36] for this case is slightly weaker than our Assumption W, since it holds only for sub-exponentially small distances to the spectrum rather than for any $\eta > 0$. (The fact it only holds for scales L large enough, uniformly in the interval I_0, does not affect the results). But in this case one can also prove a starting hypothesis with sub-exponentially decaying probabilities of bad events [7, 36], i.e., the starting hypothesis (5.9) of Theorem 5.7. The proof of this theorem only requires this weaker Assumption W, so our results are also valid for Bernoulli or alloy potentials. Another application of this work leads to strong dynamical localization for the random dimer model [16].*

4 Decay of the kernel and dynamical localization

In this section we assume Theorem 3.4 and prove Theorem 3.8, Corollary 3.10, and Theorem 3.11. We start with a preliminary lemma which translates the exponential decay of the resolvent of finite boxes at energy λ, as given by the multiscale analysis, in terms of an exponential decay of the kernel of the "generalized eigenprojector" $P_\omega(\lambda)$ defined before (2.27). We note that Lemma 2.5, with the uniform polynomial bound (2.30) that it provides, is a crucial tool for Lemma 4.1 below.

Lemma 4.1 *Let H_ω be a random operator satisfying assumptions EDI and GEE in some compact interval I_0. Given $I \subset I_0$, $m > 0$, $L \in 6\mathbb{N}$, and $x, y \in \mathbb{Z}^d$, let $R(m, L, I, x, y)$ be as in (3.5). If $\omega \in R(m, L, I, x, y)$, we have*

$$\|\chi_x P_\omega(\lambda)\chi_y\|_2 \le C\,\mathrm{e}^{-mL/4}(1 + |x|^2)^\nu(1 + |y|^2)^\nu, \tag{4.1}$$

for μ_ω-almost all $\lambda \in I$, with $C = C(m, d, \nu, \tilde{\gamma}_{I_0}) < +\infty$.

Proof: It follows from (2.27) that

$$\|\chi_x P_\omega(\lambda)\chi_y\|_2 = \|\chi_y P_\omega(\lambda)\chi_x\|_2\ ,$$

for μ_ω-almost every λ, so the roles played by x and y are symmetric.

Let $\omega \in R(m, L, I, x, y)$. Then for any $\lambda \in I$, either $\Lambda_L(x)$ or $\Lambda_L(y)$ is (m, λ)-regular for H_ω, let's say $\Lambda_L(x)$. Let $\phi \in \mathcal{H}$. Since for μ_ω-almost all λ and all $y \in \mathbb{Z}^d$, the vector $P_\omega(\lambda)\chi_y\phi$ is a generalized eigenfunction of H_ω with generalized eigenvalue λ, it follows from the EDI (see (2.15)), using $\chi_x = \chi_{x,L/3}\,\chi_x$, that

$$\|\chi_x P_\omega(\lambda)\chi_y\phi\| \le \tilde{\gamma}_{I_0}\|\Gamma_{x,L} R_{x,L}(\lambda)\chi_{x,L/3}\|_{x,L}\|\Gamma_{x,L} P_\omega(\lambda)\chi_y\phi\|. \tag{4.2}$$

Since $\Lambda_L(x)$ is (m, λ)-regular, we have, using also Lemma 2.5 and the definition of the HS norm, that

$$\begin{aligned} \|\chi_x P_\omega(\lambda)\chi_y\|_2 &\le \tilde{\gamma}_{I_0}\mathrm{e}^{-mL/2}\|\Gamma_{x,L} P_\omega(\lambda)\chi_y\|_2 && (4.3)\\ &\le C(\nu)\tilde{\gamma}_{I_0} d L^{d-1}\mathrm{e}^{-mL/2}(1 + (|x| + \tfrac{L}{2})^2)^\nu(1 + |y|^2)^\nu && (4.4)\\ &\le C(m, d, \nu, \tilde{\gamma}_{I_0})\mathrm{e}^{-mL/4}(1 + |x|^2)^\nu(1 + |y|^2)^\nu. && (4.5) \end{aligned}$$

$\square$

Remark 4.2 *The estimate (4.1) may be compared to the criterion WULE introduced in [28]. Indeed $P_\omega(\lambda)$ can be seen as the projection operator on the set of the generalized eigenfunctions $\widetilde{\varphi}_\lambda^\omega$ in $\mathcal{H}_-$ with energy λ. Hence (4.1) above provides, at a finite scale L, the exponential decay of the key quantity $\sum |\widetilde{\varphi}_\lambda^\omega(x)\widetilde{\varphi}_\lambda^\omega(y)|$. As in [28, 30], the fact that the eigenfunctions $\widetilde{\varphi}_\lambda^\omega$ are uniformly polynomially bounded ($\|\widetilde{\varphi}_\lambda^\omega\|_{\mathcal{H}_-} \leq 1$) is crucial for our approach.*

We are now in position to prove Theorem 3.8.

Proof of Theorem 3.8:

Let $0 < \xi < 1$. We will apply Theorem 3.4 together with the generalized eigenfunction expansion (2.29) to show that

$$\mathbb{E}\left(\sup_{\|f\| \leq 1} \|\chi_x f(H_\omega) E_{H_\omega}(I(\delta_0))\chi_0\|_2^2 \right) \leq C_\xi \, \mathrm{e}^{-|x|^\xi}, \tag{4.6}$$

for all $x \in \mathbb{Z}^d$, where $I(\delta_0) \subset I_0$ is given by Theorem 3.4. Since our random operator is $\mathbb{Z}^d$-ergodic, probabilities are translation invariant, so there is no loss of generality in taking $y = 0$.

Given $0 < \xi < 1$, we pick ζ such that $\zeta^2 < \xi < \zeta < 1$ (always possible) and set $\alpha = \frac{\zeta}{\xi}$, note $\alpha < \zeta^{-1}$. Theorem 3.4 then provides us with a scale L_0 and a mass $m_\zeta > 0$, such that, if we set $L_{k+1} = [L_k^\alpha]_{6\mathbb{N}}$, $k = 0, 1, \ldots$, then for each k we have the estimate (3.4) with $y = 0$ and $x \in \mathbb{Z}^d$ such that $|x| > L_k + \varrho$.

Let us now fix $x \in \mathbb{Z}^d$ and k such that $L_{k+1} + \varrho \geq |x| > L_k + \varrho$. In this case Lemma 4.1 asserts that if $\omega \in R(m_\zeta, L_k, I(\delta_0), x, 0)$, then

$$\sup_{\lambda \in I(\delta_0)} \|\chi_x P_\omega(\lambda)\chi_0\|_2 \leq C_1 \, \mathrm{e}^{-m_\zeta L_k/4}(1 + |x|^2)^\nu \leq C_1 C_2 \, \mathrm{e}^{-L_k^\zeta}, \tag{4.7}$$

with $C_1 = C_1(m_\zeta, d, \nu, \tilde{\gamma}_{I_0})$, $C_2 = C_2(\nu, \varrho, \zeta, \xi, m_\zeta)$. We split the expectation in (4.6) in two pieces: where (4.7) holds, and over the complementary event, which has probability less than $\mathrm{e}^{-L_k^\zeta}$ by (3.4). From (2.29) we have (note $E_{H_\omega}(I(\delta_0))P_{H_\omega}^\perp = E_{H_\omega}(I(\delta_0))$ in the case of equation (1.2), since in this case $I_0 \subset (0, \infty)$),

$$\begin{aligned}
&\sup_{\|f\| \leq 1} \|\chi_x f(H_\omega) E_{H_\omega}(I(\delta_0))\chi_0\|_2 \\
&\qquad \leq \sup_{\|f\| \leq 1} \int_{I(\delta_0)} |f(\lambda)| \, \|\chi_x P_\omega(\lambda)\chi_0\|_2 \, \mathrm{d}\mu_\omega(\lambda) && (4.8)\\
&\qquad \leq \int_{I(\delta_0)} \|\chi_x P_\omega(\lambda)\chi_0\|_2 \, \mathrm{d}\mu_\omega(\lambda). && (4.9)
\end{aligned}$$

Thus, it follows from (4.7) that [with $\mathbb{E}(F(\omega); A) \equiv \mathbb{E}(F(\omega)\chi_A(\omega))$]

$$\begin{aligned}
&\mathbb{E}\left(\sup_{\|f\| \leq 1} \|\chi_x f(H_\omega) E_{H_\omega}(I)\chi_0\|_2^2 \, ; R(m_\zeta, L_k, I(\delta_0), x, 0) \right) \\
&\qquad \leq C_1^2 C_2^2 \, \mathbb{E}((\mu_\omega(I(\delta_0)))^2) \, \mathrm{e}^{-2L_k^\zeta}. && (4.10)
\end{aligned}$$

To estimate the second term, note that using (2.25) we have

$$\begin{aligned}\|\chi_x f(H_\omega) E_{H_\omega}(I(\delta_0))\chi_0\|_2^2 &\leq \|\!| f |\!\|^2 \, \|E_{H_\omega}(I(\delta_0))\chi_0\|_2^2 \\ &\leq 4^\nu \|\!| f |\!\|^2 \mu_\omega(I(\delta_0))\,, \end{aligned} \tag{4.11}$$

so, using the Schwarz's inequality and (3.4) ,

$$\begin{aligned}\mathbb{E}\left(\sup_{\|\!| f |\!\| \leq 1} \|\chi_x f(H_\omega) E_{H_\omega}(I)\chi_0\|_2^2 \,; \omega \notin R(m_\zeta, L_k, I(\delta_0), x, 0)\right) \\ \leq 4^\nu \, [\mathbb{E}((\mu_\omega(I(\delta_0)))^2)]^{\frac{1}{2}} \, \mathrm{e}^{-\frac{1}{2}L_k^\zeta}\,. \end{aligned} \tag{4.12}$$

Since $C_3 = C_1^2 C_2^2 \; \mathbb{E}((\mu_\omega(I(\delta_0)))^2) + 4^\nu \; [\mathbb{E}((\mu_\omega(I(\delta_0)))^2)]^{\frac{1}{2}} < \infty$ in view of (2.34), we conclude from (4.10) and (4.12) that (recall $\alpha = \frac{\zeta}{\xi}$)

$$\mathbb{E}\left(\sup_{\|\!| f |\!\| \leq 1} \|\chi_x f(H_\omega) E_{H_\omega}(I(\delta_0))\chi_0\|_2^2 \right) \tag{4.13}$$

$$\leq C_5 \, \mathrm{e}^{-\frac{1}{2}L_k^\zeta} \leq C_5 \, \mathrm{e}^{-\frac{1}{2}L_{k+1}^\xi} \leq C_5 \mathrm{e}^{-\frac{1}{2}(|x|-\varrho)^\xi} \leq C_5 \mathrm{e}^{\frac{1}{2}\varrho^\xi} \mathrm{e}^{-\frac{1}{2}|x|^\xi}$$

for all $|x| \geq L_0 + \varrho$. Thus (4.6) follows (for a slightly smaller ξ), and Theorem 3.8 is proved. □

Proof of Corollary 3.10. Let $q > 0$, $y \in \mathbb{Z}^d$. We have

$$\begin{aligned} &\left\| \, |X|^{\frac{q}{2}} f(H_\omega) E_{H_\omega}(I(\delta_0))\chi_y \right\|_2^2 & (4.14) \\ &= \mathrm{tr}\; [\chi_y f(H_\omega) E_{H_\omega}(I(\delta_0)) |X|^q f(H_\omega) E_{H_\omega}(I(\delta_0))\chi_y] \\ &\leq \sum_{x\in\mathbb{Z}^d} (|x|+1)^q \,\mathrm{tr}\; [\chi_y f(H_\omega) E_{H_\omega}(I(\delta_0))\chi_x f(H_\omega) E_{H_\omega}(I(\delta_0))\chi_y] \\ &= \sum_{x\in\mathbb{Z}^d} (|x|+1)^q \; \|\chi_x f(H_\omega) E_{H_\omega}(I(\delta_0))\chi_y\|_2^2 \,. & (4.15) \end{aligned}$$

Corollary 3.10 now follows from (3.7). □

Remark 4.3 *Note that we proved strong HS-dynamical localization, and hence strong dynamical localization. without proving first Anderson localization or resorting to centers of localization (required in [29, 14]). This is because of our better use of the Assumption GEE (with Lemma 2.5), as in [28, 30], and Assumption SGEE. However, once Anderson localization is proven, one can use more refined properties of orthonormal sets of eigenfunctions as in [53], and bypass the explicit use of the generalized eigenfunctions as well as the discussion in [14] with the centers of localization. Nevertheless we point out that: 1) Assumption GEE is needed anyway to establish Anderson localization; 2) the analysis in this paper (and [28, 30]) shows that $P_\omega(\lambda)$ enters the game in a natural way.*

Proof of Theorem 3.11. The proof mimics the one in [29], taking into account the subexponential decay of the probabilities of bad events. Given $\varepsilon > 0$, we pick ζ, such that $\zeta^2 < (1+\varepsilon)^{-1} < \zeta < 1$ (always possible), and choose α, $1 < \alpha < \zeta(1+\varepsilon)$.

(Note $\alpha < \zeta^{-1}$). Applying Theorem 3.4 yields a sequence of scales with $L_{k+1} = [L_k^\alpha]_{6\mathbb{N}}$, $k = 0, 1, \ldots$, such that (3.4) holds. Following [29] and the notations therein, one defines

$$F_k = \bigcup_{x_0;\, |x_0| \le \exp\left(L_{k+1}^{(1+\varepsilon)^{-1}}\right)} E_k(x_0), \tag{4.16}$$

where $E_k(x_0)$ is the complement of the event $\bigcup_{x \in A_{L_{k+1}}(x_0)} R(m_\zeta, L_k, I(\delta_0), x_0, x)$, and $A_{L_{k+1}}(x_0)$ is an annulus as in [20, 29] ($A_{L_{k+1}}(x_0) \approx \Lambda_{L_{k+1}}(x_0) \backslash \Lambda_{L_k}(x_0)$). Using (3.4), one estimates the probability of F_k as follows:

$$\mathbb{P}(F_k) \le C\ L_k^{\alpha\, d}\ \exp(-L_k^\zeta + d L_k^{\frac{\alpha}{1+\varepsilon}}), \tag{4.17}$$

where C is a finite constant. Since $\alpha/(1+\varepsilon) < \zeta$, the Borel-Cantelli Lemma applies, and proceeding as in [29], it follows that for any $\varepsilon > 0$, there exists a mass $m_\varepsilon > 0$, and for a.e. ω there is a constant $C_{\varepsilon,\omega} < \infty$, such that, if we let $\{\phi_{n,\omega}\}_{n\in\mathbb{N}}$ be the normalized eigenfunctions of H_ω with energies $\{E_{n,\omega}\}_{n\in\mathbb{N}}$ in $I(\delta_0)$, there exist $\{x_{n,\omega}\}_{n\in\mathbb{N}}$, so for any $n \in \mathbb{N}$ and $x \in \mathbb{Z}^d$, we have (3.8).

For the sake of completeness we now show that it follows from GEE and (3.8) that the centers of localization $\{x_{n,\omega}\}_{n\in\mathbb{N}}$ can be reordered in such a way that $|x_{n,\omega}|$ increases with n and we have the lower bound (3.9). The spirit of the proof goes back to [18] (see also [29, 14]). Give $L > 0$, if $|x_{n,\omega}| \le L$ and $|x| \ge 2L$, we have $|x - x_{n\omega}| \ge \frac{|x|}{3} + \frac{L}{3}$, and it follows from (3.8) that for a.e. ω

$$\|\chi_x \phi_{n,\omega}\| \le C_{\varepsilon,\omega}\, \mathrm{e}^{-m_\varepsilon\left(\frac{L}{3} - (\log L)^{1+\varepsilon}\right)}\, \mathrm{e}^{-m_\varepsilon \frac{|x|}{3}} \le C_{\varepsilon,\omega}\, \mathrm{e}^{-m_\varepsilon \frac{|x|}{3}} \tag{4.18}$$

if $L \ge 3(\log L)^{1+\varepsilon}$. Thus for L sufficiently large (depending on ε and ω), if $|x_{n,\omega}| \le L$ we have $\|\chi_{0,2L}\phi_{n,\omega}\|^2 \ge \frac{1}{2}$, so if $N(L)$ is the cardinal of the set $\{n,\, E_{n,\omega} \in I(\delta_0),\, |x_{n,\omega}| \le L\}$, we conclude that

$$\begin{aligned}
\frac{1}{2} N(L) &\le \sum_{n, E_{n,\omega} \in I(\delta_0)} \|\chi_{0,2L}\phi_{n,\omega}\|^2 = \|\chi_{0,2L} E_{H_\omega}(I(\delta_0))\|_2^2 \\
&\le (1 + 4L^2)^{2\nu} \left\| T^{-1} E_{H_\omega}(I(\delta_0)) \right\|_2^2 \\
&= (1 + 4L^2)^{2\nu} \operatorname{tr}_{\mathcal{H}} \left(T^{-1} E_{H_\omega}(I(\delta_0)) T^{-1} \right) \\
&= (1 + 4L^2)^{2\nu} \mu_\omega\left(I(\delta_0)\right),
\end{aligned} \tag{4.19}$$

where $\mu_\omega(I(\delta_0)) < \infty$ for a.e. ω by (2.26) (recall $E_{H_\omega}(I(\delta_0)) = E_{H_\omega}(I(\delta_0)) P_{H_\omega}^\perp$). It follows that $N(L) < \infty$ for all $L > 0$, so we may reorder the centers $x_{n,\omega}$ in such a way that $|x_{n,\omega}|$ increases with n, so we have $N(|x_{n,\omega}|) \ge n$. Thus, with probability one, for n large enough, depending on ω (so that $|x_{n,\omega}| \ge 1$), we have

$$n \le N(|x_{n,\omega}|) \le 2\mu_\omega\left(I(\delta_0)\right)(1 + 4|x_{n,\omega}|^2)^{2\nu} \le 2\, 5^{2\nu} \mu_\omega\left(I(\delta_0)\right) |x_{n,\omega}|^{4\nu}, \tag{4.20}$$

and the lower bound (3.9) follows. $\square$

5 Multiscale analyses

In this section we discuss the four multiscale analyses we will need for the bootstrap multiscale analysis (i.e., for the proof of Theorem 3.4). They can be classified by either the resulting estimate on the probabilities of bad events, or by the type of growth of length scales. We will state them according to the first classification, and then present the proofs conforming to the second.

5.1 Polynomially decaying probabilities

We use two multiscale analyses that yield polynomially decaying probabilities for bad events.

Theorem 5.1 ([24, Lemma 36]) *Let H_ω be a random operator satisfying assumptions SLI, IAD and W in some compact interval I_0. Let $E_0 \in I_0$ and $\theta > bd$. Given an odd integer $Y \geq 11$, for any p with $0 < p < \theta - bd$ we can find $\mathcal{Z} = \mathcal{Z}(d, \varrho, Q_{I_0}, \gamma_{I_0}, b, \theta, p, Y)$, such that if for some $L_0 > \mathcal{Z}$, $L_0 \in 6\mathbb{N}$, we have*

$$\mathbb{P}\{\Lambda_{L_0}(0) \text{ is } (\theta, E_0)\text{-suitable}\} > 1 - (3Y-4)^{-2d}, \tag{5.1}$$

then, setting $L_{k+1} = YL_k$, $k = 0, 1, 2, \ldots$, we have that

$$\mathbb{P}\{\Lambda_{L_k}(0) \text{ is } (\theta, E_0)\text{-suitable}\} \geq 1 - \frac{1}{L_k^p} \tag{5.2}$$

for all $k \geq \mathcal{K}$, where $\mathcal{K} = \mathcal{K}(p, Y, L_0) < \infty$.

The value of Theorem 5.1 is that it requires a very weak starting hypothesis, in which the probability of the bad event is independent of the scale, and its conclusion, in view of Remark 3.3, gives the starting hypothesis of a modified form of the usual multiscale analysis, as given in the following theorem. We stated Theorem 5.1 in a slightly different form than in [24, Lemma 36]; it is adapted to our assumptions and definitions.

Theorem 5.2 ([24, Theorem 32]) *Let H_ω be a random operator satisfying assumptions SLI, IAD, NE and W in some compact interval I_0. Let $E_0 \in I_0$, $\theta > bd$ and $0 < p' < p < \theta - bd$. Then for $1 < \alpha < 1 + \frac{p'}{p'+2d}$, there is $\mathcal{B} = \mathcal{B}(d, \varrho, Q_{I_0}, C_{I_0}, \gamma_{I_0}, b, \theta, p, p', \alpha)$, such that, if at some finite scale $L_0 \geq \mathcal{B}$ we verify that*

$$\mathbb{P}\{\Lambda_{L_0}(0) \text{ is } (2\theta \tfrac{\log L_0}{L_0}, E_0)\text{-regular}\} \geq 1 - \frac{1}{L_0^p}\ , \tag{5.3}$$

then there exists $\delta_1 = \delta_1(d, \varrho, Q_{I_0}, \gamma_{I_0}, \theta, p, p', \alpha, L_0) > 0$, such that if we set $I(\delta_1) = [E_0 - \delta_1, E_0 + \delta_1] \cap I_0$, $m_0 = 2\theta \frac{\log L_0}{L_0}$, and $L_{k+1} = [L_k^\alpha]_{6\mathbb{N}}$, $k = 0, 1, \ldots$, we have

$$\mathbb{P}\{\Lambda_{L_k}(0) \text{ is } (\tfrac{m_0}{2}, E)\text{-regular}\} \geq 1 - \frac{1}{L_k^{p'}} \quad \text{for all } E \in I(\delta_1), \tag{5.4}$$

and

$$\mathbb{P}\ \left[R\left(\tfrac{m_0}{2}, L_k, I(\delta_1), x, y\right)\right] \geq 1 - \frac{1}{L_k^{2p'}} \text{ for } x, y \in \mathbb{Z}^d\,, |x-y| > L_k + \varrho\ , \tag{5.5}$$

for all $k = 0, 1, \ldots$.

Theorem 5.2 is quite close to the usual multiscale analysis result [20]. The crucial difference is that Theorem 5.2 allows the mass to go to zero as the inital scale L_0 goes to infinity, which may seem very surprising at the first sight. Indeed, in the usual versions of the MSA (e.g., [26, 27, 19, 20, 9, 45, 38, 29, 52]), the mass *has* to be fixed first in order to know how large L_0 has to be chosen. It turns out that one can handle a mass depending on the scale, as in (5.3) above, i.e., a mass proportional to $\log L_0/L_0$ [24, Theorem 32]. Thus the starting hypothesis (5.3) only requires the decay of the resolvent on finite boxes to be polynomially small in the scale (see Remark 3.3), not exponentially small. Note also that by using the SLI as in (2.14), so we only move between cells, we only need to require $p > 0$ as in [38], not $p > d$ as in [20] (we need to consider only the $\left(3\frac{L}{\ell}\right)^d$ cells that are cores of boxes of side ℓ inside the bigger box of side L, instead of L^d boxes as in [20]).

We will only need the weaker conclusion (5.4) for the bootstrap multiscale analysis; we also stated (5.5) because it is the usual conclusion of this multiscale analysis.

Remark 5.3 *In Theorem 5.2 the length scale $\mathcal{B}$ is increasing in $p' \in (0,p)$, and the half-interval length δ_1 depends on p and p'. This should be compared with Theorem 3.4, where the length scale $\overline{\mathcal{L}}$ and the half-interval length δ_0, while depending on θ, are independent of the parameters in the conclusion (3.4).*

5.2 Sub-exponentially decaying probabilities

Previous multiscale analyses only yielded polynomially decaying probabilities for bad events. We now provide new versions of two multiscale analyses that give sub-exponential decay for the probabilities of bad events. We believe our method can yield any decay strictly slower than exponential.

Definition 5.4 *Given $\zeta \in (0,1)$, $E \in \mathbb{R}$, $x \in \mathbb{Z}^d$, and $L \in 6\mathbb{N}$, we say that the box $\Lambda_L(x)$ is (ζ, E)-sub-exponentially-suitable, if $E \notin \sigma(H_{x,L})$ and*

$$\|\Gamma_{x,L} R_{x,L}(E) \chi_{x,L/3}\|_{x,L} \le \mathrm{e}^{-L^{\zeta}} . \tag{5.6}$$

Remark 5.5 *A box $\Lambda_L(x)$ is (ζ, E)-sub-exponentially-suitable if and only if it is $(2L^{\zeta-1}, E)$-regular.*

The multiscale analysis with multiplicative growth of length scales has the following sub-exponential version (compare with Theorem 5.1).

Theorem 5.6 *Let H_ω be a random operator satisfying assumptions SLI, IAD and W in some compact interval I_0. Let $E_0 \in I_0$ and $\zeta_0 \in (0,1)$. Given an odd integer $Y \ge 11^{\frac{1}{1-\zeta_0}}$, for any ζ_1 with $0 < \zeta_1 < \zeta_0$ we can find $\mathcal{Z} = \mathcal{Z}(d, \varrho, Q_{I_0}, \gamma_{I_0}, b, \zeta_0, \zeta_1, Y)$, such that if for some $L_0 > \mathcal{Z}$, $L_0 \in 6\mathbb{N}$, we have*

$$\mathbb{P}\{\Lambda_{L_0}(0) \ \textit{is } (\zeta_0, E_0)\textit{-sub-exponentially-suitable}\} > 1 - (3Y-4)^{-2d}, \tag{5.7}$$

then, setting $L_{k+1} = Y L_k$, $k = 0, 1, 2, \ldots$, we have that

$$\mathbb{P}\{\Lambda_{L_k}(0) \ \textit{is } (\zeta_0, E_0)\textit{-sub-exponentially-suitable}\} \ge 1 - \mathrm{e}^{-L_k^{\zeta_1}} \tag{5.8}$$

for all $k \ge \mathcal{K}$, where $\mathcal{K} = \mathcal{K}(\zeta_0, \zeta_1, Y, L_0) < \infty$.

The well known multiscale analysis with exponential growth of length scales has the following sub-exponential version (compare with Theorem 5.2). In order to get sub-exponential decay of probabilities, our proof allows the number of bad boxes to grow with the scale.

Theorem 5.7 *Let H_ω be a random operator satisfying assumptions SLI, IAD, NE and W in some compact interval I_0. Let $E_0 \in I_0$, $0 < \zeta_2 < \zeta_1 < \zeta_0 < 1$. Then for $1 < \alpha < \zeta_0/\zeta_1$, there is $\mathcal{C} = \mathcal{C}(d, \varrho, Q_{I_0}, C_{I_0}, \gamma_{I_0}, b, \zeta_0, \zeta_1, \zeta_2, \alpha)$, such that, if at some finite scale $L_0 \geq \mathcal{C}$, $L_0 \in 6\mathbb{N}$, we verify that*

$$\mathbb{P}\{\Lambda_{L_0}(0) \text{ is } (2L_0^{\zeta_0 - 1}, E_0)\text{-regular}\} \geq 1 - \mathrm{e}^{-L_0^{\zeta_1}}, \tag{5.9}$$

then there exists $\delta_2 = \delta_2(d, \varrho, Q_{I_0}, \gamma_{I_0}, \zeta_0, \zeta_1, \zeta_2, \alpha, L_0) > 0$ such that, if we set $I(\delta_2) = [E_0 - \delta_2, E_0 + \delta_2] \cap I_0$, $m_0 = 2L_0^{\zeta_0 - 1}$, and $L_{k+1} = [L_k^\alpha]_{6\mathbb{N}}$, $k = 0, 1, \ldots$, we have

$$\mathbb{P}\left[R\left(\frac{m_0}{4}, L_k, I(\delta_2), x, y\right)\right] \geq 1 - \mathrm{e}^{-L_k^{\zeta_2}} \tag{5.10}$$

for all $k = 0, 1, 2, \ldots$ and $x, y \in \mathbb{Z}^d$ with $|x - y| > L_k + \varrho$.

We took $\frac{m_0}{4}$ in (5.10) for convenience, we may take any mass $m' = \beta m_0$ with $0 < \beta < 1$, but $\mathcal{C}$ and δ_2 will also depend on β. Note also that we allow the mass m_0 in the starting hypothesis (5.9) to decay with the scale L_0.

The equivalent to (5.4) holds in the context of Theorem 5.7, but it will not be needed.

5.3 Multiplicative growth of length scales: proofs

We now prove Theorems 5.1 and 5.6, along the lines of [24, Proof of Lemma 36].

We start by introducing some notations to facilitate the simultaneous proof of both theorems. For Theorem 5.1, we will say that a box is good if it is (θ, E_0)-suitable. Pick s such that

$$p + bd < s < \theta, \tag{5.11}$$

and set

$$q_L = L^{-p}, \quad t_L = L^{-s}, \quad u_L = L^{-\theta}. \tag{5.12}$$

For Theorem 5.6, we say that a box is good if it is (ζ_0, E_0)-sub-exponentially-suitable. Pick ξ such that

$$\zeta_1 < \xi < \zeta_0, \tag{5.13}$$

and set

$$q_L = \mathrm{e}^{-L^{\zeta_1}}, \quad t_L = \mathrm{e}^{-L^{\xi}}, \quad u_L = \mathrm{e}^{-L^{\zeta_0}}. \tag{5.14}$$

In both cases, u_L is the decay of the finite volume resolvent we want to propagate, t_L or t_L^{-1} is a control term to be used with Assumption W (the Wegner estimate), and q_L is the probability decay of a bad event we want to end up with. A box is bad if it is not good. We set p_L to be the probability of a bad box at scale L, i.e.,

$$p_L = \mathbb{P}\{\Lambda_L(0) \text{ is bad}\}. \tag{5.15}$$

Note that the conclusions (5.2) and (5.8) may now be restated as $p_{L_k} \leq q_{L_k}$ for $k \geq \mathcal{K}$.

The proof will proceed by induction. For the induction step, let $\ell \in 6\mathbb{N}$, $\ell > 3\varrho$, $Y \in \mathbb{N}$ odd, and $L = Y\ell$. Knowing p_ℓ, we will estimate p_L.

We set

$$
\begin{aligned}
\Xi_{L,\ell}(x) &= \Lambda_L(x) \cap \frac{\ell}{3}\mathbb{Z}^d \subset \mathbb{Z}^d \ , \ \ \Xi_{L,\ell} = \Xi_{L,\ell}(0)\,, && (5.16)\\
\mathcal{C}_{L,\ell}(x) &= \{\Lambda_\ell(y);\ \ y \in \Xi_{L,\ell}(x)\,, \Lambda_\ell(y) \sqsubset \Lambda_L(x)\}\,,\ \ \mathcal{C}_{L,\ell} = \mathcal{C}_{L,\ell}(0)\,, && (5.17)\\
\Xi'_{L,\ell}(x) &= \Lambda_L(x) \cap \frac{\ell}{6}\mathbb{Z}^d \subset \mathbb{Z}^d \ , \ \ \Xi'_{L,\ell} = \Xi'_{L,\ell}(0)\,. && (5.18)
\end{aligned}
$$

Note $|\Xi_{L,\ell}| = (3Y)^d$, $|\Xi'_{L,\ell}| = (6Y)^d$, $\Xi_{L,\ell} \subset \Xi'_{L,\ell}$. By a *cell* we will now mean a closed box $\overline{\Lambda}_{\ell/3}(y)$ with $y \in \Xi_{L,\ell}$, the *core* of the box $\Lambda_\ell(y)$. Thus $\mathcal{C}_{L,\ell}(x)$ is the collection of boxes of side ℓ whose core is a cell and are inside the boundary belt $\tilde{\Upsilon}_L(x)$ of the big box $\Lambda_L(x)$; we have $|\mathcal{C}_{L,\ell}| = (3Y-4)^d$. Note that the big box is divided into cells: $\overline{\Lambda}_L(0) = \bigcup_{x \in \Xi_{L,\ell}} \overline{\Lambda}_{\ell/3}(x)$.

The induction step proceeds as in [24, Lemma 36], it is based on the SLI, but only boxes in $\mathcal{C}_{L,\ell}$ are allowed. The basic idea is that, if all boxes in $\mathcal{C}_{L,\ell}$ were good in scale ℓ, then it would follow from applying the SLI (2.14) repeatedly that the big box is also good in scale L. To obtain an improvement in the probability of having a good box, we need to admit the possibility of the existence of bad boxes, to be controlled by Assumption W, but we only need to allow for a fixed number of bad boxes [19, 20].

One starts in a cell inside the core of the big box $\Lambda_L(0)$, i.e., in $\Lambda_{L/3}(0)$, apply the SLI (2.14) repeatedly, and stops just before hitting the boundary belt. Each time the SLI is performed with a good box of size ℓ, one gains the small factor u_ℓ, and moves to an adjacent cell (see Remark 2.2). Each time we must perform the SLI with a bad box, we enlarge the box slightly, so the SLI moves us to the core of a good box, where we also perform the SLI. The small factor from the latter SLI is used to control the bad factor (estimated by (2.18)) coming from the former SLI (see (5.24) below).

To make this discussion rigorous, let S denote be the maximum number of ϱ-nonoverlapping bad boxes in $\mathcal{C}_{L,\ell}$ that we shall allow. Thus at most S boxes, which must be ϱ-nonoverlapping, may be bad, out of a total of $(3Y-4)^d$ boxes, and we will control the probability of such an event. The bad boxes produce bad regions, such that any box in $\mathcal{C}_{L,\ell}$ outside these bad regions must be good. If one has one bad box, then to be sure that another box of size ℓ is ϱ-nonoverlapping, it suffices to add to the bad box an exterior belt of size $2\ell/3$ (recall $\ell > 3\varrho$), and consider boxes of size ℓ with cores outside this region. So the bad region will have size $\ell + 4\ell/3 = 7\ell/3$. If one has j, $j \leq S$, bad boxes which are clustered in the worst possible way (their exterior belts of size $2\ell/3$ just touch), then the size of the bad region will be $2(2\ell/3) + j\ell + (j-1)4\ell/3 = 7j\ell/3$. Note that the bad region has center either in $\Xi_{L,\ell}$, if j is odd, or in $\Xi'_{L,\ell}$, if j is even. Since after using the SLI with a box $\Lambda_{\ell'}$ one ends up *inside* this box, on its boundary belt, we shall use slightly bigger boxes of size $\ell' = 7j\ell/3 + 2\ell/3 = (7j+2)\ell/3$, $j \leq S$, so one gets out of the bad region after executing this procedure. The bad regions are inside the big box, so we require $(7S+2)\ell/3 < L$, i.e.,

$$
Y > (7S+2)/3\,. \tag{5.19}
$$

Now let $\mathcal{F}_{L,\ell}$ denote the event that either there are at least $S+1$ ϱ-nonoverlapping bad boxes in $\mathcal{C}_{L,\ell}$, or $\operatorname{dist}(\sigma(H_{x,\ell'}, E_0)) \leq t_L$ for some $x \in \Xi'_{L,\ell}$ and ℓ' of the form

$(7j+2)\ell/3$, with $j=1,2,\cdots,S$, or $\mathrm{dist}(\sigma(H_{0,L},E_0)) \le t_L$. If $\beta(Y,S+1)$ denotes the number of possible choices of $S+1$ ϱ-nonoverlapping boxes in $\mathcal{C}_{L,\ell}$, and $S \ge 1$, we have

$$\beta(Y,S+1) \le \frac{(3Y-4)^{d(S+1)}}{(S+1)!} \le \frac{1}{2}(3Y-4)^{d(S+1)}. \tag{5.20}$$

As in [24, Lemma 36], we will show that, for L and Y large (in a sense to be specified later),

$$\{\Lambda_L(0) \text{ is bad}\} \subset \mathcal{F}_{L,\ell}, \tag{5.21}$$

so

$$\begin{aligned} p_L \le \mathbb{P}(\mathcal{F}_{L,\ell}) &\le \frac{1}{2}(3Y-4)^{d(S+1)} p_\ell^{S+1} + \left[S(6Y)^d+1\right] Q_{I_0} L^{bd} t_L && (5.22)\\ &\le \frac{1}{2}(3Y-4)^{d(S+1)} p_\ell^{S+1} + \frac{1}{2} q_L\,, && (5.23)\end{aligned}$$

where we used (2.18) to obtain the last term in (5.22). To obtain (5.23), we take L large enough: $L > \mathcal{Z}_1$, for some $\mathcal{Z}_1 = \mathcal{Z}_1(d, Q_{I_0}, b, S, Y, p, s)$ for Theorem 5.1, and $\mathcal{Z}_1 = \mathcal{Z}_1(d, Q_{I_0}, b, S, Y, \zeta_1, \xi)$ for Theorem 5.6.

To prove (5.21), note that if the event $\mathcal{F}_{L,\ell}$ does not happen (i.e., $\omega \notin \mathcal{F}_{L,\ell}$), we can find ϱ-nonoverlapping boxes Λ_{ℓ_i}, $i=1,\ldots,r \le S$, with $\ell_i \in \{7j\ell/3;\ j=1,2,\cdots,S\}$, and $\sum_{i=1}^r \ell_i \le 7S\ell/3$, such that if $x \in \Xi_{L,\ell}$, $x \notin \bigcup_{i=1}^r \Lambda_{\ell_i}$, the box $\Lambda_\ell(x)$ is good. We control the "bad region" Λ_{ℓ_i} as follows: we apply the SLI (2.14) twice as in [20, Lemma 4.2], first with the extended box $\Lambda_{\ell_i'}$ ($\ell_i' = \ell_i + 2\ell/3$), followed by a good box in $\mathcal{C}_{L,\ell}$. We require that the product of these two factors give rise to a number strictly smaller than one, so that if one keeps visiting a bad region infinitely often, it yields zero. In other words, taking into account that $\omega \notin \mathcal{F}_{L,\ell}$, we require

$$[\gamma_{I_0}(7S+2)^d t_L^{-1}][\gamma_{I_0} 3^d u_\ell] < 1, \tag{5.24}$$

which is true for ℓ large enough (how large depending on γ_{I_0}, Y, S, and on either θ, s or ζ_0, ξ). Thus repeated use of the SLI (2.14) yields

$$\begin{aligned} \|\Gamma_{0,L} R_{0,L}(E_0)\chi_{L/3}\|_{0,L} &\le \sum_{x\in\Xi_{L/3,\ell}} \|\Gamma_{0,L} R_{0,L}(E_0)\chi_{x,\ell/3}\|_{0,L} \\ &\le \left(\frac{L}{\ell}\right)^d \sup_{x\in\Xi_{L/3,\ell}} \|\Gamma_{0,L} R_{0,L}(E_0)\chi_{x,\ell/3}\|_{0,L} \\ &\le Y^d \left[\gamma_{I_0} 3^d u_\ell\right]^{N(Y)} t_L^{-1}, && (5.25)\end{aligned}$$

where $N(Y)$ is the number of times we are able to perform the SLI on good boxes, without using the result for the control of a "bad region" as in (5.24). (We cannot get trapped in the bad regions, if we keep getting back to a bad region after performing the SLI to control a bad region, the estimate (5.24) would drive the left-hand-side of (5.25) to zero.) To estimate $N(Y)$, note that one goes from a cell *inside* the core of the big box $\Lambda_L(0)$ to its boundary. Each time we perform the SLI on a good box in $\mathcal{C}_{L,\ell}$, one moves to an adjacent cell. The last good box that can be used has its core cell two cells away from the boundary of $\Lambda_L(0)$ (because of the boundary belt of size 3/2 of the latter); the shortest (thus the worst for our purposes) possible way is then made of $(L/3)/(\ell/3) - 1 = Y - 1$ cells. In addition to that one has to subtract the number of cells where one did not gain anything due to the bad regions, which is, in the worst case, $(7+1)S = 8S$ cells. We thus have

$$N(Y) \geq Y - 8S - 1\,. \tag{5.26}$$

Thus for $\Lambda_L(0)$ to be good, it suffices, in view of (5.25), to require

$$Y^d \left[\gamma_{I_0} 3^d u_\ell\right]^{Y-8S-1} t_L^{-1} \leq u_L\ , \tag{5.27}$$

which is true if we fix Y such that

$$\begin{aligned} &Y - 8S - 1 \geq 2 \quad \text{for Theorem 5.1,} \\ &Y - 8S - 1 \geq 2Y^{\zeta_0} \ \text{ for Theorem 5.6,} \end{aligned} \tag{5.28}$$

and then take ℓ large enough, large enough depending on γ_{I_0}, Y and on either θ, s or ζ_0, ξ. Thus (5.21) is proven.

So far we did not specify the value of S. Roughly, S has to be large enough so that the term p_ℓ^{S+1} in (5.23) can be converted into p_L. It turns out that $S = 1$ is sufficient for Theorem 5.1, as in [24, Proof of Lemma 36]. For Theorem 5.6 we take $S = [Y^{\zeta_0}]$, where $[M]$ denotes the largest integer $\leq M$.

We now set

$$\begin{aligned} &\mathcal{Y} = 11 \quad \text{for Theorem 5.1,} \\ &\mathcal{Y} = 11^{\frac{1}{1-\zeta_0}} \ \text{ for Theorem 5.6.} \end{aligned} \tag{5.29}$$

We now require Y to be an odd integer such that $Y \geq \mathcal{Y}$, so with our choice of S conditions (5.19) and (5.28) are satisfied, and we require that ℓ is large enough to obtain (5.27).

So if we pick $L_0 \geq \mathcal{Z}_2$, where $\mathcal{Z}_2$ is chosen so $\mathcal{Z}_2 \geq \mathcal{Z}_1$ and is large enough so (5.27), and hence (5.21), holds, and set $L_{k+1} = YL_k$, $k = 0, 1, 2, \cdots$, $p_k = p_{L_k}$ and $q_k = q_{L_k}$, it follows from (5.23) that

$$p_{k+1} \leq \frac{1}{2}\left((3Y-4)^d p_k\right)^{S+1} + \frac{1}{2} q_{k+1} \ \text{ for } \ k = 0, 1, 2, \ldots. \tag{5.30}$$

We are now in position to finish the argument. Notice first that if $p_k < q_k$, then $\left((3Y-4)^d p_k\right)^{S+1} \leq q_{k+1}$ for $\mathcal{Z}_2$ large enough (depending on the constants Y and on p for Theorem 5.1, ζ_0, ζ_1 for Theorem 5.6). This is clear in the first case. In the second, it comes from the choice of S, which satisfies $S + 1 \geq Y^{\zeta_0} > Y^{\zeta_1}$. With this choice of $\mathcal{Z}_2$, $p_k < q_k$ leads to

$$p_{k+1} \leq \frac{1}{2} q_{k+1} + \frac{1}{2} q_{k+1} = q_{k+1}. \tag{5.31}$$

It thus suffices to show that $p_k < q_k$ must occur at some scale. Suppose $p_{k+1} \geq q_{k+1}$ for $k = 0, 1, 2, \ldots, n-1$. It then follows from (5.30) that we have $\left((3Y-4)^d p_k\right)^{S+1} \geq q_{k+1}$ for $k = 0, 1, 2, \ldots, n-1$, so, using again (5.30), we conclude that $p_{k+1} \leq \left((3Y-4)^d p_k\right)^{S+1}$ for $k = 0, 1, \ldots, n-1$, obtaining

$$q_n \leq p_n \leq (3Y-4)^{-\frac{d(S+1)}{S}} \left((3Y-4)^{\frac{d(S+1)}{S}} p_0\right)^{(S+1)^n}. \tag{5.32}$$

We now require p_0 to be such that $(3Y-4)^{\frac{d(S+1)}{S}} p_0 < 1$. Note that in both cases

$$(3Y-4)^{-\frac{d(S+1)}{S}} \leq (3Y-4)^{2d}. \tag{5.33}$$

Thus taking p_0 so that

$$p_0 < (3Y-4)^{-2d}, \tag{5.34}$$

the right hand side of (5.32) decays much faster than than q_n, so we get a contradiction. This is clear in the case of Theorem 5.1, where q_n decays exponentially in n. For Theorem 5.6, we have

$$q_n = \exp\left(-(Y^{\zeta_1})^n L_0^{\zeta_1}\right), \tag{5.35}$$

and the contradiction comes from having chosen $S = [Y^{\zeta_0}]$ and $\zeta_0 > \zeta_1$. Thus there must be $\mathcal{K}$ depending on Y, L_0, d, and on either p or ζ_0, ζ_1, so $p_k \leq q_k$ for all $k \geq \mathcal{K}$.

Theorems 5.1 and 5.6 are proven. □

5.4 Exponential growth of length scales: proofs

The proofs of Theorems 5.2 and 5.7 may be done simultaneously, as we did for Theorems 5.1 and 5.6. For simplicity, we will only give the proof of Theorem 5.7, adapting the methods of [20] to get sub-exponential decay of the probabilities of bad events, rather than the usual polynomial decay. The modifications in the proof to obtain Theorem 5.2 will be apparent to the reader. (We refer to [20] and [24, Theorem 32] for the proof of Theorem 5.2.)

We start by deriving from (5.9) the initial step of the inductive process, i.e., (5.10) with $k = 0$, but with $\frac{m_0}{2}$ substituted for $\frac{m_0}{4}$. We recall $0 < \zeta_2 < \zeta_1 < \zeta_0 < 1$, $m_0 = 2L_0^{\zeta_0 - 1}$, and pick $\zeta_2 < \xi_1 < \zeta_1$. As in [20, p. 287], if $\Lambda_{L_0}(0)$ is (m_0, E_0)-regular, $\operatorname{dist}(\sigma(H_{0,L_0}), E_0) > \mathrm{e}^{-L_0^{\xi_1}}$, and we set

$$\delta_2 = \delta_2(m_0, L_0, \xi_1) = \frac{\mathrm{e}^{-2L_0^{\xi_1}}}{2}\left[\mathrm{e}^{-\frac{m_0}{2}\frac{L_0}{2}} - \mathrm{e}^{-m_0\frac{L_0}{2}}\right], \tag{5.36}$$

it follows from the resolvent identity that $\Lambda_{L_0}(0)$ is $(\frac{m_0}{2}, E)$-regular for all $E \in I(\delta_2) = [E_0 - \delta_2, E_0 + \delta_2] \cap I_0$. Thus, it is a consequence of (5.9) and (2.18) that

$$\begin{aligned}&\mathbb{P}\{\text{for every } E \in I(\delta_2),\ \Lambda_{L_0}(0) \text{ is } (\tfrac{m_0}{2}, E)\text{-regular}\} \\ &\qquad \geq 1 - \mathrm{e}^{-L_0^{\zeta_1}} - Q_{I_0} L_0^{bd} \mathrm{e}^{-L_0^{\xi_1}} \geq 1 - \mathrm{e}^{-\frac{1}{2}L_0^{\zeta_2}},\end{aligned} \tag{5.37}$$

provided L_0 is large enough (depending only on the parameters d, Q_{I_0}, b, ζ_1, ζ_2, ξ_1). Combining with Assumption IAD, we get that

$$\mathbb{P}\left[R\left(\tfrac{m_0}{2}, L_0, I(\delta_2), x, y\right)\right] \geq 1 - \mathrm{e}^{-L_0^{\zeta_2}} \tag{5.38}$$

for all $x, y \in \mathbb{Z}^d$ with $|x - y| > L_0 + \varrho$.

The theorem is now proven by induction. The induction step goes from scale $\ell \geq L_0$ to scale $L = [\ell^\alpha]_{6\mathbb{N}}$, with $1 < \alpha < \zeta_0/\zeta_1$. We assume that for some mass m, $\frac{m_0}{4} < m \leq \frac{m_0}{2}$, we have

$$\mathbb{P}\left[R\left(m, \ell, I(\delta_2), x, y\right)\right] \geq 1 - \mathrm{e}^{-\ell^{\zeta_2}} \tag{5.39}$$

for all $x, y \in \mathbb{Z}^d$ with $|x-y| > L_k + \varrho$. We will show that, if ℓ is large enough (in a sense to be specified), the same statement holds at scale L with a new mass m', and we will estimate $m - m'$.

We proceed as in the proof of Theorems 5.1 and 5.6; the basic idea is the same. But in order to propagate such a strong decay of the bad probabilities as in (5.39), it does not suffice to allow for a *fixed* number of bad (i.e., non regular) boxes of size ℓ inside a bigger box $\Lambda_L(x)$ of sixe L. We must allow the number of bad boxes to grow with the scale. We fix $\zeta_2 < \zeta' < \zeta_1$, and allow at most

$$S_\ell = 2[\ell^{(\alpha-1)\zeta'}] - 1 \tag{5.40}$$

ϱ-nonoverlapping bad boxes. That will produce, as in the proof of Theorems 5.1 and 5.6, bad regions Λ_{ℓ_i}, $i = 1, \ldots, r \leq S_\ell$, with $\ell_i \in \{7j\ell/3,\ j = 1, 2, \cdots, S_\ell\}$, with centers in $\Xi'_{L,\ell}(x)$ (see (5.18)). Note that

$$\sum_{i=1}^{r} \ell_i \leq \frac{14}{3} \ell^{1+(\alpha-1)\zeta_1} < 5\, \ell^{\alpha-(\alpha-1)(1-\zeta_1)}, \tag{5.41}$$

and $(\alpha-1)(1-\zeta_1) > 0$, so the sum of the sizes of the bad regions grows slower than ℓ^α.

The effect of the bad regions will be controlled as follows. We pick ξ,

$$\zeta_2 < \xi < \zeta_1,$$

and require that in a bad region of size $\ell'_i = \ell_i + 2\ell/3$ we have $\operatorname{dist}(\sigma(H_{\ell'_i}), E) > \mathrm{e}^{-L^\xi}$, so the corresponding resolvent will be estimated by a factor e^{L^ξ}. (The price we will have to pay, in terms of probabilities, is then given by (2.18) in Assumption W, with $\eta = \mathrm{e}^{-L^\xi}$.) By the same reasoning that lead us to (5.24) (we also specify $\ell > 3\varrho$), we require

$$[\gamma_{I_0}(7S_\ell + 2)^d \mathrm{e}^{L^\xi}] \left[\gamma_{I_0} 3^d \mathrm{e}^{-\frac{m}{2}\ell}\right] < 1\ . \tag{5.42}$$

We have

$$[\gamma_{I_0}(7S_\ell + 2)^d \mathrm{e}^{L^\xi}] \left[\gamma_{I_0} 3^d \mathrm{e}^{-\frac{m}{2}\ell}\right] = \gamma_{I_0}^2 [3(7S_\ell + 2)]^d \mathrm{e}^{\ell^{\alpha\xi}} \mathrm{e}^{-\frac{m}{2}\ell} \leq \mathrm{e}^{\frac{1}{4}\ell^{\zeta_0} - \frac{m}{2}\ell} \tag{5.43}$$

for ℓ (and thus L_0) large enough, depending on $\gamma_{I_0}, d, \alpha, \xi$ and ζ_0 (but not on m_0), provided

$$\alpha\xi < \zeta_0\ , \tag{5.44}$$

which is true since we picked $\alpha < \frac{\zeta_0}{\zeta_1}$. Moreover, recalling $m_0 = 2L_0^{\zeta_0 - 1}$, we have

$$m > \frac{m_0}{4} = \frac{1}{2} L_0^{\zeta_0 - 1} \geq \frac{1}{2} \ell^{\zeta_0 - 1}, \tag{5.45}$$

so (5.42) follows from (5.43).

Once we have (5.42), and assume $\operatorname{dist}(\sigma(H_{0,L}), E) > \mathrm{e}^{-L^\xi}$, the same argument used to derive (5.25) leads to

$$\|\Gamma_{0,L} R_{0,L}(E_0) \chi_{L/3}\|_{0,L} \leq \ell^{d(\alpha-1)} \left[3^d \gamma_{I_0} \mathrm{e}^{-\frac{m}{2}\ell}\right]^{N_\ell} \mathrm{e}^{L^\xi}, \tag{5.46}$$

where N_ℓ is the number of times we are guaranteed to be able to perform the SLI on good boxes, without using the result for the control of a "bad region" as in (5.42). Similarly to (5.26), we have

$$N_\ell \;\geq\; \frac{L}{\ell} - 8S_\ell - 1 \geq \frac{L}{\ell}\left(1 - 32\ell^{(\alpha-1)(\zeta'-1)}\right), \tag{5.47}$$

if, say $\ell \geq 12$ (so $1 - 6\ell^{-\alpha} \geq 1/2$; recall $L = [\ell^\alpha]_{6\mathbb{N}} \geq \ell^\alpha - 6$). Thus

$$\|\Gamma_{0,L} R_{0,L}(E_0)\chi_{L/3}\|_{0,L} \leq \mathrm{e}^{-m'\frac{L}{2}}, \tag{5.48}$$

with, using $L \geq \ell^\alpha - 6$ and (5.45),

$$\begin{aligned} m' &\geq m\left(1 - 32\ell^{(\alpha-1)(\zeta'-1)}\right) - 2\left[\frac{(\alpha-1)d\log\ell}{\ell^\alpha - 6} + \frac{\log(3^d\gamma_{I_0})}{\ell} + \frac{1}{(\ell^\alpha-6)^{1-\xi}}\right] \\ &\geq \; m\left\{1 - 32\ell^{(\alpha-1)(\zeta'-1)} - \frac{4}{\ell^{\zeta_0}}\left[\frac{(\alpha-1)d\,\ell\log\ell}{\ell^\alpha-6} + \log(3^d\gamma_{I_0}) + \frac{\ell}{(\ell^\alpha-6)^{1-\xi}}\right]\right\} \\ &\geq \; m\left(1 - \frac{C}{\ell^\tau}\right), \end{aligned} \tag{5.49}$$

for some finite constant $C = C(d, \gamma_{I_0}, \alpha) > 0$ and

$$\tau = \min\{\zeta_0, (\alpha-1)(1-\zeta'), \zeta_0 - (\alpha(\xi-1)+1)\} > 0\;; \tag{5.50}$$

note $\zeta_0 - (\alpha(\xi-1)+1) = \alpha - 1 + \zeta_0 - \alpha\xi > 0$ by (5.44).

We still need to assure that $\frac{m_0}{4} < m' \leq \frac{m_0}{2}$. This cannot be done in a single induction step, because we would need to take ℓ large depending on m. But we can do it in a way that applies to all inductive steps. Given L_0 large enough for the inductive step, $1 < \alpha < \zeta_1/\zeta_0$, we construct the sequence of length scales $L_{k+1} = L_k^\alpha$, $k = 0, 1, \ldots$. Applying the inductive step from scale L_k to scale L_{k+1}, we obtain a decreasing sequence of masses m'_k, with $m'_0 = \frac{m_0}{2}$, satisfying (5.48) and (5.49) at scale L_k. We then have

$$0 \leq \sum_{k=0}^{+\infty}(m'_k - m'_{k+1}) \leq \frac{m_0}{2}\sum_{k=0}^{+\infty}\frac{C}{L_k^\tau} < \frac{m_0}{4}, \tag{5.51}$$

provided L_0 is large enough, depending on $d, \gamma_{I_0}, \alpha, \zeta_0, \zeta', \xi$, but not on m_0. (Note that the fact that how large L_0 has to be is independent of m_0, possible in view of (5.45) and (5.49), is quite important, since in (5.9) we have $m_0 = 2L_0^{\zeta_0-1}$.) It follows that

$$\frac{m_0}{4} < m'_k \leq \frac{m_0}{2}, \quad k = 0, 1, 2, \ldots. \tag{5.52}$$

We now turn to the probability estimates of the inductive step; here we follow [20, Lemma 4.1]. To apply the just discussed deterministic argument in a given box $\Lambda_L(x)$, for a fixed energy $E \in I_0$, it suffices to require:

(i) $\operatorname{dist}(\sigma(H_{0,L}), E)) > \mathrm{e}^{-L^\xi}$.

(ii) $\operatorname{dist}(\sigma(H_{y,\ell'}), E)) > \mathrm{e}^{-L^\xi}$ for all $\ell' \in \{(7j+2)\ell/3,\ j = 1, 2, \cdots, S_\ell\}$ and $y \in \Xi'_{L,\ell}(x)$.

(iii) There are at most S_ℓ ϱ-nonoverlapping bad boxes in $\mathcal{C}_{L,\ell}(x)$.

It follows that

$$\mathbb{P}\left[R\left(m', L, I(\delta_2), x, y\right)\right] \geq \tag{5.53}$$
$$\mathbb{P}\left[\text{for any } E \in I(\delta_2),\ \text{(i), (ii) and (iii) hold for either } \Lambda_L(x) \text{ or } \Lambda_L(y)\right].$$

Thus, to complete the inductive step, it suffices to show that the the right-hand-side of (5.53) is bigger than $1 - \mathrm{e}^{-L^{\zeta_2}}$ for any $x, y \in \mathbb{Z}^d$ with $|x - y| > L + \varrho$.

Let $\tilde{I}_0 \supset I_0$ be the open interval in Assumptions NE and W, and let $\tilde{\sigma}(A) = \sigma(A) \cap \tilde{I}_0$ for any operator A. If $\Lambda_{\ell_1}(x_1)$ and $\Lambda_{\ell_2}(x_2)$ are ϱ-nonoverlapping boxes, then it follows from Assumptions IAD, NE and W that

$$\mathbb{P}\left[\operatorname{dist}\left(\tilde{\sigma}(H_{x_1,\ell_1}), \tilde{\sigma}(H_{x_2,\ell_2})\right) \leq \eta\right] \leq C_{I_0} Q_{I_0} \eta \ell_1^d \ell_2^{bd}\ , \tag{5.54}$$

by the same argument as in [20, p. 293], using (2.18) and (2.17) (see [29] for an argument using (2.19)). Thus, if we fix $x, y \in \mathbb{Z}^d$ with $|x - y| > L + \varrho$, we have

$$\begin{aligned}
&\mathbb{P}\Big[\operatorname{dist}(\tilde{\sigma}(H_{x_1,\ell_1}), \tilde{\sigma}(H_{x_2,\ell_2})) \leq 2\mathrm{e}^{-L^{\xi}} \text{ for some} \\
&\qquad x_1 \in \Xi'_{L,\ell}(x), x_2 \in \Xi'_{L,\ell}(y), \ell_1, \ell_2 \in \{L, (7j+2)\ell/3,\ j = 1, 2, \cdots, S_\ell\}\Big] \\
&\qquad \leq 2C_{I_0} Q_{I_0} (S_\ell + 1)^2 \left(\frac{6L}{\ell}\right)^{2d} L^{(b+1)d} \mathrm{e}^{-L^{\xi}} \\
&\qquad \leq 8 \cdot 36^d C_{I_0} Q_{I_0} \ell^{(\alpha-1)(\zeta'+2d)+(b+1)\alpha d} \mathrm{e}^{-L^{\xi}} \leq \frac{1}{2} \mathrm{e}^{-L^{\zeta_2}}\ ,
\end{aligned} \tag{5.55}$$

for ℓ sufficiently large, depending on $d, C_{I_0}, Q_{I_0}, b, \zeta_2, \alpha, \zeta', \xi$.

Now, let $E \in I(\delta_2)$, and suppose there exist $x_1 \in \Xi'_{L,\ell}(x)$, $\ell_1 \in \{L, (7j+2)\ell/3,\ j = 1, 2, \cdots, S_\ell\}$, with $\operatorname{dist}(\sigma(H_{x_1,\ell_1}), E)) \leq \mathrm{e}^{-L^{\xi}}$. If L is large enough, depending only on $\tilde{I}_0 \backslash I_0$, we must also have $\operatorname{dist}(\tilde{\sigma}(H_{x_1,\ell_1}), E)) \leq \mathrm{e}^{-L^{\xi}}$. If the event whose probability is estimated in (5.55) does not occur, we must have $\operatorname{dist}(\tilde{\sigma}(H_{x_2,\ell_2}), E) > \mathrm{e}^{-L^{\xi}}$, and hence also $\operatorname{dist}(\sigma(H_{x_2,\ell_2}), E) > \mathrm{e}^{-L^{\xi}}$, for all $x_2 \in \Xi'_{L,\ell}(y)$ and $\ell_2 \in \{L, (7j+2)\ell/3,\ j = 1, 2, \cdots, S_\ell\}$. Since we can interchange x and y in this argument, we can conclude that if ℓ is large enough,

$$\begin{aligned}
&\mathbb{P}\left[\text{for any } E \in I(\delta_2),\ \text{(i) and (ii) hold for either } \Lambda_L(x) \text{ or } \Lambda_L(y)\right] \\
&\qquad \geq 1 - \tfrac{1}{2} \mathrm{e}^{-L^{\zeta_2}}\ .
\end{aligned} \tag{5.56}$$

On the other hand, since we chose S_ℓ to be an odd integer, and using Assumption IAD, we have

$$\begin{aligned}
&\mathbb{P}[\text{for some } E \in I(\delta_2) \text{ there are at least } S_\ell + 1\ \varrho\text{-nonoverlapping} \\
&\qquad \text{bad boxes in } \mathcal{C}_{L,\ell}(0)] \\
&\qquad \leq \mathbb{P}[\text{for some } E \in I(\delta_2) \text{ there are at least two } \varrho\text{-nonoverlapping} \\
&\qquad\ \text{bad boxes in } \mathcal{C}_{L,\ell}(0)]^{\frac{S_\ell+1}{2}} \\
&\qquad \leq \left[\left(\frac{3L}{\ell}\right)^{2d} \mathrm{e}^{-\ell^{\zeta_2}}\right]^{\frac{S_\ell+1}{2}} \leq \left[9^d \ell^{2(\alpha-1)d} \mathrm{e}^{-\ell^{\zeta_2}}\right]^{[\ell^{(\alpha-1)\zeta'}]} &(5.57) \\
&\qquad \leq \tfrac{1}{4} \mathrm{e}^{-L^{\zeta_2}}\ , &(5.58)
\end{aligned}$$

where we used the induction hypothesis (5.39) to get (5.57). The final estimate (5.58) holds for ℓ sufficiently large, depending on $d, \zeta_2, \alpha, \zeta'$, since $\zeta_2 + (\alpha - 1)\zeta' > \alpha\zeta_2$ as $\zeta' > \zeta_2$. We can thus conclude that

$$\mathbb{P}\left[\text{for some } E \in I(\delta_2), \text{(iii) does not hold for either } \Lambda_L(x) \text{ or } \Lambda_L(y)\right] \\ \leq \tfrac{1}{2}\mathrm{e}^{-L^{\zeta_2}} \ . \tag{5.59}$$

Combining (5.53), (5.56) and (5.59), we get that

$$\mathbb{P}\left[R\left(m', L, I(\delta_2), x, y\right)\right] \geq 1 - \mathrm{e}^{-L^{\zeta_2}} \ , \tag{5.60}$$

for ℓ sufficiently large, the desired result.

Thus, if L_0 is large enough, how large depending only on the parameters d, ϱ, Q_{I_0}, $C_{I_0}, \gamma_{I_0}, b, \zeta_0, \zeta_1, \zeta_2, \alpha$, we construct the sequence of length scales $L_{k+1} = L_k^\alpha$, $k = 0, 1, \ldots$, and we may apply the inductive step from scale L_k to scale L_{k+1}, starting from (5.38) for $k = 0$, obtaining (5.60) with $L = L_k$ and $m' = m'_k$, and hence, using (5.52), the conclusion (5.10) for all $k = 0, 1, 2, \ldots$.

This finishes the proof of Theorem 5.7. □

6 Bootstrap Multiscale Analysis

We now prove Theorem 3.4. This will be done by a bootstrapping argument, making successive use of Theorems 5.1, 5.2, 5.6, and 5.7.

We start by giving an outline of the proof:

Prologue: Under the hypotheses of Theorem 3.4, we note that hypothesis (5.1) of Theorem 5.1 is the same as hypothesis (3.3) for appropriate choices of the parameters.

Act 1: We apply Theorem 5.1 obtaining a sequence of length scales satisfying conclusion (5.2), with its polynomial decay estimate of the probability of bad events.

Act 2: In view of Remark 3.3, it follows that hypothesis (5.3) of Theorem 5.2 is now satisfied at suitably large scale. (We have bootstrapped from hypothesis (3.3) to hypothesis (5.3)!). Thus we can apply Theorem 5.2 with appropriate parameters, getting $\delta_1 > 0$ and a sequence of length scales satisfying conclusion (5.4) for all $E \in I(\delta_1)$. We set $\delta_0 = \delta_1$.

Act 3: We fix ζ and α as in Theorem 3.4, and pick $\zeta_0, \zeta_1, \zeta_2$ such that $0 < \zeta < \zeta_2 < \zeta_1 < \zeta_0 < 1 < \alpha < \zeta_0\zeta_1^{-1} < \zeta_2^{-1} < \zeta^{-1}$. We note that we have bootstrapped again: hypothesis (5.7) of Theorem 5.6 is satisfied at all energies $E \in I(\delta_0)$ at appropriately large scale (the same for all E). Applying Theorem 5.6, we obtain a sequence of length scales for which conclusion (5.8) holds for all $E \in I(\delta_0)$, with its sub-exponential decay estimate of the probability of bad events.

Act 4: Using Remark 5.5, we can see that we have bootstrapped to Theorem 5.7: for any $0 < \zeta_2 < \zeta_1 < \zeta_0 < 1$, hypothesis (5.9) is satisfied at all energies $E \in I(\delta_1)$ at sufficiently large scale (depending on $\zeta_0, \zeta_1, \zeta_2$ but independent of E). We apply Theorem 5.7, obtaining $\delta_2 > 0$ and and an exponentially growing sequence of length scales, depending on $\zeta_0, \zeta_1, \zeta_2$, but independent of E, such that conclusion (5.10) holds for all $E \in I(\delta_1)$.

Epilogue: We have constructed in Act 4 a sequence of length scales for which (5.10) holds for all $E \in I(\delta_0)$. Since the interval $I(\delta_0)$ (which is independent of ζ) can be covered by $[\frac{\delta_0}{\delta_2}] + 1$ closed intervals of length δ_2, we note that the desired conclusion (3.4) now follows from (5.10), at the energies that are the centers of the $[\frac{\delta_1}{\delta_2}] + 1$ covering intervals, if we take L_0 appropriately large.

We now give the detailed proof of Theorem 3.4: Given $\theta > bd$, we pick p, $0 < p < \theta - bd$; to fixate ideas we take $p = \frac{\theta - bd}{2}$. We choose $Y = 11$, and let $\mathcal{Z} = \mathcal{Z}(d, \varrho, Q_{I_0}, \gamma_{I_0}, b, \theta, p, Y = 11)$ be as in Theorem 5.1. We take $\overline{\mathcal{L}} = \overline{\mathcal{L}}(d, \varrho, Q_{I_0}, \gamma_{I_0}, b, \theta) = \mathcal{Z}$, and note that hypothesis (5.1) of Theorem 5.1 is now the same as hypothesis (3.3) with $L_0 = \mathcal{L}$ and $(3Y - 4)^{2d} = 841^d$.

We now fix $E_0 \in I_0$ and assume (3.3) for this E_0 with $\mathcal{L} > \overline{\mathcal{L}}$. We set $L_0^{(1)} = \mathcal{L}$, and define a sequence of length scales $L_k^{(1)}$ by $L_{k+1}^{(1)} = Y L_k^{(1)}$, $k = 0, 1, 2, \ldots$. We apply Theorem 5.1, and conclude that (5.2) holds for these length scales for all $k \geq \mathcal{K}^{(1)} = \mathcal{K}^{(1)}(d, \gamma_{I_0}, b, \theta, \mathcal{L})$. In view of Remark 3.3, we have that

$$\mathbb{P}\left\{\Lambda_{L_k^{(1)}}(0) \text{ is } \left(2\theta \frac{\log L_k^{(1)}}{L_k^{(1)}}, E_0\right)\text{-regular}\right\} \geq 1 - \frac{1}{\left(L_k^{(1)}\right)^p} \tag{6.1}$$

for all $k \geq \mathcal{K}^{(1)}$.

Note that we have bootstrapped to hypothesis (5.3) of Theorem 5.2, since (6.1) is the same as (5.3) at scale $L_k^{(1)}$. We take $p' = \frac{\theta - bd}{4}$ and $\alpha_1 = 1 + \frac{p'}{2(p' + 2d)}$, and take $\mathcal{B} = \mathcal{B}(d, \varrho, Q_{I_0}, C_{I_0}, \gamma_{I_0}, b, \theta, p, p', \alpha_1)$ as in Theorem 5.2. Letting k_1 be the smallest $k \geq \mathcal{K}^{(1)}$ such that $L_k^{(1)} > \mathcal{B}$, we define length scales $L_0^{(2)} = L_{k_1}^{(1)}$, $L_{k+1}^{(2)} = \left[\left(L_k^{(2)}\right)^{\alpha_1}\right]_{6\mathbb{N}}$ for $k = 0, 1, 2, \ldots$. We apply Theorem 5.2 with $L_0 = L_0^{(2)}$ in (5.3), and conclude that (5.4) holds for these length scales with $\delta_1 = \delta_1(d, \varrho, Q_{I_0}, \gamma_{I_0}, \theta, p, p', \alpha_1, L_0^{(2)}) > 0$. Letting

$$\delta_0 = \delta_0(d, \varrho, Q_{I_0}, C_{I_0}, \gamma_{I_0}, \theta, \mathcal{L}) = \delta_1(d, \varrho, Q_{I_0}, \gamma_{I_0}, \theta, p, p', \alpha_1, L_0^{(2)}) > 0, \tag{6.2}$$

we proved that for all $k = 0, 1, 2, \ldots$ we have

$$\mathbb{P}\{\Lambda_{L_k^{(2)}}(0) \text{ is } (m_1, E)\text{-regular}\} \geq 1 - \frac{1}{\left(L_k^{(2)}\right)^{p'}} \quad \text{for all } E \in I(\delta_0), \tag{6.3}$$

with $I(\delta_0) = (E_0 - \delta_0, E_0 + \delta_0) \cap I_0$ and $m_1 = \theta \frac{\log L_0^{(2)}}{L_0^{(2)}}$.

Now let us fix ζ and α as in Theorem 3.4, so $0 < \zeta < 1 < \alpha < \zeta^{-1}$. To be definite, we take

$$\zeta_2 = \sqrt{\zeta \alpha^{-1}}\ , \quad \zeta_1 = \sqrt{\zeta_2 \alpha^{-1}}\ , \quad \zeta_0 = \sqrt{\zeta_1 \alpha}\ , \tag{6.4}$$

so we have

$$0 < \zeta < \zeta_2 < \zeta_1 < \zeta_0 < 1 < \alpha < \zeta_0 \zeta_1^{-1} < \zeta_2^{-1} < \zeta^{-1}\ . \tag{6.5}$$

Next, we apply Theorem 5.6. To do so, let Y_1 be the first odd integer bigger than $11^{\frac{1}{1-\zeta_0}}$ and let $\mathcal{Z}_1 = \mathcal{Z}(d, \varrho, Q_{I_0}, \gamma_{I_0}, b, \zeta_0, \zeta_1, Y_1)$ be as in Theorem 5.6. Let $L_0^{(3)} = L_{k_2}^{(2)}$, where k_2 is the smallest integer k such that:

$$L_k^{(2)} > \mathcal{Z}_1\ , \quad \left(L_k^{(2)}\right)^{p'} > (3Y_1 - 4)^{2d}\ , \quad 2\left(L_k^{(2)}\right)^{\zeta_0 - 1} < m_1\ . \tag{6.6}$$

Then, recalling Remark 5.5, it follows from (6.3) that for all $E \in I(\delta_0)$ we have

$$\mathbb{P}\{\Lambda_{L_0^{(3)}}(0) \ \text{ is } (\zeta_0, E)\text{-sub-exponentially-suitable}\} > 1 - (3Y_1 - 4)^{-2d}\,, \tag{6.7}$$

and we have bootstrapped to hypothesis (5.7) of Theorem 5.6 for all $E \in I(\delta_0)$, uniformly in $E \in I(\delta_0)$. We now set $L_{k+1}^{(3)} = Y_1 L_k^{(3)}$, $k = 0, 1, 2, \ldots$, so it follows from Theorem 5.6 that for all $E \in I(\delta_0)$,

$$\mathbb{P}\{\Lambda_{L_k^{(3)}}(0) \ \text{ is } (\zeta_0, E)\text{-sub-exponentially-suitable}\} \geq 1 - \mathrm{e}^{-\left(L_k^{(3)}\right)^{\zeta_1}} \tag{6.8}$$

for all $k \geq \mathcal{K}^{(3)}$, where $\mathcal{K}^{(3)} = \mathcal{K}(\zeta_1, Y_1, L_0^{(3)}) < \infty$.

To complete our final bootstrap, we use Remark 5.5 to rewrite (6.8) as

$$\mathbb{P}\left\{\Lambda_{L_k^{(3)}}(0) \ \text{ is } \left(2\left(L_k^{(3)}\right)^{\zeta_0 - 1}, E\right)\text{-regular}\right\} \geq 1 - \mathrm{e}^{-\left(L_k^{(3)}\right)^{\zeta_1}} \tag{6.9}$$

for all $E \in I(\delta_0)$ and $k \geq \mathcal{K}^{(3)}$. Note that (6.9) is just hypothesis (5.9) of Theorem 5.7 at scale $L_k^{(3)}$ for each $E \in I(\delta_0)$. Thus we set $L_0^{(4)} = L_{k_3}^{(3)}$, where k_3 is the smallest integer $k \geq \mathcal{K}^{(3)}$ such that $L_k^{(3)} > \mathcal{C}$, where the constant $\mathcal{C} = \mathcal{C}(d, \varrho, Q_{I_0}, C_{I_0}, \gamma_{I_0}, b, \zeta_0, \zeta_1, \zeta_2, \alpha)$ is as in Theorem 5.7, with the parameters from (6.4). Note the crucial fact that $L_0^{(4)}$ is the same for all $E \in I(\delta_0)$. Theorem 5.7 provides us with $\delta_2 = \delta_2(d, \varrho, Q_{I_0}, \gamma_{I_0}, \zeta_0, \zeta_1, \zeta_2, \alpha, L_0^{(4)}) > 0$, so, setting $I(\delta_2, E) = [E - \delta_2, E + \delta_2] \cap I_0$, $m_\zeta = \frac{1}{2}(L_0^{(4)})^{\zeta_0 - 1}$, and $L_{k+1}^{(4)} = \left[\left(L_k^{(4)}\right)^{\alpha}\right]_{6\mathbb{N}}$, $k = 0, 1, \ldots$, we have

$$\mathbb{P}\left[R\left(m_\zeta, L_k^{(4)}, I(\delta_2, E), x, y\right)\right] \geq 1 - \mathrm{e}^{-\left(L_k^{(4)}\right)^{\zeta_2}} \tag{6.10}$$

for all $E \in I(\delta_0)$, $k = 0, 1, 2, \ldots$, and $x, y \in \mathbb{Z}^d$ with $|x - y| > L_k^{(4)} + \varrho$. Since $I(\delta_0)$ can be covered by intervals $I(\delta_2, E_i)$, $i = 1, 2, \ldots, [\frac{\delta_0}{\delta_2}] + 1$, with each $E_i \in I(\delta_0)$, we can conclude from (6.10) that

$$\begin{aligned} \mathbb{P}\left[R\left(m_\zeta, L_k^{(4)}, I(\delta_0), x, y\right)\right] \ &\geq \ 1 - \left([\tfrac{\delta_0}{\delta_2}] + 1\right) \mathrm{e}^{-\left(L_k^{(4)}\right)^{\zeta_2}} \\ &\geq \ 1 - \mathrm{e}^{-\left(L_k^{(4)}\right)^{\zeta}}, \end{aligned} \tag{6.11}$$

for all $x, y \in \mathbb{Z}^d$ with $|x - y| > L_k^{(4)} + \varrho$, and $k \geq k_4$, where k_4 is the smallest k such that the last inequality in (6.11) holds. Note $L_{k_4}^{(4)}$ depends only on $\delta_0, \delta_2, \alpha, L_0^{(4)}, \zeta, \zeta_2$, and hence only on $d, \varrho, Q_{I_0}, C_{I_0}, \gamma_{I_0}, \theta, \zeta, \alpha, \mathcal{L}$. To conclude the proof of Theorem 3.4, we set $L_0 = L_{k_4}^{(4)}$, so $L_k = L_{k_4 + k}^{(4)}$, $k = 0, 1, \ldots$, and note that (3.4) now follows from (6.11).

The proof of Theorem 3.4 is complete. □

Acknowledgement: F. G. thanks Peter Stollman for enjoyable discussions. F. G. also thanks the University of California, Irvine for its hospitality. A. K. thanks Alex Figotin and Svetlana Jitomirskaya for enjoyable discussions.

References

[1] Aizenman, M.: Localization at weak disorder: some elementary bounds. Rev. Math. Phys. **6**, 1163-1182 (1994)

[2] Aizenman, M., Molchanov, S.: Localization at large disorder and extreme energies: an elementary derivation. Commun. Math. Phys. **157**, 245-278 (1993)

[3] Aizenman, M., Schenker, J., Friedrich, R. , Hundertmark, D.: Finite-volume criteria for Anderson localization. Commun. Math. Phys., to appear

[4] Barbaroux, J.M., Combes, J.M., Hislop, P.D.: Localization near band edges for random Schrödinger operators. Helv. Phys. Acta **70**, 16-43 (1997)

[5] Barbaroux, J.M., Fischer, W., Müller, P.: Dynamical properties of random Schrödinger operators. Preprint (1999)

[6] Berezanskii, Ju.M.: *Expansions in eigenfunctions of selfadjoint operators.* Providence: Amer. Mat. Soc., 1968

[7] Carmona, R., Klein, A., Martinelli, F.: Anderson localization for Bernoulli and other singular potentials. Commun. Math. Phys. **108**, 41-66 (1987)

[8] Carmona, R, Lacroix, J.: *Spectral theory of random Schrödinger operators.* Boston: Birkhaüser, 1990

[9] Combes, J.M., Hislop, P.D.: Localization for some continuous, random Hamiltonian in d-dimension. J. Funct. Anal. **124**, 149-180 (1994)

[10] Combes, J.M., Hislop, P.D.: Landau Hamiltonians with random potentials: localization and the density of states. Commun. Math. Phys. **177**, 603-629 (1996)

[11] Combes, J.M., Hislop, P.D., Nakamura, S.: The L^p-theory of the spectral shift function, the Wegner estimate and the integrated density of states for some random operators. Commun. Math. Phys. **218**, 113-130 (2001)

[12] Combes, J.M., Hislop, P.D., Tip, A.: Band edge localization and the density of states for acoustic and electromagnetic waves in random media. Ann. Inst. H. Poincare Phys. Theor. **70** , 381-428 (1999)

[13] Cycon, H.L., Froese, R.G., Kirsch, W., Simon, B.: *Schrödinger operators.* Heidelberg: Springer-Verlag, 1987

[14] Damanik, D., Stollman, P.: Multi-scale analysis implies strong dynamical localization. Geom. Funct. Anal. **11**, 11-29 (2001)

[15] Damanik, D., Sims, R., Stolz, G.: Localization for one dimensional, continuum, Bernoulli-Anderson models. Preprint (2000)

[16] De Bièvre, S., Germinet, F.: Dynamical localization for random dimer Schrödinger operator. J. Stat. Phys. **98**, 1135-1147 (2000)

[17] Del Rio, R., Jitomirskaya, S., Last, Y., Simon, B.: What is Localization? Phys. Rev. Lett. 75, 117-119 (1995)

[18] Del Rio, R., Jitomirskaya, S., Last, Y., Simon, B.: Operators with singular continuous spectrum IV: Hausdorff dimensions, rank one pertubations and localization. J. d'Analyse Math. **69**, 153-200 (1996)

[19] von Dreifus, H.: *On the effects of randomness in ferromagnetic models and Schrödinger operators.* Ph.D. thesis, New York University (1987)

[20] von Dreifus, H., Klein, A.: A new proof of localization in the Anderson tight binding model. Commun. Math. Phys. **124**, 285-299 (1989)

[21] von Dreifus, H., Klein, A.: Localization for random Schrodinger operators with correlated potentials. Commun. Math. Phys. **140**, 133-147 (1991)

[22] Figotin, A., Klein, A.: Localization phenomenon in gaps of the spectrum of random lattice operators. J. Stat. Phys. **75**, 997-1021 (1994)

[23] Figotin, A., Klein, A.: Localization of electromagnetic and acoustic waves in random media. Lattice model. J. Stat. Phys. **76**, 985-1003 (1994)

[24] Figotin, A., Klein, A.: Localization of classical waves I: Acoustic waves. Commun. Math. Phys. **180**, 439-482 (1996)

[25] Figotin, A., Klein, A.: Localization of classical waves II: Electromagnetic waves. Commun. Math. Phys. **184**, 411-441 (1997)

[26] Fröhlich, J., Spencer, T.: Absence of diffusion with Anderson tight binding model for large disorder or low energy. Commun. Math. Phys. **88**, 151-184 (1983)

[27] 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)

[28] Germinet, F.: Dynamical localization II with an application to the almost Mathieu operator. J. Stat Phys. **95**, 273-286 (1999)

[29] Germinet, F, De Bièvre, S.: Dynamical localization for discrete and continuous random Schrödinger operators. Commun. Math. Phys. **194**, 323-341 (1998)

[30] Germinet, F, Jitomirskaya, S.: Strong dynamical localization for the almost Mathieu model. Rev. Math. Phys., to appear

[31] Germinet, F, Klein, A.: Finite volume criterium for localization in random media. In preparation

[32] Germinet, F, Klein, A.: Order parameter for the Anderson metal-insulator transport transition. In preparation

[33] Holden, H., Martinelli, F.: On absence of diffusion near the bottom of the spectrum for a random Schrödinger operator. Commun. Math. Phys. **93**, 197-217 (1984)

[34] Jitomirskaya, S., Last, Y.: Anderson localization for the almost Mathieu equation, III. Semi-uniform localization, continuity of gaps, and measure of the spectrum. Commun. Math. Phys. **195**, 1-14 (1998)

[35] Jona-Lasinio, G., Martinelli, F., Scoppola, E.: Multiple tunnelings in d-dimensions: a quantum particle in a hierarchical potential. Ann. Inst. Henri Poincaré **vol. 42**, 73-108 (1985)

[36] Lacroix, J., Klein., A., Speis, A.: Localization for the Anderson model on a strip with singular potentials. J. Funct. Anal. **94**, 135-155 (1990)

[37] Kirsch, W., Martinelli, F. : On the ergodic properties of the spectrum of general random operators. J. Reine Angew. Math. **334**, 141-156 (1982)

[38] Kirsch, W., Stolz, G., Stollman, P.: Localization for random perturbations of periodic Schrödinger operators. Random Oper. Stochastic Equations **6**, 241-268 (1998)

[39] Kirsch, W., Stolz, G., Stollman, P.: Anderson localization for random Schrödinger operators with long range interactions. Commun. Math. Phys. **195**, 495-507 (1998)

[40] Klein, A.: Localization in the Anderson model with long range hopping. Brazilian Journal of Physics **23**, 363-371 (1993)

[41] Klein, A.: Extended states in the Anderson model on the Bethe lattice. Advances in Math. **133**, 163-184 (1998)

[42] Klein, A., Koines, A.: A general framework for localization of classical waves: I. Inhomogeneous media and defect eigenmodes. Math. Phys. Anal. Geom. To appear

[43] Klein, A., Koines, A.: A general framework for localization of classical waves: II. Random media. In preparation

[44] Klein, A., Koines, A., Seifert, M.: Generalized eigenfunctions for waves in inhomogeneous media. J. Funct. Anal., to appear

[45] Klopp, F.: Localization for continuous random Schrödinger operators. Commun. Math. Phys. **167**, 553-569 (1995)

[46] Kunz, H., Souillard, B.: Sur le spectre des operateurs aux differences finies aleatoires. Commun. Math. Phys. **78**, 201-246 (1980)

[47] Martinelli, F., Scoppola, E.: Introduction to the mathematical theory of Anderson localization. Riv. Nuovo Cimento **10**, 1-90 (1987)

[48] Pastur, L.: Spectral properties of disordered systems in one-body approximation. Commun. Math. Phys. **75**, 179-196 (1980)

[49] Pastur, L., Figotin, A.: *Spectra of Random and Almost-Periodic Operators.* Heidelberg: Springer-Verlag, 1992

[50] Simon, B.: Schrödinger semi-groups. Bull. Amer. Math. Soc. **Vol.7**, 447-526 (1982)

[51] Spencer, T. : Localization for random and quasiperiodic potentials. J. Stat. Phys. **51**, 1009-1019 (1988)

[52] Stollmann, P.: *Caught by disorder.* To appear

[53] Tcheremchantsev, S.: How to prove dynamical localization. Commun. Math. Phys., to appear

[54] Wang, W.-M.: Microlocalization, percolation, and Anderson localization for the magnetic Schrödinger operator with a random potential. J. Funct. Anal. **146**, 1-26 (1997)

[55] Wang, W.-M.: Localization and universality of Poisson statistics for the multidimensional Anderson model at weak disorder. Invent. Math., to appear