# Moment analysis for localization in random Schrödinger operators

Invent. math. 163, 343–413 (2006)
DOI: 10.1007/s00222-005-0463-y

*Inventiones mathematicae*

# Moment analysis for localization in random Schrödinger operators

**Michael Aizenman**[1,⋆], **Alexander Elgart**[2], **Serguei Naboko**[3], **Jeffrey H. Schenker**[4], **Gunter Stolz**[5]

[1] Department of Mathematics, Princeton University, Princeton, NJ 08544-1000, USA (e-mail: aizenman@princeton.edu)
[2] Department of Mathematics, Stanford University, CA 94305-2125, USA (e-mail: elgart@math.stanford.edu)
[3] Department of Mathematics, St. Petersburg State University, 198904 St. Petersburg, Russia (e-mail: naboko@math.su.se)
[4] Theoretische Physik, ETH Zürich, 8093 Zürich, Switzerland (e-mail: jschenker@itp.phys.ethz.ch)
[5] Department of Mathematics, University of Alabama, Birmingham, AL 35294-1170, USA (e-mail: stolz@math.uab.edu)

Oblatum 22-VIII-2003 & 30-V-2005
Published online: 26 October 2005 – 

**Abstract.** We study localization effects of disorder on the spectral and dynamical properties of Schrödinger operators with random potentials. The new results include exponentially decaying bounds on the transition amplitude and related projection kernels, including in the mean. These are derived through the analysis of fractional moments of the resolvent, which are finite due to the resonance-diffusing effects of the disorder. The main difficulty which has up to now prevented an extension of this method to the continuum can be traced to the lack of a uniform bound on the Lifshitz-Krein spectral shift associated with the local potential terms. The difficulty is avoided here through the use of a weak-$L^1$ estimate concerning the boundary-value distribution of resolvents of maximally dissipative operators, combined with standard tools of relative compactness theory.

## Contents



---

⋆ also in the Physics Department, Princeton University

## 1. Introduction

**1.1. Random Schrödinger operators.** The addition of disorder can have a profound effect on the spectral and dynamical properties of a self adjoint differential operator. We consider here such phenomena for a class of operators in $L^2(\mathbb{R}^d)$ of the form

$$H_\omega \;:=\; H_0 \;+\; \lambda V_\omega \;, \tag{1.1}$$

with the disorder expressed through a random potential $V_\omega$. In the prototypical example $H_0$ is the Schrödinger operator

$$H_0 \;=\; -\Delta \;+\; V_0(q) \tag{1.2}$$

with $\Delta$ the Laplacian and $V_0(q)$ a bounded periodic background potential. The random term $V_\omega$ is given by a sum of local non-negative "bumps", $U_\alpha(q) = U(q-\alpha)$, centered at the lattice sites $\alpha \in \mathcal{I} = \mathbb{Z}^d$,

$$V_\omega(q) \;:=\; \sum_{\alpha\in\mathcal{I}} \eta_{\alpha;\omega}\, U_\alpha(q) \;, \tag{1.3}$$

with $\{\eta_\alpha\}_{\alpha\in\mathcal{I}}$ a collection of independent random variables uniformly distributed in $[0, 1]$. It will be assumed that the space $\mathbb{R}^d$ is covered by the supports of $\{U_\alpha(\cdot)\}$ so that $\inf_q \sum U_\alpha(q) \geq 1$, with the parameter $\lambda \geq 0$ controlling the strength of the disorder. The subscript $\omega$ indicates a point in a probability space $(\Omega, \mathrm{Prob}(d\omega))$ and often will be dropped when it is clear from context we are discussing a random variable.

More generally, the initial term $H_0$ may incorporate a magnetic field, i.e., take the form

$$H_0 \;:=\; \boldsymbol{D_A}\cdot\boldsymbol{D_A} \;+\; V_0(q) \tag{1.4}$$

where $\boldsymbol{D_A} = \mathrm{i}\nabla - \boldsymbol{A}(q)$ with $\boldsymbol{A}(q)$ the magnetic vector potential, and the periodicity of $V_0$ and of the bump potentials may be replaced by more relaxed assumptions. The required technical conditions, $\mathcal{A}$, are listed in Sect. 1.7.

Our objective is to present tools for the study of the phenomenon known as Anderson localization [8], which concerns the potentially drastic effect of the disorder on the dynamical and spectral properties of the perturbed operator. In general terms, the effect is that in certain energy ranges the absolutely continuous spectrum of the unperturbed operator may be modified to consist of a random dense set of eigenvalues associated with localized eigenfunctions, and scattering solutions of the time-dependent Schrödinger equation may become dynamically localized wave packets.

A convenient tool is provided by the Green function $G_E(x, y)$, which is the kernel of the resolvent operator $(H_\omega - E - \mathrm{i}0)^{-1}$. This kernel is well known to decay exponentially in $|x - y|$ when $E$ is in the resolvent set [20]. The hallmark of localization is rapid (even exponential) decay of $G_E(x, y)$ at energies in the spectrum, though in this case it occurs with pre-factors which are not uniform in space and diverge at a dense countable set of eigenvalues. Rapid decay of the Green function is related to the non-spreading of wave packets supported in the corresponding energy regimes and various other manifestations of localization whose physical implications have been extensively studied in regards to the conductive properties of metals [8,55,73,1,54] and in particular to the quantum Hall effect [36,57,10,12,3].

**1.2. Dynamical localization through Green function moment estimates.** In presenting our results let us start with a statement which shows that dynamical localization can be deduced from suitable bounds on the moments of the Green function. This relation shows that moment estimates form a natural and useful tool. For reasons which will be made apparent later, moments with power $s \geq 1$ diverge in regimes of localization, however, we shall see that this problem does not affect moments in the fractional range $s \in (0, 1)$, with which we shall work.

We denote here by $H^{(\Lambda)}$ the restrictions of $H$ to open sets $\Lambda \subset \mathbb{R}^d$. The default boundary conditions are Dirichlet, however much of what is said is rather insensitive to the boundary conditions and can easily be adapted to other choices, including Neumann, periodic, or quasi-periodic boundary conditions. The latter play a role in our discussion of the application of density of states bounds (Sect. 5.3).

Throughout we denote the characteristic function of a set $\Lambda$ by $\mathbf{1}_\Lambda$. It is convenient to set the distance unit to $r$ – the size of the "bumps" $U_\alpha$, as described in assumption $\mathcal{A}2$ below. Thus, for $x \in \mathbb{R}^d$ we let $\chi_x = \mathbf{1}_{B_x^r}$, where $B_x^r$ is the ball of radius $r$ centered at $x$.

Decay rates will be expressed below through a distance function $\mathrm{dist}(x, y) = |x - y|$ for which the choice of the norm on $\mathbb{R}^d$ does not affect our analysis. It is convenient to interpret it as $|x| = \sup_j |x_j|$, in which case "balls" $B_x^r$ are hypercubes. We shall also use the domain-adapted distance

$$\mathrm{dist}_\Lambda(x, y) \;=\; \min\left\{|x - y|, \mathrm{dist}(x, \Lambda^c) + \mathrm{dist}(y, \Lambda^c)\right\} \,, \tag{1.5}$$

for which the boundary of $\Lambda \subset \mathbb{R}^d$ is in effect regarded as a single point. As explained in [6] within the context of discrete operators, the use of the modified distance enables the analysis to cover also the cases where exponential localization in the bulk may possibly coincide with the occurrence of extended boundary states in certain subdomains.

**Theorem 1.1.** *Let $H$ be a random Schrödinger operator which satisfies the regularity assumptions $\mathcal{A}$ (formulated in Sect. 1.7). Let $\Omega$ be an open subset of $\mathbb{R}^d$, and $\Lambda_n$ an increasing sequence of bounded open subsets of $\Omega$ with $\cup\Lambda_n = \Omega$. Suppose that for some $0 < s < 1$ and an open bounded interval $\mathcal{J}$ there are constants $A < \infty$ and $\mu > 0$ such that*

$$\int_{\mathcal{J}} \mathbb{E}\left( \left\| \chi_x \frac{1}{H^{(\Lambda_n)} - E} \chi_y \right\|^s \right) \mathrm{d}E \;\le\; A e^{-\mu \operatorname{dist}_{\Lambda_n}(x,y)} \tag{1.6}$$

*for all $n \in \mathbb{N}$, $x, y \in \Lambda_n$. Then for every $v < 1/(2-s)$ there exists $A_v < \infty$ such that, for all $x, y \in \Omega$,*

$$\mathbb{E}\left( \sup_{g:\,|g|\le 1} \left\| \chi_x g(H^{(\Omega)}) P_{\mathcal{J}}(H^{(\Omega)}) \chi_y \right\| \right) \le A_v\, e^{-v\mu \operatorname{dist}_\Omega(x,y)} , \tag{1.7}$$

*where the supremum is taken over all Borel measurable functions $g$ which satisfy $|g| \le 1$ pointwise and $P_{\mathcal{J}}(H^{(\Omega)})$ is the spectral projection for $H^{(\Omega)}$ associated to the interval $\mathcal{J}$.*

The constant $A_v$ also depends on $s$, $\lambda$, and $E_+ = \sup \mathcal{J}$, as can be seen in the proof, which is in Sect. 2. In particular, the dependence on $E_+ = \sup \mathcal{J}$ is polynomial with degree slightly larger than $d/2$. One can also see from the proof that the above result holds for $s \ge 1$ as well, in which case one can choose $v = 1/s$. However, as explained in Sect. 1.3 below, for $s \ge 1$ the assumption (1.6) will not be satisfied within the pure point spectrum.

Of special interest are the following three implications of (1.7).

*(1) Dynamical localization:* With $g(H) = \mathrm{e}^{-\mathrm{i}tH}$, eq. (1.7) yields for the unitary evolution operator:

$$\mathbb{E}\left( \sup_t \left\| \chi_x e^{-\mathrm{i}tH^{(\Omega)}} P_{\mathcal{J}}(H^{(\Omega)}) \chi_y \right\| \right) \le A_v\, e^{-v\mu \operatorname{dist}_\Omega(x,y)} , \tag{1.8}$$

which is a strong form of dynamical localization. The result established here through this criterion is new for continuum models and has not been obtained with other methods. (The relation with previous results is discussed further in Sect. 1.6.)

*(2) Spectral localization:* For $\Omega = \mathbb{R}^d$ the bound (1.7) permits one to further conclude (using the RAGE theorem as in [35]) that the spectrum of $H$ in $\mathcal{J}$ is almost surely pure point with exponentially decaying eigen-projections, i.e. for every $\nu < \mu/(2-s)$ and $E \in \mathcal{J}$,

$$\left\| \chi_x \delta_E(H) \chi_y \right\| \;=\; \mathrm{O}(e^{-\nu|x-y|}) , \tag{1.9}$$

where $\delta_E(x) = 1$ if $E = x$ and 0 otherwise. An argument provided in [18] shows that almost surely all eigenvalues of $H$ in $\mathcal{J}$ are finitely degenerate. This allows to deduce exponential decay of eigenfunctions from (1.9). The proofs of these results are included at the end of Sect. 2.5. Such spectral localization can also be directly deduced from (1.6) using the Simon-Wolff criterion [69] as adapted to continuum operators in [18].

*(3) Decay of the Fermi-projection kernel:* Another example which plays an important role in physics applications of the model involves the Fermi projection $P_{(-\infty,E_F)}(H^{(\Omega)})$ for $E_F \in \mathcal{J}$. Although not necessarily of the form $g(H^{(\Omega)})P_{\mathcal{J}}(H^{(\Omega)})$ since the projection range may be larger than the interval $\mathcal{J}$, these operators nonetheless satisfy

$$\mathbb{E}\left(\sup_{E_F\in\mathcal{J}} \left\| \chi_x P_{(-\infty,E_F)}(H^{(\Omega)})\chi_y \right\| \right) \le \widetilde{A} e^{-\widetilde{\mu}\,\mathrm{dist}_\Omega(x,y)} , \tag{1.10}$$

with constants $\widetilde{A} < \infty$ and $\widetilde{\mu} > 0$ whenever eq. (1.6) holds. This may be proved by combining eqs. (1.6, 1.7) and the Helffer-Sjöstrand formula [38] which is presented in Appendix A (Remark 11), or using the argument of [3] – where the issue is discussed in the context of lattice operators.

Theorem 1.1 is proven below in Sect. 2. Beyond that, the bulk of our article deals with the derivation of finite volume criteria which permit to establish the condition (1.6) for localization.

**1.3. The reason for fractional moments.** The criterion (1.6) will be of use to us only with the fractional exponents $s < 1$. For $s \ge 1$ the integral over $E$ on the left-hand side of (1.6) will diverge even before the average over the disorder. It is relevant here to note that in the presence of point spectrum the Lebesgue measure of the set of energies at which $\|\chi_x \frac{1}{H^{(\Omega)}-E}\chi_x\|$ is larger than t exhibits 1/t tails.

For instance, if $\Omega$ is an open subset of $\mathbb{R}^d$ and $H^{(\Omega)}$ has only pure-point spectrum in $\mathcal{J}$, then for any $\phi \in L^2(\Omega)$:

$$\left|\left\{E \in \mathbb{R} : \left|\left(\phi, \chi_x \frac{1}{H^{(\Omega)} - E} P_{\mathcal{J}}(H^{(\Omega)})\chi_x\phi\right)\right| \ge t\right\}\right| = \frac{\text{Const.}}{t} \tag{1.11}$$

with Const. $= 2(\phi, \chi_x\, P_{\mathcal{J}}(H^{(\Omega)})\, \chi_x\, \phi)$. (To see that, one may use the spectral representation and the Theorem of Boole [16,2].) Since:

$$\int_{\mathbb{R}} |Y(E)|^s \, dE = \int_0^\infty |\{\, E : |Y(E)| \ge t\}| \; d(t^s) \tag{1.12}$$

it follows that the $s$-moments of the Green function seen in (1.11) diverge for all $s \ge 1$, whenever $(\phi, \chi_x\, P_{\mathcal{J}}(H^{(\Omega)})\, \chi_x\, \phi) \neq 0$, but are finite for $0 < s < 1$.

Thus, an important step for our analysis is to show that the left-hand side of (1.6) is finite. A highly instructive statement is the following estimate, which is formulated in a simplified setting.

**Proposition 1.1.** *Let $H = -\Delta + V$, with a bounded potential $V$. Then, for any $0 < s < 1$, and $a, b \in \mathbb{R}$, $x, y \in \mathbb{R}^d$:*

$$\int_a^b \left\| \chi_x \frac{1}{H - E - \mathrm{i}0} \chi_y \right\|^s dE \;\le\; C(a, b, d, \|V\|_\infty, s) \;<\; \infty\,, \qquad (1.13)$$

*where the upper bound holds uniformly in x, y and depends only on the explicitly listed quantities.*

We will not use Proposition 1.1 in the given form, but rather a related result discussed in the next subsection below. Still, the following sketch of proof of (1.13) may serve to introduce the main ideas behind establishing finiteness of fractional moments:

As is discussed in Sect. 3.1, weak $L^1$ bounds as in (1.11) do not have a direct and useful extension to quantities such as the operator norms considered in (1.13). However the following result, which is valid for any maximally dissipative operator $A$ and Hilbert-Schmidt operators $T_1$, $T_2$, will serve as a key element:

$$\left|\left\{E : \left\|T_1(A - E - \mathrm{i}0)^{-1} T_2\right\|_{HS} > t\right\}\right| \;\le\; \frac{C}{t}\, \|T_1\|_{HS}\, \|T_2\|_{HS}\,, \qquad (1.14)$$

see Sect. 3.2 and Appendix C. Such a bound implies finiteness of the $s < 1$ moments by means of the "layer-cake" representation (1.12) of the integral. However, first some further work needs to be done since $\chi_x$ and $\chi_y$ are not Hilbert-Schmidt operators. To this end, write

$$\begin{aligned}&\chi_x(H - E - \mathrm{i}0)^{-1}\chi_y && (1.15)\\ &\quad = \chi_x(-\Delta + 1)^{-1}\chi_y + \chi_x(-\Delta + 1)^{-1}(E - V + 1)(H - E - i0)^{-1}\chi_y\,.\end{aligned}$$

The first term on the right is trivial for the proof of (1.13). The factor $\chi_x(-\Delta + 1)^{-1}$ in the second term is Hilbert-Schmidt if $d \le 3$. As $E - V + 1$ is bounded, the weak bound (1.14) becomes applicable after a similar argument is used on the right of the resolvent. If $d > 3$ one can iterate this construction until eventually $(\chi_x(-\Delta + 1)^{-1})^n$ is Hilbert-Schmidt. This also implicitly justifies the existence of the boundary value in (1.13), which is known to exist for Lebesgue-a.e. $E$ in the corresponding expression in (1.14).

**1.4. Finite-volume criteria.** Our next result deals with finite volume sufficiency criteria for the localization bounds (1.6). The basic idea is that if in some ball $B$ the fractional moments from the center to $\partial B$ are "small enough" then the input criteria of Theorem 1.1 are satisfied.

This will require an initial step in which the techniques mentioned above in the context of integrals over $E$ are applied also to the averages over disorder parameters at fixed energy. By such means we show that the independent variation of a local parameter can resolve a singularity in $\|\chi_x(H+\eta_\alpha U_\alpha - E)^{-1}\chi_y\|$ which may be present due to the proximity of the given energy $E$ to an eigenvalue whose eigenfunction has significant support near $x$ and/or $y$. Instrumental for the analysis is the Birman-Schwinger relation:

$$U_\alpha^{1/2} \frac{1}{H + \eta_\alpha U_\alpha - z} U_\alpha^{1/2} = \left[ \left[ U_\alpha^{1/2} \frac{1}{H - z} U_\alpha^{1/2} \right]^{-1} + \eta_\alpha \right]^{-1} \tag{1.16}$$

where $[\dots]^{-1}$ is to be interpreted as operator inverse in $L^2(\operatorname{supp} U_\alpha)$.

The Birman-Schwinger relation makes (1.14) applicable for averages over individual disorder parameters, which take the role of a "local energy parameter". This strategy motivates two of our technical assumptions, the covering condition (1.22) on the single site potentials $U_\alpha$, as well as the required absolute continuity of the distribution of the random parameters $\eta_\alpha$. Detailed statements and proofs of these results are given in Sects. 3.3–3.4. In the following, these preliminary bounds serve as worst-case estimates, somewhat reminiscent of the role of Wegner estimates in multi-scale analysis.

Let $r_0$ be the independence length introduced below next to the assumption $\mathcal{IAD}$. Also define the boundary layer of a set $\Lambda$ to be the (open) set

$$\delta\Lambda := \{q \ : \ r < \operatorname{dist}(q, \Lambda^c) < 23r\}, \tag{1.17}$$

where the choice of the depth is somewhat arbitrary, but convenient for our argument.

**Theorem 1.2.** *Let $H$ be a random Schrödinger operator which satisfies the assumptions $\mathcal{A}$ as well as $\mathcal{IAD}$. For each $s \in (0, 1/3)$, $\lambda > 0$ and $E \in \mathbb{R}$, there exists $M(s, \lambda, E) < \infty$, such that if for some $L > r_0 + 23r$,*

$$\begin{aligned} e^{-\gamma} := {} & M(s, \lambda, E)(1 + L)^{2(d-1)} \\ & \times \sup_{\alpha \in \mathcal{I}} \limsup_{\varepsilon \downarrow 0} \mathbb{E}\left( \left\| \chi_\alpha \frac{1}{H^{(B_\alpha^L)} - E - i\varepsilon} \mathbf{1}_{\delta B_\alpha^L} \right\|^s \right) \\ < {} & 1 \, , \end{aligned} \tag{1.18}$$

*then there exists $A(s, \lambda, E)$ such that for any open $\Omega \subset \mathbb{R}^d$ and any $x, y \in \Omega$*

$$\limsup_{\varepsilon \downarrow 0} \mathbb{E}\left( \left\| \chi_x \frac{1}{H^{(\Omega)} - E - i\varepsilon} \chi_y \right\|^s \right) \le e^{\gamma} A(s, \lambda, E) \, e^{-\gamma \operatorname{dist}_\Omega(x,y)/2L} \, . \tag{1.19}$$

*Here, the constants $M(s, \lambda, E)$, $A(s, \lambda, E)$ can be chosen polynomially bounded: in $E$, in $1/\lambda$ as $\lambda \to 0$, and in $\lambda$ as $\lambda \to \infty$; if $H$ satisfies also $\mathcal{A}3'$ then they can be chosen uniformly bounded in $\lambda > 1$. Furthermore, for any bounded region $\Omega$ eq. (1.19) holds also with $\varepsilon = 0$.*

The above result serves as a finite-volume criterion for localization. Eq. (1.19) reflects the fact that the scale $L$ at which (1.18) is verified indeed determines the localization length. The locally uniform $E$-bound of the constant in (1.19) allows one to deduce integral bounds as required in (1.6) of Theorem 1.1. Theorem 1.2 is proven in Sect. 4 with methods similar to those developed in [6] in the proof of a corresponding result for lattice operators. This uses what is frequently called the *geometric resolvent identity* (Lemma 4.2) as well as decoupling and re-sampling arguments to factorize expectations.

**1.5. Applications.** In Sect. 5 we show how the general framework provided here can be used to prove localization in specific disorder regimes. This includes the familiar large disorder (Theorem 5.2) and band edge or Lifshitz tail regimes. For the latter we provide two results, one based on smallness of the finite volume density of states (Theorem 5.3) and another – less traditional – result using smallness of the infinite volume density of states (Theorem 5.4). We also show in Theorem 5.1 that the "output" of a multi-scale analysis can be used to provide the "input" for Theorem 1.2, thus proving that the stronger results found by our methods hold throughout the multi-scale analysis regime. A useful technical result (Lemma 5.1) is a continuity property of fractional resolvent moments. It shows that in applications it suffices to check the bound (1.18) at a single energy.

This observation also leads to a proof of the following complementary criterion which rounds off our discussion.

**Theorem 1.3.** *Let $H$ be a random Schrödinger operator which satisfies $\mathcal{A}$ and $\mathcal{IAD}$. Suppose that for some $A < \infty$, $\mu > 0$ and $E \in \mathbb{R}$*

$$\limsup_{\varepsilon \downarrow 0} \mathbb{E}\left( \left\| \chi_\alpha \frac{1}{H - E - \mathrm{i}\varepsilon} \chi_\beta \right\|^s \right) \le A e^{-\mu|\alpha-\beta|} \tag{1.20}$$

*for all $\alpha, \beta \in \mathcal{I}$. Then, for sufficiently large $L$, eq. (1.18) is satisfied uniformly for all $E'$ in an open neighborhood of $E$.*

Combined with Theorems 1.1 and 1.2 this shows that exponential decay of the Green function as in eq. (1.20) provides a necessary and sufficient condition for eq. (1.19), and thus eq. (1.7), to hold in a neighborhood of an energy $E$. Thus, in principle the entire regime of localization in the sense of eq. (1.19) may be mapped out using the criterion provided by Theorem 1.2.

The applications of Sect. 5 are important to tie our general method with concrete examples. But we stress that these examples are somewhat secondary to the main goal of this work, which is the outline of a general framework for studying localization properties consisting of the three interrelated Theorems 1.1, 1.2, and 1.3 and based on the preliminary bounds obtained in Sect. 3.

As such, we do not attempt to give an exhaustive list of applications here, but rather try to illustrate how known methods may be combined with the arguments developed in this paper. Further developments based on the new techniques will be left to future work. For example, the work [17] will use a variant of the techniques used here to study continuum random surface models.

**1.6. Relation with past works.** Mathematical analysis of localization for random operators has been a very active field. The continuum operators have analogs in the discrete setting, obtained by replacing (1.1) with analogous operators on the $\ell^2$ space of a graph such as $\mathbb{Z}^d$; i.e., replacing the differential operator $\boldsymbol{D_A} \cdot \boldsymbol{D_A}$ by a "hopping matrix" and the potential $V_\omega(q)$ by a multiplication operator with $\{V_\omega(x)\}_{x\in\mathbb{Z}^d}$ iid random variables. Localization phenomena are rather similar in the two setups, in broad terms, but the discrete case is simplified by the fact that the random coefficients affect terms of finite rank, often just rank-1 or rank-2. Naturally, various analytical arguments were initially developed in that setup, although this was not true for the analysis in one dimension which provided the first rigorous results, initiated in [34], and where special tools are available.

The existing results on localization in the continuum of dimension larger than one have been based on the multiscale analysis, first obtained for discrete operators [30, 25, 29, 69, 74] and then extended to the continuum [50, 18, 32]. We do not attempt to give an exhaustive survey of the vast related literature. The reader is referred to the recent book of P. Stollmann [71] for a review of the history of the subject and a gentle introduction to the multi-scale analysis – which is not used here. Our work presents a continuum version of the fractional-moment method which was developed for discrete systems in [4, 2, 3, 6]. The fractional-moment method was applied already to certain continuum models in which some crucial features from the discrete case persist, in particular the single site perturbations are rank one [39, 27].

It seems appropriate to make a brief comparison of the two approaches which have been developed to handle multidimensional localization, the multi-scale analysis (MSA) and the fractional moment method (FMM). Both have now been found to apply to discrete as well as continuum models. They lead to similar results: spectral and dynamical localization, though expressed through somewhat different estimates, and apply to essentially the same disorder and energy regimes.

Two significant differences lie in: *i)* the iterative schemes which are used in the two methods for the derivation of results for the infinite volume from finite volume characteristics of localization, and *ii)* the tools used to express localization. MSA uses a KAM-type strategy through an infinite collection of scales, whereas FMM is a *single scale* method – that of the localization length. In MSA, the quantity to be controlled is the probability of rare events, whereby the random configuration locally manifests traits which may inhibit localization. In FMM the localization is expressed through rapid decay of the Green function's suitable (fractional) moments.

While MSA is a multiscale method, in FMM once a scale is reached at which the finite volume localization criterion is met, all spatial correlators are shown to decay exponentially on that length scale, including in the mean, i.e., error estimates are also exponentially small. The technical parts of the proof become more involved in the continuum, as is also true for other methods, but the basic mechanism of working with just one scale remains. A particular consequence of this is that FMM yields exponential decay of the Green function fractional moments in (1.19) and, subsequently, also in the dynamical localization bound (1.7). This is a bit better than the best result of this kind obtained through MSA, which is a bound of the form $C_\zeta \exp(-|x-y|^\zeta)$ for any $\zeta < 1$, see [33]. The limiting factor there is the estimate for the probability of *a-typical* configurations, for which the multiscale scheme yields a fast, yet suboptimal decay rate.

Possible directions for the extension of the analysis presented here, which does not cover all the results which were derived using MSA, include: removal of the condition that the random potential bumps fully cover the space, which is required in (1.22) (this will be addressed in [17]), and the relaxation of the regularity of the distribution of the random parameters $\eta_\alpha$. The case which may be well beyond the reach of the averaging methods used here, even under some natural improvements, is that of $\eta_\alpha$ having the discrete "Bernoulli" distribution.

**1.7. Assumptions.** Following are the regularity conditions required for our results. The condition $\mathcal{A}$ refers to the collection $\mathcal{A}1$–$\mathcal{A}3$.

$\mathcal{A}1$ The components of the vector potential $\boldsymbol{A}$, its first derivatives $\partial_i \boldsymbol{A}$ for $i = 1, ..., d$, and $V_{o,+}$ (the positive part of $V_0$) are locally bounded on $\mathbb{R}^d$. The negative part $V_{o,-}$ of $V_0$ is bounded.

$\mathcal{A}2$ Each function $U_\alpha$ is bounded, non-negative, supported in a ball of radius $r$ around $\alpha$ for fixed $r > 0$ and some $\alpha \in \mathcal{I}$, with $\mathcal{I}$ a discrete set of points in $\mathbb{R}^d$. The number of points $\alpha$ falling within any unit cube is uniformly bounded by some $N < \infty$, and the function

$$F(q) \ := \sum_{\alpha \in \mathcal{I}} U_\alpha(q) \tag{1.21}$$

satisfies uniform bounds

$$1 \ \le \ \inf_q F(q) \ \le \ \sup_q F(q) \ =: \ b_+ \ < \ \infty \, . \tag{1.22}$$

Moreover, $|\partial(\operatorname{supp} U_\alpha)| = 0$.

$\mathcal{A}3$ The random variables $\eta_\alpha$, $\alpha \in \mathcal{I}$ take values in $[0, 1]$ and the conditional distribution of $\eta_\alpha$ at specified values of $\{\eta_\zeta\}_{\zeta \neq \alpha}$ has a density, denoted $\rho_\alpha(\eta \mid \omega)$, which is uniformly bounded:

$$D \ := \ \sup_\alpha \|\rho_\alpha(\cdot|\cdot)\|_\infty \ < \ \infty \tag{1.23}$$

where $\|\cdot\|_\infty$ indicates the essential supremum over $\eta$ and $\omega$.

Assumption $\mathcal{A}$ is sufficiently general to cover many important examples, such as the two-dimensional Landau hamiltonian with random potential, where $\boldsymbol{A}(q_1, q_2) = \frac{1}{2}(-Bq_2, Bq_1)$. For both $\boldsymbol{A}$ and $V_0$ one could allow suitable, dimension-dependent, $L^p$-type singularities, but we prefer to avoid the additional technicalities which are caused by this.

A key requirement for Theorem 1.2 is "independence at a distance," namely the random functions obtained by restricting $V_\omega$ to well separated regions are pairwise independent.

$\mathcal{IAD}$: There exists $r_0 > 0$ such that if $\Lambda, \Lambda' \subset \mathbb{R}^d$ with $\mathrm{dist}(\Lambda, \Lambda') > r_0$ then the collections of random variables $\eta_\Lambda := \{\eta_\zeta : \zeta \in \Lambda \cap \mathcal{I}\}$ and $\eta_{\Lambda'} := \{\eta_\zeta : \zeta \in \Lambda' \cap \mathcal{I}\}$ are independent, i.e.,

$$\mathbb{E}\left(f(\eta_\Lambda)g(\eta_{\Lambda'})\right) = \mathbb{E}\left(f(\eta_\Lambda)\right)\mathbb{E}\left(g(\eta_{\Lambda'})\right) , \tag{1.24}$$

for arbitrary bounded measurable functions $f, g$ on $\mathbb{R}^{\Lambda\cap\mathcal{I}}$ and $\mathbb{R}^{\Lambda'\cap\mathcal{I}}$, respectively. Without loss of generality we assume that $r_0 \geq 2r$.

We note that this assumption is not required for the proof of the boundedness of the fractional moments in Sect. 3 and also not for the derivation of localization from the Green function decay (Theorem 1.1).

The restriction of $\eta_\alpha$ to range over [0, 1] is not essential; through the adjustment of the background potential and the disorder parameter that range can be replaced by any other bounded interval. However, the normalization becomes relevant when one considers the strong disorder regime.

Some of the bounds derived below – c.f., Lemmas 3.3 and 3.4 – exhibit coefficients which grow with increasing $\lambda$. For applications of these bounds to the "large disorder regime" ($\lambda \gg 1$) it is useful to break the coupling $\lambda\eta_\alpha$ into a sum of variables $\eta_{1;\lambda} + \eta_{2;\lambda}$ such that $\eta_{1;\lambda}$ is of order one and obeys $\mathcal{A}3$. This could be accomplished in a number of ways – e.g., let $\eta_{1;\lambda}$ be the fractional part of $\lambda\eta_\alpha$. However, without an additional assumption these decompositions might become more and more singular as $\lambda$ increases. Following is a useful notion.

**Definition 1.1.** A real valued random variable $X$, with an absolutely continuous probability measure of density $\rho(\cdot)$ on $\mathbb{R}$, is *blow-up regular* if there exist two sequences of real valued random variables, $\{X^{(n)} : n \geq 1\}$ and $\{Y^{(n)} : n \geq 1\}$ such that

(1) $X^{(n)}$ takes values in [0, 1].
(2) For each $n \geq 1$,

$$nX = Y^{(n)} + X^{(n)} . \tag{1.25}$$

(3) The conditional distribution of $X^{(n)}$, at a specified value $Y^{(n)} = y$, has a bounded density, $\rho_n(\cdot|y)$, and

$$D \equiv \sup_{n\geq 1;\, x,y\in\mathbb{R}} |\rho_n(x|y)| < \infty . \tag{1.26}$$

The *blow-up norm* of the probability density $\rho$, denoted $D_\rho$, is the infimum of the above quantity $D$ taken over all sequences $\{X^{(n)}, Y^{(n)} : n \geq 1\}$ satisfying (1) and (2).

The following assumption, in lieu of $\mathcal{A}3$, allows better bounds for the strong disorder regime.

$\mathcal{A}3'$ The random variables $\eta_\alpha$, $\alpha \in \mathcal{I}$ take values in $[0, 1]$, and for each $\alpha$ the conditional distribution of $\eta_\alpha$ at specified values of $\{\eta_\zeta\}_{\zeta \neq \alpha}$ is blow-up regular, with the blow-up norms bounded by a common $D < \infty$.

The above condition is satisfied for independent uniform in $[0, 1]$ random variables, and also for i.i.d. variables with a common density $\rho$ provided $\ln \rho$ is Lipshitz-continuous in $[0, 1]$ (see Appendix A).

**1.8. Outline of contents.** The contents of Sects. 2, 3, 4 and 5 were discussed in Sects. 1.2, 1.3, 1.4 and 1.5, in that order. Thus we focus here on briefly describing the contents of the four Appendices of this paper:

The assumptions and results outlined above deserve a number of more technical comments. In order to keep the introduction relatively free of technicalities, further discussion of these is postponed to Appendix A.

Appendix B contains a short description of the Birman-Schwinger relation, which is used throughout this work.

The weak $L^1$ bound (1.14), central to the extension of the fractional moment method to the continuum, has not been used previously in the literature on random operators. We thus present a self-contained proof based on properties of the vector-valued Hilbert transform in Appendix C. Much of the main argument is taken from [56].

In Appendix D we show how the methods of Sect. 3 can be used to derive a bound for the disorder averaged spectral shift which is locally uniform in energy. As discussed in Sect. 3.1, such bounds do not hold without averaging over the disorder. We do not use this result in our main argument, however it may be of independent interest.

## 2. From resolvent bounds to eigenfunction correlators and dynamical localization

In this section we prove Theorem 1.1. For discrete models such results were derived in refs. [2,6] employing the observation that the eigenfunctions of the operator $H$ play the role of Green functions for a re-sampled operator $\hat{H}$, at other values of the coupling variables. Key in that analysis were properties of rank one perturbations. Use was also made of a very convenient interpolation argument which permits to extract bounds on the off-diagonal matrix elements of the spectral projections of $H$ from fractional moment bounds on the Green function of $\hat{H}$.

We find that the approach of refs. [2,6] can also be applied to continuum operators, with the arguments which were based on rank one perturbation replaced, or generalized, by considerations of the Birman-Schwinger operator.

**2.1. Correlators – eigenfunction and other.** In the discussion of dynamical localization for discrete random operators in ref. [6], estimates were developed for the spectral measures $\mu_{x,y}$ which are defined through the Riesz theorem by

$$\int f(E)\mu_{x,y}(\mathrm{d}E) \;=\; \langle x|f(H)|y\rangle\,, \qquad f \in C_0(\mathbb{R})\,. \tag{2.1}$$

Exponential decay (in $|x-y|$) of the *total variation* of these measures provided a strong description of dynamical localization. For analogous bounds in the present context, it is convenient to work with the "operator valued measures,"

$$f \mapsto \chi_x f(H^{(\Omega)})\chi_y\,, \qquad f \in C_c(\mathbb{R})\,. \tag{2.2}$$

We introduce also the "total variation" of these measures,

$$Y_\Omega(\mathcal{J}; x, y) \;:=\; \sup_{\substack{f\in C_c(\mathcal{J})\\ \|f\|_\infty\le 1}} \left\|\chi_x f(H^{(\Omega)})\chi_y\right\|\,, \tag{2.3}$$

defined for bounded open intervals $\mathcal{J}$ where $C_c(\mathcal{J})$ denotes the continuous functions compactly supported inside $\mathcal{J}$.

For a finite region $\Lambda$ and fixed $\alpha \in \mathcal{I}$, we use the Birman-Schwinger relation of Appendix B to study the dependence of $H^{(\Lambda)}$ on the single random parameter $\eta_\alpha$, keeping $\{\eta_\beta\}_{\beta\neq\alpha}$ fixed. We express $H_{\eta_\alpha} \equiv H^{(\Lambda)}$ in terms of a re-sampled reference operator $H_{\hat\eta_\alpha}$ as

$$H_{\eta_\alpha} = H_{\hat\eta_\alpha} - \lambda(\hat\eta_\alpha - \eta_\alpha)U_\alpha, \tag{2.4}$$

where we will take the re-sampled variable $\hat\eta_\alpha$ to have the same conditional distribution as $\eta_\alpha$. The family $H_{\eta_\alpha}$ has the form of the one-parameter family eq. (B.5) of Appendix B, with $H_0 = H_{\hat\eta_\alpha}$, $V = U_\alpha$ and $\xi = \lambda(\hat\eta_\alpha - \eta_\alpha)$.

For $0 \le v \le 2$ we define the fractional "eigenfunction correlators" as

$$Q_v(\mathcal{J}; x, \alpha) \;=\; \sum_{n:\,E_n\in\mathcal{J}} \langle \chi_x\psi_n, \psi_n\rangle^{v/2}\, \langle U_\alpha\psi_n, \psi_n\rangle^{1-v/2}\,, \tag{2.5}$$

where $n$ labels the eigenvalues $E_n = E_n(\xi)$ and the corresponding orthonormal eigenfunctions $\psi_n = \psi_n(\xi)$ of $H_{\eta_\alpha}$, choosing the labeling so that these are holomorphic in $\xi$, as in Appendix B.

For $v = 1$ the eigenfunction correlators provide bounds for $Y_\Lambda(\mathcal{J}; x, y)$: If $f \in C_c(\mathcal{J})$, then $f(H^{(\Lambda)}) = \sum_{n:E_n \in \mathcal{J}} f(E_n) P_{\psi_n}$, where $P_{\psi_n}$ is the orthogonal projector onto $\psi_n$. Thus, by the "covering condition" (1.22),

$$\begin{aligned} Y_\Lambda(\mathcal{J}; x, y) &\le \sum_{\substack{\alpha \in \mathcal{I} \\ |y-\alpha| \le 2r}} \sup_{\substack{f \in C_c(\mathcal{J}) \\ |f| \le 1}} \left\| \chi_x f(H^{(\Lambda)}) U_\alpha^{1/2} \right\| \\ &\le \sum_{\substack{\alpha \in \mathcal{I} \\ |y-\alpha| \le 2r}} \sum_{n:E_n \in \mathcal{J}} \left\| \chi_x P_{\psi_n} U_\alpha^{1/2} \right\| \\ &= \sum_{\substack{\alpha \in \mathcal{I} \\ |y-\alpha| \le 2r}} Q_1(\mathcal{J}; x, \alpha) \ . \end{aligned} \tag{2.6}$$

Here it was used that $\|\chi_x P_{\psi_n} U_\alpha^{1/2}\| = \|\chi_x \psi_n\| \, \|U_\alpha^{1/2} \psi_n\|$.

For all $0 < v < 2$ the quantity $Q_v(\mathcal{J}; x, \alpha)$ reflects the overlap between the eigenfunctions at $x$ and $\alpha$. That, however, is not the case at the endpoints $v = 0, 2$ for which the corresponding values of $Q_v$ depend only on the density of states:

$$\begin{aligned} Q_0(\mathcal{J}; x, \alpha) &= \operatorname{Tr} U_\alpha \, P_{\mathcal{J}}(H^{(\Lambda)}) \\ Q_2(\mathcal{J}; x, \alpha) &= \operatorname{Tr} \chi_x \, P_{\mathcal{J}}(H^{(\Lambda)}) \end{aligned} \tag{2.7}$$

with $P_{\mathcal{J}}(H^{(\Lambda)})$ the spectral projection operator. For the Schrödinger operators considered here, both $Q_0(\mathcal{J}; x, \alpha)$ and $Q_2(\mathcal{J}; x, \alpha)$ are of order one, for a finite interval $\mathcal{J}$, in the sense that they are finite and do not decay for increasing $\operatorname{dist}(x, \alpha)$. In particular, if $\mathcal{J} \subset (-\infty, E)$ and $p > d/2$ we have

$$\begin{aligned} Q_2(\mathcal{J}; x, \alpha) &\le \operatorname{Tr} \chi_x P_{\le E}(H^{(\Lambda)}) \\ &\le (|E - E_0| + 1)^p \operatorname{Tr} \chi_x (H^{(\Lambda)} + E_0 + 1)^{-p} \\ &\le C(|E - E_0| + 1)^p \ , \end{aligned} \tag{2.8}$$

where here and in the following we set $E_0 = \inf \sigma(H_0)$. This bound is deterministic, and thus holds also for $\mathbb{E}(Q_2(\mathcal{J}; x, \alpha))$.

**Lemma 2.1.** *$Q_v(\mathcal{J}; x, \alpha)$ is log convex in $v$, and for any $v \in (0, 1)$:*

$$\mathbb{E}\left(Q_1(\mathcal{J})\right) \ \le \ \mathbb{E}\left(Q_v(\mathcal{J})\right)^{1/(2-v)} \ \mathbb{E}\left(Q_2(\mathcal{J})\right)^{(1-v)/(2-v)} \ . \tag{2.9}$$

*at any value of the (omitted) argument $(x, \alpha)$ of $Q_v$.*

*Proof.* The log convexity of $Q_v$ in $v$ is a standard observation for a function of the form $F(v) = \sum_n A_n B_n^v$. In particular, for $v < 1 < 2$, writing 1 as a convex combination of the other values: $1 = av + (1 - a)2$, one gets via the Hölder inequality: $F(1) \le F(v)^a F(2)^{1-a}$ (with $a = 1/(2 - v)$). In the present context that yields

$$Q_1(\mathcal{J}) \ \le \ Q_v(\mathcal{J})^{1/(2-v)} \ Q_2(\mathcal{J})^{(1-v)/(2-v)} \ . \tag{2.10}$$

One more application of the Hölder inequality, this time to the average over the randomness ($\mathbb{E}(\cdot)$), yields eq. (2.9). □

**2.2. Eigenfunction correlators and resolvent moments.** Here we will relate the $Q_v$ to fractional resolvent moments by applying the results of Lemma B.2 to the family eq. (2.4). By eq. (B.6) and eq. (B.7) we have

$$\frac{d}{d\xi}E_n = -\langle U_\alpha \psi_n, \psi_n \rangle \tag{2.11}$$

and for $\Gamma_n(E)$, the inverse function of $E_n$,

$$\frac{d}{dE}\Gamma_n(E) = -\frac{1}{\langle U_\alpha \psi_n(\Gamma_n(E)), \psi_n(\Gamma_n(E))\rangle} . \tag{2.12}$$

For $E \notin \sigma(H_{\eta_\alpha})$ define $K_{\eta_\alpha,E} := U_\alpha^{1/2}(H_{\eta_\alpha} - E)^{-1}U_\alpha^{1/2}$. If $E \notin \sigma(H_{\hat{\eta}_\alpha})$ then, by Lemma B.2, $\{\Gamma_n(E)\}$ are the repeated eigenvalues and $\phi_n(E) := U_\alpha^{1/2}\psi_n(\Gamma_n(E))$ corresponding complete (non-normalized) eigenvectors for the unbounded self-adjoint operator $K^{-1}_{\hat{\eta}_\alpha,E}$. With this notation we have

**Theorem 2.1.** *If $\eta_\alpha \neq \hat{\eta}_\alpha$ and $\sigma(H_{\eta_\alpha}) \cap \sigma(H_{\hat{\eta}_\alpha}) \cap \mathcal{J} = \emptyset$, then the eigenfunction correlator $Q_v(\mathcal{J}; x, \alpha)$ for $H^{(\Lambda)}$ admits the following representation:*

$$\begin{aligned} Q_v(\mathcal{J}; x, \alpha) \;&=\; \sum_n \int_{\mathcal{J}} dE\; \delta(\Gamma_n(E) + \lambda(\eta_\alpha - \hat{\eta}_\alpha))\; |\Gamma_n(E)|^v \\ &\quad \times \left\| \chi_x\, (H_{\hat{\eta}_\alpha} - E)^{-1}\, U_\alpha^{1/2}\, \phi_n(E) \right\|^v \Big/ \, \|\phi_n(E)\|^v \; . \end{aligned} \tag{2.13}$$

*Furthermore, for any $E$ and $a < b$ such that $E$ is not an eigenvalue of $H_a$ or $H_b$,*

$$\int_a^b \mathrm{d}\eta_\alpha \sum_n \; \delta(\Gamma_n(E) + \lambda(\eta_\alpha - \hat{\eta}_\alpha)) = \left[\operatorname{Tr} P_{\le E}(H_a) - \operatorname{Tr} P_{\le E}(H_b)\right] / \lambda \; . \tag{2.14}$$

*Remark.* Using Lemma B.2 it is easy to see that the condition $\sigma(H_{\eta_\alpha}) \cap \sigma(H_{\hat{\eta}_\alpha}) \cap \mathcal{J} = \emptyset$ holds for Lebesgue almost every $\eta_\alpha$.

*Proof.* As $\Gamma_n(E) = E_n^{-1}(E)$ we have for arbitrary $\eta_\alpha$ that $E \in \sigma(H_{\eta_\alpha})$ if and only if $\Gamma_n(E) = \xi = \lambda(\hat{\eta}_\alpha - \eta_\alpha)$ for some $n$. Thus one may express sums over eigenvalues as integrals with suitably weighted $\delta$-functions:

$$\sum_{n:\, E_n \in \mathcal{J}} \ldots \;=\; \int_{\mathcal{J}} \mathrm{d}E \sum_n \delta\left(\Gamma_n(E) + \lambda \Delta\eta_\alpha\right) \left| \frac{\mathrm{d}}{\mathrm{d}E}\Gamma_n(E) \right| \ldots , \tag{2.15}$$

where $\Delta\eta_\alpha = \eta_\alpha - \hat{\eta}_\alpha$. In particular, eq. (2.12) implies that

$$\begin{aligned} Q_0(\mathcal{J}; x, \alpha) \;&=\; \sum_{n:\, E_n \in \mathcal{J}} \langle U_\alpha \psi_n(\Gamma_n(E)), \psi_n(\Gamma_n(E))\rangle \\ &= \int_{\mathcal{J}} \mathrm{d}E \sum_n \delta\left(\Gamma_n(E) + \lambda\, \Delta\eta_\alpha\right) \; , \end{aligned} \tag{2.16}$$

and for other values of $v$:

$$
\begin{aligned}
Q_v(\mathcal{J}; x, \alpha) &= \sum_{n:\, E_n \in \mathcal{J}} \langle U_\alpha \psi_n, \psi_n \rangle \left( \frac{\langle \chi_x \psi_n, \psi_n \rangle}{\langle U_\alpha \psi_n, \psi_n \rangle} \right)^{v/2} \\
&= \int_{\mathcal{J}} \mathrm{d}E \sum_n \delta\left(\Gamma_n(E) + \lambda\, \Delta\eta_\alpha\right) \frac{\langle \chi_x \psi_n, \psi_n \rangle^{v/2}}{\|\phi_n(E)\|^v} \,.
\end{aligned}
\tag{2.17}
$$

For $\Gamma_n(E) = -\lambda \Delta\eta_\alpha = \xi$, i.e. $E = E_n(\xi)$, it follows that

$$
\psi_n(\xi) = \xi (H_{\hat{\eta}_\alpha} - E)^{-1} U_\alpha^{1/2} \phi_n(E) \,, \tag{2.18}
$$

since $K_{\hat{\eta}_\alpha}^{-1} \phi_n(E) = \Gamma_n(E) U_\alpha^{1/2} \psi_n(\xi)$ . Thus eq. (2.13) follows from eq. (2.17).

For fixed $E$, the left-hand side of eq. (2.14), up to a factor $1/\lambda$, counts the number of $n$ for which $\Gamma_n(E) = \lambda(\hat{\eta}_\alpha - \eta_\alpha)$ has a solution with $a < \eta_\alpha < b$. This number is exactly the decrease in the number of eigenvalues of $H_{\eta_\alpha}$ below $E$ as $\eta_\alpha$ is moved from $a$ to $b$. That yields eq. (2.14). □

**2.3. Spectral shift bounds.** In applying the interpolation argument seen in Lemma 2.1, we shall need bounds on the *spectral shift function*, defined by

$$
S_{\alpha,\lambda}(H^{(\Lambda)}; E) \;:=\; \mathrm{Tr}\left[ P_{\le E}\big(H^{(\Lambda)}_{\eta_\alpha = 0}\big) - P_{\le E}\big(H^{(\Lambda)}_{\eta_\alpha = 1}\big) \right] , \tag{2.19}
$$

and expressing how many energy levels are pushed over $E$ when the value of the parameter $\eta_\alpha$ is increased by 1. Note that $S_{\alpha,\lambda}(H^{(\Lambda)}; E)$ is non-negative and bounded (since $|\Lambda| < \infty$):

$$
0 \le S_{\alpha,\lambda}(H^{(\Lambda)}; E) \le \mathrm{Tr}\, P_{\le E}\big(H^{(\Lambda)}_{\eta_\alpha = 0}\big) \le C_p (1 + |E - E_0|)^p |\Lambda| \,, \tag{2.20}
$$

for any $p > d/2$.

In the discrete setup, where the role of $U_\alpha$ is taken by a rank-one operator, the shift is at most 1. For continuum operators there is no such uniform bound independent of the volume (see ref. [44]), however $S_{\alpha,\lambda}(H^{(\Lambda)}; E)$ has locally bounded $L^p$ norms as a function of the energy $E$, as was shown in ref. [19].

**Lemma 2.2** **($L^p$ boundedness of the spectral shift).** *Let $1 \le p < \infty$ and $n \in \mathbb{N}$ such that $n > dp/2$. Then there exists a constant $C_{p,\lambda,n} < \infty$ such that*

$$
\int_{-\infty}^{E_+} S_{\alpha,\lambda}(H^{(\Lambda)}; E)^p \mathrm{d}E \;\le\; C_{p,\lambda,n} (1 + |E_+ - E_0|)^{n+1} \tag{2.21}
$$

*uniformly in the domain $\Lambda$ as well as in the choice of $\alpha$ and the random parameters $\eta_\beta$ ($\beta \ne \alpha$).*

**2.4. From fractional moment bounds to localization.** We shall now put together the elements introduced in the previous sections and prove that rapid decay of the resolvent moments implies localization in its various manifestations.

**Theorem 2.2.** *Let $H_\omega$ be a random Schrödinger operator which satisfies $\mathcal{A}$. Let $0 < s \le 1$ and $\mathcal{J} \subset (-\infty, E_+]$ be a bounded open interval. Suppose that for some $C < \infty$, $\mu > 0$ and a bounded region $\Lambda$,*

$$\mathbb{E}\left(\int_{\mathcal{J}} \mathrm{d}E \left\| \chi_x \frac{1}{H^{(\Lambda)} - E} \chi_y \right\|^s\right) \le Ce^{-\mu \operatorname{dist}_\Lambda(x,y)}, \tag{2.22}$$

*for all $x, y \in \Lambda$. Then, for any $v < 1/(2-s)$, there exists a volume independent constant $C_{s,v}(E_+, \lambda) < \infty$ such that*

$$\mathbb{E}\left(\sup_{f \in C_c(\mathcal{J}):\, |f| \le 1} \left\| \chi_x\, f(H^{(\Lambda)})\, \chi_y \right\|\right) \le C_{s,v}(E_+, \lambda)\, e^{-v\mu \operatorname{dist}_\Lambda(x,y)}, \tag{2.23}$$

*for all $x, y \in \Lambda$.*

*Remark.* The operators $H^{(\Lambda)}$ have finite spectrum in $\mathcal{J}$. Thus, in eq. (2.23) one could equivalently take the supremum over *all* functions on $\mathcal{J}$ as long as they pointwise satisfy $|f| \le 1$. The only reason for stating eq. (2.23) in terms of functions in $C_c(\mathcal{J})$ is to prepare for a limiting argument in the proof of Theorem 1.1.

*Proof.* Assume that for some open interval $\mathcal{J}$ the fractional moment bound eq. (2.22) holds. By eq. (2.6), to prove Theorem 2.2 it suffices to establish a related bound on $\mathbb{E}(Q_1(\mathcal{J}; x, \alpha))$ for $\alpha \in \mathcal{I}$ with $\operatorname{dist}(\alpha, y) \le 2r$. This in turn is controlled through Lemma 2.1 by an interpolating product of $\mathbb{E}(Q_2(\mathcal{J}; x, \alpha))$ and $\mathbb{E}(Q_v(\mathcal{J}; x, \alpha))$, with $v < 1$. Leaving room for one more interpolation, we choose $v < s$.

To estimate $\mathbb{E}(Q_v(\mathcal{J}; x, \alpha))$, we start from the representation eq. (2.13) and average first over $\eta_\alpha$ at specified values of $\eta_\zeta$ for $\zeta \ne \alpha$. For almost every choice of $\eta_\zeta$ $(\zeta \ne \alpha)$, $E$ is neither an eigenvalue of $H^{(\Lambda)}_{\eta_\alpha = 0}$ nor of $H^{(\Lambda)}_{\eta_\alpha = 1}$. Also, for fixed $\hat{\eta}_\alpha$ and almost every $\eta_\alpha$ we have $\sigma(H^{(\Lambda)}_{\eta_\alpha}) \cap \sigma(H^{(\Lambda)}_{\hat{\eta}_\alpha}) \cap \mathcal{J} = \emptyset$. Thus we can apply eq. (2.13) and eq. (2.14) to conclude

$$\begin{aligned} &\mathbb{E}(Q_v(\mathcal{J}; x, \alpha)) \\ &\quad \le 2^v \lambda^{v-1} D\, \mathbb{E}\left(\int_{\mathcal{J}} \mathrm{d}E\, S_{\alpha,\lambda}(H^{(\Lambda)}; E) \left\| \chi_x (H^{(\Lambda)} - E)^{-1} U_\alpha^{1/2} \right\|^v\right). \end{aligned} \tag{2.24}$$

Here we have used $\mathcal{A}3$, the bound $|\Gamma_n(E)| \le 2\lambda$ (imposed by the $\delta$ functions in eq. (2.13)), and we chose the re-sampled variable $\hat{\eta}_\alpha$ to have the same conditional distribution as $\eta_\alpha$.

Using Hölder's inequality we can further estimate the right hand side of eq. (2.24) to get

$$\begin{aligned}\mathbb{E}(Q_v(\mathcal{J}; x, \alpha)) &\le 2^v \lambda^{v-1} D \left( \mathbb{E} \int_{\mathcal{J}} \mathrm{d}E \; S_{\alpha,\lambda}(H^{(\Lambda)}; E)^{\frac{s}{s-v}} \right)^{\frac{s-v}{s}} \\ &\quad \times \left( \mathbb{E} \int_{\mathcal{J}} \mathrm{d}E \; \left\| \chi_x (H^{(\Lambda)} - E)^{-1} U_\alpha^{1/2} \right\|^s \right)^{v/s} \\ &\le C_{s,v,\lambda,n} (1 + |E_+ - E_0|)^{(n+1)(s-v)/s} e^{-\mu v \operatorname{dist}_\Lambda(x,\alpha)/s} \end{aligned} \tag{2.25}$$

for any integer $n > \frac{sd}{2(s-v)}$. In the final estimate we have used the spectral shift bound from Lemma 2.2 and the assumption eq. (2.22).

Collecting all our bounds as well as the uniform bound for $\mathbb{E}(Q_2(\mathcal{J}; x, \alpha))$ provided by eq. (2.8), we arrive at

$$\mathbb{E}(Y_\Lambda(\mathcal{J}; x, y)) \le C_{v,s}(E_+, \lambda) e^{-\frac{\mu v \operatorname{dist}_\Lambda(x,\alpha)}{s(2-v)}} . \tag{2.26}$$

After reorganizing the exponent this yields eq. (2.23). □

**2.5. Infinite volumes – proof of Theorem 1.1.** The eigenfunction correlator methods used above are most easily implemented for the finite volume operators $H^{(\Lambda)}$. However, it is important to note that the bounds obtained in this way do not depend on the size of the volume $\Lambda$ (except for the presence of the modified distance $\operatorname{dist}_\Lambda(x, \alpha)$, which for fixed $x, \alpha$ is equal to the usual distance $|x - \alpha|$ for sufficiently large $\Lambda$). In this section we take the infinite volume limit, proving Theorem 1.1 and deduce the results on spectral localization, as discussed in Sect. 1.2. The fundamental analytic tools are 1) strong resolvent convergence and 2) the RAGE theorem. The infinite region $\Omega$ being fixed throughout this section, we write $H = H^{(\Omega)}$.

*Proof of Theorem 1.1.* For $g \in C_c(\mathcal{J})$, $g(H^{(\Lambda_n)})$ converges strongly to $g(H)$ since $H^{(\Lambda_n)}$ converges to $H$ in the strong resolvent sense. Thus

$$\|\chi_x g(H) \chi_y\| \le \liminf_{n\to\infty} \left\| \chi_x g(H^{(\Lambda_n)}) \chi_y \right\| . \tag{2.27}$$

Taking suprema and applying Fatou's lemma yields

$$\begin{aligned}&\mathbb{E}\left( \sup_{g:\, |g| \le 1} \|\chi_x g(H) P_{\mathcal{J}}(H) \chi_y\| \right) \\ &\qquad \le \liminf_{n\to\infty} \mathbb{E} \left( \sup_{f \in C_c(\mathcal{J}):\, |f| \le 1} \left\| \chi_x f(H^{(\Lambda_n)}) \chi_y \right\| \right) . \end{aligned} \tag{2.28}$$

The supremum on the left-hand side is initially only taken over $g \in C_c(\mathcal{J})$, but can be extended to all Borel functions $g$ without changing its value. To see this, let $T_g := \chi_x g(H) P_{\mathcal{J}}(H) \chi_y$ for a fixed Borel function

$g$ with $|g| \le 1$. For an orthonormal basis $(\phi_k)$, $k = 1, 2, \dots$, set $P_N = \sum_{k=1}^{N} \langle \cdot, \phi_k \rangle \phi_k$. As $T_g$ is compact we have

$$\lim_{N\to\infty} \left\| T_g P_N - T_g \right\| = 0 \,. \tag{2.29}$$

The spectral measures $d\mu_k(\lambda) = d \left\| E(\lambda)\chi_y\phi_k \right\|^2$ associated with $H$ are Borel measures on $\mathbb{R}$ and thus regular (e.g. [61]). Thus it follows from elementary considerations that there exist $g_n \in C_c(\mathcal{J})$ with $|g_n| \le 1$ and $\int_{\mathcal{J}} |g_n - g|^2 \, d\mu_k \to 0$ as $n \to \infty$ simultaneously for $k = 1, \dots, N$. We have

$$\left\| T_{g_n} P_N - T_g P_N \right\| \le \sum_{k=1}^{N} \left\| (g_n(H) - g(H)) P_{\mathcal{J}}(H) \chi_y \phi_k \right\| \,. \tag{2.30}$$

By the above, this implies that $T_{g_n} P_N \to T_g P_N$ as $n \to \infty$. Together with (2.29) this implies that for every $\varepsilon > 0$ there exist $N$ and $n$ such that $\left\| T_g \right\| \le \left\| T_{g_n} P_N \right\| + \varepsilon \le \left\| T_{g_n} \right\| + \varepsilon$. As $\varepsilon > 0$ was arbitrary, we conclude that the left-hand side of eq. (2.28) does not change if the supremum is taken over all Borel functions with $|g| \le 1$.

Equation (1.7) now follows through the uniform bounds eq. (2.23) provided by Theorem 2.2 under the assumption eq. (1.6). □

We now turn to the spectral type of $H$, where $\Omega = \mathbb{R}^d$ is assumed, and therefore $\mathrm{dist}_\Omega(x, y) = \mathrm{dist}(x, y)$. The absence of continuous spectrum can be demonstrated from our estimates using the RAGE theorem (e.g., see ref. [21] for discussion and references) which implies that the projection $P_c(H)$ onto the continuous spectrum of $H$ satisfies

$$\|P_c(H)\psi\|^2 \;=\; \lim_{R\to\infty} \lim_{T\to\infty} \int_0^T \frac{\mathrm{d}t}{T} \; \left\| \chi_{\{|x-x_0| \ge R\}} \, \mathrm{e}^{-\mathrm{i}tH} \psi \right\|^2 \,. \tag{2.31}$$

As a consequence,

**Theorem 2.3 (RAGE theorem for random operators).** *For the random operators considered here, if for some open interval $\mathcal{J}$*

$$\mathbb{E}\left( \left\| \chi_{\{|x-x_0| \ge R\}} \, e^{-\mathrm{i}tH} \, P_{\mathcal{J}}(H) \, \chi_{x_0} \right\| \right) \;\le\; g(R) \tag{2.32}$$

*with $g(R) \to 0$ as $R \to \infty$ uniformly in $t$ and $x_0$, then $P_c(H) P_{\mathcal{J}}(H) = 0$ almost surely.*

*Proof.* If $\phi \in L^2(\mathbb{R}^d)$ is compactly supported with $\operatorname{supp} \phi \subset B_{x_0}^r$ for some $x_0 \in \mathbb{R}^d$, then the RAGE theorem implies

$$\begin{aligned} &\|P_c(H) P_{\mathcal{J}}(H)\phi\|^2 \\ &\qquad \le \liminf_{R\to\infty} \liminf_{T\to\infty} \int_0^T \frac{\mathrm{d}t}{T} \left\| \chi_{|x-x_0| \ge R} e^{-\mathrm{i}tH} P_{\mathcal{J}}(H) \chi_{x_0} \right\|^2 \|\phi\|^2 \,. \end{aligned} \tag{2.33}$$

Upon taking expectations, Fatou's lemma and Fubini's theorem yield

$$\mathbb{E}\left(\|P_c(H)P_{\mathcal{J}}(H)\phi\|^2\right) \leq \liminf_{R\to\infty}\liminf_{T\to\infty}\int_0^T \frac{dt}{T} g(R)\|\phi\|^2 = 0\,. \quad (2.34)$$

Thus $P_c(H)P_{\mathcal{J}}(H)\phi = 0$ almost surely. This implies the theorem since there exists a *countable* total set of $\phi$ with each $\phi$ supported in $B^r_{x_0}$ for suitable $x_0$. □

To complete the proof of pure point spectrum in $\mathcal{J}$ we note that eq. (1.7), proven above, yields the assumption of Theorem 2.3: Covering $\{|x - x_0| \geq R\}$ with balls $B^r_\alpha$, $\alpha \in \mathcal{I}$, we get for every $\nu \in (0, \mu/(2-s))$ that $\mathbb{E}\left(\|\chi_{\{|x-x_0|\geq R\}}e^{-\mathrm{i}tH}P_{\mathcal{J}}(H)\chi_{x_0}\|\right) \leq Ce^{-\nu R}$.

Another consequence of eq. (1.7) is that

$$\mathbb{E}\left(\sum_{x,y}\frac{\mathrm{e}^{\nu|x-y|}}{1+|y|^{d+1}}\sup_{g:|g|\leq 1}\left\|\chi_x g(H)P_{\mathcal{J}}(H)\chi_y\right\|\right) < \infty\,, \quad (2.35)$$

which implies

$$\sup_{g:|g|\leq 1}\left\|\chi_x g(H)P_{\mathcal{J}}(H)\chi_y\right\| \leq \text{const.}\ (1+|y|^{d+1})\ \mathrm{e}^{-\nu|x-y|} \qquad \text{a.s.} \quad (2.36)$$

Since the Kronecker delta functions $\delta_E(x)$ are Borel measurable, this implies that with probability one all eigen-projections of $H$ satisfy (1.9).

Almost surely all eigenvalues of $H$ in $\mathcal{J}$ are finitely degenerate. An argument for this (based on compactness considerations and spectral averaging) which applies to our model was provided in the proof of Theorem 3.2 of ref. [18].

Given this and using (1.9), we can proceed as follows. We have, for almost every configuration $\omega$ and $E \in \mathcal{J}$, (*i.*) The range of $\delta_E(H)$, denoted $\mathcal{R}(\delta_E(H))$, is finite dimensional and (*ii.*) $\|\chi_x\delta_E(H)\chi_{|x|\leq R}\| = \mathrm{O}(e^{-\nu|x|})$ for every $x \in \mathbb{R}^d$ and $R > 0$.

From (*i.*) it follows that $\mathcal{R}(\delta_E(H)) = \mathcal{R}(\delta_E(H)\chi_{|x|\leq R})$ for $R \geq R_0(\omega, E)$. Thus (*ii.*) implies $\|\chi_x\phi\| = \mathrm{O}(e^{-\nu|x|})$ for every $\phi$ in the range of $\delta_E(H)$, i.e. all eigenfunctions. An $L^\infty$ version of exponential decay follows from well know facts regarding the smoothness of eigenfunctions of Schrödinger operators ("elliptic regularity", e.g. see ref. [67]).

This completes the proof of spectral localization properties.

## 3. Finiteness of the fractional-moments

In this section we prove two technical results, Lemmas 3.3 and 3.4, which permit to bound disorder averages of resolvent norms (raised to a fractional power) at fixed energy. These are the analogues of Proposition 1.1 required in the proof of Theorem 1.2 in Sect. 4, where energy averaging is replaced by disorder averaging in "local environments." An improved bound under the stronger assumption $\mathcal{A}3'$ is presented in Proposition 4.3 below.

We begin with an overview of the argument and then present the two lemmas. Finally we give a proof of the two main technical results.

**3.1. Why are disorder averages finite?** The analysis presented below has its genesis in the discrete setup, where the finiteness of disorder averages of the Green function is quickly implied by a rank-one perturbation argument. However, for the continuum operators considered here, that short argument requires thorough remaking since the local potential term ($U_\alpha$) is now an operator of infinite rank.

To contrast the two cases, discrete and continuum, it is instructive to compare them in a unified framework provided by the Birman-Schwinger relation (see eq. (1.16), eq. (B.4), and eq. (B.9)):

$$U^{1/2} \frac{1}{\widehat{H} + \eta U - E} U^{1/2} = \left[ \left[ \widehat{K}_E \right]^{-1} + \eta \right]^{-1} , \tag{3.1}$$

with

$$\widehat{K}_E \;=\; U^{1/2} \frac{1}{\widehat{H} - E} U^{1/2} \,. \tag{3.2}$$

In the discrete case, with $U = |o\rangle \langle o|$ a rank-one operator, $\widehat{K}_E^{-1}$ is simply a complex number, and eq. (3.1) readily implies

$$\left| \left\{ \eta \;:\; \left| \langle o| \frac{1}{\widehat{H} + \eta\, |o\rangle \langle o| - E} |o\rangle \right| > t \right\} \right| \;=\; \frac{2}{t} \,, \tag{3.3}$$

and thus finiteness of fractional $\eta$-moments by the "layer-cake" representation eq. (1.12).

For general $U$, eq. (3.1) implies the weak $L^1$ bound

$$\left| \left\{ \eta \in [0,1] \;:\; \left\| U^{1/2} \frac{1}{\widehat{H} + \lambda \eta U - E} U^{1/2} \right\| > t \right\} \right| \;\leq\; \frac{2(1 + \xi_{E,\lambda})}{\lambda t} \tag{3.4}$$

where $\xi_{E,\lambda}$ denotes the number of eigenvalues of $\widehat{K}_E^{-1}$ in the interval $(-\lambda, 0)$ (see eq. (B.9)). For trace class perturbations $U$, the simple bound $\xi_{E,\lambda} \leq \operatorname{Tr} U$ allows the analysis to proceed from eq. (3.4) much as in the rank-one case, however for the Schrödinger operators considered here, with $U$ one of the "bumps" $U_\alpha$, the perturbation $\lambda\eta U$ is not trace class. Nonetheless, it is *relatively compact* with respect to $\widehat{H}$ due to the kinetic term $\boldsymbol{D_A} \cdot \boldsymbol{D_A}$. It follows that for $E \notin \sigma(\widehat{H})$ the operator $\widehat{K}_E$ is *compact* and therefore $\xi_{E,\lambda}$ is *finite*. However, there is no reason to expect a bound on $\xi_{E,\lambda}$ which is uniform in $E$ and in the various parameters implicit in $\widehat{H}$.

To examine the factor $\xi_{E,\lambda}$ more closely, note that $E$ is an eigenvalue of $H = \widehat{H} + \lambda\eta U$ precisely when $\lambda\eta$ is an eigenvalue of $(-\widehat{K}_E^{-1})$. Hence, we may equate $\xi_{E,\lambda}$ with the number of times that $E$ becomes an eigenvalue of $H$ as $\eta$ is moved from 0 to 1. By the monotonicity of $H$ in $\eta$ (implied by the positivity of $U$), we get

$$\xi_{E,\lambda} \leq \mathrm{Tr}\,[P(\widehat{H} < E) - P(\widehat{H} + \lambda U < E)]\,, \tag{3.5}$$

with equality unless $E$ is a degenerate eigenvalue for some $\eta \in (0, 1)$.

The quantity on the right side of eq. (3.5) is the 'Krein spectral shift' at energy $E$ between $\widehat{H}$ and $\widehat{H} + \lambda U$, and has recently been the subject several of studies [19,40]. A key fact is that for Schrödinger operators the Krein spectral shift is locally integrable as a function of $E$ – indeed, it has been shown to be locally $L^p$ for every $p \in [1, \infty)$ with explicit bounds [19], a fact which we used in Sect. 2 in the proof of Theorem 1.1. However, in general the spectral shift is not locally bounded; examples exist of Schrödinger operators for which it is arbitrarily large at certain energies [44].

Thus, the moments of continuum Green functions differ from their discrete counterparts in an essential way – after averaging over a *single* coupling one still does not get a uniform bound. Nevertheless, it is natural to guess that for random operators the spectral shift has a finite expectation value, since averaging over disorder may play a role somewhat similar to averaging over energy.

In the derivation of our fractional moment bounds (see Lemma 3.3 below), we do not consider directly the average of the spectral shift. Instead, we find it more convenient to derive an analogue of eq. (3.3) for continuum operators (from which fractional moment bounds follow quite easily). Nonetheless, the notion that a disorder averaged spectral shift is bounded is one of the driving ideas behind this work, and we find that the techniques developed in this section imply a result of this type (Theorem D.1, stated and proved in Appendix D below). In the end however, the fractional moment bounds obtained here are somewhat stronger than those which follow from Theorem D.1.

**3.2. The $1/t$-tails.** A useful result for moment bounds is a general weak-$L_1$ type bound for the boundary values of resolvents of maximally dissipative operators. We state this as the second part of the following lemma, which is essentially a special case of results proven in [56]. For completeness, we give a proof based on properties of the Hilbert transform in Appendix C. The first part of the lemma is a well known result in scattering theory, e.g. [24].

**Lemma 3.1 (Weak $L_1$ bound).** *Let $\mathcal{H}$ and $\mathcal{H}_1$ be separable Hilbert spaces, $A$ a maximally dissipative operator in $\mathcal{H}$, as well as $M_1 : \mathcal{H} \to \mathcal{H}_1$ and $M_2 : \mathcal{H}_1 \to \mathcal{H}$ Hilbert-Schmidt operators. Then*

(1) *The "boundary values"*

$$M_1 \frac{1}{A - v + i0} M_2 \;=\; \lim_{\epsilon \downarrow 0} M_1 \frac{1}{A - v + i\epsilon} M_2 \tag{3.6}$$

*exist as Hilbert-Schmidt operators for almost every $v \in \mathbb{R}$, with convergence in the Hilbert-Schmidt norm.*

(2) *There exists $C_W < \infty$ (independent of $A$, $M_1$, $M_2$) such that*

$$\left|\left\{v \ : \ \left\| M_1 \frac{1}{A - v + \mathrm{i}0} M_2 \right\|_{HS} > t \right\}\right| \ \leq \ C_W \ \|M_1\|_{HS} \ \|M_2\|_{HS} \ \frac{1}{t} \ . \tag{3.7}$$

*Remark.* Recall that a densely defined operator $A$ is called *dissipative* if, for each $\varphi \in \mathcal{D}(A)$, we have $\mathrm{Im}\langle \varphi, A\varphi \rangle \geq 0$. It is called *maximally dissipative* if it has no proper dissipative extension, which is equivalent to contractivity of the semi-group $e^{\mathrm{i}tA}$ generated by $\mathrm{i}A$. If a dissipative operator $A$ has strictly positive imaginary part, such as a Birman-Schwinger operator $A_{BS}$ of the type consider in Lemma B.1, then the $i0$ in (3.7) is not needed since $(\xi - A_{BS})^{-1}$ is norm-continuous for $\xi$ in the closed lower half plane.

In our applications of Lemma 3.1 we will use the following "off-diagonal" version which follows from the lemma via the Birman-Schwinger identity and a simple change of variables. For ease of presentation, we state this version with the additional assumption that the operator $A$ has strictly positive imaginary part, thus avoiding the issue of whether certain limits exist.

**Proposition 3.2.** *Let $A$ be a maximally dissipative operator with strictly positive imaginary part on a Hilbert space $\mathcal{H}$, let $M_1, M_2$ be Hilbert-Schmidt operators, and let $U_1$, $U_2$ be non-negative operators. Then*

$$\left|\left\{\langle v_1, v_2\rangle \in [0,1]^2 \ : \ \left\| M_1 U_1^{1/2} \frac{1}{A - v_1 U_1 - v_2 U_2} U_2^{1/2} M_2 \right\|_{HS} > t \right\}\right| \\ \leq \ 2\, C_W \, \|M_1\|_{HS} \, \|M_2\|_{HS} \, \frac{1}{t} \ . \tag{3.8}$$

*Proof.* The key to the proof is the change of variables $v_\pm = \frac{1}{2}(v_1 \pm v_2)$, so that $v_1 U_1 + v_2 U_2 = v_+(U_1 + U_2) + v_-(U_1 - U_2)$. From the Birman-Schwinger identity, eq. (B.4),

$$\left[(U_1 + U_2)^{1/2} \frac{1}{A - v_1 U_1 - v_2 U_2} (U_1 + U_2)^{1/2}\right] = \left[[K(v_-)]^{-1} - \ v_+\right]^{-1} \ , \tag{3.9}$$

where $[\cdot]$ denotes the restriction of an operator to $\ker(U_1 + U_2)^\perp$ and

$$K(v_-) = (U_1 + U_2)^{1/2} \frac{1}{A - v_-(U_1 - U_2)} (U_1 + U_2)^{1/2} \ . \tag{3.10}$$

Thus,

$$M_1 U_1^{1/2} \frac{1}{A - v_1 U_1 - v_2 U_2} U_2^{1/2} M_2 = \widetilde{M}_1 \left[[K(v_-)]^{-1} - v_+\right]^{-1} \widetilde{M}_2 \tag{3.11}$$

with $\widetilde{M}_1 = M_1 U_1^{1/2} P[U_1 + U_2]^{-1/2} P$ and $\widetilde{M}_2 = P[U_1 + U_2]^{-1/2} P U_2^{1/2} M_2$ where $P$ denotes orthogonal projection onto $\ker(U_1 + U_2)^{\perp}$.

Since $U_1$ and $U_2$ are positive, $\|U_j^{1/2} P[U_1 + U_2]^{-1/2} P\| \leq 1$ and therefore

$$\left\| \widetilde{M}_j \right\|_{HS} \leq \left\| M_j \right\|_{HS} . \tag{3.12}$$

Furthermore $[K(v_-)]^{-1}$ is maximally dissipative as shown in Lemma B.1. Thus for each fixed $v_-$ we are in the situation governed by Lemma 3.1, and Prop. 3.2 follows via the Fubini Theorem. The factor of 2 on the right hand side results from the Jacobian of the transformation $\langle v_1, v_2 \rangle \mapsto \langle v_+, v_- \rangle$. □

We will use Prop. 3.2 to conclude that for the random operator $H$,

$$\begin{aligned} &\left| \left\{ \langle \eta_\alpha, \eta_\beta \rangle \in [0,1]^2 \ : \ \left\| M_1 U_\alpha^{1/2} \frac{1}{H - z} U_\beta^{1/2} M_2 \right\|_{HS} > t \right\} \right| \\ &\qquad\qquad \leq \ 2\, C_W \ \|M_1\|_{HS} \ \|M_2\|_{HS} \ \frac{1}{\lambda t} . \end{aligned} \tag{3.13}$$

Several comments are in order:

(1) When the energy $z$ lies in the lower half plane (3.13) follows directly from Prop. 3.2 with $v_1 = \eta_\alpha$, $v_2 = \eta_\beta$ and $U_1 = \lambda U_\alpha$, $U_2 = \lambda U_\beta$.
(2) Eq. (3.13) also holds for $z$ in the upper half plane, as can be seen by taking conjugates.
(3) For real energies ($z = E \in \mathbb{R}$), eq. (3.13) holds also in the limits $z \to E \pm \mathrm{i}0$, as follows from the bound at complex energies and Fatou's lemma (compare the argument for proving Lemma 3.1 in Appendix C).

The above results display that once the resolvent of the continuum operator is bracketed with a pair of Hilbert-Schmidt operators $M_1$, $M_2$ its fractional moments can be handled similarly to the discrete case. In the proofs of Lemmas 3.3 and 3.4 we use an argument which shows that $\chi_x (H - z)^{-1} \chi_y$ can be presented as a sum of a *bounded* operator and one of the form

$$M_x \mathbf{1}_{B_x} (H - z)^{-1} \mathbf{1}_{B_y} M_y , \tag{3.14}$$

with $M_\sharp$ Hilbert-Schmidt operators and $B_x$, $B_y$ sets somewhat larger than the balls of radius $r$ around $x$ and $y$. This will allow us to conclude finiteness of fractional moments after "averaging over the local environment," i.e., averaging over all $\eta_\alpha$ with $U_\alpha$ non-zero in $B_x$ or $B_y$.

**3.3. A pair of fractional-moment lemmas.** We now present the two basic technical results which give finiteness of the fractional moments, as well as a "decoupling argument" to be used in the proof of Theorem 1.2.

**Definition 3.1.** For $\alpha \in \mathcal{I}$, we denote

$$\mathcal{I}_\alpha := \{\zeta \in \mathcal{I} : \mathrm{dist}(\alpha, \zeta) < 3r\} \tag{3.15}$$

and set $\mathcal{I}_{\alpha,\beta} = \mathcal{I}_\alpha \cup \mathcal{I}_\beta$. Likewise, for any subset $\mathcal{L} \subset \mathcal{I}$, $\mathcal{I}_\mathcal{L} := \cup_{\alpha\in\mathcal{L}} \mathcal{I}_\alpha$. By $\mathcal{F}^c_\mathcal{L}$ we denote the $\sigma$-algebra generated by all $\eta_\alpha$ with $\alpha$ *not* in $\mathcal{I}_\mathcal{L}$. Thus, $\mathbb{E}\left(\cdot|\mathcal{F}^c_\mathcal{L}\right)$ represents averaging over the "local environment" of $\mathcal{L}$.

Our first bound yields the finiteness of the $s$-moments after averaging over the local-environment.

**Lemma 3.3.** *Let $H$ be a random Schrödinger operator which satisfies $\mathcal{A}$ with disorder strength $\lambda > 0$. Then there exists $C_\lambda < \infty$ such that the restriction of $H$ to a region $\Omega$ obeys*

$$\mathrm{Prob}\left( \left\| U_\alpha \frac{1}{H^{(\Omega)} - E - \mathrm{i}\varepsilon} U_\beta \right\| > t \,\middle|\, \mathcal{F}^c_{\alpha,\beta} \right) \leq C_\lambda (1 + |E - E_0|)^{d+2} \, \frac{D^2}{t} \,, \tag{3.16}$$

*for any $\alpha, \beta \in \mathcal{I}$, any $E \in \mathbb{R}$ and $\varepsilon > 0$, where the coefficient $C_\lambda$ can be chosen such that*

$$C_\lambda \leq \text{const.}\, (1 + \lambda^{-1})(1 + \lambda)^{d+2} \,. \tag{3.17}$$

*Remarks:*

(1) Recall that $E_0 = \inf \sigma(H_0)$ and that $D$ is a bound on the conditional densities for $\eta_\alpha$ – see eq. (1.23).
(2) In Prop. 4.3 below we show that if $H$ satisfies $\mathcal{A}3'$ then Lemma 3.3 can be improved so that eq. (3.16) holds with $\sup_{\lambda>1} C_\lambda < \infty$.
(3) For fractional moments we use the "layer-cake" representation – $\mathbb{E}\,(X^s) = \int_0^\infty \mathrm{Prob}(X > t^{1/s})\mathrm{d}t$ – to conclude that

$$\mathbb{E}\left( \left\| U_\alpha \frac{1}{H^{(\Omega)} - E - \mathrm{i}\varepsilon} U_\beta \right\|^s \,\middle|\, \mathcal{F}^c_{\alpha,\beta} \right) \leq \frac{C_\lambda^s}{1-s} (1 + |E - E_0|)^{s(d+2)} D^{2s}. \tag{3.18}$$

Of course, this bound implies a similar estimate for the average over all variables.
(4) Using condition $\mathcal{A}2$, we obtain the following bound from eq. (3.18): For all measurable $\Lambda, \Lambda' \subset \mathbb{R}^d$,

$$\begin{aligned} \mathbb{E}\left( \left\| \mathbf{1}_\Lambda \frac{1}{H^{(\Omega)} - E - \mathrm{i}\varepsilon} \mathbf{1}_{\Lambda'} \right\|^s \,\middle|\, \mathcal{F}^c_{\mathcal{L}(\Lambda\cup\Lambda')} \right) & \\ \leq \frac{C_\lambda^s}{1-s} b_-^{-2s} (1 + |E - E_0|)^{s(d+2)} D^{2s} N_\Lambda \, N_{\Lambda'} & \end{aligned} \tag{3.19}$$

where $\mathcal{L}(\Lambda) := \{\gamma \in \mathcal{I} : \mathbf{1}_\Lambda U_\gamma \neq 0\}$ and $N_\Lambda$ indicates the number of points in $\mathcal{L}(\Lambda)$.

The second lemma (Lemma 3.4) focuses on the average of $\|U_\alpha(H-z)^{-1}U_\beta\|$ with respect to the local environment of *one* of the sites $\alpha, \beta$. The idea underlying this result is "re-sampling:" we compare the distribution of $(H-z)^{-1}$ with that of a reference operator $(\widehat{H}-z)^{-1}$, where $\widehat{H}$ is obtained from $H$ by redrawing the coupling variables $\eta_\zeta$ for $\zeta$ near $\beta$. The basic result is an estimate of the form

$$\mathbb{E}\left(\left\|U_\alpha \frac{1}{H_\omega - z} U_\beta\right\|^s \middle| \mathcal{F}_\beta^c\right) \leq \text{const.} \left\|\chi_\alpha \frac{1}{\widehat{H}_\omega - z} \mathbf{1}_S\right\|^s , \tag{3.20}$$

with $S$ an appropriate neighborhood of $\beta$. However, the result permits a number of variations, and the estimate (3.21) stated below is slightly complicated because it is tailored to the required application in the proof of Theorem 1.2.

For that application it is convenient to use a smooth function in place of $U_\beta$. Thus we fix a choice of a partition of unity, $\{\Theta_\alpha : \alpha \in \mathcal{I}\}$, with the following properties:

(1) Each function $\Theta_\alpha$ is non-negative and smooth with compact support in the ball of radius $4r/3$ centered at $\alpha$.
(2) The collection $\{\Theta_\alpha\}$ is a partition of unity: $\sum_\alpha \Theta_\alpha(q) = 1$.
(3) $\sup_\alpha(\|\nabla\Theta_\alpha\|_\infty , \|\Delta\Theta_\alpha\|_\infty) < \infty$.

Any choice with these properties will do, although the constant $\widetilde{C}$ in eq. (3.21) below will depend on the supremum in (3).

**Lemma 3.4.** *Let $H$ be a random Schrödinger operator which satisfies $\mathcal{A}$ with disorder strength $\lambda > 0$. Then there exists $\widetilde{C}_\lambda < \infty$ such that any restriction of $H$ to a region $\Omega$ obeys*

$$\begin{aligned}\text{Prob}&\left(\left\|\chi_x \frac{1}{H^{(\Omega)} - z}\Theta_\beta(1 + H_0 - E_0)^{1/2}\right\| > t \middle| \mathcal{F}_{\beta,\gamma}^c\right) \\ &\leq \widetilde{C}_\lambda\, (1 + |z - E_0|)^{(d+3)} \left\|\chi_x \frac{1}{\widehat{H}^{(\Omega)} - z}\mathbf{1}_{S_{\beta,\gamma}}\right\| \frac{D^2}{t} ,\end{aligned} \tag{3.21}$$

*for any $z \in \mathbb{C}\setminus\mathbb{R}$, $x \in \mathbb{R}^d$, and $\beta, \gamma \in \mathcal{I}$ with* $\mathrm{dist}(x, \beta)$, $\mathrm{dist}(x, \gamma) > 6r$. *The coefficient $\widetilde{C}_\lambda$ obeys*

$$\widetilde{C}_\lambda \leq \text{const.}\, (1 + \lambda)^{d+4} . \tag{3.22}$$

*Here $S_{\beta,\gamma} = \{q : \mathrm{dist}(q, \{\beta, \gamma\}) \leq 5r\}$ and $\widehat{H}$ is obtained from $H$ by replacing $\left\{\eta_\zeta : \zeta \in \mathcal{I}_{\beta,\gamma}\right\}$ with arbitrary values $\hat{\eta}_\zeta \in [0, 1]$:*

$$\widehat{H} = H + \lambda \sum_{\zeta\in\mathcal{I}_{\beta,\gamma}} (\hat{\eta}_\zeta - \eta_\zeta)U_\zeta(q) . \tag{3.23}$$

*Remark. i.* Note that we have resampled $H$ at a second site $\gamma \in \mathcal{I}$ as well as $\beta$, with the result that $\Theta_\beta$ is replaced by the characteristic function of a neighborhood of $\gamma$ and $\beta$. This additional resampling is required in the application of this lemma in the proof of Theorem 1.2, but does not play a key role in the present section. *ii.* The factor $(1 + H_0 - E_0)^{1/2}$ is included here to control various commutators $[H, \Theta]$ which appear in applications of the lemma. It is to bound this factor that we choose to work with the smooth function $\Theta_\beta$.

In the applications of Lemma 3.4 we will choose $\widehat{\eta}_\zeta$ to have the same probability distribution as $\eta_\zeta$. This will provide us with a "decoupling argument," as follows: If $X$ is a quantity for which $\mathbb{E}(X^s | \mathcal{F}^c_{\beta,\gamma}) \le A_s < \infty$, then eq. (3.21) and the Cauchy-Schwarz inequality imply that

$$
\begin{aligned}
&\mathbb{E}\left( X^{s/2} \left\| \chi_x \frac{1}{H^{(\Omega)} - z} \Theta_\beta (1 + H_0 - E_0)^{1/2} \right\|^{s/2} \middle| \mathcal{F}^c_{\beta,\gamma} \right) \\
&\quad \le A_s^{1/2}\, \mathbb{E}\left( \left\| \chi_x \frac{1}{H^{(\Omega)} - z} \Theta_\beta (1 + H_0 - E_0)^{1/2} \right\|^{s} \middle| \mathcal{F}^c_{\beta,\gamma} \right)^{1/2} \\
&\quad \le \text{const. } A_s^{1/2} \left\| \chi_x \frac{1}{\widehat{H}^{(\Omega)} - z} \mathbf{1}_S \right\|^{s/2} .
\end{aligned}
\tag{3.24}
$$

If $\widehat{\eta}_\zeta$ is distributed identically to $\eta_\zeta$ then, upon taking expectations, we obtain

$$
\begin{aligned}
&\mathbb{E}\left( X^{s/2} \left\| \chi_x \frac{1}{H^{(\Omega)} - z} \Theta_\beta (1 + H_0 - E_0)^{1/2} \right\|^{s/2} \right) \\
&\qquad\qquad \le \text{const. } A_s^{1/2}\, \mathbb{E}\left( \left\| \chi_x \frac{1}{H^{(\Omega)} - z} \mathbf{1}_S \right\|^{s/2} \right) .
\end{aligned}
\tag{3.25}
$$

In the final expression, we have replaced $\widehat{H}$ by $H$ since the two are identically distributed.

**3.4. Averaging over local environments.** We now derive the two technical lemmas stated above. This subsection is essential for our analysis, but the reader may wish to skip it at first reading, as the arguments here are not required elsewhere in the article.

*Proof of Lemma 3.3.* We will write $z = E + i\varepsilon$ and assume without loss that $\varepsilon < 1$. Thus $1 + |E - E_0| \sim 1 + |z - E_0|$. Before applying Prop. 3.2 and the associated eq. (3.13), we must introduce Hilbert-Schmidt operators to the left and the right of $U_\alpha (H - z)^{-1} U_\beta$. The procedure we use to insert these operators is somewhat involved, so let us first outline the argument:

(1) We replace $U_\alpha,\ U_\beta$ by the upper bounds $b_+\Theta^2,\ b_+\Psi^2$,

$$\left\| U_\alpha \frac{1}{H-z} U_\beta \right\| \le b_+^2 \left\| \Theta^2 \frac{1}{H-z} \Psi^2 \right\| , \tag{3.26}$$

where $\Theta$, $\Psi$ denote smoothed characteristic functions for the balls $B_\alpha^r$, $B_\beta^r$ respectively.

(2) The smoothness of $\Theta$ allows us to prove an *identity*

$$\Theta^2 \frac{1}{H-z} = B + T\widetilde{\Theta}^2 \frac{1}{H-z} , \tag{3.27}$$

where $\widetilde{\Theta}$ is a "fattened" version of $\Theta$ and $T$ is Hilbert-Schmidt with HS-norm which is uniformly bounded with respect to the disorder. The operator $B$ is norm bounded (uniformly with respect to the disorder).

(3) We repeat this procedure to the right of the resolvent to obtain

$$\Theta^2 \frac{1}{H-z} \Psi^2 = T_\alpha \widetilde{\Theta}^2 \frac{1}{H-z} \widetilde{\Psi}^2 T_\beta + B' , \tag{3.28}$$

where $B'$ is norm bounded.

(4) We introduce the partition of unity $\sum U_\alpha/F$ between $\widetilde{\Theta}^2$, $\widetilde{\Psi}^2$ and the resolvent, and then apply Prop. 3.2 to each term in the resulting sum. (The actual argument is complicated somewhat by the fact that $T_\alpha$, $T_\beta$ carry some dependence on the randomness.)

We now turn to specifics. A good deal rests on the proof of eq. (3.27). This identity is a consequence of the following: Let $\Theta_j$ be any sequence of smooth functions such that $\Theta_j$ takes value 1 on the support of $\Theta_{j-1}$. Then for each $n \ge 1$

$$\Theta_1^2 \frac{1}{H-z} \;=\; T_n\, \Theta_{n+1}^2 \frac{1}{H-z} \;+\; B_n \tag{3.29}$$

where

$$T_n \;=\; A_1 \cdots A_n \tag{3.30}$$

with

$$A_j \;=\; \Theta_j \frac{1}{H_0+a} \left( \Theta_j (a+z-\lambda V_\omega) \;-\; \left[\Theta_j, H_0\right]\right) , \tag{3.31}$$

and

$$B_n \;=\; \Theta_1 \frac{1}{H_0+a}\Theta_1 \;+\; T_1 \Theta_2 \frac{1}{H_0+a}\Theta_2 \;+\; \cdots \;+\; T_{n-1}\Theta_n \frac{1}{H_0+a}\Theta_n . \tag{3.32}$$

Here $a := 1 - E_0$, so $H_0 + a \ge 1$.

To verify (3.29), use induction on $n$ with the induction step provided by the following "commutator argument:"

$$\begin{aligned}\Theta_j^2 \frac{1}{H-z} &= \Theta_j \frac{1}{H_0+a}(H_0+a)\Theta_j \frac{1}{H-z} \\ &= \Theta_j \frac{1}{H_0+a}\Theta_j(H_0+a)\frac{1}{H-z} - \Theta_j \frac{1}{H_0+a}\left[\Theta_j, H_0\right]\frac{1}{H-z} \\ &= \Theta_j \frac{1}{H_0+a}\Theta_j \\ &\qquad + \Theta_j \frac{1}{H_0+a}\left(\Theta_j(a+z-\lambda V_\omega) - \left[\Theta_j, H_0\right]\right)\frac{1}{H-z}\,. \end{aligned} \tag{3.33}$$

The crucial point here is that

$$\Theta_j(a+z-\lambda V_\omega) - \left[\Theta_j, H_0\right] = \left(\Theta_j(a+z-\lambda V_\omega) - \left[\Theta_j, H_0\right]\right)\Theta_{j+1}^2\,, \tag{3.34}$$

since $\Theta_{j+1}$ is identically one throughout the support of $\Theta_j$ and $\Theta_j(-a+z-\lambda V_\omega) - \left[\Theta_j, H_0\right]$ is a differential – hence local – operator.

The representation provided by eq. (3.29) has two key features:

(1) The norm of $B_n$ is bounded uniformly in $\omega$ and locally uniformly in energy.
(2) $T_n$ is in the Schatten class $\mathcal{I}_p$ (see Remark (2) in Appendix A) for arbitrary $p > d/n$ with $\|\cdot\|_p$-norm bounded uniformly in $\omega$ and locally uniformly in energy.

To see this we first observe that $A_j \in \mathcal{I}_p$ for $p > d$ with uniform $\|\cdot\|_p$-bounds. This follows since $\theta_j(H_0+a)^{-1} \in \mathcal{I}_p$ for $p > d/2$, $\theta_j(H_0+a)^{-1/2} \in \mathcal{I}_p$ for $p > d$ and $(H_0+a)^{-1/2}[\theta_j, H_0] = (H_0+a)^{-1/2}(-2\mathrm{i}\boldsymbol{D_A}\cdot(\nabla\theta_j) - (\Delta\theta_j))$ is bounded. Thus (2) follows from the Hölder property of Schatten classes and (1) from

$$B_n = B_{n-1} + T_{n-1}\theta_n \frac{1}{H_0+a}\theta_n. \tag{3.35}$$

Once $n > d/2$, property (2) above implies that the operator $T_n$ is Hilbert-Schmidt. Henceforth, we fix $n$ to be the least integer greater than $d/2$ – in any case $n \le (d+2)/2$. This provides the representation eq. (3.27) with $\widetilde{\Theta} = \Theta_n$.

There is much flexibility in the definition of $\Theta_j$ which, in turn, affects the norm estimates for $T_n$ and $B_n$. For our purposes, it is sufficient to let $\Theta_j$ be supported in $\{q : \mathrm{dist}(q,\alpha) \le r + [j/n]r\}$ so that $\widetilde{\Theta} = \Theta_n$ is supported in $\{q : \mathrm{dist}(q,\alpha) \le 2r\}$. The specific choice of $\Theta_j$ is not so important. We note, however, that the choice may be made so that

$$\left\|\nabla\Theta_j\right\|_\infty \le \mathrm{O}(d), \ \left\|\Delta\Theta_j\right\|_\infty \le \mathrm{O}(d^2)\,, \tag{3.36}$$

since the gradient of $\Theta_j$ is supported on a set of width $r/n \approx 2r/d$.

Having fixed the sequence $\Theta_j$ we obtain a Hilbert-Schmidt operator $T_n$. More precisely, $T_n$ is a *random* Hilbert-Schmidt operator since $A_j$ depend on the random potential through the terms $\Theta_j V_\omega$. However, $T_n$ is a product of $n$ terms each of which is linear in $\Theta_j V_\omega$. Since

$$\Theta_j V_\omega = \sum_{\zeta \in \mathcal{I}_\alpha} \eta_\zeta \Theta_j U_\zeta \ , \tag{3.37}$$

we conclude that $T_n$ is a polynomial of degree $n$ in the variables $\{\eta_\zeta : \zeta \in \mathcal{I}_\alpha\}$ with Hilbert-Schmidt valued coefficients. Recall that $\mathcal{I}_\alpha$ is the set of lattice sites within distance $3r$ of $\alpha$.

We now repeat this procedure to the right of the resolvent to obtain the two sided representation:

$$\Theta \frac{1}{H-z} \Psi \ = \ B \ + \ T_\alpha \widetilde{\Theta}^2 \frac{1}{H-z} \widetilde{\Psi}^2 T_\beta \ . \tag{3.38}$$

The above discussion shows that we may obtain the following properties for the terms in this identity:

(1) There is $C_d < \infty$, which depends on the parameters of the model and the choice of $\Theta_j$ (but not on the coupling variables $\eta_\zeta$), such that

$$\begin{aligned} \|B\| \ &\le C_d \ (1+\lambda+|z-E_0|)^{2n} \\ &\le C_d \ (1+\lambda)^{2n}(1+|z-E_0|)^{2n} \ . \end{aligned} \tag{3.39}$$

(2) Each $T_\sharp$ is a polynomial of degree $n \le (d+2)/2$ in the variables $\eta_\zeta$ for $\zeta \in \mathcal{I}_\sharp$ with coefficients which are (non-random) Hilbert-Schmidt operators:

$$T_\sharp \ = \ \sum_{(k)} T_\sharp^{(k)} \lambda^{|k|} \prod_{\zeta \in \mathcal{I}_\sharp} \eta_\zeta^{k_\zeta} \ , \tag{3.40}$$

where the summation is over multi-indices $(k) \in \mathbb{N}^{\mathcal{I}_\sharp}$ with $|k| := \sum_\zeta k_\zeta \le n$.

(3) There is $\widetilde{C}_d < \infty$ such that

$$\left\| T_\sharp^{(k)} \right\|_{HS} \ \le \ \widetilde{C}_d (1+|z-E_0|)^n \ , \tag{3.41}$$

for $\sharp = \alpha, \beta$ and each $(k) \in \mathbb{N}^{\mathcal{I}_\sharp}$ with $|k| \le n$.

(4) $\widetilde{\Theta}$ and $\widetilde{\Psi}$ are bounded by one and supported in $\{q : \mathrm{dist}(q,\alpha) < 2r\}$ and $\{q : \mathrm{dist}(q,\beta) < 2r\}$ respectively.

With this representation in hand we insert the partition of unity $\sum U_\zeta / F$ between each factor $\widetilde{\Theta}$, $\widetilde{\Psi}$ and the resolvent $(H-z)^{-1}$. Upon taking norms and applying the triangle inequality this yields

$$\left\| U_\alpha \frac{1}{H-z} U_\beta \right\| \le \|B\| + \sum_{\substack{\zeta\in\mathcal{I}_\alpha \\ \zeta'\in\mathcal{I}_\beta}} \sum_{\substack{(k)\in\mathbb{N}^{\mathcal{I}_\alpha} \\ (l)\in\mathbb{N}^{\mathcal{I}_\beta}}} \lambda^{|k|+|l|} \left\| T_\alpha^{(k)} \widetilde{\Theta}^2 \frac{U_\zeta}{F} \frac{1}{H-z} \frac{U_{\zeta'}}{F} \widetilde{\Psi}^2 T_\beta^{(l)} \right\|, \tag{3.42}$$

where we have used that $\left|\eta_\zeta\right| \le 1$.

Consider now the probability, conditioned on $\mathcal{F}^c_{\alpha,\beta}$, that $\left\| U_\alpha \frac{1}{H-z} U_\beta \right\| > t$. This may be bounded from above by the probability that *one* of the terms on the right hand side of eq. (3.42) is greater than $t/M$, where $M$ is the number of terms. In turn this may be bounded by the sum of the individual probabilities:

$$\begin{aligned} \operatorname{Prob}\left( \left\| U_\alpha \frac{1}{H-z} U_\beta \right\| > t \,\middle|\, \mathcal{F}^c_{\alpha,\beta} \right) \;&\le\; \operatorname{Prob}\left( \|B\| > t/M \,|\, \mathcal{F}^c_{\alpha,\beta} \right) \\ + \sum_{\substack{\zeta\in\mathcal{I}_\alpha \\ \zeta'\in\mathcal{I}_\beta}} \sum_{\substack{(k)\in\mathbb{N}^{\mathcal{I}_\alpha} \\ (l)\in\mathbb{N}^{\mathcal{I}_\beta}}} &\operatorname{Prob}\left( \lambda^{|k|+|l|} \left\| T_\alpha^{(k)} \widetilde{\Theta} \frac{U_\zeta}{F} \frac{1}{H-z} \frac{U_{\zeta'}}{F} \widetilde{\Psi} T_\beta^{(l)} \right\| > t/M \,\middle|\, \mathcal{F}^c_{\alpha,\beta} \right) . \end{aligned} \tag{3.43}$$

Applying Prop. 3.2 – via eq. (3.13) – and the bound on $\|T_\sharp^{(\cdot)}\|_{HS}$ provided by eq. (3.41), we see that each term of the summation is bounded:

$$\begin{aligned} &\operatorname{Prob}\left( \lambda^{|k|+|l|} \left\| T_\alpha^{(k)} \widetilde{\Theta} \frac{U_\zeta}{F} \frac{1}{H-z} \frac{U_{\zeta'}}{F} \widetilde{\Psi} T_\beta^{(l)} \right\| > t/M \,\middle|\, \mathcal{F}^c_{\alpha,\beta} \right) \\ &\qquad\qquad \le\; 2\, C_W\, b_+\, \widetilde{C}_d^2 (1+|z-E_0|)^{2n} D^2 \frac{(1+\lambda)^{2n} M}{\lambda t} , \end{aligned} \tag{3.44}$$

while by eq. (3.39)

$$\operatorname{Prob}\left( \|B\| > t/M \,|\, \mathcal{F}^c_{\alpha,\beta} \right) \;\le\; C_d\, (1+\lambda)^{2n} (1+|z-E_0|)^{2n} \frac{M}{t} . \tag{3.45}$$

The factor $D^2$ in eq. (3.44) is an upper bound for the joint density of $\eta_\zeta$ and $\eta_{\zeta'}$.

Putting this all together we find that

$$\begin{aligned} &\operatorname{Prob}\left( \left\| U_\alpha \frac{1}{H-z} U_\beta \right\| > t \,\middle|\, \mathcal{F}^c_{\alpha,\beta} \right) \\ &\le M \left( C_d \;+\; (M-1) \times \left[ 2\, C_W\, \widetilde{C}_d^2\, b_+ \right] \times \frac{D^2}{\lambda} \right) \\ &\qquad\qquad \times (1+\lambda)^{2n}\, (1+|z-E_0|)^{2n}\, \frac{1}{t} \end{aligned} \tag{3.46}$$

$$\leq M \left(C_d + (M-1) \times \left[2\, C_W\, \widetilde{C}_d^2\, b_+\right]\right) \\ \times (1+\lambda)^{2n}\, (1+|z-E_0|)^{2n}\, (1+\lambda^{-1})\, \frac{D^2}{t}\,,$$

since $D \geq 1$. By $\mathcal{A}2$ the number of terms $M$ is bounded independent of $\alpha, \beta$. This completes the proof of Lemma 3.3. □

*Proof of Lemma 3.4.* This proof follows rather closely that of Lemma 3.3. The main new ingredients are (1) controlling the (un-bounded) factor $(a + H_0)^{1/2}$ and (2) the "re-sampling."

The unbounded factor is controlled by a one-step "commutator argument," similar to that used to generate the Hilbert-Schmidt operators in the proof of Lemma 3.3. The key is the following identity

$$\chi_x \frac{1}{H-z} \Theta_\beta (a+H_0)^{1/2} \\ = \chi_x \frac{1}{H-z}\, \widetilde{\Theta}_\beta^2\, \left((z+a-\lambda V_\omega)\Theta_\beta + \left[\Theta_\beta, H_0\right]\right) \frac{1}{(a+H_0)^{1/2}}\,, \tag{3.47}$$

where we have used that $\chi_x \Theta_\beta = 0$. Here $\widetilde{\Theta}_\beta$ may be any function which is one throughout the support of $\Theta_\beta$. As a result,

$$\left\| \chi_x \frac{1}{H-z} \Theta_\beta (a+H_0)^{1/2} \right\| \;\leq\; c\,(1+\lambda)(1+|z-E_0|)\, \left\| \chi_x \frac{1}{H-z}\, \widetilde{\Theta}_\beta^2 \right\|\,, \tag{3.48}$$

with $c < \infty$ (depending on the size of $\nabla \Theta_\beta$). We choose $\widetilde{\Theta}_\beta$ to have support in the ball of radius $5r/3$ centered at $\beta$.

Thus it suffices to prove eq. (3.21) with $\Theta_\beta (a + H_0)^{1/2}$ replaced by $\widetilde{\Theta}_\beta^2$ provided the power of $(1+|z - E_0|)$ on the right hand side and the permitted power growth of $\widetilde{C}_\lambda$ are each reduced by one, i.e., we must show

$$\mathrm{Prob}\left( \left\| \chi_x \frac{1}{H^{(\Omega)}-z} \widetilde{\Theta}_\beta^2 \right\| > t \,\middle|\, \mathcal{F}^c_{\beta,\gamma} \right) \\ \leq\; C(1+\lambda)^{d+3}\, (1+|z-E_0|)^{(d+2)}\, \left\| \chi_x \frac{1}{\widehat{H}^{(\Omega)}-z} \mathbf{1}_{S_{\beta,\gamma}} \right\|\, \frac{D^2}{t}\,. \tag{3.49}$$

We want to average over $\zeta \in \mathcal{I}_{\beta,\gamma}$, using the Birman Schwinger identity together with the weak $L^1$ inequality, and compare the result to the "re-sampled" operator:

$$\chi_x \frac{1}{H-z} \widetilde{\Theta}_\beta^2 \;=\; \chi_x \frac{1}{\widehat{H}-z} \widetilde{\Theta}_\beta^2 \;+\; \chi_x \frac{1}{\widehat{H}-z} (\widehat{H} \;-\; H) \frac{1}{H-z} \widetilde{\Theta}_\beta^2\,, \tag{3.50}$$

where

$$\widehat{H} \;=\; H \;+\; \lambda \sum_{\zeta\in\mathcal{I}_{\beta,\gamma}} (\widehat{\eta}_\zeta - \eta_\zeta) U_\zeta(q)\,. \tag{3.51}$$

We use the construction presented in the proof of Lemma 3.3 – see eq. (3.29) – to introduce the necessary Hilbert-Schmidt operators.

We begin by noting that

$$\chi_x \frac{1}{\widehat{H}-z}(\widehat{H}-H) \;=\; \chi_x \frac{1}{\widehat{H}-z}\Psi^2(\widehat{H}-H) \tag{3.52}$$

where $\Psi$ is identically one throughout $\{q : \mathrm{dist}(q, \{\beta, \gamma\}) < 4r\}$ and smooth. Thus the commutator argument in the proof of Lemma 3.3 yields the identity

$$\chi_x \frac{1}{\widehat{H}-z}(\widehat{H}-H) \;=\; \chi_x \frac{1}{\widehat{H}-z}\widetilde{\Psi}^2\widehat{T}(\widehat{H}-H) \tag{3.53}$$

where $\widetilde{\Psi}$ is one on the support of $\Psi$ and supported in the set $\{q : \mathrm{dist}(q, \{\beta, \gamma\}) \le 5r\}$, and $\widehat{T}$ is Hilbert-Schmidt with uniformly bounded norm: $\|\widehat{T}\|_2 \le \widetilde{C}_d(1+\lambda)^n(1+|z-E_0|)^n$ ($n$ the smallest integer greater than $d/2$). The bounded term, "$B$", drops out of this representation because $\chi_x\widetilde{\Psi}^2 = 0$ and multipliers of support no larger than that of $\widetilde{\Psi}$ appear on the left of the terms of $B$. Note that $\widehat{T}$ is independent of $\{\eta_\zeta : \zeta \in \mathcal{I}_{\beta,\gamma}\}$ since it is constructed from $\widehat{H}$.

Next, we apply the commutator argument to the left of $(H-z)^{-1}$ to obtain

$$\frac{1}{H-z}\widetilde{\Theta}_\beta^2 \;=\; \widehat{\Theta}_\beta B \;+\; \frac{1}{H-z}\widehat{\Theta}_\beta^2 T \tag{3.54}$$

with $B$ bounded, $\|B\| \le C_d(1+\lambda+|z-E_0|)^{2n}$, and $T$ a polynomial in $\{\eta_\zeta : \zeta \in \mathcal{I}_\beta\}$,

$$T \;=\; \sum_{\substack{(k)\in\mathbb{N}^{\mathcal{I}_\beta}\\ |k|\le n}} \lambda^{|k|}T^{(k)} \prod_{\zeta\in\mathcal{I}_\beta} \eta_\zeta^{k_\zeta}\ , \tag{3.55}$$

with Hilbert-Schmidt coefficients,

$$\|T^{(k)}\|_{HS} \le \widetilde{C}_d(1+|z-E_0|)^n\ . \tag{3.56}$$

Here the function $\widehat{\Theta}_\beta$ is bounded by one and supported in $\{q : \mathrm{dist}(q, \beta) < 2r\}$. Note that $U_\zeta\widehat{\Theta}_\beta = 0$ for $\zeta \notin \mathcal{I}_\beta$.

Putting this all together, we obtain

$$\left\| \chi_x \frac{1}{H-z} \widetilde{\Theta}_\beta^2 \right\| \le \left\| \chi_x \frac{1}{\widehat{H}-z} \widetilde{\Psi} \right\| \times \Bigg( 1 + \left\| (\widehat{H} - H) \widehat{\Theta}_\beta B \right\| \\ + \sum_{(k) \in \mathbb{N}^{\mathcal{I}_\beta}} \lambda^{|k|} \left\| \widetilde{\Psi}\, \widehat{T}\, (\widehat{H} - H) \frac{1}{H-z} \widehat{\Theta}_\beta^2 T^{(k)} \right\| \Bigg) . \quad (3.57)$$

Thus,

$$\left\| \chi_x \frac{1}{H-z} \widetilde{\Theta}_\beta^2 \right\| \le \left\| \chi_x \frac{1}{\widehat{H}-z} \widetilde{\Psi} \right\| \times \Bigg( 1 \;+\; 2\, b_+ \lambda \,\|B\| \;+ \\ + \sum_{\substack{\zeta \in \mathcal{I}_{\beta,\gamma}, \zeta' \in \mathcal{I}_\beta \\ (k) \in \mathbb{N}^{\mathcal{I}_\beta}}} 2\lambda^{1+|k|} \left\| \widetilde{\Psi}\, \widehat{T} U_\zeta \frac{1}{H-z} \frac{U_{\zeta'}}{F} \widehat{\Theta}_\beta^2 T^{(k)} \right\| \Bigg) . \quad (3.58)$$

Based on (3.58), using Prop. 3.2 and the arguments from the proof of Lemma 3.3, it follows that

$$\begin{aligned} \mathrm{Prob} & \left( \left\| \chi_x \frac{1}{H-z} \widetilde{\Theta}_\beta^2 \right\| > t \,\middle|\, \mathcal{F}^c_{\beta,\gamma} \right) \\ & \le \left\| \chi_x \frac{1}{\widehat{H}-z} \widetilde{\Psi} \right\| \times \big( 1 + 2 b_+ C_d \lambda (1+\lambda)^{2n} (1 + |z - E_0|)^{2n} \\ & \qquad + \; (M-2)\, 4 C_W \widetilde{C}_d^2 (1+\lambda)^{2n} (1 + |z - E_0|)^{2n}\, b_+\, D^2 \big) \, \frac{M}{t} \\ & \le \left\| \chi_x \frac{1}{\widehat{H}-z} \widetilde{\Psi} \right\| \times M \big( 1 \;+\; 2 b_+ C_d (M-2)\, 4 C_W \widetilde{C}_d^2 \big) \\ & \qquad \times (1+\lambda)^{2n+1} (1 + |z - E_0|)^{2n} \frac{D^2}{t} \,, \end{aligned} \quad (3.59)$$

where $M$ is the number of terms between the brackets in (3.58). By $\mathcal{A}2$, $M$ is bounded uniformly in $x$, $\beta$, $\gamma$. This completes the proof of Lemma 3.4. □

## 4. Finite-volume criteria

**4.1. The finite-volume inequality.** With the bounds provided by Lemmas 3.3 and 3.4, the proof of Theorem 1.2 proceeds according to a set of arguments familiar from the fractional moment method for discrete operators [6], and related to the multi-scale analysis of random Schrödinger operators, e.g. [74, 18], as well as the analysis of a number of lattice models in statistical mechanics, e.g. [37, 66, 52, 7, 5, 26]. We begin by proving a "correlation inequality" – eq. (4.2) below – and then iterate this inequality to obtain exponential decay of the bulk Green-function.

**Lemma 4.1.** *Let $H$ be a random Schrödinger operator with disorder parameter $\lambda > 0$ which obeys $\mathcal{A}$ and $\mathcal{IAD}$. Then, for each $s < 1/3$ and $z \in \mathbb{C} \setminus \mathbb{R}$ there exists $\widetilde{C}_{\lambda,s,z} < \infty$ such that if we define, for $L > 23r$,*

$$a(x; L) := L^{d-1} \sum_{\zeta \in \mathcal{S}_{x,L}} \mathbb{E}\left( \left\| \chi_x \frac{1}{H^{(B_x^L)} - z} \chi_\zeta \right\|^s \right) \tag{4.1}$$

*where $\mathcal{S}_{x,L} = \{\zeta \in \mathcal{I} : L - 23r < |\zeta - x| < L - 3r\}$, then for any region $\Omega \supset B_x^L$,*

$$\mathbb{E}\left( \left\| \chi_x \frac{1}{H^{(\Omega)} - z} \chi_y \right\|^s \right) \leq \widetilde{C}_{\lambda,s,z}\, a(x; L) \sum_{\zeta \in \mathcal{S}'_{x,L}} \mathbb{E}\left( \left\| \chi_\zeta \frac{1}{H^{(\Omega)} - z} \chi_y \right\|^s \right) \tag{4.2}$$

*for any $y$ with $\operatorname{dist}(x, y) > L + r_0 + 23r$, where $\mathcal{S}'_{x,L} = \{\zeta \in \mathcal{I} : L + r_0 - 13r < |x - \zeta| < L + r_0 + 23r\}$ (recall that $r_0$ is the length scale appearing in $\mathcal{IAD}$), with*

$$\widetilde{C}_{\lambda,s,z} \leq \frac{\text{Const.}}{1 - 3s} (1 + \lambda^{-1})^{2s} (1 + \lambda)^{5s(d+4)} (1 + |z - E_0|)^{5s(d+2)} D^{10s}\ . \tag{4.3}$$

*If $H$ satisfies the stronger condition $\mathcal{A}3'$ then*

$$\widetilde{C}_{\lambda,s,z} \ \leq\ \frac{\text{Const.}}{1 - 3s} (1 + \lambda^{-1})^{2s} (1 + |z - E_0|)^{5s(d+2)} D^{10s}\ . \tag{4.4}$$

The main tools in the proof are the moment bounds presented above and a well known analog of the resolvent expansion commonly used for discrete Schrödinger operators, sometimes called the *geometric resolvent identity*.

**Lemma 4.2.** *Let $H$ be a Schrödinger operator. Consider a sequence of three open sets $\Lambda_0 \subset \Lambda \subset \Omega$ with $\operatorname{dist}(\Lambda_0, \Lambda^c) > 0$ and let $\Theta$ be a smooth function which is identically $1$ in a neighborhood of $\Lambda_0$ and identically zero in a neighborhood of $\Lambda^c$. Given any restrictions $H^{(\Omega)}$ and $H^{(\Lambda)}$ of $H$ to $\Omega$ and $\Lambda$ respectively,*

$$\mathbf{1}_{\Lambda_0} \frac{1}{H^{(\Omega)} - z} = \mathbf{1}_{\Lambda_0} \frac{1}{H^{(\Lambda)} - z} \Theta + \mathbf{1}_{\Lambda_0} \frac{1}{H^{(\Lambda)} - z} [H, \Theta] \frac{1}{H^{(\Omega)} - z} \tag{4.5}$$

*for any $z$ at which both resolvents exist.*

*If furthermore, $\Lambda'_0 \subset \Lambda' \subset \Omega$ with $\operatorname{dist}(\Lambda'_0, \Lambda'^c) > 0$ and $\operatorname{dist}(\Lambda', \Lambda) > 0$, and $\Theta'$ is a function which is identically $1$ in a neighborhood of $\Lambda'_0$ and identically $0$ in a neighborhood of $\Lambda'^c$ then*

$$\begin{aligned} &\mathbf{1}_{\Lambda_0} \frac{1}{H^{(\Omega)} - z} \mathbf{1}_{\Lambda'_0} \\ &\quad = -\mathbf{1}_{\Lambda_0} \frac{1}{H^{(\Lambda)} - z} [H, \Theta] \frac{1}{H^{(\Omega)} - z} \left[H, \Theta'\right] \frac{1}{H^{(\Lambda')} - z} \mathbf{1}_{\Lambda'_0}\ , \end{aligned} \tag{4.6}$$

*with $H^{(\Lambda')}$ any restriction of $H$ to $\Lambda'$ and $z$ such that all three inverses exist.*

*Proof.* The second identity, eq. (4.6), is a consequence of the first, eq. (4.5), and its transpose (eq. (4.8) below). A number of terms drop out because $\Lambda \cap \Lambda' = \emptyset$. To verify eq. (4.5) use the identity

$$[H, \Theta] = (H^{(\Lambda)} - z)\Theta - \Theta(H^{(\Omega)} - z) \tag{4.7}$$

on $\mathcal{D}(H^{(\Omega)})$, which follows from the fact that $H^{(\Lambda)} f = H^{(\Omega)} f$ if the support of $f$ is strictly contained in $\Lambda$. Multiplying on the left by $\mathbf{1}_{\Lambda_0}(H^{(\Lambda)} - z)^{-1}$ and on the right by $(H^{(\Omega)} - z)^{-1}$ yields eq. (4.5). □

*Remarks:*

(1) The identity,

$$\frac{1}{H^{(\Omega)} - z}\mathbf{1}_{\Lambda_0} \; = \; \Theta\frac{1}{H^{(\Lambda)} - z}\mathbf{1}_{\Lambda_0} \; - \; \frac{1}{H^{(\Omega)} - z}\,[H, \Theta]\,\frac{1}{H^{(\Lambda)} - z}\mathbf{1}_{\Lambda_0} \; , \tag{4.8}$$

follows from the transpose of eq. (4.5) (at conjugate $z$).

(2) Eq. (4.6) holds in a number of other contexts. In particular, it is true for discrete Schrödinger operators and in that case gives the usual geometric resolvent expansion (see [6, eq. (2.16)]).

*Proof of Lemma 4.1.* We shall assume $\mathcal{A}3$. For the extension of the argument to couplings which satisfy $\mathcal{A}3'$ see Prop. 4.4 below.

We start by using the geometric resolvent identity eq. (4.6) of Lemma 4.2 with the sets

$$\begin{aligned} \Lambda_0 \; &= \; B_x^{L-11r} \; , & \Lambda \; &= \; B_x^{L} \; , \\ \Lambda_0' \; &= \; \Omega \setminus B_x^{L+r_0+12r} \; , & \Lambda' \; &= \; \Omega \setminus B_x^{L+r_0} \; , \end{aligned} \tag{4.9}$$

and $\nabla\Theta$, $\nabla\Theta'$ supported in $\{q : L - 11r < |q - x| < L - 10r\}$, $\{q : L + r_0 + 11r < |q - x| < L + r_0 + 12r\}$ respectively.

The operators $[H, \Theta^\sharp]$ are *local* and supported in the set where $\nabla\Theta$ is non-zero. A particular consequence of this observation is that

$$\begin{aligned} [H, \Theta] &= \sum_{\zeta_1 \in \mathcal{S}_0} \sum_{\substack{\zeta_2 \\ U_{\zeta_2}\Theta_{\zeta_1} \neq 0}} \Theta_{\zeta_1}\,[H, \Theta]\,\frac{U_{\zeta_2}}{F} \\ \left[H, \Theta'\right] &= \sum_{\zeta_1' \in \mathcal{S}_0'} \sum_{\substack{\zeta_2' \\ U_{\zeta_2'}\Theta_{\zeta_1'} \neq 0}} \frac{U_{\zeta_2'}}{F}\left[H, \Theta'\right]\Theta_{\zeta_1'} \; , \end{aligned} \tag{4.10}$$

with $\mathcal{S}_0 = \{q \in \mathcal{I} : L - 13r < |q - x| < L - 8r\}$ and $\mathcal{S}_0' = \{q \in \mathcal{I} : L + r_0 + 9r < |q - x| < L + r_0 + 14r\}$. Here $\Theta_\zeta$ is the smooth partition of unity used in Lemma 3.4.

Recall that

$$\left\| \frac{1}{(a+H_0)^{1/2}} \left[H, \Theta^\sharp\right] \right\| \le \left( \left\| \Delta\Theta^\sharp \right\|_\infty + \sqrt{2} \left\| \nabla\Theta^\sharp \right\|_\infty \right), \tag{4.11}$$

which is bounded for smooth $\Theta$. We may choose $\Theta$ so that the right hand side is bounded uniformly in $L$, being no larger than (const.) $d^2/r^2$.

With these observations, we obtain from eq. (4.6) the inequality:

$$\begin{aligned} &\left\| \chi_x \frac{1}{H^{(\Omega)} - z} \chi_y \right\| \\ &\quad \le \text{(const.)} \sum_{\substack{\zeta_1 \in \mathcal{S}_0 \\ \zeta_2 : |\zeta_2 - \zeta_1| < 3r}} \sum_{\substack{\zeta_1' \in \mathcal{S}_0' \\ \zeta_2' : |\zeta_2' - \zeta_1'| < 3r}} \left\| \chi_x \frac{1}{H^{(B_x^L)} - z} \Theta_{\zeta_1} (a+H_0)^{1/2} \right\| \\ &\qquad \times \left\| U_{\zeta_2} \frac{1}{H^{(\Omega)} - z} U_{\zeta_2'} \right\| \\ &\qquad \times \left\| (a+H_0)^{1/2} \Theta_{\zeta_1'} \frac{1}{H^{(\Omega \setminus B_x^{L+r_0})} - z} \chi_y \right\| . \end{aligned} \tag{4.12}$$

We now raise eq. (4.12) to the power $s < 1/3$ and take expectation values, using the inequality $(\sum a_n)^s \le \sum a_n^s$. Consider each term on the right-hand side separately: first estimate the expectation conditioned on $\mathcal{F}^c_{\zeta_1,\zeta_2,\zeta_1',\zeta_2'}$, using the Hölder inequality to separate factors. The central factor may be estimated with Lemma 3.3 (see eq. (3.18)), to yield

$$\begin{aligned} &\mathbb{E}\left( \left\| \chi_x \frac{1}{H^{(\Omega)} - z} \chi_y \right\|^s \right) \le \text{(const.)} \frac{C_\lambda^s}{(1-3s)^{1/3}} (1 + |z - E_0|)^{s(d+2)} D^{2s} \\ &\quad \times \sum_{\substack{\langle \zeta_1, \zeta_2 \rangle \\ \langle \zeta_1', \zeta_2' \rangle}} \mathbb{E}\Bigg[ \mathbb{E}\left( \left\| \chi_x \frac{1}{H^{(B_x^L)} - z} \Theta_{\zeta_1} (a+H_0)^{1/2} \right\|^{3s} \middle| \mathcal{F}^c_{\zeta_1,\zeta_2} \right)^{1/3} \\ &\quad \times \mathbb{E}\left( \left\| (a+H_0)^{1/2} \Theta_{\zeta_1'} \frac{1}{H^{(\Omega \setminus B_x^{L+r_0})} - z} \chi_y \right\|^{3s} \middle| \mathcal{F}^c_{\zeta_1',\zeta_2'} \right)^{1/3} \Bigg] . \end{aligned} \tag{4.13}$$

Here we have noted that $H^{(B_x^L)}$ ($H^{(\Omega \setminus B_x^{L+r_0})}$) does not depend on the variables $\eta_\alpha$ with $\alpha \in \mathcal{I}_{\zeta_1',\zeta_2'}$ ($\alpha \in \mathcal{I}_{\zeta_1,\zeta_2}$).

The remaining two factors in each term are i) independent by $\mathcal{I}AD$ and ii) of the correct form to be estimated using Lemma 3.4. Thus,

$$
\begin{aligned}
&\mathbb{E}\left(\left\| \chi_x \frac{1}{H^{(\Omega)} - z} \chi_y \right\|^s\right) \\
&\quad \le \ (\text{const.}) \left[ \frac{C_\lambda^s}{(1-3s)^{1/3}} (1 + |z - E_0|)^{s(d+2)} \ D^{2s} \right] \\
&\qquad \times \left[ \frac{\widetilde{C}_\lambda^s}{(1-3s)^{1/3}} (1 + |z - E_0|)^{s(d+3)} D^{2s} \right]^2 \\
&\qquad\quad \times \sum_{\substack{\langle \zeta_1, \zeta_2 \rangle \\ \langle \zeta_1', \zeta_2' \rangle}} \mathbb{E}\left(\left\| \chi_x \frac{1}{\widehat{H}^{(B_x^L)}_{\zeta_1, \zeta_2} - z} \mathbf{1}_{S_{\zeta_1, \zeta_2}} \right\|^s\right) \\
&\qquad\qquad \times \mathbb{E}\left(\left\| \mathbf{1}_{S_{\zeta_1', \zeta_2'}} \frac{1}{\widehat{H}^{(\Omega \setminus B_x^{L+r_0})}_{\zeta_1', \zeta_2'} - z} \chi_y \right\|^s\right) .
\end{aligned}
\tag{4.14}
$$

Here $\widehat{H}_{\zeta_1,\zeta_2}$ ($\widehat{H}_{\zeta_1',\zeta_2'}$) denotes a version of $H$ which is re-sampled over $\zeta \in \mathcal{I}_{\zeta_1,\zeta_2}$ ($\zeta \in \mathcal{I}_{\zeta_1',\zeta_2'}$) in the sense of eq. (3.23). The assumptions $L > 23r$ and $\operatorname{dist}(x, y) > L + r_0 + 23r$ in Lemma 4.1 are used to satisfy the distance requirements of Lemma 3.4.

We now pick the re-sampled variables with the same distribution as the $\{\eta_\zeta\}$, and include averaging over these variables in the expectations. Since $S_{\zeta_1,\zeta_2} \subset B^{8r}_{\zeta_1}$, $S_{\zeta_1',\zeta_2'} \subset B^{8r}_{\zeta_1'}$ and for each $\zeta_1$, $\zeta_1'$ there is only a fixed finite number of values for $\zeta_2$, $\zeta_2'$ (by $\mathcal{A}2$), we infer from eq. (4.14), after adjusting the constant, that

$$
\begin{aligned}
\mathbb{E}\left(\left\| \chi_x \frac{1}{H^{(\Omega)} - z} \chi_y \right\|^s\right) &\le \ (\text{const.}) \ \frac{C_\lambda^s \widetilde{C}_\lambda^{2s}}{1 - 3s} (1 + |z - E_0|)^{3s(d+3)} \ D^{6s} \\
&\quad \times \sum_{\substack{\zeta_1 \in S_0 \\ \zeta_1' \in S_0'}} \mathbb{E}\left(\left\| \chi_x \frac{1}{H^{(B_x^L)} - z} \mathbf{1}_{B^{8r}_{\zeta_1}} \right\|^s\right) \\
&\qquad \times \mathbb{E}\left(\left\| \mathbf{1}_{B^{8r}_{\zeta_1'}} \frac{1}{H^{(\Omega \setminus B_x^{L+r_0})} - z} \chi_y \right\|^s\right) .
\end{aligned}
\tag{4.15}
$$

To complete the proof of eq. (4.2) we need a bound for the resolvent of $H^{(\Omega \setminus B_x^{L+r_0})}$ in terms of the resolvent of $H^{(\Omega)}$. For this we will apply eq. (4.5) of Lemma 4.2 with

$$
\Lambda = \Omega \setminus B_x^{L+r_0}, \qquad \Lambda_0 = \Omega \setminus B_x^{L+r_0+r} \tag{4.16}
$$

and a smooth function $\Theta''$ such that $\nabla \Theta''$ is supported in $\{q : L + r_0 < |q - x| < L + r_0 + r\}$. Note that $\Theta'' \chi_y = \chi_y$ and $B^{8r}_{\zeta_1'} \subset \Lambda_0$ for every

$\zeta_1' \in \mathcal{S}_0'$. In place of (4.12) we obtain

$$\left\| \mathbf{1}_{B^{8r}_{\zeta_1'}} \frac{1}{H^{(\Omega \setminus B_x^{L+r_0})} - z} \chi_y \right\| \leq \left\| \mathbf{1}_{B^{8r}_{\zeta_1'}} \frac{1}{H^{(\Omega)} - z} \chi_y \right\| + (\text{const.}) \sum_{\substack{\zeta_1'' \in \mathcal{S}_0'' \\ \zeta_2'' : |\zeta_2'' - \zeta_1''| < 3r}} \left\| \mathbf{1}_{B^{8r}_{\zeta_1'}} \frac{1}{H^{(\Omega \setminus B_x^{L+r_0})} - z} U_{\zeta_2''} \right\| \times \left\| (a + H_0)^{1/2} \Theta_{\zeta_1''} \frac{1}{H^{(\Omega)} - z} \chi_y \right\| , \tag{4.17}$$

where $\mathcal{S}_0'' := \{q \in \mathcal{I} : \ L + r_0 - 2r < |q - x| < L + r_0 + 3r\}$.

This yields the expectation bound

$$\mathbb{E}\left( \left\| \mathbf{1}_{B^{8r}_{\zeta_1'}} \frac{1}{H^{(\Omega \setminus B_x^{L+r_0})} - z} \chi_y \right\|^s \right) \leq \mathbb{E}\left( \left\| \mathbf{1}_{B^{8r}_{\zeta_1'}} \frac{1}{H^{(\Omega)} - z} \chi_y \right\|^s \right) + (\text{const.}) \sum_{\langle \zeta_1'', \zeta_2'' \rangle} \sum_{\substack{\zeta \in \mathcal{I} \\ |\zeta - \zeta_1'| < 9r}} \mathbb{E}\left( \left\| U_\zeta \frac{1}{H^{(\Omega \setminus B_x^{L+r_0})} - z} U_{\zeta_2''} \right\|^s \times \left\| (a + H_0)^{1/2} \Theta_{\zeta_1''} \frac{1}{H^{(\Omega)} - z} \chi_y \right\|^s \right) . \tag{4.18}$$

For the terms in the sum of eq. (4.18) we first take the expectation conditioned on $\mathcal{F}^c_{\zeta_1'', \zeta_2'', \zeta}$ and use the Cauchy-Schwarz inequality on the product. The resulting term

$$\mathbb{E}\left( \left\| U_\zeta \frac{1}{H^{(\Omega \setminus B_x^{L+r_0})} - z} U_{\zeta_2''} \right\|^{2s} \middle| \mathcal{F}^c_{\zeta_1'', \zeta_2'', \zeta} \right) , \tag{4.19}$$

is treated by eq. (3.18), while

$$\mathbb{E}\left( \left\| (a + H_0)^{1/2} \Theta_{\zeta_1''} \frac{1}{H^{(\Omega)} - z} \chi_y \right\|^{2s} \middle| \mathcal{F}^c_{\zeta_1'', \zeta_2'', \zeta} \right) \tag{4.20}$$

can be estimated through a straightforward variant of Lemma 3.4, with re-sampling done over all variables in $\mathcal{I}_{\zeta_1'', \zeta_2'', \zeta}$.

Averaging over the re-sampled variables yields

$$\mathbb{E}\left( \left\| \mathbf{1}_{B^{8r}_{\zeta_1'}} \frac{1}{H^{(\Omega \setminus B_x^{L+r_0})} - z} \chi_y \right\|^s \right) \leq \mathbb{E}\left( \left\| \mathbf{1}_{B^{8r}_{\zeta_1'}} \frac{1}{H^{(\Omega)} - z} \chi_y \right\|^s \right) + (\text{const.}) \frac{C_\lambda^s \widetilde{C}_\lambda^s}{1 - 2s} (1 + |z - E_0|)^{2s(d+3)} D^{4s} \times \sum_{\langle \zeta_1'', \zeta_2'', \zeta \rangle} \mathbb{E}\left( \left\| \mathbf{1}_{S_{\zeta_1'', \zeta_2'', \zeta}} \frac{1}{H^{(\Omega)} - z} \chi_y \right\|^s \right) . \tag{4.21}$$

After inserting this into eq. (4.15) we arrive at eq. (4.2) by covering all the sets $B^{8r}_{\zeta_1}$, $B^{8r}_{\zeta'_1}$ and $S_{\zeta''_1,\zeta''_2,\zeta}$ with balls $B^r_\alpha$, $\alpha \in \mathcal{I}$. The factor $L^{d-1}$ appears in this calculation since by $\mathcal{A}2$ the sets $S'_0$ and $S''_0$ contain $CL^{d-1}$ points in $\mathcal{I}$. This completes the proof of eq. (4.2).

The resulting bound for the coefficient $\widetilde{C}_{s,\lambda}$ is

$$\widetilde{C}_{s,\lambda,z} \;\le\; (\text{const.})\frac{C_\lambda^{2s}\widetilde{C}_\lambda^{3s}}{(1-2s)(1-3s)}(1+|z-E_0|)^{5s(d+2)}D^{10s}\;, \tag{4.22}$$

and the growth bounds for $C_\lambda$ and $\widetilde{C}_\lambda$ provided by Lemmas 3.3 and 3.4 easily yield eq. (4.3). For the proof of eq. (4.4), we refer to Prop. 4.4 in the following section. □

**4.2. Large disorder and blow-up regularity.** In this section we consider the large disorder regime ($\lambda > 1$) and discuss the use of assumption $\mathcal{A}3'$ to improve the bounds provided by Lems. 3.3 and 4.1 – i.e., eq. (3.18) and eq. (4.2). The basic idea is to apply Lemmas 3.3 and 3.4 to fluctuations of $\lambda\eta_\zeta$ which are of order one.

It is instructive first to consider independent identically distributed random couplings $\eta_\zeta$. Given $\lambda > 1$, we decompose the interval [0, 1] as a union of a finite number of intervals of length less than $1/\lambda$ with disjoint interiors, $I_j = [a_j, a_{j+1}]$, and consider $\eta_\zeta$ as a super-position of its "integer" and "fractional" parts with respect to this decomposition:

$$\lambda\eta_\zeta \;=\; \lambda a_\zeta \;+\; f_\zeta\;. \tag{4.23}$$

Here the random variable $a_\zeta$ takes value $a_j$ when $\eta_\zeta$ falls in the interval $I_j$, and $f_\zeta = \lambda(\eta_\zeta - a_\zeta)$ is a random variable which takes values in the interval [0, 1]. The conditional distribution of $f = f_\alpha$, at a specified value of $a_\alpha$, is given by the following expression:

$$\frac{\rho(f/\lambda + a_\alpha)}{\lambda\int_{I_j}\rho(\eta)\mathrm{d}\eta}\,\mathrm{d}f\;, \tag{4.24}$$

where $\rho$ is the common density for $\eta_\zeta$. The denominator is just the probability that $\eta_\alpha$ falls in the interval $I_j$.

To apply Lemmas 3.3 and 3.4 to conditional averages with respect to $f_\zeta$, we would need to use the following value for $D$:

$$D \;= D(\lambda) = \;\sup_j\,\underset{f\in I_j}{\text{ess-sup}}\,\frac{\rho(f/\lambda + a_j)}{\lambda\int_{I_j}\rho(\eta)\mathrm{d}\eta}\;. \tag{4.25}$$

For fixed $\lambda > 1$, this is certainly finite. However, $D(\lambda)$ depends on $\lambda$ as well as the choice of $I_j$. Blow-up regularity of the distribution $\rho$, which is guaranteed by assumption $\mathcal{A}3'$, is precisely the requirement that $I_j = I_j(\lambda)$ may be chosen so that $D(\lambda)$ remains bounded for large $\lambda$.

The situation for general couplings $\{\eta_\zeta\}$ obeying $\mathcal{A}3'$ is somewhat more complicated since we must admit *random* decompositions of the interval $[0, 1]$. Specifically, $\mathcal{A}3'$ guarantees that the conditional distributions $\rho_\alpha(\eta_\zeta|\omega)$ of $\eta_\alpha$ at specified values of the remaining couplings are almost-surely blow-up regular with uniformly bounded blow-up norm: $D_{\rho_\alpha(\cdot|\omega)} \leq D$. Thus for each $\alpha$ we may choose $I_j^\alpha = I_j^\alpha(\omega) = [a_j^\alpha(\omega), a_{j+1}^\alpha(\omega)]$ which are measurable with respect to $\mathcal{F}_\alpha^c$, satisfy $|I_j^\alpha| \leq 1/\lambda$, and such that

$$\sup_\alpha \operatorname*{ess-sup}_\omega \max_j \operatorname*{ess-sup}_{f \in I_j^\alpha(\omega)} \frac{\rho_\alpha(f/\lambda + a_j^\alpha(\omega)|\omega)}{\lambda \int_{I_j^\alpha(\omega)} \rho_\alpha(\eta|\omega) \mathrm{d}\eta} \leq D\,. \tag{4.26}$$

We now define $a_\zeta$ and $f_\zeta$ as above: $a_\zeta = a_j^\alpha(\omega)$ if $\eta_\zeta$ falls in $I_j^\alpha(\omega)$ and $f_\zeta = \lambda\eta_\zeta - \lambda a_\zeta$. The density of the distribution of $f = f_\alpha$ conditioned on $a_\alpha$ and $\mathcal{F}_\alpha^c$ is

$$\frac{\rho_\alpha(f/\lambda + a_\alpha|\omega)}{\lambda \int_{I_j^\alpha(\omega)} \rho_\alpha(\eta|\omega) \mathrm{d}\eta}\,, \tag{4.27}$$

which is a density supported in $[0, 1]$ and bounded by $D$.

With these notions, it is an easy exercise in conditional expectations to prove the following extension of Lemma 3.3:

**Proposition 4.3.** *Let $H$ be a random Schrödinger operator which satisfies $\mathcal{A}1$, $\mathcal{A}2$, and $\mathcal{A}3'$. Then eq. (3.16) of Lemma 3.3 holds with a coefficient $C_\lambda$ which satisfies $\sup_{\lambda\geq 1} C_\lambda < \infty$, provided we use for $D$ the constant appearing in $\mathcal{A}3'$ – i.e, the blow-up norm of the distribution.*

Turning now to Lemma 3.4, we note that the disorder strength plays a role in the re-sampling procedure in addition to the averaging. Hence, it is most natural to use the lemma as stated and re-sample only the variables $f_\zeta$ with the $a_\zeta$ fixed. An argument along these lines will be used to complete the proof of eq. (4.4) from Lemma 4.1:

**Proposition 4.4.** *Let $H$ be a random Schrödinger operator which satisfies $\widetilde{\mathcal{I}AD}$, $\mathcal{A}1$, $\mathcal{A}2$, and $\mathcal{A}3'$. Then eq. (4.2) of Lemma 4.1 holds with a coefficient $\widetilde{C}_{s,\lambda}$ which obeys eq. (4.4).*

*Proof.* The proof follows closely the derivation of eq. (4.2) given above with a few modifications which we now indicate.

Let $\mathcal{A}$ denote the sigma-algebra generated by $a_\zeta$, $\zeta \in \mathcal{I}$. Conditioning on $\mathcal{A}$, we obtain a random Schrödinger operator with $\lambda = 1$ and couplings $f_\zeta, \zeta \in \mathcal{I}$ which obey $\mathcal{I}AD$ and $\mathcal{A}3$. Thus, following the proof of Lemma 4.1 we obtain the following analog of eq. (4.14):

$$\mathbb{E}\left(\left\|\chi_x \frac{1}{H^{(\Omega)}-z}\chi_y\right\|^s \middle| \mathcal{A}\right) \leq \text{(const.)}\ \frac{1}{1-3s}(1+|z-E_0|)^{3s(d+3)}\ D^{6s}$$
$$\times \sum_{\substack{\zeta_1 \in S_0 \\ \zeta_1' \in S_0'}} \mathbb{E}\left(\left\|\chi_x \frac{1}{H^{(B_x^L)}-z}\mathbf{1}_{B^{8r}_{\zeta_1}}\right\|^s \middle| \mathcal{A}\right)$$
$$\times\, \mathbb{E}\left(\left\|\mathbf{1}_{B^{8r}_{\zeta_1'}} \frac{1}{H^{(\Omega\setminus B_x^{L+r_0})}-z}\chi_y\right\|^s \middle| \mathcal{A}\right)\ . \tag{4.28}$$

We now average with respect to $\mathcal{A}$, noting that $H^{(B_x^L)}$ and $H^{(\Omega\setminus B_x^{L+r_0})}$ are independent by $\mathcal{I}AD$, and this independence is inherited by conditional averages over $\mathcal{A}$. The resulting bound is identical to eq. (4.14) with $\lambda = 1$. To complete the proof, use a similar argument to prove the analog of eq. (4.21), first conditioning on $\mathcal{A}$ to obtain coefficients with $\lambda = 1$. $\square$

**4.3. The criteria – proof of Theorem 1.2.** Let

$$M(s,\lambda,E+\mathrm{i}\delta)\ :=\ \sup_{0\le\varepsilon\le\delta} \widetilde{C}_{\lambda,s,E+\mathrm{i}\varepsilon}\ \sup_{\substack{\alpha\in\mathcal{I}\\ L>23r}} \frac{\#S'_{\alpha,L}}{L^{d-1}}\ , \tag{4.29}$$

with the sets $\mathcal{S}'_{\alpha,L}$ as in Lemma 4.1, and suppose eq. (1.18) holds for some $s \in (0,1/3)$ and $E \in \mathbb{R}$.

For each open set $\Omega \subset \mathbb{R}^d$ and $\delta > 0$, we define

$$G^\delta_\Omega(\alpha,\beta) := \sup_{|\varepsilon|<\delta} \mathbb{E}\left(\left\|\chi_\alpha \frac{1}{H^{(\Omega)}-E-\mathrm{i}\varepsilon}\chi_\beta\right\|^s\right) \tag{4.30}$$

with $\alpha,\beta \in \mathcal{I}\cap\Omega$. Whenever $\alpha,\beta \in \mathcal{I}\cap\Omega$ with $B^L_\alpha \subset \Omega$ and $|\alpha-\beta| > L + r_0 + 23r$, Lemma 4.1 implies that

$$G^\delta_\Omega(\alpha,\beta)\ \le\ \mathrm{e}^{-\gamma(\delta)} \frac{1}{\#S'_{\alpha,L}} \sum_{\zeta\in S'_{\alpha,L}} G^\delta_\Omega(\zeta,\beta)\ , \tag{4.31}$$

with

$$\mathrm{e}^{-\gamma(\delta)}\ :=\ M(s,\lambda,E+\mathrm{i}\delta) \sup_{\varepsilon<\delta}\sup_{\alpha\in\mathcal{I}} \mathbb{E}\left(\left\|\chi_\alpha\,(H^{(B^L_\alpha)}-E-i\varepsilon)^{-1}\,\mathbf{1}_{\delta B^L_\alpha}\right\|^s\right)\ . \tag{4.32}$$

Note that $\gamma(\delta)$ is an increasing function of $\delta$ and the assumed finite volume bound eq. (1.18) shows that $\lim_{\delta\to 0}\gamma(\delta) > 0$. In particular, there is $\delta_0 > 0$ such that $\gamma(\delta) > 0$ for $\delta < \delta_0$.

Let us define,

$$F^\delta_\Omega(\nu)\ :=\ \sup_{\alpha,\beta\in\mathcal{I}\cap\Omega} \mathrm{e}^{\nu\,\mathrm{dist}_\Omega(\alpha,\beta)/2L} G^\delta_\Omega(\alpha,\beta)\ . \tag{4.33}$$

The proof will proceed as follows. First, for $\nu < \gamma(\delta)$ and bounded $\Omega$, we derive an $\Omega$ independent bound on $F^\delta_\Omega(\nu)$. Second, unbounded $\Omega$ are handled via finite volume approximations. Finally, we pass to the limits $\delta \to 0$ and $\nu \to \gamma$.

Now, consider a bounded region $\Omega$ and $\alpha, \beta \in \mathcal{I} \cap \Omega$ with $\text{dist}_\Omega(\alpha, \beta) > 2L$. Observe that either eq. (4.31) or its conjugate holds for this pair, since $|\alpha - \beta| > 2L > L + r_0 + 23r$ and one of $B^L_\alpha$ or $B^L_\beta$ is contained entirely in $\Omega$. If, say $B^L_\alpha \subset \Omega$, then eq. (4.31) implies

$$\begin{aligned} e^{\nu \,\text{dist}_\Omega(\alpha,\beta)/2L} G^\delta_\Omega(\alpha, \beta) &\le e^{-\gamma(\delta)} \frac{1}{\# S'_{\alpha,L}} \sum_{\zeta \in S'_{\alpha,L}} e^{\nu \,\text{dist}_\Omega(\alpha,\beta)/2L} G^\delta_\Omega(\zeta, \beta) \\ &\le e^{\nu} e^{-\gamma(\delta)} \frac{1}{\# S'_{\alpha,L}} \sum_{\zeta \in S'_{\alpha,L}} e^{\nu \,\text{dist}_\Omega(\zeta,\beta)/2L} G^\delta_\Omega(\zeta, \beta)\ , \end{aligned} \tag{4.34}$$

where we have used the triangle inequality for $\text{dist}_\Omega$ and observed that $\text{dist}_\Omega(\alpha, \zeta) \le 2L$ for $\zeta \in S'_{\alpha,L}$. If instead $B^L_\beta \subset \Omega$, the conjugate version of eq. (4.31) implies the conjugate version of this bound. Either way, the end result is that

$$e^{\nu \,\text{dist}_\Omega(\alpha,\beta)/2L} G^\delta_\Omega(\alpha, \beta) \;\le\; e^{\nu} e^{-\gamma(\delta)} F^\delta_\Omega(\nu)\ . \tag{4.35}$$

If $\nu < \gamma(\delta)$, eq. (4.35) implies that $F^\delta_\Omega(\nu)$ may be found by restricting the supremum in eq. (4.33) to "nearby" pairs (here we use that $F^\delta_\Omega(\nu)$ is finite due to boundedness of $\Omega$):

$$F^\delta_\Omega(\nu) = \sup_{\alpha,\beta:\text{dist}_\Omega(\alpha,\beta)<2L} e^{\nu \,\text{dist}_\Omega(\alpha,\beta)/2L} G^\delta_\Omega(\alpha, \beta) \le e^{\nu} A(s, \lambda, E)\ , \tag{4.36}$$

where $A(s, \lambda, E) < \infty$ is the *a priori* bound on fractional moments provided by Lemma 3.3, so

$$A(s, \lambda, E) \;\le\; \frac{C^s_\lambda}{1 - s}\, (1 + |E - E_0|)^{s(d+2)}\, D^{2s} \tag{4.37}$$

by eq. (3.18).

To complete the proof, we must extend eq. (4.36) to unbounded regions. For this purpose, fix an open set $\Omega \subset \mathbb{R}_d$ and let $\Omega_j = [-j, j]^d \cap \Omega$. One may verify, for example using the geometric resolvent identity eq. (4.5) and the Combes-Thomas estimate [20], that for any $\alpha, \beta \in \mathcal{I} \cap \Omega$ and $z \in \mathbb{C} \setminus \mathbb{R}$

$$\lim_{j\to\infty} \left\| \chi_\alpha \frac{1}{H^{(\Omega_j)} - z} \chi_\beta - \chi_\alpha \frac{1}{H^{(\Omega)} - z} \chi_\beta \right\| \;=\; 0\ . \tag{4.38}$$

Thus

$$\sup_{|\varepsilon|\le\delta} \left\| \chi_\alpha \frac{1}{H^{(\Omega)} - E - i\varepsilon} \chi_\beta \right\| \;\le\; \liminf_{j\to\infty} \sup_{|\varepsilon|\le\delta} \left\| \chi_\alpha \frac{1}{H^{(\Omega_j)} - E - i\varepsilon} \chi_\beta \right\|\ , \tag{4.39}$$

and by Fatou's Lemma,

$$G^{\delta}_{\Omega}(\alpha, \beta) \ \leq \ \liminf_{j\to\infty} G^{\delta}_{\Omega_j}(\alpha, \beta) \ . \tag{4.40}$$

Since $\mathrm{dist}_{\Omega_j}(\alpha, \beta) \to \mathrm{dist}_{\Omega}(\alpha, \beta)$, this implies eq. (4.36) holds for $\Omega$ when $\nu < \gamma(\delta)$.

In the limit $\delta \to 0$, eq. (4.36) implies that

$$\limsup_{\varepsilon\to 0} \mathbb{E}\left( \left\| \chi_\alpha \frac{1}{H^{(\Omega)} - E - \mathrm{i}\varepsilon} \chi_\beta \right\|^s \right) \ \leq \ \mathrm{e}^{\nu} A(s, \lambda, E) \mathrm{e}^{-\nu \, \mathrm{dist}_{\Omega}(\alpha,\beta)/2L} \tag{4.41}$$

for any $\nu < \gamma$. Finally, taking $\nu \to \gamma$, we obtain eq. (1.19). □

## 5. Applications: results for distinct energy ranges

To apply the above results on localization we need to verify the sufficiency criteria of Theorem 1.2 in specific disorder regimes and energy ranges. In this section we present several such results for regimes of interest, including the familiar large disorder and band edge (Lifshitz tail) regimes. The mechanisms which allow us to check eq. (1.18) are essentially those used to verify the initial length scale estimates for a multiscale analysis. As such, we do not attempt to give an exhaustive list of applications here, but rather try to illustrate how known methods may be combined with the arguments developed in this paper.

A useful feature of our results is that it suffices to check eq. (1.18) at a single energy $E \in \mathbb{R}$, since the following continuity result allows us to extend the obtained bounds onto an interval containing $E$ (as well as off the real axis).

**Lemma 5.1.** *Let $H$ satisfy $\mathcal{A}$. Fix $\Lambda$, and subsets $\Lambda_1$ and $\Lambda_2$ of $\Lambda$. Let $G(z) := \mathbf{1}_{\Lambda_1}(H^{(\Lambda)} - z)^{-1}\mathbf{1}_{\Lambda_2}$. Then, for any $0 < s \leq 1/2$, $z \mapsto \mathbb{E}(\|G(z)\|^s)$ is locally Hölder-continuous with exponent $s$ for all $z \in \mathbb{C}$. More precisely, there exists a constant $A_{s,\lambda} < \infty$ such that*

$$\begin{aligned} \left|\mathbb{E}\left(\|G(z)\|^s\right) - \mathbb{E}\left(\|G(w)\|^s\right)\right| &\leq A_{s,\lambda} N_\Lambda N_{\Lambda_1}^{1/2} N_{\Lambda_2}^{1/2} \\ &\quad \times (1 + |z - E_0|)^{s(d+2)} (1 + |w - E_0|)^{s(d+2)} |z - w|^s \end{aligned} \tag{5.1}$$

*for all $z, w \in \mathbb{C}$. Here $N_\Lambda = \#(\mathcal{I} \cap \Lambda)$ is the "volume" defined in the discussion following Lemma 3.3 and*

$$A_{s,\lambda} \ \leq \ \mathrm{const.}\, \frac{(1 + 1/\lambda)^{2s}}{1 - 2s} \times \begin{cases} (1 + \lambda)^{2s(d+2)} & \text{if } \mathcal{A}3 \text{ holds,} \\ 1 & \text{if } \mathcal{A}3' \text{ holds.} \end{cases} \tag{5.2}$$

*Proof.* By the resolvent identity, noting that $\mathbf{1}_\Lambda = \mathbf{1}$ on $L^2(\Lambda)$,

$$G(z) - G(w) = (z-w)\mathbf{1}_{\Lambda_1}(H^{(\Lambda)} - z)^{-1}\mathbf{1}_\Lambda(H^{(\Lambda)} - w)^{-1}\mathbf{1}_{\Lambda_2}\,. \tag{5.3}$$

Thus

$$\begin{aligned}
&\left|\mathbb{E}\left(\|G(z)\|^s\right) - \mathbb{E}\left(\|G(w)\|^s\right)\right| \le \mathbb{E}\left(\|G(z) - G(w)\|^s\right) \\
&\le |z-w|^s\mathbb{E}\left(\left\|\mathbf{1}_{\Lambda_1}\frac{1}{H^\Lambda - z}\mathbf{1}_\Lambda\frac{1}{H^{(\Lambda)} - w}\mathbf{1}_{\Lambda_2}\right\|^s\right) \qquad (5.4)\\
&\le |z-w|^s\mathbb{E}\left(\left\|\mathbf{1}_{\Lambda_1}\frac{1}{H^{(\Lambda)} - z}\mathbf{1}_\Lambda\right\|^{2s}\right)^{1/2}\mathbb{E}\left(\left\|\mathbf{1}_{\Lambda}\frac{1}{H^{(\Lambda)} - w}\mathbf{1}_{\Lambda_2}\right\|^{2s}\right)^{1/2}.
\end{aligned}$$

The result now follows from Lemma 3.3 – see eq. (3.19) – combined with Prop. 4.3. □

A preliminary application of the continuity provided by Lemma 5.1 is the proof of Theorem 1.3.

*Proof.* Due to Lemma 5.1, we need only verify eq. (1.18) for the energy $E$ appearing in eq. (1.20). For this we argue along the lines of the proof of Lemma 4.1, using the geometric resolvent identity eq. (4.5) to obtain for $L > 23r$,

$$\chi_\alpha\frac{1}{H^{(B^L_\alpha)} - z}\mathbf{1}_{\delta B^L_\alpha} = \chi_\alpha\frac{1}{H - z}\mathbf{1}_{\delta B^L_\alpha} + \chi_\alpha\frac{1}{H - z}\left[H, \Theta_L\right]\frac{1}{H^{(B^L_\alpha)} - z}\mathbf{1}_{\delta B^L_\alpha} \tag{5.5}$$

where $\Theta_L$ is any smooth function equal to one on $B^{L-r}_\alpha$ and zero on $\mathbb{R}^d \setminus B^L_\alpha$ – recall that $\delta B^L_\alpha = \{q : r < \mathrm{dist}(q, \partial\Lambda) < 23r\} \subset B^{L-r}_\alpha$. From this we conclude, as in the proof of eq. (4.21), that

$$\mathbb{E}\left(\left\|\chi_\alpha\frac{1}{H^{(B^L_\alpha)} - z}\mathbf{1}_{\delta B^L_\alpha}\right\|^s\right) \le \mathrm{const.}L^{2(d-1)}\sum_{\substack{\zeta \\ \mathrm{dist}(\zeta, \delta B^L_\alpha) < r}}\mathbb{E}\left(\left\|\chi_\alpha\frac{1}{H - z}\chi_\zeta\right\|^s\right). \tag{5.6}$$

By assumption the right hand side is $\mathcal{O}(L^{3(d-1)}\mathrm{e}^{-\mu L})$ uniformly in $\alpha$. Thus eq. (1.18) is satisfied for sufficiently large $L$ and the theorem follows. □

**5.1. The multi-scale analysis regime.** The region in which a multiscale analysis may be carried out provides another characterization of the localization regime. An important observation is that the "output" of the multiscale analysis implies the "input" for Theorem 1.2, i.e., eq. (1.18). Thus for the operators considered here the stronger results proved by our methods hold throughout the multiscale regime. A precise formulation of this statement is the following.

**Theorem 5.1.** *Let $H$ be a random Schrödinger operator which satisfies $\mathcal{A}$ and $\mathcal{I}AD$. Suppose that for some $A < \infty$, $\mu > 0$, $\xi > 2(d-1)$, $C < \infty$ and $E \in \mathbb{R}$ it holds for $L$ sufficiently large that*

$$\sup_{\alpha\in\mathcal{I}} \mathrm{Prob}\left[\left\| \chi_\alpha \frac{1}{H^{(B^L_\alpha)} - E} \mathbf{1}_{\delta B^L_0} \right\| > Ae^{-\mu L}\right] \le CL^{-\xi} . \tag{5.7}$$

*Then, there exist $0 < s < 1/3$, $A' < \infty$, $\mu' > 0$ and an open interval $\mathcal{J}$ containing $E$ such that*

$$\limsup_{\varepsilon\downarrow 0} \mathbb{E}\left(\left\| \chi_x \frac{1}{H^{(\Omega)} - E' - \mathrm{i}\varepsilon} \chi_y \right\|^s\right) \le A' e^{-\mu' \operatorname{dist}_\Omega(x,y)} \tag{5.8}$$

*for all open sets $\Omega \subset \mathbb{R}^d$, $x, y \in \Omega$ and $E' \in \mathcal{J}$.*

Bounds on the probability of exceptional behavior of the type of eq. (5.7) are a characterization of the localized regime mapped by the multi-scale analysis, where $\xi$ can be made arbitrarily large, e.g. [18]. Using a "bootstrap" approach in which the output of one multi-scale analysis serves as the input for another, Germinet and Klein have shown that this probability is $\mathcal{O}(\mathrm{e}^{-L^\alpha})$ with any $\alpha < 1$ throughout the multi-scale regime [32]. *A posteriori*, we conclude from Theorem 1.2 that this bound may be replaced by $\mathcal{O}(\mathrm{e}^{-\nu L})$ with $\nu > 0$ for the operators considered here.

For ergodic random Schrödinger operators (see $\mathcal{A}$4 in Sect. 5), there is a notion of "strong localization region," introduced in [33], analogous to Dobrushin and Shlosman's regime of complete analyticity in statistical mechanics [26]. The strong localization region may be characterized by any of a number of criteria, one of which is the applicability of the multi-scale analysis, and throughout this region all of those criteria hold (see [33, Theorem 4.4]). In the words of the authors of [33], it is a region "possessing every possible virtue we can imagine!" For the random operators considered here, one may add to that list of virtues exponential localization of Green function moments and dynamical localization in the (stronger) sense of Corollary 1.1.

*Proof.* Following the argument for the lattice case from Sect. 4.4 of ref. [6], we define complementary "good" and "bad" subsets of the probability space $\Omega$ by

$$\Omega_G := \left\{\omega \,\middle|\, \left\|\chi_\alpha (H^{(B^L_\alpha)} - E)^{-1} \mathbf{1}_{\delta B^L_\alpha}\right\| \le Ae^{-\mu L}\right\} \tag{5.9}$$

and $\Omega_B := \Omega_G^c$. Then

$$\begin{aligned} \mathbb{E}&\left(\left\|\chi_\alpha (H^{(B^L_\alpha)} - E)^{-1} \mathbf{1}_{\delta B^L_\alpha}\right\|^s\right) \\ &= \mathbb{E}\left(\left\|\chi_\alpha (H^{(B^L_\alpha)} - E)^{-1} \mathbf{1}_{\delta B^L_\alpha}\right\|^s I[\omega \in \Omega_G]\right) \\ &\quad + \mathbb{E}\left(\left\|\chi_\alpha (H^{(B^L_\alpha)} - E)^{-1} \mathbf{1}_{\delta B^L_\alpha}\right\|^s I[\omega \in \Omega_B]\right) . \end{aligned} \tag{5.10}$$

The first term is bounded by $A^s e^{-s\mu L}$, while we may apply the Hölder inequality to bound the second term for any $s < t < 1$ by

$$\mathbb{E}\left(\left\| \chi_\alpha (H^{(B_\alpha^L)} - E)^{-1} \mathbf{1}_{\delta B_\alpha^L} \right\|^t\right)^{s/t} \mathbb{E}\left(I[\omega \in \Omega_B]\right)^{1-s/t} . \tag{5.11}$$

We may further estimate this by $C(s, E, d) L^{(s(d-1)-\xi(t-s))/t}$ using (1) the fractional moment bound eq. (3.19) and (2) the assumed bound eq. (5.7) on $\mathrm{Prob}(\Omega_B)$. Here we have used that $\delta B_0^L$ contains $\mathcal{O}(L^{d-1})$ points in $\mathcal{I}$.

In summary, we get the bound

$$\mathbb{E}\left(\left\| \chi_\alpha (H^{(B_\alpha^L)} - E)^{-1} \mathbf{1}_{\delta B_\alpha^L} \right\|^s\right) \le A^s e^{-s\mu L} + C(t, E, \lambda)^{s/t} L^{(s(d-1)-\xi(t-s))/t} , \tag{5.12}$$

uniformly in $\alpha$. Since $\xi > 2(d-1)$, we can choose $s$ sufficiently close to 0 and $t$ close to 1 to guarantee $(\xi(t-s) - s(d-1))/t > 2(d-1)$. Thus

$$\limsup_{L\to\infty} L^{2(d-1)} \sup_{\alpha \in \mathcal{I}} \mathbb{E}\left(\left\| \chi_\alpha (H^{(B_\alpha^L)} - E)^{-1} \mathbf{1}_{\delta B_\alpha^L} \right\|^s\right) = 0 , \tag{5.13}$$

and we may choose $L$ large enough that eq. (1.18) holds at $E$.

Working in the finite volumes $B_\alpha^L$, we can use the continuity given by Lemma 5.1 and the continuity of $b_s(\lambda, E)$ in $E$ to conclude the existence of a complex neighborhood $\mathcal{U}$ of $E$ such that eq. (1.18) holds for every $E' + i\varepsilon \in \mathcal{U}$. Then Theorem 1.2 applies at all real $E' \in \mathcal{U}$, which concludes the proof by Theorem 5.1. □

**5.2. Large disorder.** Perhaps the easiest localization regime to understand is that induced at the bottom of the spectrum by large disorder. If we fix an energy $E > E_0$ and a length scale $L$ we may adjust $\lambda$ so that for any $x \in \mathbb{R}^d$ it is overwhelmingly likely that $E$ is far below the bottom of the spectrum of the "local Hamiltonian," i.e., $H^{(B_x^L)}$. Heuristically, this suggests that $E$ lies in the localization regime, since the resolvent $(H^{(B_x^L)} - E)^{-1}$ is typically bounded with small norm.

In the discrete setting, the bound $\mathbb{E}(|G(x,x)|^s) \lesssim 1/\lambda^s$ provides the basis for the 'single-site' criterion of [4]. For the operators considered here, by taking $\lambda$ large enough one may directly verify the localization condition eq. (1.18) at any fixed finite scale $L$ allowed in Theorem 1.2. In fact, eq. (1.18) may be satisfied uniformly for all energies in an arbitrary finite interval. For this result, we assume blow-up regularity $\mathcal{A}3'$ to ensure that the constants which appear in eq. (1.18) remain bounded as $\lambda$ increases.

**Theorem 5.2.** *Let $H$ satisfy $\mathcal{I}AD$, $\mathcal{A}1$, $\mathcal{A}2$ and $\mathcal{A}3'$. Then to every $E_1 \in \mathbb{R}$ and $0 < s < 1/3$ there exists $\lambda_0 = \lambda_0(E_1, s)$ such that for every $\lambda > \lambda_0$ there are constants $A < \infty$ and $\mu > 0$ such that for any open set $\Omega \subset \mathbb{R}^d$*

$$\limsup_{\varepsilon \downarrow 0} \mathbb{E}\left(\left\| \chi_x \frac{1}{H^{(\Omega)} - E - i\varepsilon} \chi_y \right\|^s\right) \le A e^{-\mu \operatorname{dist}_\Omega(x,y)} \tag{5.14}$$

*for all $E \in [E_0, E_1]$ and $x, y \in \Omega$.*

*Remark.* The proof will show that we can take $\lambda_0 \propto (1+|E_1-E_0|)^{2+3d/2}$.

*Proof.* We cannot use eq. (3.18) to verify eq. (1.18), since the r.h.s. does not approach 0 as $\lambda \to \infty$, even assuming $\mathcal{A}3'$. Instead we fix any $L > 23r$ and use the method discussed in Sect. 3.1 to show that $\mathbb{E}(\|\chi_\alpha(H^{(B_\alpha^L)} - E)^{-1}\mathbf{1}_{\delta B_\alpha^L}\|^s)$ approaches 0 as $\lambda \to \infty$. This method gives volume dependent bounds on fractional moments, but as $L$ is fixed in the present argument that is of no importance.

By the covering condition in $\mathcal{A}2$ it suffices to show that

$$\mathbb{E}\left(\left\| U_\beta \frac{1}{H^{(B_\alpha^L)} - E} U_\zeta \right\|^s\right) \to 0 \quad \text{as } \lambda \to \infty \tag{5.15}$$

for all $\beta, \zeta \in \mathcal{I}$. Convergence needs to be shown uniformly with respect to $\alpha, \beta, \zeta \in \mathcal{I}$ and $E \in [E_0, E_1]$. One may then switch to complex energy using Lemma 5.1 and complete the proof with Theorem 1.2.

To derive eq. (5.15), we write $H = H^{(B_\alpha^L)}$, understanding henceforth that all operators are restricted to $L^2(B_\alpha^L)$. Let $\widehat{H} = H - \lambda\eta_\beta U_\beta - \lambda\eta_\zeta U_\zeta$ and define $\eta_\pm = \frac{1}{2}(\eta_\beta \pm \eta_\zeta)$, so that $\eta_\beta U_\beta + \eta_\zeta U_\zeta = \eta_+(U_\beta + U_\zeta) + \eta_-(U_\beta - U_\zeta)$. Assumption $\mathcal{A}2$ and the Birman-Schwinger argument yield

$$\begin{aligned} \left\| U_\beta \frac{1}{H-E} U_\zeta \right\| &\leq b_+ \left\| (U_\beta + U_\zeta)^{1/2} \frac{1}{H-E} (U_\beta + U_\zeta)^{1/2} \right\| \\ &= b_+ \left\| \left[ \widehat{K}_{E,\eta_-}^{-1} + \lambda\eta_+ \right]^{-1} \right\| , \end{aligned} \tag{5.16}$$

where $\widehat{K}_{E,\eta_-} = (U_\beta + U_\zeta)^{1/2}(\hat{H} + \lambda\eta_-(U_\beta - U_\zeta) - E)^{-1}(U_\beta + U_\zeta)^{1/2}$ in $L^2((\operatorname{supp} U_\beta \cup \operatorname{supp} U_\zeta) \cap B_\alpha^L)$.

Now $\eta_+ \in [0,\infty)$ and with $\eta_-$ fixed one can use the argument described in Sect. 3.1 to show that

$$\left| \left\{ \eta_+ \in [0,\infty) \ : \ \left\| \left[ \widehat{K}_{E,\eta_-}^{-1} + \lambda\eta_+ \right]^{-1} \right\| > t \right\} \right| \leq \frac{(1 + \xi_{E,\eta_-})}{\lambda t} , \tag{5.17}$$

where $\xi_{E,\eta_-}$ may be expressed as a spectral shift function:

$$\begin{aligned} \xi_{E,\eta_-} \ \leq \ & \operatorname{Tr}\left[ P(\widehat{H} + \lambda\eta_-(U_x - U_y) < E) \right] \\ & - \operatorname{Tr}\left[ P(\widehat{H} + \lambda\eta_-(U_x - U_y) + \lambda(U_x + U_y) < E) \right] . \end{aligned} \tag{5.18}$$

Since both Schrödinger operators appearing in eq. (5.18) are positive perturbations of $H_0$, we may bound each trace by a Weyl-type bound (see eq. (2.8)) to yield

$$\xi_{E,\eta_-} \ \leq \ \operatorname{Tr} P(H_0 < E) \ \leq \ \text{const.}\ (1 + |E - E_0|)^{d/2} L^d . \tag{5.19}$$

Averaging over $\eta_+, \eta_-$ as in the proof of Prop. 3.2 we obtain the following bound:

$$\mathrm{Prob}\left(\left\| U_\beta \frac{1}{H^{(B_\alpha^L)} - E} U_\zeta \right\| > t | \mathcal{F}^c_{\beta,\zeta}\right) \le \mathrm{const.} D^2 L^d \frac{(1 + |E_1 - E_0|)^{d/2}}{\lambda t}\ , \tag{5.20}$$

for any $\alpha, \beta, \zeta \in \mathcal{I}$ and $E \in [E_0, E_1]$, from which eq. (5.15) follows. □

**5.3. Localization via density of states bounds.** To state the bounds in this section, which are based on the notion of the *density of states*, it is necessary to assume that the random operator $H$ is *ergodic*, i.e. its distribution is invariant under a sufficiently large group of translations. In particular this implies that the spectrum of $H$ is a non-random set, c.f. [58, 71]. Hence, throughout we make the additional assumption

$\mathcal{A}4$: $\mathcal{I}$ is a lattice containing the origin, $U_\alpha = U(\cdot - \alpha)$, and $H$ is ergodic with respect to shifts in $\mathcal{I}$.

We will drop the supremum over $\alpha$ in eq. (1.18) and work with $\alpha = 0$, since if $\mathcal{A}4$ holds eq. (1.18) is equivalent to

$$M(s, \lambda, E)(1 + L)^{2(d-1)} \limsup_{\varepsilon \downarrow 0} \mathbb{E}\left(\left\| \chi_0 \frac{1}{H^{(B_0^L)} - E - i\varepsilon} \mathbf{1}_{\delta B_0^L} \right\|^s\right) < 1\ . \tag{5.21}$$

For simplicity, we shall assume that the lattice $\mathcal{I} = \mathbb{Z}^d$ and work throughout with the '$\ell^\infty$-norm' on $\mathbb{R}^d$: $|x| = \max_j |x_j|$. Thus balls are cubes with sides parallel to the co-ordinate axes and unit balls are fundamental cells for the lattice $\mathcal{I}$. The reader should have no problem extending the results below to more general lattices, in which case it is natural to work with the '$\ell^\infty$-norm' induced on $\mathbb{R}^d$ by the decomposition of a vector into components parallel to lattice generators.

The density of states measure for an ergodic random Schrödinger operator is a Borel measure $\kappa$ on the real line defined through a limiting procedure via its action on compactly supported continuous functions:

$$\int f(t)\kappa(dt) := \lim_{L\to\infty} \frac{1}{(2L)^d} \mathrm{Tr}\, f(H^{(\Lambda_L)}) \tag{5.22}$$

where $\Lambda_L = B_0^L = [-L, L]^d$. For the operators considered here, $\mathcal{A}4$ implies that this limit exists almost surely, is non-random [58], and is equal to

$$\int f(t)\kappa(\mathrm{d}t) = \lim_{L\to\infty} \frac{1}{(2L)^d} \mathrm{Tr}\, \mathbb{E}\left(f(H^{(\Lambda_L)})\right) = \frac{1}{|\mathcal{C}|} \mathbb{E}\left(\mathrm{Tr}\, \mathbf{1}_{\mathcal{C}} f(H)\right)\ , \tag{5.23}$$

where $\mathcal{C}$ is any unit cell of the lattice $\mathcal{I}$, e.g., $[0, 1]^d$. In addition, the random operators studied here satisfy a Wegner estimate which shows that the density of states measure is absolutely continuous with a density which is in $L^p_{loc}$ for all finite $p$ [19].

We also introduce the finite volume measures

$$\int_A \kappa_L(\mathrm{d}t) \;=\; \mathbb{E}\left(\operatorname{Tr} P_A(H^{(\Lambda_L)})\right) \,. \tag{5.24}$$

Again $\kappa_L$ is absolutely continuous with a density in $L^p_{loc}$ [19]. Note that $\lim_L \kappa_L = \kappa$ in the sense of weak convergence in the dual space of the compactly supported continuous functions.

If $E$ is a band edge of the almost sure spectrum of $H$, e.g. if $E = \inf \sigma(H)$, then it is generally expected that $\kappa[E, E + \Delta E]$ vanishes to very high order in $\Delta E$. This phenomenon was first observed by I.M. Lifshitz [53] and the regions where such estimates hold are called 'Lifshitz tails'.

In the Lifshitz tail regime one expects it is very rare to find an eigenvalue of a finite volume operator near the band edge $E$. With favorable estimates this suggests localization should hold near $E$ since the Combes-Thomas estimate [20, 11] may typically be used to obtain resolvent decay across finite volumes.

One result in this vein shows that smallness bounds for the finite volume measures $\kappa_L$ may be used to verify the localization criterion in Theorem 1.2 in a similar way as done for initial length scale estimates in proofs of band edge localization via multiscale analysis.

**Theorem 5.3.** *Let $H$ satisfy $\mathcal{I}AD$, $\mathcal{A}1{-}4$. Suppose that for some $\beta \in (0, 2)$, $\xi > 2(d-1)$, $C_1 > 0$, $C_2 > 0$, and $E \in \mathbb{R}$ it holds for sufficiently large $L$ that*

$$\kappa_L\left(\left[E - C_1 L^{-\beta}, E + C_1 L^{-\beta}\right]\right) < C_2 L^{-\xi - d} \,. \tag{5.25}$$

*Then the conclusion of Theorem 5.1 holds.*

*Proof.* The proof is very similar to the proof of Theorem 5.1. This time, define $\Omega_G := \{\omega : \operatorname{dist}(\sigma(H^{(\Lambda_L)}), E) > C_1 L^{-\beta}\}$ and $\Omega_B = \Omega_G^c$. Then, using eq. (5.25)

$$\begin{aligned} \operatorname{Prob}(\Omega_B) &\le \mathbb{E}\left(\operatorname{Tr} P_{[E - C_1 L^{-\beta}, E + C_1 L^{-\beta}]}(H^{(\Lambda_L)})\right) \\ &= (2L)^d \kappa_L\left(\left[E - C_1 L^{-\beta}, E + C_1 L^{-\beta}\right]\right) \le 2^d C_2 L^{-\xi} \,. \end{aligned} \tag{5.26}$$

The improved Combes-Thomas estimate due to ref. [11] implies that there are $\eta > 0$ and $C < \infty$ such that for $\omega \in \Omega_G$,

$$\left\| \chi_0 \frac{1}{H^{(\Lambda_L)} - E} \mathbf{1}_{\delta\Lambda_L} \right\| \le C L^{\beta} e^{-\eta L^{1-\beta/2}} \,. \tag{5.27}$$

This gives a bound for the analogue of the first term on the right hand side of eq. (5.10). From here on the proof is identical to the proof of Theorem 5.1. $\square$

There are two well-known situations in which bounds of the form (5.25) can be derived. One is the aforementioned Lifshitz-tail regime, which generally holds if the background operator $H_0$ is $\mathcal{I}$-periodic and $E$ is the infimum of the almost sure spectrum. In this case one gets eq. (5.25) for arbitrary $\xi > 0$, see e.g. [18,46]. Conditions on the periodic background operator for the appearance of this regime at more general band edges were given by F. Klopp [47].

It is not known if the conditions used in ref. [47] hold for general periodic $H_0$. But one always has the second option of "forcing" a bound like eq. (5.25) to hold by assuming the distribution density $\rho$ of the random coupling constants $\eta_\alpha$ has small tails near the boundary of their support. For example, one gets eq. (5.25) with $\xi \in (0, 2\tau - d)$ at a lower band edge $E_0$ of the almost sure spectrum if $\int_0^h \rho(q)\, \mathrm{d}q = \mathcal{O}(h^\tau)$ for $h \to 0$, see [46, Prop. 4.1]. Thus Theorem 5.3 is applicable if $\tau > (4d-3)/2$.

To complement the above discussion, it is interesting to note that one may prove localization directly from an estimate on the infinite volume density of states.

**Theorem 5.4.** *Let $H$ satisfy $\mathcal{I}AD$, $\mathcal{A}1-4$. For each $\xi > 3d-2$ and $E \in \mathbb{R}$, there exists $C = C(E,\xi,\lambda) > 0$ and $\delta_0 = \delta(E,\xi,\lambda)$ such that if for some $\delta < \delta_0$*

$$\kappa([E-\delta, E+\delta]) \;\le\; C\,\delta^\xi\,, \tag{5.28}$$

*then there exist $s \in (0,1)$, $A < \infty$, and $\mu > 0$ such that for any open set $\Omega \subset \mathbb{R}^d$*

$$\limsup_{\varepsilon \downarrow 0} \mathbb{E}\left( \left\| \chi_x \frac{1}{H^{(\Omega)} - E' - \mathrm{i}\varepsilon} \chi_y \right\|^s \right) \;\le\; A\, \mathrm{e}^{-\mu\, \mathrm{dist}_\Omega(x,y)}\,, \tag{5.29}$$

*for every $E' \in [E - \delta/4, E + \delta/4]$.*

*Proof.* Klopp has shown that the densities of states for periodic approximations to a random Schrödinger operator approach the infinite volume density of states extremely rapidly, c.f. refs. [47,48]. We shall adapt his argument to use eq. (5.28) to bound the probability that a certain finite volume operator with random quasi-periodic boundary conditions has spectrum in the interval $[E-\delta/2, E+\delta/2]$. Together with arguments in the proof of Theorems 5.1 and 5.3 this will imply the theorem via an analogue of Theorem 1.2 in which the smallness criterion is satisfied for a quasi-periodic finite volume Hamiltonian. That result may be proved in exactly the same way as Theorem 1.2 since, as remarked in the introduction, quasi-periodic boundary conditions preserve all the properties implied by $\mathcal{A}1$ which were needed in the proof.

We begin by introducing a sequence of periodic approximations to $H$. For each $\ell \in \mathbb{N}$ greater than $r_0$, define a random potential $V_\ell^P$ which is periodic under translations in $\ell\mathcal{I}$:

$$V_\ell^P(q) \;:=\; \sum_{\alpha\in\mathcal{I}\cap\Lambda_{\ell/2}} \eta_\alpha \sum_{\zeta\in\ell\mathcal{I}} U(q-\alpha-\zeta)\;, \tag{5.30}$$

and note that for any set $\Lambda \subset \mathbb{R}^d$ with diameter less than $\ell - r_0$ the random functions $V_\ell^P(q)\mathbf{1}_\Lambda(q)$ and $V(q)\mathbf{1}_\Lambda(q)$ are identically distributed. We define

$$H_\ell^P \;:=\; H_0 \;+\; V_\ell^P\;. \tag{5.31}$$

Then $H_\ell^P$ is periodic under shifts in $\ell\mathcal{I}$ and its distribution is invariant under shifts in $\mathcal{I}$.

The averaged density of states measure for $H_\ell^P$, denoted $\kappa_\ell^P$, is defined to be

$$\kappa_\ell^P(A) \;:=\; \frac{1}{\ell^d}\mathbb{E}\left(\mathrm{Tr}\,\mathbf{1}_{\Lambda_{\ell/2}} P_A\big(H_\ell^P\big)\right)\;. \tag{5.32}$$

By translation invariance of the distribution of $H_\ell^P$, this is also given by

$$\kappa_\ell^P(A) \;=\; \mathbb{E}\left(\mathrm{Tr}\,\mathbf{1}_{\mathcal{C}} P_A\big(H_\ell^P\big)\right)\;, \tag{5.33}$$

with $\mathcal{C} = [0,1]^d$.

In the present situation, the arguments of ref. [48] may be adapted to show:

*For each* $n > 0$ *there exists* $C_n = C_n(E)$ *such that*

$$\kappa_\ell^P([E-\delta/2, E+\delta/2]) \;\le\; \kappa([E-\delta, E+\delta]) \;+\; C_n\,\ell^{d+1-n}\delta^{-n}\;. \tag{5.34}$$

Since our assumptions differ somewhat from those of ref. [48], let us describe the proof of eq. (5.34). In the following we use $C(E)$ to denote a generic energy and dimension dependent parameter whose value does not depend on the length scale $\ell$ but may change from line to line. Choose a $C^\infty$ function $f$ with $\mathbf{1}_{\mathcal{J}_{\delta/2}} \le f \le \mathbf{1}_{\mathcal{J}_\delta}$ where $\mathcal{J}_t := [E-t, E+t]$. Then

$$\kappa_\ell^P(\mathcal{J}_{\delta/2}) \le \mathbb{E}\left(\mathrm{Tr}\,\mathbf{1}_{\mathcal{C}} f\big(H_\ell^P\big)\right) \le \kappa(\mathcal{J}_\delta) + \mathbb{E}\left(\mathrm{Tr}\,\mathbf{1}_{\mathcal{C}}\big(f\big(H_\ell^P\big) - f(H)\big)\mathbf{1}_{\mathcal{C}}\right)\;. \tag{5.35}$$

Using the Helffer-Sjöstrand formula [38,22], write

$$\begin{aligned}&\mathbf{1}_{\mathcal{C}}\big(f\big(H_\ell^P\big) - f(H)\big)\mathbf{1}_{\mathcal{C}}\\ &\quad = \int_{\mathcal{J}_\delta\times[-1,1]} \big(\partial_{\bar z}\widetilde{f}_n(z)\big)\;\mathbf{1}_{\mathcal{C}}\left(\frac{1}{z-H_\ell^P} - \frac{1}{z-H}\right)\mathbf{1}_{\mathcal{C}}\,\mathrm{d}x\mathrm{d}y\;,\end{aligned} \tag{5.36}$$

where $z = x + \mathrm{i}y$ and $\widetilde{f}_n$ is an "almost-analytic" extension of $f$ which vanishes to order $n$ at $y = 0$:

$$\widetilde{f}_n(x + \mathrm{i}y) = \left[\sum_{j=0}^{n} \frac{1}{j!} f^{(j)}(x)(\mathrm{i}y)^j\right] \sigma(y) . \tag{5.37}$$

Here $\sigma$ is a fixed cut-off function supported in $[-1, 1]$ and identically one in a neighborhood of zero. The key property here is that $\partial_{\bar{z}} \widetilde{f}(x + \mathrm{i}y) = \mathcal{O}(|y|^n/\delta^{n+1})$ as may be verified directly.

The difference of resolvents appearing in eq. (5.36) can be expressed in terms of the geometric resolvent identity eq. (4.5):

$$\mathbf{1}_{\mathcal{C}} \left(\frac{1}{z - H_\ell^P} - \frac{1}{z - H}\right) \mathbf{1}_{\mathcal{C}} = \mathbf{1}_{\mathcal{C}} \frac{1}{z - H_\ell^P} [H, \Theta] \frac{1}{z - H} \mathbf{1}_{\mathcal{C}} \tag{5.38}$$

with $\Theta$ any function which is identically one in a neighborhood of $\mathcal{C}$ and supported inside $\Lambda_{\ell/2-r_0}$. We choose $\Theta$ with $\nabla\Theta$ supported in a strip of width one near the boundary of $\Lambda_{\ell/2-r_0}$.

Using a "commutator argument" similar to that in the proof of Lemmas 3.3 and 3.4, we may show that for $z \in \mathcal{J} \times \{[-1, 0) \cup (0, 1]\}$,

$$[H, \Theta] \frac{1}{z - H} \mathbf{1}_{\mathcal{C}} = T \Psi \frac{1}{z - H} \mathbf{1}_{\mathcal{C}} , \tag{5.39}$$

where $T$ is a trace class operator with

$$\operatorname{Tr} |T| \leq C(E)\, \ell^{d-1} \tag{5.40}$$

and $\Psi$ is a function which is identically one on the support of $\nabla\Theta$ and also supported in a strip of width of order one near the boundary of $\Lambda_{\ell/2-r_0}$. Therefore

$$\begin{aligned} &\left|\operatorname{Tr} \mathbf{1}_{\mathcal{C}} \frac{1}{z - H_\ell^P} [H, \Theta] \frac{1}{z - H} \mathbf{1}_{\mathcal{C}}\right| \\ &\qquad \leq C(E) \ell^{d-1} \left\| \mathbf{1}_{\mathcal{C}} \frac{1}{z - H_\ell^P} \Psi \right\| \left\| \Psi \frac{1}{z - H} \mathbf{1}_{\mathcal{C}} \right\| \\ &\qquad \leq C(E)\, \ell^{d-1} \frac{1}{y^2} \mathrm{e}^{-\mu(E) y \ell} , \end{aligned} \tag{5.41}$$

where we have used the Combes-Thomas bound [20, 11] to estimate the resolvent norms.

Putting this into eq. (5.36) we find that

$$\begin{aligned} \left|\operatorname{Tr} \mathbf{1}_{\mathcal{C}} \big(f\big(H_\ell^P\big) - f(H)\big) \mathbf{1}_{\mathcal{C}}\right| &\leq C(E) \int_{\mathcal{J}_\delta \times [-1,1]} \frac{y^n}{\delta^{n+1}} \ell^{d-1} \frac{1}{y^2} \mathrm{e}^{-\mu(E) y \ell}\, \mathrm{d}x \mathrm{d}y \\ &\leq C(E, n)\, \ell^{d-n} \delta^{-n} , \end{aligned} \tag{5.42}$$

which completes the proof of eq. (5.34).

We now recall some standard facts from the Bloch/Floquet theory of periodic operators, c.f. [60]. For each $\mathbf{k} \in BZ$ with $BZ = [0, 2\pi]^d$ define the restriction of $H_\ell^P$ to $\Lambda_{\ell/2}$ with quasi-periodic boundary conditions at quasi-momentum $\mathbf{k}/\ell$ to be the self-adjoint operator $H_{\ell;\mathbf{k}}$ on $L^2(\Lambda_{\ell/2})$ that agrees with $H$ when applied to functions compactly supported in the interior of $\Lambda_{\ell/2}$ and whose domain includes all functions on $\Lambda_{\ell/2}$ of the form $\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot x/\ell}\phi(x)$ with $\phi(x)$ smooth and periodic.

It is well known that the periodic density of states $\kappa_\ell^P$ may be obtained as an average of the densities for $H_{\ell;\mathbf{k}}$:

$$\kappa_\ell^P(\mathcal{J}_{\delta/2}) \;=\; \frac{1}{(2\pi)^d}\int_{BZ} \kappa_{\ell;\mathbf{k}}(\mathcal{J}_{\delta/2})\,\mathrm{d}\mathbf{k}\;, \tag{5.43}$$

where the quasi-periodic densities $\kappa_{\ell;\mathbf{k}}$ are defined by

$$\kappa_{\ell;\mathbf{k}}(A) \;:=\; \frac{1}{\ell^d}\mathbb{E}\left(\mathrm{Tr}\,P_A(H_{\ell;\mathbf{k}})\right)\;. \tag{5.44}$$

Now consider the probability space $\Omega' = \Omega \times BZ$ with associated measure $\mathrm{Prob}' = \mathrm{Prob} \times \mathrm{d}\mathbf{k}/(2\pi)^d$, and let $H_\ell = H_{\ell;\omega'}$ be a random Schrödinger operator with $\omega' = (\omega, \mathbf{k}) \in \Omega'$ distributed according to $\mathrm{Prob}'$. As in the proof of Theorems 5.1 and 5.3, we define complementary good and bad sets, $\Omega_G := \{\omega' : \mathrm{dist}(\sigma(H_{\ell;\omega'}), E) > \delta/2\}$ and $\Omega_B = \Omega' \setminus \Omega_G$.

Using eq. (5.43) and eq. (5.34) we see that

$$\begin{aligned}\mathrm{Prob}'(\Omega_B) \;\le\; \mathbb{E}'\left(\mathrm{Tr}\,P_{\mathcal{J}_{\delta/2}}(H_{\ell;\omega'})\right) &= \ell^d\kappa_\ell^P(\mathcal{J}_{\delta/2})\\ &\le \ell^d\kappa(\mathcal{J}) \;+\; C_n\ell^{2d-n}\delta^{-n}\;.\end{aligned} \tag{5.45}$$

On the other hand, the improved Combes-Thomas bound [11] shows that there are $A < \infty$ and $\mu > 0$ such that for $\omega' \in \Omega_G$ and $E' \in \mathcal{J}_{\delta/4}$

$$\left\|\chi_0\frac{1}{H_{\ell;\omega'} - E'}\delta\Lambda_\ell\right\| \;\le\; A\,\delta^{-1}\,\mathrm{e}^{-\mu\delta^{1/2}\ell}\;. \tag{5.46}$$

Arguing as in the proof of Theorem 5.1 this implies

$$\begin{aligned}\mathbb{E}'\left(\left\|\chi_0\frac{1}{H_{\ell;\omega'} - E'}\delta\Lambda_\ell\right\|^s\right) \;&\le\; A^s\,\delta^{-s}\,\mathrm{e}^{-s\mu\delta^{1/2}\ell}\\ &+\; C(t,\lambda,E)^{s/t}\ell^{(2+s)(d-1)}\left(\ell^d\kappa(\mathcal{J}) \;+\; C_n\ell^{2d-n}\delta^{-n}\right)^{1-s/t}\;,\end{aligned} \tag{5.47}$$

for any $s < t < 1$. Thus

$$\ell^{2(d-1)}\mathbb{E}'\left(\left\|\chi_0\frac{1}{H_{\ell;\omega'} - E'}\delta\Lambda_\ell\right\|^s\right) \;\le\; C\,\max(A_1, A_2, A_3) \tag{5.48}$$

with

$$\begin{aligned} A_1 &= \delta^{-s}\ell^{2(d-1)}\, \mathrm{e}^{-s\mu\delta^{1/2}\ell} \ , \\ A_2 &= \ell^{(2+s)(d-1)+(1-s/t)d}\kappa(\mathcal{J}_\delta)^{1-s/t} \ , \\ A_3 &= \ell^{(2+s)(d-1)+(1-s/t)(2d-n)}\delta^{-n(1-s/t)} \ . \end{aligned} \tag{5.49}$$

By the analogue of Theorem 1.2 discussed above, in which the Dirichlet operator $H^{(\Lambda_\ell)}$ is replaced by the random quasi-periodic operator $H_\ell$, we see that there is a fixed quantity $B = B(s, \lambda, E)$ such that if $\max_j(A_j) < B$ then the conclusion of the present theorem holds, i.e., the interval $\mathcal{J}_{\delta/4}$ is contained in the localization regime. Note that the infinite volume operator does not depend on the quasi-momentum **k**. Nonetheless, the finite volume quasi-periodic operators may be used because locally, i.e., for functions supported in the interior of the cube $\Lambda_\ell$, they agree with the infinite volume operator.

To obtain a concrete result, we let $\ell \approx \delta^{-r}$ for some $r > 1$. Then $A_1 < B$ and $A_3 < B$ for all sufficiently small $\delta$ provided we choose $n$ sufficiently large. Thus, there is $\delta_0 > 0$ such that for $\delta < \delta_0$ the condition $\max_j(A_j) < B$ is equivalent to requiring that

$$A_2 = \delta^{-\xi(1-s/t)}\kappa(\mathcal{J}_\delta)^{1-s/t} \ < \ B \ , \tag{5.50}$$

where $\xi = rd + r(d-1)(2+s)/(1-s/t)$. For $s = 0$, $t = 1$, and $r = 1$, the expression for $\xi$ reduces to $3d-2$, and hence any value $\xi > 3d-2$ can be attained with some permissible selection of $0 < s < t < 1$ and $r > 1$. This completes the proof of Theorem 5.4. □

## Appendix A. Technical comments

Following are some comments of technical nature, which are intended to supplement the discussion of the assumptions and results stated in the introduction.

(1) *(The operator nature of $H_\omega$)* Under the stated assumptions, $H_\omega$ is essentially self adjoint on $C_0^\infty$ [51]. The random potential $V_\omega(q)$ is non-negative and uniformly bounded by $\lambda b_+$. The operator $H_\omega$ is bounded below, with $\sigma(H_\omega) \subset [E_0, \infty)$.

(2) For $1 \le p < \infty$ let $\mathcal{I}_p$ denote the Schatten class of order $p$, i.e., the ideal of bounded operators $A$ on $L^2(\mathbb{R}^d)$ with $\|A\|_p := (\mathrm{Tr}\,|A|^p)^{1/p} < \infty$. Then for all $z$ in the resolvent set of $H$ and all $f \in L^p(\mathbb{R}^d)$ with $p > d/2$,

$$f(q)(H_\omega - z)^{-1} \in \mathcal{I}_p \ . \tag{A.1}$$

Also, for any $a > -E_0$,

$$f(q)(H_\omega + a)^{-r} \in \mathcal{I}_p \tag{A.2}$$

if $f \in L^p(\mathbb{R}^d)$ and $r > d/2p$. Proofs can be found, for example, in refs. [65, 67]. Under the assumptions $\mathcal{A}2$ and $\mathcal{A}3$, which imply bounds on the random potential $V_\omega(q)$, the Schatten class bounds hold uniformly in $\omega$.

(3) Weak solutions of $H\varphi = z\varphi$, $z \in \mathbb{C}$, have the *unique continuation property*, i.e. if $\varphi$ vanishes on a non-empty open set, then $\varphi$ vanishes identically. This follows from ref. [42] which requires only local form-boundedness of $V_0^2$, $\boldsymbol{A}^2$, and $(\boldsymbol{\nabla} \wedge \boldsymbol{A})^2$ with respect to the Laplacian. There is quite a bit of literature on unique continuation allowing more general $\boldsymbol{A}$ and $V$; c.f., [13, 67, 41, 49, 75, 76].

(4) The above three properties hold also for the restrictions of $H$ to any open set $\Omega$ with Dirichlet boundary conditions and for the restriction to any cube $\Lambda_{x,L} := x + [-L/2, L/2]^d$ with Neumann or quasi-periodic boundary conditions (see Sect. 5.3).

(5) *(The regularization by* i0*)* For *unbounded* regions $\Omega$ one knows that the operator norm limit

$$\chi_x \big( H_\omega^{(\Omega)} - E - \mathrm{i}0 \big)^{-1} \chi_y := \lim_{\varepsilon \downarrow 0} \chi_x \big( H_\omega^{(\Omega)} - E - \mathrm{i}\varepsilon \big)^{-1} \chi_y \tag{A.3}$$

exists almost surely for almost every $E \in \mathbb{R}$. This follows from Fubini's theorem and the fact that, for fixed $\omega$, $\chi_x (H_\omega^{(\Omega)} - E - \mathrm{i}0)^{-1} \chi_y$ exists for almost every $E$. To prove the latter one writes

$$\begin{aligned} \chi_x \big( H_\omega^{(\Omega)} - E - \mathrm{i}\varepsilon \big)^{-1} \chi_y = {} & \chi_x \big( H_\omega^{(\Omega)} - E - \mathrm{i}\varepsilon \big)^{-1} P_I \big( H_\omega^{(\Omega)} \big) \chi_y \\ & + \chi_x \big( H_\omega^{(\Omega)} - E - \mathrm{i}\varepsilon \big)^{-1} P_{\mathbb{R} \setminus I} \big( H_\omega^{(\Omega)} \big) \chi_y \,, \end{aligned} \tag{A.4}$$

where $I$ is a bounded interval and we denote the spectral projection onto a measurable set $M$ for a self-adjoint operator $H$ by $P_M(H)$. If $E$ is in the interior of $I$, then the second term in the sum trivially has a limit. For the first term, a polarization argument shows that it suffices to consider $G_A(E + \mathrm{i}\varepsilon) := \mathbf{1}_A (H_\omega^{(\Omega)} - E - \mathrm{i}\varepsilon)^{-1} P_I (H_\omega^{(\Omega)}) \mathbf{1}_A$ for bounded regions $A$. The operator function $z \mapsto G_A(z)$ defined for $z$ in the upper half plane $\mathbb{C}_+$ is trace class valued (see eq. (2.8)) and analytic with non-negative imaginary part. It follows from a result of de Branges [24], also used in Appendix C, that the limit exists for almost every $E$. To conclude one exhausts $\mathbb{R}$ with bounded intervals $I$.

By eq. (A.3) we can extend the bound (3.16) of Lemma 3.3 to $\mathrm{Prob}(\| \chi_x (H^{(\Omega)} - E - \mathrm{i}0)^{-1} \chi_y \| > t)$ for almost every $E$. However, it is convenient for us to have eq. (3.16) for *all* $E$, which is why we prefer to work with $\varepsilon > 0$.

(6) *(The removal of* i0 *for bounded domains)* If we fix a *bounded* region $\Omega$, then any fixed $E \in \mathbb{R}$ is almost surely not in $\sigma(H^{(\Omega)})$, as follows from the assumptions on the distribution of the random couplings stated in Sect. 1.2

(via analytic perturbation theory and unique continuation of eigenfunctions). Thus, $\|\chi_x(H^{(\Omega)} - E)^{-1}\chi_y\|$ is an almost surely finite random variable, and Fatou's lemma shows that the bound eq. (3.16) on Prob($\|\chi_x(H^{(\Omega)} - E - \mathrm{i}0)^{-1}\chi_y\| > t$) implies also such a bound for $\varepsilon = 0$.

(7) *(Results for large disorder regime)* Readers familiar with fractional-moment methods for discrete random operators will likely note the bound we obtain in Lemma 3.3 is a bit weaker than the *a-priori* bound derived in that context, which falls off like $\lambda^{-1}$ for large $\lambda$. We showed – in the proof of Theorem 5.2 – that the $s$-moments of the resolvent of $H_\omega^{(\Omega)}$ for a bounded region $\Omega$ are $\mathcal{O}(|\Omega|^s/\lambda^s)$. Coupled with Theorem 1.2 this allowed us to conclude localization at "large disorder." However, we do not show here that the fractional-moments of the infinite volume resolvent tend to zero for large $\lambda$ (although this may still be true).

(8) *(Possible coexistence of bulk localization with extended boundary states)* An operator may exhibit localization in the bulk (in terms of transition amplitudes) along with extended *boundary states* occurring in certain geometries. Such situations have been studied and are of particular interest for the Quantum Hall Effect, with $H_o$ the Landau operator [23,28]. Our use of the domain-adapted metric, $\mathrm{dist}_\Omega$ – in which exponential decay is compatible with the above picture – allows the analysis of localization to proceed even in such cases. However, it is also possible to formulate other finite volume criteria which rule out extended surface states. The input conditions need to be more restrictive and involve propagators between boundary regions in arbitrary geometries. For this purpose one may present a modified version of Theorem 1.1, changed in a manner similar to what was done for discrete models in ref. [6, Theorem 1.1]. A key point is that the domain-adapted metric can be replaced in eq. (1.6) by the usual distance, in which case the conclusions – eq. (1.7) and pure point spectrum – hold in any sufficiently regular region and in particular, under the stronger assumptions, rule out also extended boundary states.

(9) *(Energy dependence of the bounds)* We note that Theorem 1.1 only requires exponential decay of the energy-averaged Green function. However, typically this will be established, for example by Theorem 1.2, through a bound which is uniform in energy.

(10) *(Other norms in* (1.7)*)* We have used operator norms in stating (1.7) and its various consequences, but they extend to arbitrary Schatten norms. This follows from the fact that the operators involved are "super-trace class" in the sense that for every $p > 0$ we have $\mathrm{Tr}\left|\chi_x g(H^{(\Omega)}) P_{\mathcal{I}}(H^{(\Omega)})\chi_y\right|^p \le C_p < \infty$ independent of $x$, $y$, $\Omega$, $g$ and the disorder (this follows from remark (2) above). Considering, for example, the trace norm $\|\cdot\|_1$ we find, picking

$p < 1$,

$$\mathbb{E}\left(\sup_g \left\| \chi_x g(H^{(\Omega)}) P_{\mathcal{J}}(H^{(\Omega)}) \chi_y \right\|_1\right)$$
$$\leq C_p \mathbb{E}\left(\sup_g \left\| \chi_x g(H^{(\Omega)}) P_{\mathcal{J}}(H^{(\Omega)}) \chi_y \right\|^{1-p}\right)$$
$$\leq C_p \left[\mathbb{E}\left(\sup_g \left\| \chi_x g(H^{(\Omega)}) P_{\mathcal{J}}(H^{(\Omega)}) \chi_y \right\|\right)\right]^{1-p} \leq \widetilde{A} \mathrm{e}^{-\widetilde{\mu}\, \mathrm{dist}_\Omega(x,y)}\,, \qquad \text{(A.5)}$$

for appropriate constants $\widetilde{A} < \infty$ and $\widetilde{\mu} > 0$.

(11) *(The bound on Fermi projections)* In fact, the hypothesis of Theorem 1.1 – namely eq. (1.6) – implies a result somewhat stronger than eq. (1.7), namely

$$\mathbb{E}\left(\sup_g \| \chi_x g(H^{(\Omega)}) \chi_y \|\right) \leq \widetilde{A} e^{-\widetilde{\mu}\, \mathrm{dist}_\Omega(x,y)}\,, \qquad \text{(A.6)}$$

where the supremum is over all Borel measurable functions $g$, with $|g| \leq 1$ pointwise and constant on $\mathcal{J}_< = \left\{E : E \leq \inf \mathcal{J}\right\}$ and $\mathcal{J}_> = \{E : E \geq \sup \mathcal{J}\}$. The collection of sets consisting of $\mathcal{J}_{<,>}$ and the Borel subsets of $\mathcal{J}$ is a sigma algebra $\Sigma$, and the supremum is taken over the unit ball $B_1(\Sigma)$ of the bounded $\Sigma$ measurable functions. A special case, corresponding to $g(H) = P_{(-\infty, E_F)}(H)$, is the bound (1.10) on Fermi projections, which may also be derived using a contour integral representation of the projection operator as in [3].

To verify eq. (A.6), we fix a $C^\infty$ function $h$, $0 \leq h \leq 1$, supported in $[E_0 - 2, E_+]$, with $E_+ = \sup \mathcal{J}$, and identically equal to 1 on $[E_0 - 1, \inf \mathcal{J}]$. Note that, for $g \in B_1(\Sigma)$, $g(H^{(\Omega)})$ may be decomposed as $\alpha \mathbf{1} + \beta h(H^{(\Omega)}) + \widetilde{g}(H_\omega)$ with $\widetilde{g}$ supported in $\mathcal{J}$. The contribution from $\widetilde{g}$ may be estimated by eq. (1.7), and the contribution from $\alpha\mathbf{1}$ is bounded and zero for $|x - y| \geq 2r$. To estimate the contribution from $h(H^{(\Omega)})$ we write it, using the Helffer-Sjöstrand formula [38], as

$$h(H^{(\Omega)}) \;=\; \frac{1}{2\pi} \int_{\mathbb{C}} F(z) \frac{1}{H^{(\Omega)} - z} \mathrm{d}x \mathrm{d}y\,, \qquad \text{(A.7)}$$

where $z = x + \mathrm{i}y$ and the bounded function $F(x + \mathrm{i}y)$ satisfies $F(x + \mathrm{i}y) = \mathcal{O}(y^n)$ for some $n \geq 1$ ($n = 1$ will do) and is supported in the union of the sets $\mathcal{J} + \mathrm{i}[-1, 1]$, $[E_0 - 2, E_+] + \mathrm{i}[-1, -\frac{1}{2}]$, $[E_0 - 2, E_+] + \mathrm{i}[\frac{1}{2}, 1]$ and $[E_0 - 2, E_0 - 1] + \mathrm{i}[-1, 1]$. ($F(x + \mathrm{i}y) = \partial_{\bar{z}} \widetilde{h}(z)$ with $\widetilde{h}(z)$ an almost analytic extension of $h$; see eq. 5.36.) Using that $\|(H - z)^{-1}\| \leq 1/\mathrm{Im} z$, we get

$$\mathbb{E}\left(\left\| \chi_x h(H^{(\Omega)}) \chi_y \right\|\right) \;\leq\; \frac{1}{2\pi} \int_{\mathbb{C}} \frac{|F(z)|}{|\mathrm{Im} z|^{1-s}} \mathbb{E}\left(\left\| \chi_x \frac{1}{H^{(\Omega)} - z} \chi_y \right\|^s\right)\,. \qquad \text{(A.8)}$$

Since the support of $F$ approaches the spectrum of $H$ only in the interval $\mathcal{J}$, the Combes-Thomas estimate and the fractional moment bound eq. (1.6) (which holds with $\Omega$ in place of $\Lambda_n$ by strong resolvent convergence) together show that the integrand here is exponentially small in $\mathrm{dist}_\Omega(x, y)$. This in turn gives (A.6).

The above argument combined with the previous remark (10) shows that the estimate (A.6) also holds in trace norm under the restriction that $|x - y| \geq 2r$ or if $g$ is required to vanish on $\mathcal{J}_>$.

(12) *(The assumption $\mathcal{A}3'$)* The condition of blow-up regularity of a random variable $X$ requires its probability density to be absolutely continuous, with a bounded density $\rho(\cdot)$, since eq. (1.25) implies (for arbitrary $n$)

$$\rho(x) \;=\; n \int \rho_n(nx - y|y)\mu_n(\mathrm{d}y) \;, \tag{A.9}$$

where $\mu_n$ is the probability distribution of $Y^{(n)}$. The converse is not true: there exist bounded densities such that the associated probability measure has infinite blow up norm. However, if $\ln \rho$ is Lipschitz-continuous then $D_\rho < \infty$. In this case, a particularly simple decomposition of the random variable $X$ with distribution $\rho(x)\mathrm{d}x$ is obtained with $X^{(n)}$ the fractional part of $nX$, so that $Y^{(n)}$ is integer valued and

$$\rho_n(x|j) \;:=\; \frac{\rho(j/n + x/n)}{n\int_{j/n}^{(j+1)/n} \rho(x)\mathrm{d}x} \;. \tag{A.10}$$

The Lipschitz condition guarantees the uniform boundedness of $\rho_n$ defined in this way. A particularly simple example, for which the results are already of interest, is provided by the uniform distribution in [0, 1].

## Appendix B. The Birman-Schwinger relation

Central to our analysis is the consideration of one-parameter operator families of the form

$$A_\xi = A_0 - \xi V \tag{B.1}$$

with a non-negative operator $V$ and $\xi \in \mathbb{R}$. Such families arise here when all but one of the random couplings in the random operator (1.1) are considered fixed. Their study replaces the rank one perturbation arguments which have played a key role in the analysis of discrete operators, e.g. in [69,4].

The family (B.1) formally satisfies the *Birman-Schwinger relation*:

$$V^{1/2}A_\xi^{-1}V^{1/2} = \left(\left(V^{1/2}A_0^{-1}V^{1/2}\right)^{-1} - \xi\mathbf{1}\right)^{-1} \tag{B.2}$$

to be interpreted as equality of operators in $(\ker V)^\perp$, subject to issues of invertibility. In the "classical" version of this relation [14,63], $A_0 = -\Delta + \gamma\mathbf{1}$

for some $\gamma > 0$, and $V$ is a non-negative relatively compact potential perturbation. It is shown, e.g. in [64, Sect. 8], that the number of eigenvalues of $-\Delta - V$ less than $-\gamma < 0$ equals the number of eigenvalues of $V^{1/2}(-\Delta + \gamma\mathbf{1})^{-1}V^{1/2}$ greater than one. This can be understood through the observation that the two sides of (B.2) become singular for the same values of the parameters $\gamma$ and $\xi$.

We use here two non-classical, though certainly not new, versions of the Birman-Schwinger relation which are described below. In our applications they arise through the two standard procedures for regularizing the Green function at energies in the infinite volume spectrum: (1) adding a small imaginary part to the energy or (2) using finite volume approximations.

The case of complex energy is covered by the following.

**Lemma B.1.** *Let $A_0 = B + iC$ be an operator on a separable Hilbert space $\mathcal{H}$, where $B$ is self-adjoint and $C$ is bounded with $C \geq \delta\mathbf{1}$ for some $\delta > 0$. Also, let $V$ be a bounded non-negative operator in $\mathcal{H}$. Then the Birman-Schwinger operator*

$$A_{BS} := \left(V^{1/2}A_0^{-1}V^{1/2}\right)^{-1} \tag{B.3}$$

*is maximally dissipative in $(\ker V)^{\perp}$, with $\mathcal{D}(A_{BS}) = \mathcal{R}(V^{1/2}A_0^{-1}V^{1/2})$, the range of $V^{1/2}A_0^{-1}V^{1/2}$. Moreover, its resolvent set $\rho(A_{BS})$ includes the closed lower half plane $\overline{\mathbb{C}}_-$, and the Birman-Schwinger relation*

$$(A_{BS} - \xi\mathbf{1})^{-1} = V^{1/2}(A_0 - \xi V)^{-1}V^{1/2} \tag{B.4}$$

*holds in $(\ker V)^{\perp}$ for every $\xi \in \overline{\mathbb{C}}_-$.*

*Remark.* In our applications $V$ appears as a non-negative potential and $A_0 = z - H$ with $\operatorname{Im} z > 0$ and $H$ a self-adjoint Schrödinger operator on $L^2(\Omega)$. In this case $(\ker V)^{\perp} = L^2(\{x : V(x) > 0\})$.

*Proof.* Note that $A_0$ is boundedly invertible with $\|A_0^{-1}\| \leq 1/\delta$. To show that the restriction of $V^{1/2}A_0^{-1}V^{1/2}$ to $(\ker V)^{\perp}$ is invertible, suppose we are given $f \in (\ker V)^{\perp}$ such that $V^{1/2}A_0^{-1}V^{1/2}f = 0$. Then $g := A_0^{-1}V^{1/2}f \in \mathcal{D}(A_0)$ with $V^{1/2}g = 0$. Thus $0 = \langle V^{1/2}g, f\rangle = \langle g, V^{1/2}f\rangle = \langle g, A_0 g\rangle$. Taking imaginary parts and using $C \geq \delta\mathbf{1}$ we find that $g = 0$. Thus $V^{1/2}f = A_0 g = 0$ and therefore $f \in (\ker V) \cap (\ker V)^{\perp} = \{0\}$.

We conclude that $A_{BS}$ exists as an operator in $(\ker V)^{\perp}$. It is densely defined, since if $\langle f, V^{1/2}A_0^{-1}V^{1/2}g\rangle = 0$ for all $g \in (\ker V)^{\perp}$, then $V^{1/2}(A_0^{-1})^{*}V^{1/2}f = 0$ and, by the same argument as above, $f = 0$.

For $\xi \in \overline{\mathbb{C}}_-$ one verifies explicitly that

$$(A_{BS} - \xi\mathbf{1})V^{1/2}A_0^{-1}V^{1/2} = \mathbf{1} - \xi V^{1/2}A_0^{-1}V^{1/2}$$

is the inverse of $\mathbf{1} + \xi V^{1/2}(A_0 - \xi V)^{-1}V^{1/2}$ in $(\ker V)^{\perp}$. In particular, for $\xi = -\mathrm{i}$, this shows that $(A_{BS} + \mathrm{i}\mathbf{1})\mathcal{D}(A_{BS}) = (\ker V)^{\perp}$, proving that $A_{BS}$

is maximally dissipative (see [72]). Using the resolvent identity we also get (B.4). $\square$

The argument in the above proof which showed that $V^{1/2}A_0^{-1}V^{1/2}$ is invertible does not generally carry over to the case where $A_0$ is an invertible self-adjoint operator. If $A_0$ has fixed sign (as in the classical BS-principle) it does, but not if 0 is in a spectral gap for $A$. However, in the case which is of interest to us, namely that $A_0 = H - E$ with $H$ a finite volume Schrödinger operator, we can make use of the fact that $H$ is *local* in the sense that $H\varphi$ vanishes on an open set $O$ if $\varphi \in \mathcal{D}(H)$ vanishes on $O$.

Thus, let $\Lambda \subset \mathbb{R}^d$ be open and bounded and let $H_0$ be the Dirichlet restriction of a Schrödinger operator $(i\nabla - A)^2 + V_0$ onto $L^2(\Lambda)$, where $A$ and $V_0$ satisfy the general assumptions of Sect. 1. Let $V \geq 0$ be a bounded non-zero potential of compact support with $|\partial(\mathrm{supp}\, V)| = 0$, and for $\xi \in \mathbb{R}$ let

$$H_\xi = H_0 - \xi V\ . \tag{B.5}$$

Then $H_\xi$ is self-adjoint with compact resolvent. Thus, by Theorem VII.3.9 of [43], its repeated eigenvalues $E_n(\xi)$, $n \in \mathbb{N}$, and corresponding complete set of orthonormal eigenfunctions $\psi_n(\xi)$ can be labeled such that $E_n(\cdot)$ and $\psi_n(\cdot)$ are holomorphic (note that crossings are possible, that is, the $E_n$ may be degenerate and are not necessarily in increasing order). By the Feynman-Hellmann Theorem and the unique continuation property of eigenfunctions (e.g., remark (3) in Appendix A),

$$E_n'(\xi) = -\langle V\psi_n(\xi), \psi_n(\xi)\rangle < 0\ . \tag{B.6}$$

Thus $\Gamma_n := E_n^{-1}$ exists on the range of $E_n$ and

$$\Gamma_n'(E) = -\frac{1}{\langle V\psi_n(\Gamma_n(E)), \psi_n(\Gamma_n(E))\rangle}\ . \tag{B.7}$$

For real $E \notin \sigma(H_\xi)$, we define

$$K_{\xi,E} = V^{1/2}(H_\xi - E)^{-1}V^{1/2} \tag{B.8}$$

as an operator in $L^2(\mathrm{supp}\, V)$, where it is compact and self-adjoint. We claim that $\ker K_{\xi,E} = \{0\}$. Indeed, if $V^{1/2}(H_\xi - E)^{-1}V^{1/2}f = 0$ for $f \in L^2(\mathrm{supp}\, V)$, then $(H_\xi - E)^{-1}V^{1/2}f = 0$ on $\mathrm{supp}\, V$ and, since $H_\xi$ is local, $V^{1/2}f = 0$ in the interior of $\mathrm{supp}\, V$. This implies that $f = 0$ since $|\partial(\mathrm{supp}\, V)| = 0$.

We conclude that $K_{\xi,E}^{-1}$ exists in $L^2(\mathrm{supp}\, V)$ as an unbounded self-adjoint operator with discrete spectrum. By the arguments used in the proof of Lemma B.1 it may be shown that for $E \notin \sigma(H_\xi) \cup \sigma(H_0)$,

$$K_{\xi,E} = \left(K_{0,E}^{-1} - \xi\mathbf{1}\right)^{-1}\ . \tag{B.9}$$

**Lemma B.2.** *Let $E \notin \sigma(H_0)$ and $\xi \neq 0$. Then*

(1) *$\phi$ is a normalized eigenfunction of $K_{0,E}^{-1}$ with eigenvalue $\xi$ if and only if $\psi := \xi(H_0 - E)^{-1} V^{1/2}\phi$ is an eigenfunction of $H_\xi$ with eigenvalue $E$ and $\langle \psi, V\psi \rangle = 1$.*

(2) *the repeated eigenvalues of $K_{0,E}^{-1}$ are given by $\Gamma_n(E)$, $n \in \mathbb{N}$, with (non-normalized) complete eigenvectors $V^{1/2}\psi_n(\Gamma_n(E))$.*

*Proof.* The second claim follows from the first and the definition of $\Gamma_n(E)$.

To prove the first claim, first suppose that that $K_{0,E}^{-1}\phi = \xi\phi$ and let $\psi := \xi(H_0 - E)^{-1} V^{1/2}\phi$. Then $V^{1/2}\psi = \xi K_{0,E}\phi = \phi$ and $(H_\xi - E)\psi = \xi(H_\xi - E)(H_0 - E)^{-1} V^{1/2}\phi = \xi V^{1/2}\phi - \xi V\psi = 0$.

Conversely, if $H_\xi\psi = E\psi$ with $\langle V\psi, \psi\rangle = 1$, then $\psi = \xi(H_0 - E)^{-1} V\psi$ and thus $\phi := V^{1/2}\psi \in \mathcal{R}(V^{1/2}(H_0 - E)^{-1} V^{1/2}) = \mathcal{D}(K_{0,E}^{-1})$ and is normalized. Moreover, $K_{0,E}^{-1}\phi = K_{0,E}^{-1} V^{1/2}\psi = \xi K_{0,E}^{-1} K_{0,E} V^{1/2}\psi = \xi\phi$. $\square$

## Appendix C. A "weak $L^1$" bound for resolvents of dissipative operators

The goal of this appendix is to provide a proof of Lemma 3.1. All the main arguments are taken from [56].

We will use here that a maximally dissipative operator $A$ in $\mathcal{H}$ has a selfadjoint dilation $L$ in a Hilbert space $\tilde{\mathcal{H}}$ which contains $\mathcal{H}$ as a subspace, i.e.

$$(A - \xi)^{-1} = P(L - \xi)^{-1} P^* \tag{C.1}$$

for every $\xi \in \mathbb{C}$ with $\operatorname{Im}\xi < 0$. Here $P$ is the orthogonal projection onto $\mathcal{H}$ in $\tilde{\mathcal{H}}$. For this and much more on the general theory of dissipative operators see the survey [59] or the book [72] (where the equivalent theory of contractions and their unitary dilations is presented).

We start the proof of Lemma 3.1 with two reduction steps. First, we show that it is sufficient to deal with the case $M_1 = M = M_2^*$.

Thus, let us assume that it is proven that

$$\left|\left\{v \in \mathbb{R} : \|M^*(A - v + i0)^{-1} M\|_{HS} > t\right\}\right| \leq \frac{C}{t}\|M\|_{HS}^2 \tag{C.2}$$

for all self-adjoint $A$ in $\mathcal{H}$ and Hilbert-Schmidt operators $M : \mathcal{H}_1 \to \mathcal{H}$.

The estimate (3.7) follows from (C.2) by a polarization and a scaling argument. For this let us temporarily write $T = (A - v + i0)^{-1}$. One checks that

$$\begin{aligned} M_2 T M_1 =& \frac{1}{2}\left(M_2 + M_1^*\right)T\left(M_2^* + M_1\right) - \frac{i}{2}\left(M_2 - iM_1^*\right)T\left(M_2^* + iM_1\right) \\ & - \frac{1-i}{2} M_2 T M_2^* - \frac{1-i}{2} M_1^* T M_1 \, . \end{aligned} \tag{C.3}$$

All four terms on the r.h.s. of (C.3) are of the type which is covered by (C.2). The set $\{v : \|M_2TM_1\|_{HS} > t\}$ is contained in the union of $\{v : \|\frac{1}{2}(M_2 + M_1^*)T(M_2^* + M_1)\|_{HS} > t/4\}$ and three similar sets. Applying (C.2) to all of them gives

$$|\{v : \|M_2TM_1\|_{HS} > t\}| \leq \frac{C_1}{t}\left(\|M_2\|_{HS}^2 + \|M_1\|_{HS}^2\right) \tag{C.4}$$

with a suitable constant $C_1$. Scaling of (C.4) yields

$$\begin{aligned} &|\{v : \|M_2TM_1\|_{HS} > t\}| \\ &\quad = \left|\left\{v : \left\|\frac{M_2}{\|M_2\|_{HS}}T\frac{M_1}{\|M_1\|_{HS}}\right\|_{HS} > \frac{t}{\|M_1\|_{HS}\|M_2\|_{HS}}\right\}\right| \\ &\quad \leq \frac{2C_1}{t}\|M_1\|_{HS}\|M_2\|_{HS} \ , \end{aligned} \tag{C.5}$$

and thus (3.7).

It remains to show (C.2), which follows if we can show that

$$\left|\left\{v \in \mathbb{R} : \|M^*(A - v + i\delta)^{-1}M\|_{HS} > t\right\}\right| \leq \frac{C}{t}\|M\|_{HS}^2 \tag{C.6}$$

for all $\delta > 0$, with $C < \infty$ independent of $\delta$.

To show that (C.6) implies (C.2), consider the function $\Phi$ defined on $\mathbb{C}^- = \{\operatorname{Im}\xi < 0\}$ by $\Phi(\xi) = M^*(A-\xi)^{-1}M$. The function $\Phi$ is analytic and takes values in the trace class operators on $\mathcal{H}$ with non-negative imaginary part. By a result of de Branges [24], later proven independently in [9] and [15], $\Phi(v - i0) := \lim_{\delta\downarrow 0}\Phi(v - i\delta)$ exists in Hilbert-Schmidt norm for almost every $v \in \mathbb{R}$. Together with (C.3) this implies the existence statement in part (1) of Lemma 3.1. For $\delta \geq 0$ let $g_\delta$ denote the characteristic function of the set $\{v \in \mathbb{R} : \|M^*(A - v + i\delta)^{-1}M\|_{HS} > t\}$. One checks that $g_0(v) \leq \liminf_{\delta\downarrow 0} g_\delta(v)$ for almost every $v$. Therefore (C.2) is a consequence of (C.6) and Fatou's lemma.

Before we proceed with the remaining proof of (C.6), we state two classical facts on Hilbert transforms which will be used.

The Hilbert transform of a function $f : \mathbb{R} \to \mathbb{C}$ is defined by the principle-value integral

$$Hf(x) = \frac{1}{\pi}\lim_{\varepsilon\downarrow 0}\int_{\mathbb{R}\setminus[x-\varepsilon,x+\varepsilon]}\frac{f(y)}{x - y}\,dy \ , \tag{C.7}$$

whenever this limit exists. The same definition applies when $f$ takes values in a Hilbert space, in which case the r.h.s. of (C.7) is interpreted as a Bochner integral.

**Proposition C.1.** *Suppose that $\Phi \in H^2(\mathbb{C}^-)$, i.e. $\Phi : \mathbb{C}^- \to \mathbb{C}$ is analytic and*

$$\sup_{y>0}\int_{\mathbb{R}}|\Phi(x - iy)|^2\,dx < \infty \ . \tag{C.8}$$

*Then the boundary value* $\Phi(x) = \lim_{y \downarrow 0} \Phi(x - iy)$ *exists for almost every* $x \in \mathbb{R}$, $\Phi \in L^2(\mathbb{R})$, *and its real and imaginary parts are conjugate, i.e.*

$$\mathrm{Re}\,\Phi(x) = H(\mathrm{Im}\,\Phi)(x) \quad \textit{for a.e. } x \in \mathbb{R}\ . \tag{C.9}$$

**Proposition C.2.** *Let* $\mathcal{H}$ *be a separable Hilbert space and* $f \in L^1(\mathbb{R}, \mathcal{H})$ *in the sense of Bochner integration. Then the Hilbert transform* $Hf(y) \in \mathcal{H}$ *exists for a.e.* $y \in \mathbb{R}$, *and there exists a constant* $C < \infty$, *independent of* $f$, *such that for all* $t > 0$

$$|\{y \in \mathbb{R} :\ \|Hf(y)\|_{\mathcal{H}} > t\}| \leq \frac{C}{t} \int_{\mathbb{R}} \|f(x)\|_{\mathcal{H}}\, dx\ . \tag{C.10}$$

A modern proof of Proposition C.1 can be found in [31]. Proposition C.2, i.e. the weak-$L^1$-property of the Hilbert transform, is well known for the case $\mathcal{H} = \mathbb{C}$. A proof in the context of more general Calderon-Zygmund inequalities can be found in [70, Chap. 2], where it is remarked that the result extends to the $\mathcal{H}$-valued case. Detailed proofs of such results for vector-valued functions, which contain Proposition C.2 as a special case, can be found in [62].

We now apply these facts to $T_\delta(v) := M^*(A - v + i\delta)^{-1}M$. The trace class norm will be denoted $\|\cdot\|_1$. The real and imaginary parts of bounded operators are defined as usual by $\mathrm{Re}\, C = \frac{1}{2}(C+C^*)$ and $\mathrm{Im}\, C = \frac{1}{2i}(C-C^*)$.

**Lemma C.3.** *For every* $\delta > 0$ *it holds that*

$$\int_{\mathbb{R}} \|\mathrm{Im}\, T_\delta(v)\|_{HS}\, dv \leq \int_{\mathbb{R}} \|\mathrm{Im} T_\delta(v)\|_1\, dv = \pi \|M\|_{HS}^2\ . \tag{C.11}$$

*Proof.* The first part of (C.11) follows from $\|\cdot\|_{HS} \leq \|\cdot\|_1$. Let $E(t)$ be the spectral resolution of the selfadjoint dilation $L$ of $A$ and $\phi \in \mathcal{H}_1$. Then by (C.1), the spectral theorem, and Fubini

$$\begin{aligned}\int \langle \mathrm{Im}\, T_\delta(v)\phi, \phi\rangle_{\mathcal{H}_1}\, dv &= \int\int \frac{\delta}{(x-v)^2+\delta^2}\, d\|E(x)P^*M\phi\|_{\tilde{\mathcal{H}}}^2\, dv \\ &= \pi \|M\phi\|_{\mathcal{H}_1}^2\ .\end{aligned} \tag{C.12}$$

Let $(\phi_n)$ be an orthonormal basis in $\mathcal{H}_1$. We have $\mathrm{Im}\, T_\delta(v) \geq 0$ and thus by (C.12)

$$\begin{aligned}\int \|\mathrm{Im}\, T_\delta(v)\|_1\, dv &= \int \mathrm{Tr}(\mathrm{Im}\, T_\delta(v))\, dv \\ &= \int \sum_n \langle \mathrm{Im}\, T_\delta(v)\phi_n, \phi_n\rangle_{\mathcal{H}_1}\, dv \\ &= \pi \|M\|_{HS}^2\ .\end{aligned} \tag{C.13}$$

$\square$

**Lemma C.4.** *Let $\mathcal{H}_{HS}$ denote the separable Hilbert space of all Hilbert-Schmidt operators on $\mathcal{H}_1$.*

*(a) If $\phi \in \mathcal{H}_1$, then $\langle T_\delta(\cdot)\phi, \phi\rangle \in H^2(\mathbb{C}^-)$ for each fixed $\delta > 0$.*
*(b) For almost every $v \in \mathbb{R}$ one has*

$$\operatorname{Re} T_\delta(v) = H(\operatorname{Im} T_\delta)(v) \tag{C.14}$$

*in the sense of Hilbert transforms of $\mathcal{H}_{HS}$-valued functions.*

*Proof.* Consider arbitrary $\phi$ and $\psi$ in $\mathcal{H}_1$ and use (C.1) and the spectral theorem to estimate

$$\begin{aligned}\int |\langle T_\delta(v-iy)\phi,\psi\rangle|^2\,dv &\le \int \|(L-v-i(y+\delta))^{-1}P^*M\phi\|^2\,dv\,\|M\psi\|^2\\ &= \int\int \frac{1}{(x-v)^2+(y+\delta)^2}\,dv\,d\|E(x)P^*M\phi\|^2\|M\psi\|^2\\ &= \frac{\pi}{y+\delta}\|M\phi\|^2\|M\psi\|^2 \le \frac{\pi}{\delta}\|M\phi\|^2\|M\psi\|^2\,.\end{aligned} \tag{C.15}$$

Summing this over an orthonormal basis of vectors $\psi$ and then over an orthonormal basis of vectors $\phi$ leads to

$$\int \|T_\delta(v-iy)\|_{HS}^2\,dy \le \frac{\pi}{\delta}\|M\|_{HS}^4\,. \tag{C.16}$$

This implies (a). In fact, we have proven the stronger result that $T_\delta(\cdot)$ is a Hilbert-Schmidt-valued $H^2$-function in the lower half plane.

By Lemma C.3 and Proposition C.2, $H(\operatorname{Im} T_\delta)$ exists almost everywhere as a Hilbert-Schmidt operator. Since the strong topology is weaker than the Hilbert-Schmidt topology, this implies the existence of $H\langle \operatorname{Im} T_\delta(\cdot)\phi,\phi\rangle = \langle H(\operatorname{Im} T_\delta)(\cdot)\phi,\phi\rangle$ for every $\phi \in \mathcal{H}_1$. By (a) and Proposition C.1 the latter is equal to $\langle \operatorname{Re} T_\delta(\cdot)\phi,\phi\rangle$. We conclude (C.14) since bounded operators are determined by their quadratic form. □

We are now prepared to prove (C.6) and thereby complete the proof of Lemma 3.1. We have, using Lemma C.4(b),

$$\begin{aligned}&|\{v : \|T_\delta(v)\|_{HS} > t\}|\\ &\quad\le \left|\left\{v:\ \|\operatorname{Re} T_\delta(v)\|_{HS} > \frac{t}{2}\right\}\right| + \left|\left\{v:\|\operatorname{Im} T_\delta(v)\|_{HS} > \frac{t}{2}\right\}\right|\\ &\quad= \left|\left\{v:\ \|H(\operatorname{Im} T_\delta)(v)\|_{HS} > \frac{t}{2}\right\}\right| + \left|\left\{v:\|\operatorname{Im} T_\delta(v)\|_{HS} > \frac{t}{2}\right\}\right|\\ &\quad\le \frac{2(C+1)}{t}\int \|\operatorname{Im} T_\delta(v)\|_{HS}\,dv\,,\end{aligned} \tag{C.17}$$

where in the end Proposition C.2 and Chebychev's inequality were used. Thus (C.6) is a consequence of Lemma C.3.

## Appendix D. A disorder-averaged spectral shift bound

Here we present a new result, which amounts to boundedness at fixed energy of a fractional moment – under averaging over local disorder – of the spectral shift associated with the addition to a Schrödinger operator of a local potential.

As was explained in the introduction, some of the difficulties which have in the past impeded the extention of fractional moment methods to the continuum can be traced to the lack of uniform bounds on such spectral shifts. The following result is enabled by the methods of Sect. 3. While the analysis presented above does not proceed through this bound, the issues involved are closely related, and the result may provide a useful tool.

**Theorem D.1.** *Let* $H_t = \widehat{H} + tV$ *where* $\widehat{H}$ *satisfies* $\mathcal{A}1$ *and* $V$ *is a non-negative bounded function with compact support. Let* $U$ *be a non-negative bounded function such that* $V$ *is strictly positive throughout the set* $Q = \{q : \mathrm{dist}(q, \mathrm{supp}(U)) < \delta\}$ *with some* $\delta > 0$ *and set* $v_- = \inf_{x \in Q} V(x)$. *Then, for any* $0 < s < \min(2/d, 1/2)$ *there is* $C_{s,\delta} < \infty$ *such that the spectral shift function, defined as*

$$\xi(t, E) \;=\; \mathrm{Tr}\,[P\,(H_t < E) - P\,(H_t + U < E)]\ , \tag{D.1}$$

*satisfies, for any* $E \geq \inf \sigma(\widehat{H})$*:*

$$\int_0^1 |\xi(t, E)|^s \,\mathrm{d}t \;\leq\; C_{s,\delta}\, \|U\|_\infty\, (1 + |E - E_0| + \|V\|_\infty)^{s(2d+2)}\ , \tag{D.2}$$

*with* $E_0 = \inf \sigma(\widehat{H})$.

*Proof.* We claim that it suffices to prove eq. (D.2) for operators $H_t$ restricted to bounded regions with a constant $C_{s,\delta}$ which is independent of the region. To verify this, note that strong resolvent convergence and lower semi-continuity of the trace norm imply that

$$\xi(t, E) \;\leq\; \liminf_{L \to \infty} \xi_L(t, E) \tag{D.3}$$

where $\xi_L(t, E)$ is computed with $H_t^{(\Lambda_L)}$ in place of $H_t$ with $\Lambda_L = [-L, L]^d$. It is useful to note that, because $U$ is non-negative, the difference of projections appearing in eq. (D.1) is a positive semi-definite operator so its trace is equal to its trace norm. An application of Fatou's lemma yields eq. (D.2) for $H_t$ provided it holds for $H_t^{(\Lambda_L)}$.

Throughout the rest of the proof we fix $L > 0$ and write $H_t$ and $\widehat{H}$ for the restrictions of these operators to $[-L, L]^d$ with Dirichlet boundary conditions. We begin with the observation that

$$\xi(t, E) \;=\; \mathrm{Tr}\, P_t\ , \tag{D.4}$$

where $P_t$ is the spectral projection to the interval $(-\infty, -1]$ of the Birman-Schwinger operator

$$K_t \;=\; U^{1/2}\frac{1}{H_t - E}U^{1/2} \,. \tag{D.5}$$

This fact follows from the Birman-Schwinger representation since $E$ becomes an eigenvalue of $H_t + \eta U$ precisely when $\eta$ is equal to an eigenvalue of $-K_t^{-1}$.

Note that for any $n \geq 1$,

$$\left\| K_t^{-n} P_t \big(1 + K_t^{-1}\big)^n \right\| \leq 1 \,. \tag{D.6}$$

Thus by the Hölder inequality for trace norms

$$\xi(t, E) \leq \left\| K_t^n \right\|_{HS} \left\| \big(1 + K_t^{-1}\big)^{-n} \right\|_{HS} . \tag{D.7}$$

Noting that

$$\big(1 + K_t^{-1}\big)^{-1} = U^{1/2}\frac{1}{H_t + U - E}U^{1/2} \,, \tag{D.8}$$

we find

$$\begin{aligned} \int_0^1 \mathrm{d}t\, |\xi(t, E)|^s \leq & \left( \int_0^1 \left\| U^{1/2}\frac{1}{H_t - E}U^{1/2} \right\|_{2n}^{2ns} \mathrm{d}t \right)^{1/2} \\ & \times \left( \int_0^1 \left\| U^{1/2}\frac{1}{H_t + U - E}U^{1/2} \right\|_{2n}^{2ns} \mathrm{d}t \right)^{1/2} , \end{aligned} \tag{D.9}$$

where $\|A\|_m = (\mathrm{Tr}\,|A|^m)^{1/m}$.

Lemma 3.1 and the representation eq. (3.38) used in the proof of Lemma 3.3 may be used to show that the integrals on the right hand side of eq. (D.9) are bounded if (1) $2ns < 1$ and (2) $4n > d$ ($d$ is the dimension). We now outline the proof of this assertion.

The arguments used in Lemma 3.3 can be used to produce a representation

$$U^{1/2}\frac{1}{H_t - E}U^{1/2} \;=\; U^{1/2}T\Theta\frac{1}{H_t - E}\Theta T^\dagger U^{1/2} \;+\; B \tag{D.10}$$

with $T$, $T^\dagger$ Hilbert-Schmidt. In the proof of Lemma 3.3 it was noted that $B$ is bounded. In addition, using $\mathcal{A}1$, $B$ can be seen to be in $\mathcal{I}_p$ for any $p > d/2$, with uniform bounds on its $\mathcal{I}_p$ norm. Since $\|\cdot\|_{2n} \leq \|\cdot\|_{HS}$ we find that

$$\begin{aligned} & \int_0^1 \left\| U^{1/2}\frac{1}{H_t - E}U^{1/2} \right\|_{2n}^{2ns} \mathrm{d}t \\ & \qquad \leq \int_0^1 \left\| U^{1/2}T\Theta\frac{1}{H_t - E}\Theta T^\dagger U^{1/2} \right\|_{HS}^{2ns} + \mathrm{O}(1) \,, \end{aligned} \tag{D.11}$$

with a similar expression for the other factor in eq. (D.9). The weak $L^1$ bound can be used to bound this final integral since

$$\Theta \frac{1}{H_t - E} \Theta = \Theta V^{-1/2} \frac{1}{t + \widehat{K}^{-1}} V^{-1/2} \Theta \,, \tag{D.12}$$

with appropriate $\widehat{K}$. Note that $\Theta V^{-1/2}$ is bounded since $\Theta$ is supported in $Q$. As in Sect. 3 there is slight complication due to the fact that $T$ (and $T^\dagger$) depend on $t$. However, as before, the dependence is polynomial and may be handled in the same way. □

*Acknowledgements.* In the course of this project, the authors' work was supported in part by NSF grants PHY-9971149 (MA, and AE), DMS 0070343 and 0245210 (GS), INT-0204308 (JS), NSF postdoctoral research fellowship (JS), and a NATO Collaborative Linkage Grant PST.CLG.976441 (GS and SN). The authors also thank Caltech, Université Paris 7, and the Institute Mittag-Leffler for hospitality which has facilitated this collaboration.

## References

1. Abraham, E., Anderson, P.W., Licciardello, D.C., Ramakrishnan, T.V.: Scaling theory of localization: Absence of quantum diffusion in two dimensions. Phys. Rev. Lett. **42**, 673–676 (1979)
2. Aizenman, M.: Localization at weak disorder: some elementary bounds. Rev. Math. Phys. **6**, 1163–1182 (1994)
3. Aizenman, M., Graf, G.M.: Localization bounds for an electron gas. J. Phys. A **31**, 6783–6806 (1998)
4. Aizenman, M., Molchanov, S.: Localization at large disorder and at extreme energies: an elementary derivation. Commun. Math. Phys. **157**, 245–278 (1993)
5. Aizenman, M., Newman, C.M.: Tree graph inequalities and critical behavior in percolation models. J. Stat. Phys. **36**, 107–143 (1984)
6. Aizenman, M., Schenker, J.H., Friedrich, R.M., Hundertmark, D.: Finite volume fractional moment criteria for Anderson localization. Commun. Math. Phys. **224**, 219–253 (2001)
7. Aizenman, M., Simon, B.: Local Ward identities and the decay of correlations in ferromagnets. Commun. Math. Phys. **77**, 137–143 (1980)
8. Anderson, P.W.: Absence of diffusion in certain random lattices. Phys. Rev. **109**, 1492–1505 (1958)
9. Asano, K.: Notes on Hilbert transforms of vector valued functions in the complex plane and their boundary values. Proc. Japan Acad. **43**, 572–577 (1967)
10. Avron, J.E., Seiler, R., Simon, B.: Charge deficiency, charge transport and comparison of dimensions. Commun. Math. Phys. **159**, 399–422 (1994)
11. Barbaroux, J.-M., Combes, J.-M., Hislop, P.D.: Localization near band edges for random Schrödinger operators. Helv. Phys. Acta **70**, 16–43 (1997)
12. Bellissard, J., van Elst, A., Schulz-Baldes, H.: The noncommutative geometry of the quantum Hall effect. J. Math. Phys. **35**, 5373–5451 (1994)
13. Berthier, A.M.: Sur le spectre ponctuel de l'opérateur de Schrödinger. C. R. Acad. Sci., Paris, Sér. A-B **290**, A393–395 (1980)
14. Birman, M.Š.: On the number of eigenvalues in a quantum scattering problem. Vestn. Leningrad Univ. **16**, 163–166 (1961)
15. Birman, M.Š, Entina, S.B.: Stationary approach in abstract scattering theory. Izv. Akad. Nauk SSSR, Ser. Mat. **31**, 401–430 (1967)

16. Boole, G.: On the comparison of transcendents, with certain applications to the theory of definite integrals. Philos. Trans. R. Soc. **147**, 745–803 (1857)
17. Boutet de Monvel, A., Naboko, S., Stollmann, P., Stolz, G.: Localization for continuum random surfaces via fractional moments. In preparation
18. Combes, J.-M., Hislop, P.D.: Localization for some continuous, random Hamiltonians in $d$-dimensions. J. Funct. Anal. **124**, 149–180 (1994)
19. 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)
20. Combes, J.-M., Thomas, L.: Asymptotic behaviour of eigenfunctions for multiparticle Schrödinger operators. Commun. Math. Phys. **34**, 251–270 (1973)
21. Cycon, H.L., Froese, R.G., Kirsch, W., Simon, B.: Schrödinger operators with application to quantum mechanics and global geometry. Texts and Monographs in Physics, study edition. Berlin: Springer 1987
22. Davies, E.B.: Spectral theory and differential operators. Cambridge Studies in Advanced Mathematics, vol. 42. Cambridge: Cambridge University Press 1995
23. De Bièvre, S., Pulé J.V.: Propagating edge states for a magnetic hamiltonian. Math. Phys. Electron. J. **5**, paper 3 (1999)
24. de Branges, L.: Perturbations of self-adjoint transformations. Am. J. Math. **84**, 543–560 (1962)
25. Delyon, F., Levy, Y., Souillard, B.: Anderson localization for multidimensional systems at large disorder or low energy. Commun. Math. Phys. **100**, 463–470 (1985)
26. Dobrushin, R.L., Shlosman, S.B.: Completely analytical interactions: constructive description. J. Stat. Phys. **46**, 983–1014 (1987)
27. Dorlas, T.C., Macris, N., Pulé, J.V.: Characterization of the spectrum of the Landau Hamiltonian with delta impurities. Commun. Math. Phys. **204**, 367–396 (1999)
28. Fröhlich, J., Graf, G.M., Walcher, J.: On the extended nature of edge states of quantum hall hamiltonians. Ann. Inst. Henry Poincaré **1**, 405–444 (2000)
29. 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)
30. Fröhlich, J., Spencer, T.: Absence of diffusion in the Anderson tight binding model for large disorder or low energy. Commun. Math. Phys. **88**, 151–184 (1983)
31. Garnett, J.B.: Bounded analytic functions. New York: Academic Press Inc. [Harcourt Brace Jovanovich Publishers] 1981
32. Germinet, F., Klein, A.: Bootstrap multiscale analysis and localization in random media. Commun. Math. Phys. **222**, 415–448 (2001)
33. Germinet, F., Klein, A.: A characterization of the Anderson metal-insulator transition. Duke Math. J. **124**, 309–359 (2004)
34. Goldsheid, I.Ya., Molchanov, S.A., Pastur, L.A.: Typical one-dimensional Schrödinger operator has pure point spectrum. Funkts. Anal. Prilozh. **11**, 1–10 (1977)
35. Graf, G.-M.: Anderson localization and the space-time characteristic of continuum states. J. Stat. Phys. **75**, 337 (1994)
36. Halperin, B.I.: Quantized hall conductance, current-carrying edge states, and the existence of extended states in a two-dimensional disordered potential. Phys. Rev. B, Condens. Matter **25**, 2185–2190 (1982)
37. Hammersley, J.M.: Percolation processes. II. The connective constant. Proc. Cambridge Philos. Soc. **53**, 642–645 (1957)
38. Helffer, B., Sjöstrand, J.: Équation de Schrödinger avec champ magnétique et équation de Harper. In: Schrödinger operators (Sønderborg, 1988), Lect. Notes Phys., vol. 345, pp. 118–197. Berlin: Springer 1989
39. Hislop, P.D., Kirsch, W., Krishna, M.: Spectral and dynamical properties of random models with nonlocal and singular interactions. Math. Nachr. **278**, 627–664 (2005)
40. Hundertmark, D., Simon, B.: An optimal $L^p$-bound on the Krein spectral shift function. J. Anal. Math. **87**, 199–208 (2002)

41. Jerison, D., Kenig, C.E.: Unique continuation and absence of positive eigenvalues for Schrödinger operators. With an appendix by E.M. Stein. Ann. Math. (2) **121**, 463–494 (1985)
42. Kalf, H.: Une remarque au sujet de prolongement unique des solutions de l'équation de Schrödinger. C. R. Acad. Sci. Paris, Sér. I, Math. **295**, 579–581 (1982)
43. Kato, T.: Perturbation theory for linear operators. Classics in Mathematics. Berlin: Springer 1995. Reprint of the 1980 edition
44. Kirsch, W.: Small perturbations and the eigenvalues of the Laplacian on large bounded domains. Proc. Am. Math. Soc. **101**, 509–512 (1987)
45. Kirsch, W.: Wegner estimates and Anderson localization for alloy-type potentials. Math. Z. **221**, 507–512 (1996)
46. Kirsch, W., Stollmann, P., Stolz, G.: Localization for random perturbations of periodic Schrödinger operators. Random Oper. Stoch. Equ. **6**, 241–268 (1998)
47. Klopp, F.: Internal Lifshits tails for random perturbations of periodic Schrödinger operators. Duke Math. J. **98**, 335–396 (1999)
48. Klopp, F.: Internal Lifshitz tails for Schrödinger operators with random potentials. J. Math. Phys. **43**, 2948–2958 (2002)
49. Koch, H., Tataru, D.: Recent results on unique continuation for second order elliptic equations. In: Carleman estimates and applications to uniqueness and control theory (Cortona, 1999). Progr. Nonlinear Differential Equations Appl., vol. 46, pp. 73–84. Boston, MA: Birkhäuser 2001
50. Kotani, S., Simon, B.: Localization in general one-dimensional random systems, II. Continuum Schrödinger operators. Commun. Math. Phys. **112**, 103–119 (1987)
51. Leinfelder, H., Simader, C.G.: Schrödinger operators with singular magnetic vector potentials. Math. Z. **176**, 1–19 (1981)
52. Lieb, E.H.: A refinement of Simon's correlation inequality. Commun. Math. Phys. **77**, 127–135 (1980)
53. Lifshitz, I.M.: Energy spectrum structure and quantum states of disordered condensed systems. Sov. Phys. Usp. **7**, 549–573 (1965)
54. Martinelli, F., Scoppola, E.: Introduction to the mathematical theory of Anderson localization. Riv. Nuovo Cimento (3) **10**, 1–90 (1987)
55. Mott, N., Twose, W.: The theory of impurity conduction. Adv. Phys. **10**, 107–163 (1961)
56. Naboko, S.N.: The structure of singularities of operator functions with a positive imaginary part. Funkts. Anal. Prilozh. **25**, 1–13, 96 (1991)
57. Niu, Q., Thouless, D.J., Wu, Y.S.: Quantized Hall conductance as a topological invariant. Phys. Rev. B (3) **31**, 3372–3377 (1985)
58. Pastur, L., Figotin, A.: Spectra of random and almost-periodic operators. Berlin: Springer 1992
59. Pavlov, B.S.: Spectral analysis of a dissipative singular Schrödinger operator in terms of a functional model [MR 93c:47057]. In: Partial differential equations, VIII, pp. 87–153. Berlin: Springer 1996
60. Reed, M., Simon, B.: Methods of modern mathematical physics. IV. Analysis of operators. New York: Academic Press [Harcourt Brace Jovanovich Publishers] 1978
61. Rudin, W.: Real and complex analysis, third edition. New York: McGraw-Hill Book Co. 1987
62. Schwartz, J.: A remark on inequalities of Calderon-Zygmund type for vector-valued functions. Commun. Pure Appl. Math. **14**, 785–799 (1961)
63. Schwinger, J.: On the bound states of a given potential. Proc. Natl. Acad. Sci. USA **47**, 122–129 (1961)
64. Simon, B.: Functional Integration and Quantum Physics. New York: Academic Press 1979
65. Simon, B.: Trace ideals and their applications. Lond. Math. Soc. Lect. Note Ser., vol. 35. Cambridge: Cambridge University Press 1979
66. Simon, B.: Correlation inequalities and the decay of correlations in ferromagnets. Commun. Math. Phys. **77**, 111–126 (1980)

67. Simon, B.: Schrödinger semigroups. Bull. Am. Math. Soc., New Ser. **7**, 447–526 (1982)
68. Simon, B.: Cyclic vectors in the Anderson model. Rev. Math. Phys. **6**, 1183–1185 (1994)
69. Simon, B., Wolff, T.: Singular continuous spectrum under rank one perturbations and localization for random Hamiltonians. Commun. Pure Appl. Math. **39**, 75–90 (1986)
70. Stein, E.M.: Singular integrals and differentiability properties of functions. Princeton, NJ: Princeton University Press 1970
71. Stollmann, P.: Caught by disorder: bound states in random media. Progress in Mathematical Physics, vol. 20. Boston, MA: Birkhäuser 2001
72. Sz.-Nagy, B., Foias, C.: Harmonic analysis of operators on Hilbert space. Amsterdam: North-Holland Publishing Co. 1970
73. Thouless, D.J.: Electrons in disordered systems and the theory of localization. Phys. Rep. **13**, 93–142 (1974)
74. von Dreifus, H., Klein, A.: A new proof of localization in the Anderson tight binding model. Commun. Math. Phys. **124**, 285–299 (1989)
75. Wolff, T.H.: Unique continuation for $|\Delta u| \leq V|\nabla u|$ and related problems. Rev. Mat. Iberoam. **6**, 155–200 (1990)
76. Wolff, T.H.: A property of measures in $\mathbf{R}^N$ and an application to unique continuation. Geom. Funct. Anal. **2**, 225–284 (1992)