Commun. Math. Phys. 378, 299–328 (2020)
Digital Object Identifier (DOI) https://doi.org/10.1007/s00220-020-03740-1

Communications in Mathematical Physics

Regularity of the Density of States of Random Schrödinger Operators

Dhriti Ranjan Dolai[1], **M. Krishna**[2], **Anish Mallick**[3]

[1] Department of Mathematics, Indian Institute of Technology, Dharwad, Dharwad 580011, India. E-mail: dhriti@iitdh.ac.in
[2] Ashoka University, Plot No 2, Rajiv Gandhi Education City, Rai, Haryana 131029, India. E-mail: krishna.maddaly@ashoka.edu.in
[3] Department of Mathematics, Pontificia Universidad Católica de Chile, Vicuna Mackenna 4860, Macul, Santiago, Chile. E-mail: anish.mallick@mat.uc.cl

Received: 7 September 2019 / Accepted: 13 February 2020
Published online: 29 March 2020 –

Abstract: In this paper we solve a long standing open problem for Random Schrödinger operators on $L^2(\mathbb{R}^d)$ with i.i.d single site random potentials. We allow a large class of free operators, including magnetic potential, however our method of proof works only for the case when the random potentials satisfy a complete covering condition. We require that the supports of the random potentials cover $\mathbb{R}^d$ and the bump functions that appear in the random potentials form a partition of unity. For such models, we show that the Density of States (DOS) is m times differentiable in the part of the spectrum where exponential localization is valid, if the single site distribution has compact support and has Hölder continuous $m + 1$ st derivative. The required Hölder continuity depends on the fractional moment bounds satisfied by appropriate operator kernels. Our proof of the Random Schrödinger operator case is an extensions of our proof for Anderson type models on $\ell^2(\mathbb{G})$, $\mathbb{G}$ a countable set, with the property that the cardinality of the set of points at distance N from any fixed point grows at some rate in $N^\alpha, \alpha > 0$. This condition rules out the Bethe lattice, where our method of proof works but the degree of smoothness also depends on the localization length, a result we do not present here. Even for these models the random potentials need to satisfy a complete covering condition. The Anderson model on the lattice for which regularity results were known earlier also satisfies the complete covering condition.

1. Introduction

In the study of the Anderson Model and Random Schrödinger operators, modulus of continuity of the Integrated Density of States (IDS) is well understood, (see Kirsch and Metzger [35] for a comprehensive review). In dimension bigger than one, there are very few results on further smoothness of the IDS, even when the single site distribution is assumed to have more smoothness, except for the case of the Anderson model itself at high disorder, (see for example Campanino and Klein [9], Bovier et al. [8], Klein and Speis [39], Simon and Taylor [51]).

In this paper we will show, in Theorems 3.4 and 4.4, that the IDS is almost as smooth as the single site distribution for a large class of continuous and discrete random operators. These are

$$H^{\omega} = H_0 + \sum_{n \in \mathbb{Z}^d} \omega_n u_n, \tag{1.1}$$

on $L^2(\mathbb{R}^d)$ and

$$h^{\omega} = h_0 + \sum_{n \in \mathbb{G}} \omega_n P_n, \tag{1.2}$$

on the separable Hilbert space $\mathscr{H}$ and a countable set $\mathbb{G}$. The operator h_0 is a bounded self-adjoint operator and the $\{P_n\}$ are finite rank projection. We specify the conditions on H_0, h_0, u_n, P_n and ω_n in the following sections.

The IDS, denoted $\mathcal{N}(E)$, is the distribution function of a non-random measure obtained as the weak limit of a sequence of random atomic measures. The proof of the existence of such limits for various models of random operators has a long history. These results are well documented in the books of Carmona and Lacroix [10], Figotin and Pastur [46], Cycon et al. [18], Kirsch [33] and the reviews of Kirsch and Metzger [35], Veselić [57] and in a review for stochastic Jacobi matrices by Simon [49]. In terms of the projection valued spectral measures $E_{H^{\omega}}, E_{h^{\omega}}$ associated with the self-adjoint operators H^{ω}, h^{ω}, the function $\mathcal{N}(E)$ has an explicit expression, for the cases when h^{ω}, H^{ω} are ergodic. For the model (1.1) it is given as

$$\frac{1}{\int u_0(x)dx}\mathbb{E}\Big[tr\Big(u_0 E_{H^{\omega}}((-\infty, E])\Big)\Big]$$

and for the model (1.2) it turns out to be

$$\frac{1}{tr(P_0)}\mathbb{E}\Bigg[tr\Bigg(P_0 E_{h^{\omega}}((-\infty, E])\Bigg)\Bigg].$$

We note that by using the same symbol $\mathcal{N}$ for two different models, we are abusing notation but this abuse will not cause any confusion as the contexts are clearly separated to different sections. The first of these expressions for the IDS is often called the *Pastur-Shubin trace formula*.

In the case of the model (1.1) in dimensions $d \geq 2$, there are no results in the literature on the smoothness of $\mathcal{N}(E)$, our results are the first to show even continuity of the density of states (DOS), which is the derivative of $\mathcal{N}$ almost every E. The results of Bovier et al. in [8] are quite strong for the Anderson model at large disorder and it is not clear that their proof using supersymmetry extends to other discrete random operators.

In the one dimensional Anderson model, Simon and Taylor [51] showed that $\mathcal{N}(E)$ is C^{∞} when the single site distribution (SSD) is compactly supported and is Hölder continuous. Subsequently, Campanino and Klein [9] proved that $\mathcal{N}(E)$ has the same degree of smoothness as the SSD. In the one dimensional strip, smoothness results were shown by Speis [53,54], Klein and Speis [38,39], Klein et al. [37], Glaffig [30]. For some non-stationary random potentials on the lattice, Krishna [41] proved smoothness for an averaged total spectral measure.

There are several results showing $\mathcal{N}(E)$ is analytic for the Anderson model on $\ell^2(\mathbb{Z}^d)$. Constantinescu et al. [16] showed analyticity of $\mathcal{N}(E)$ when SSD is analytic. The result of Carmona [10, Corollary VI.3.2] improved the condition on SSD to requiring fast exponential decay to get analyticity. In the case of the Anderson model over $\ell^2(\mathbb{Z}^d)$

at large disorder the results of Bovier et al. [8] give smoothness of $\mathcal{N}(E)$ when the Fourier transform $h(t)$ of the SSD is C^∞ and the derivatives decay like $1/t^\alpha$ for some $\alpha > 1$ at infinity. They also give variants of these, in particular if the SSD is C^{n+d} then $\mathcal{N}(E)$ is C^n under mild conditions on its decay at ∞. They also obtain some analyticity results. Acosta and Klein [1] show that $\mathcal{N}(E)$ is analytic on the Bethe lattice for SSD close to the Cauchy distribution. While all these results are valid in the entire spectrum, Kaminaga et al. [32] showed local analyticity of $\mathcal{N}(E)$ when the SSD has an analytic component in an interval allowing for singular parts elsewhere, in particular for the uniform distribution. Analyticity results obtained by March and Sznitman [44] were similar to those of Campanino and Klein [9].

In all the above models, only when E varies in the pure point spectrum that regularity of $\mathcal{N}(E)$ beyond Lipshitz continuity is shown. This condition that E has to be in the pure point spectrum may not have been explicitly stated, but it turns out to be a consequence of the assumptions on disorder or assumptions on the dimension in which the models were considered. For the Cauchy distribution in the Anderson model on $\ell^2(\mathbb{Z}^d)$, Carmona and Lacroix [10] have a theorem showing analyticity in the entire spectrum. However, absence of pure point spectrum is only a conjecture in these models as of now. At the time of revision of this paper one of us Kirsch and Krishna [36] could show that in the Anderson model on the Bethe lattice analyticity of the density of states with Cauchy distribution is valid at all disorders as part of a more general result. This result in particular exhibits regularity of the density of states through the mobility edge in the Bethe lattice case.

In the case of random band matrices, with the random variables following a Gaussian distribution, Disertori and Lager [25], Disertori [22,23], Disertori et al. [24] have smoothness results for an appropriately defined density of states. Recently Chulaevsky [11] proved infinite smoothness for non-local random interactions.

For the one dimensional ergodic random operators IDS was shown to be log Hölder continuous by Craig and Simon [17]. Wegner proved Lipshitz continuity of the IDS for the Anderson model independent of disorder in the pioneering paper [58]. Subsequently there are numerous results giving the modulus of continuity of $\mathcal{N}(E)$, for independent random potential, showing its Lipschitz continuity. Combes et al. in [14] showed that for Random Scrödinger operators with independent random potentials, the modulus of continuity of $\mathcal{N}(E)$ is the same as that of the SSD. For non i.i.d potentials in higher dimensions there are some results on modulus of continuity for example that of Schlag [48] showing and by Bourgain and Klein [7] who show log Hölder continuity for the distribution functions of outer measures for a large class of random and non-random Schrödinger operators. We refer to these papers for more recent results on the continuity of $\mathcal{N}(E)$ not given in the books cited earlier.

The idea of proof of our Theorems is the following. Suppose we have a self-adjoint matrix A^ω of size N with i.i.d real valued random variables $\{\omega_1, \dots, \omega_N\}$ on the diagonal with each ω_j following the distribution $\rho(x)dx$. Then the average of the matrix elements of the resolvent of A^ω are given by

$$f(z) = \int (A^\omega - zI)^{-1}(i,i) \prod_{k=1}^{N} \rho(\omega_k) d\omega_k,$$

for any $z \in \mathbb{C}^+$. We take $z = E + i\epsilon,\ \epsilon > 0$, then we see that from the definitions, the function $(A^\omega - zI)^{-1}(i,i)$ can be written as a function of $\vec{\omega} - E\vec{1}$ and ϵ, namely,

$$F(\vec{\omega} - E\vec{1}, \epsilon) = (A^{\omega} - zI)^{-1}(i, i), \;\; \Phi(\vec{\omega}) = \prod_{i=1}^{N} \rho(\omega_i)$$
$$\vec{\omega} = (\omega_1, \omega_2, \dots, \omega_N), \;\; \vec{1} = (1, 1, \dots, 1).$$

Then it is clear that with $*$ denoting convolution of functions on $\mathbb{R}^N$ and setting $\tilde{F}_{\epsilon}(x) = F(-x, \epsilon)$,

$$\mathbb{E}\left((A^{\omega} - zI)^{-1}(i, i)\right) = (\tilde{F}_{\epsilon} * \Phi)(E\vec{1}).$$

Since convolutions are smoothing, we get the required smoothness as a function of E if one of the components $\tilde{F}_{\epsilon}$ or Φ is smooth on $\mathbb{R}^N$. Since we are assuming that each ρ has a degree of smoothness, which passes on to Φ, we get a smoothness result for operators with finitely many random variables having the above form.

Let us remark here that it is in this step, which is crucial for further analysis, that we need a *complete covering condition*, even for finite dimensional compressions of our random operators be they continuous or discrete.

If we were to replace A^{ω} by an operator with infinitely many random variables ω_i, we would encounter the problem of concluding smoothing properties of "convolutions of" functions of infinitely many variables. This is an important difficulty that needs to be solved.

One of the interesting aspects of the operator H^{ω} (or h^{ω}) we are dealing with is that there is a sequence of operators (denoted by A_k^{ω}), containing finitely many random variables ω_i, which converges to H^{ω} (or h^{ω}) in strong resolvent sense. Hence we can write the limit as a telescoping sum, namely,

$$(A^{\omega} - z)^{-1}(i, i) = (A_1^{\omega} - z)^{-1}(i, i) + \sum_{k=1}^{\infty}\Big[(A_{k+1}^{\omega} - z)^{-1}(i, i) - (A_k^{\omega} - z)^{-1}(i, i)\Big].$$

Since the operators appearing in the summands all contain finitely many ω_i their averages over the random variables can be written as convolutions of functions of finitely many variables ω_i. Then, most of the work in the proof is to show that quantities of the form

$$\left| \int \left[(A_{k+1}^{\omega} - z)^{-1}(i, i) - (A_k^{\omega} - z)^{-1}(i, i)\right] \left(\sum_{j=1}^{N_{k+1}} \frac{\partial}{\partial \omega_j}\right)^{l} \prod_{n=1}^{N_{k+1}} \rho(\omega_n) d\omega_n \right|$$

with N_k growing at most as a fixed polynomial in k, are summable in k. This is the part where we use the fact that we are working in the localized regime, where it is possible to show that they are exponentially small in k.

For the discrete case the procedure is relatively straight forward and there are no major technical difficulties to overcome, but in the continuous case, the infinite rank perturbations pose a problem, since the trace of the Borel–Stieltjes transform of the average spectral measures do not converge. We overcome this problem by renormalizing this transform appropriately. For our estimates to work, we have to use fractional moment bounds and also uniform bounds on the integrals of resolvents. Both of these are achieved because we have dissipative operators (up to a constant) whose resolvents can be written in terms of integrals of contraction semigroups.

As stated above, our proof is in the localized regime. The Anderson model was formulated by Anderson [5] who argued that there is no diffusion in these models for high disorder or at low energies. The corresponding spectral statement is that there is only pure point spectrum or only localization for these cases. In the one dimensional systems, where the results are disorder independent, localization was shown rigorously by Goldsheid et al. [31] for random Schrödinger operators and by Kunz and Souillard [43] for the Anderson model. For higher dimensional Anderson model the localization was proved simultaneously by Fröhlich et al. [26], Simon and Wolff [52], Delyon et al. [20] based on exponential decay shown by Fröhlich and Spencer [27] who introduced a tool called multiscale analysis in the discrete case. A simpler proof based on exponential decay of fractional moments was later given by Aizenman and Molchanov [4]. There are numerous improvements and extensions of localization results beyond these papers.

In the case of continuous models, Combes and Hislop [12, 14], Klopp [40], Germinet and Klein [28], Combes et al. [15], Bourgain and Kenig [6] and Germinet and De Bievre [29] provided proof of localization for different types of models. The fractional moment method was first extended to the continuous case in Aizenman et al. in [3] and later improved by Boutet de Monvel et al. [19].

We refer to Stollmann [55] for the numerous advances that followed on localization.

The rest of the article is divided into three parts. Section 2 has all the preliminary results, which will be used significantly for both the discrete and the continuous case. Section 3 will deal with the discrete case, where we use a method of proof which will be reused for the continuous case. The main result of Sect. 3 is Theorem 3.4 which in the case of Anderson tight binding model would prove the regularity of density of states. Finally in Sect. 4 we will deal with the random Schrödinger operators and the main result there is Theorem 4.4.

2. Some Preliminary Results

In this section we present some general results that are at the heart of the proofs of our theorems. These are Theorems 2.1 and 2.2. The latter theorem, stated for functions, gives a bound of the form

$$\left|\int \left(\frac{1}{x-w} - \frac{1}{x-z}\right) f(x)dx\right| \leq C_{f,s}|z-w|^s$$

for certain family of f. For operators, we need more work and need more uniformity for f.

The first theorem is quite general and is about random perturbations of self-adjoint operators and their smoothing properties of complex valued functions of the operators.

Theorem 2.1. *Consider a self-adjoint operator A on a separable Hilbert space $\mathscr{H}$ and let $\{T_n\}_{n=1}^N$, $N < \infty$ be bounded positive operators such that $\sum_{n=1}^N T_n = I$, where I denotes the identity operator on $\mathscr{H}$. Suppose $\{\omega_n, n = 1, \ldots, N\}$ are independent real valued random variables distributed according to $\rho_n(x)dx$ and consider the random operators $A^\omega = A + \sum_{n=1}^N \omega_n T_n$. If f is a complex valued function on the set of linear operators on $\mathscr{H}$, such that $f(A^\omega - EI)$ is a bounded measurable function of $(\omega_1, \ldots, \omega_n, E)$, then $h(E) = \mathbb{E}\big[f(A^\omega - EI)\big]$ satisfies $h \in C^m(\mathbb{R})$, if $\rho_n \in C^m(\mathbb{R})$ and $\rho_n^{(k)} \in L^1(\mathbb{R})$, $n = 1, 2, \ldots, N$ and $0 \leq k \leq m$.*

Proof. Using the conditions on $\{T_n\}$ we see that $A^\omega - EI = A + \sum_{n=1}^N (\omega_n - E)T_n$. Thus $f(A^\omega - EI)$ is a bounded measurable function of the variables $(\omega_1 - E, \omega_2 - E, \ldots, \omega_n - E)$, which is a point $\vec{\omega} - E\vec{1}$ in $\mathbb{R}^N$, where $\vec{1} = (1, \ldots, 1)$, we write

$F(\vec{\omega} - E\vec{1}) = f(A^\omega - EI)$. Then the expectation can be written as

$$\mathbb{E}[f(A^\omega - EI)] = \int_{\mathbb{R}^N} F(\vec{\omega} - E\vec{1})\Phi(\vec{\omega})d\vec{\omega} = \int_{\mathbb{R}^N} F(-(E\vec{1} - \vec{\omega}))\Phi(\vec{\omega})d\vec{\omega},$$

where we set $\Phi(\vec{\omega}) = \prod_{n=1}^N \rho_n(\omega_n)$. Writing now $g(\vec{x}) = F(-\vec{x})$ we see that

$$\mathbb{E}[f(A^\omega - EI)] = (g * \Phi)(E\vec{1}),$$

where $*$ denotes convolution in $\mathbb{R}^N$. The result now follows easily from the properties of convolution of functions on $\mathbb{R}^N$. □

For later use we note that if ∇ denotes the gradient operator on differentiable functions on $\mathbb{R}^N$ and $\mathbf{D}$ denotes $\mathbf{D}\Phi = \nabla\Phi \cdot \vec{1} = \sum_{j=1}^N \frac{\partial}{\partial x_i}\Phi$, then an integration by parts yields

$$\frac{d^\ell}{dE^\ell}h(E) = \frac{d^\ell}{dE^\ell}(g * \Phi)(E\vec{1}) = (g * (\mathbf{D}^\ell\Phi))(E\vec{1}). \tag{2.1}$$

Remark 2.1. This theorem clarifies why the complete covering condition is needed in main our results for the discrete and the continuous models. The covering property is needed even for obtaining smoothness of finitely many random perturbations of a self-adjoint operator, while such a property is not needed for modulus of continuity results. We are unsure at the moment if this condition can be relaxed.

Let A, B be self-adjoint operators and let F_1, F_2 be bounded non-negative operators on a separable Hilbert space $\mathscr{H}$. For $X \in \{A, B\},\ z \in \mathbb{C}^+$, set,

$$R(X, x, y, z) = (X + xF_1 + yF_2 - z)^{-1}$$

and

$$R(X, x, z) = (X + xF - z)^{-1}, \ \ F = F_1 + F_2$$

for the following Theorem. For the rest of the paper by a *smooth indicator function* on an interval (a, b) we mean a smooth function which is one in $[c, d] \subset (a, b)$ which vanishes on $\mathbb{R}\backslash(a, b)$ with $a - c + b - d$ as small as one wishes.

Theorem 2.2. *Suppose A, B, F_1, F_2, F, z and $\mathscr{H}$ be as above. Suppose ρ_1, ρ_2 are compactly supported functions on $\mathbb{R}^+$ such that their derivatives are τ-Hölder continuous and their supports are contained in* $(0, \mathbf{R})$. *Let $\chi_{\mathbf{R}}$ denote a smooth indicator function of the set* $(0, 2\mathbf{R} + 1)$ *and let* $\phi_{\mathbf{R}}(x) = \chi_{\mathbf{R}}(x + \frac{5}{2}\mathbf{R} + 1)$. *Then for any* $0 < s < \tau$ *and some constant Ξ (depending upon ρ_1, ρ_2, s, τ but independent of z, A, F_1, F_2),*

1.
$$\left\| \int F^{\frac{1}{2}}\Big(R(A, x_1, x_2, z) - R(B, x_1, x_2, z)\Big)F^{\frac{1}{2}}\rho_1(x_1)\rho_2(x_2)dx_1dx_2 \right\|$$
$$\le \Xi \int \left\| F^{\frac{1}{2}}\Big(R(A, x_1, x_2, z) - R(B, x_1, x_2, z)\Big)F^{\frac{1}{2}} \right\|^s,$$
$$\times\, \phi_{\mathbf{R}}(x_1)\phi_{\mathbf{R}}(x_2)dx_1dx_2. \tag{2.2}$$

2. Specializing to the case when $F_1 = F_2$, $x_1 = x_2 = x/2$ we have

$$\left\| \int F^{\frac{1}{2}} \Big(R(A,x,z) - R(B,x,z) \Big) F^{\frac{1}{2}} \rho_1(x) dx \right\| \leq \Xi \int \left\| F^{\frac{1}{2}} \Big(R(A,x,z) - R(B,x,z) \Big) F^{\frac{1}{2}} \right\|^s \phi_{\mathbf{R}}(x)\, dx. \tag{2.3}$$

Remark 2.3. The integrals appearing in (2.2) and (2.3) are viewed as operators in the sense of direct integrals (see [47, Theorem XIII.85]). This is the case because $X + x_1 F_1 + x_2 F_2$ is decomposable on

$$L^2\big(\mathbb{R}^2, \prod_i \rho(x_i) dx_i, \mathcal{H}\big).$$

Hence all the integrals of this operator valued function, that appear in the proof, are well-defined in the sense of direct integral representation [42].

Proof. We define

$$A^t = A + t\ F, \quad B^t = B + t\ F, \qquad \forall -2\mathbf{R} - 1 < t < -2\mathbf{R}.$$

Then, we have the equality,

$$A + x_1 F_1 + x_2 F_2 = A^t + \left(\frac{x_1 - x_2}{2} \right) (F_1 - F_2) + \left(\frac{x_1 + x_2}{2} - t \right) F. \tag{2.4}$$

Using the resolvent equation, we have, with $F_- = F_1 - F_2$,

$$R(A, x_1, x_2, z) = \left(A^t + \left(\frac{x_1 - x_2}{2} \right) F_- - z \right)^{-1} - \left(\frac{x_1 + x_2}{2} - t \right) R(A, x_1, x_2, z) F \left(A^t + \left(\frac{x_1 - x_2}{2} \right) F_- - z \right)^{-1} \tag{2.5}$$

which can be re-written (using the notation $\tilde{A}^t = A^t + \left(\frac{x_1 - x_2}{2}\right) F_-$) as

$$\sqrt{F} R(A, x_1, x_2, z) \sqrt{F} = \frac{1}{\frac{x_1+x_2}{2} - t} I - \frac{1}{\left(\frac{x_1+x_2}{2} - t\right)^2} \left(\frac{1}{\frac{x_1+x_2}{2} - t} I + \sqrt{F} \left(\tilde{A}^t - z \right)^{-1} \sqrt{F} \right)^{-1}. \tag{2.6}$$

(I is the identity operator on the range of $\sqrt{F}$) Similar relations hold for B, where B^t, $\tilde{B}^t$ are defined by replacing A with B in the Eqs. (2.4–2.6). We set

$$\tilde{R}^t_{A,z} = \sqrt{F}(\tilde{A}^t - z)^{-1}\sqrt{F}, \quad \tilde{R}^t_{B,z} = \sqrt{F}(\tilde{B}^t - z)^{-1}\sqrt{F}.$$

Then using Eq. (2.6) we get the relation,

$$\int \sqrt{F}(R(A, x_1, x_2, z) - R(B, x_1, x_2, z))\sqrt{F}\ \rho_1(x_1)\rho_2(x_2)dx_1dx_2$$
$$= \int \left[\left(\frac{1}{\frac{x_1+x_2}{2} - t} I + \tilde{R}^t_{A,z}\right)^{-1} - \left(\frac{1}{\frac{x_1+x_2}{2} - t} I + \tilde{R}^t_{B,z}\right)^{-1}\right]$$
$$\frac{1}{\left(\frac{x_2+x_2}{2} - t\right)^2}\rho_1(x_1)\rho_2(x_2)dx_1dx_2$$
$$= 2\int \left[\left(\gamma I + \tilde{R}^t_{A,z}\right)^{-1} - \left(\gamma I + \tilde{R}^t_{B,z}\right)^{-1}\right]$$
$$\rho_1\left(t + \frac{1}{\gamma} + \eta\right)\rho_2\left(t + \frac{1}{\gamma} - \eta\right)d\gamma d\eta \tag{2.7}$$

where $\gamma = \left(\frac{x_1+x_2}{2} - t\right)^{-1}$ and $\eta = \frac{x_1-x_2}{2}$. For X self-adjoint, $\tilde{R}^t_{X,z}$ is an operator valued Herglotz function and its imaginary part is a positive operator for $\Im(z) > 0$. Hence the operators $\left(\gamma I + \tilde{R}^t_{X,z}\right)$ generate a strongly continuous one parameter semi-group, and we can apply the Lemma A.3 for the γ integral, and then do the η integral to get

$$\int \left[\left(\gamma I + \tilde{R}^t_{A,z}\right)^{-1} - \left(\gamma I + \tilde{R}^t_{B,z}\right)^{-1}\right]$$
$$\rho_1\left(t + \frac{1}{\gamma} + \eta\right)\rho_2\left(t + \frac{1}{\gamma} - \eta\right)d\gamma d\eta$$
$$= -\int \left[\int_0^\infty \left(e^{iw\left(\gamma I + \tilde{R}^t_{A,z}\right)} - e^{iw\left(\gamma I + \tilde{R}^t_{B,z}\right)}\right)dw\right]$$
$$\rho_1\left(t + \frac{1}{\gamma} + \eta\right)\rho_2\left(t + \frac{1}{\gamma} - \eta\right)d\gamma d\eta$$
$$= -\int\int_0^\infty \left(e^{iw\tilde{R}^t_{A,z}} - e^{iw\tilde{R}^t_{B,z}}\right)\ e^{i\gamma w}\rho_1\left(t + \frac{1}{\gamma} + \eta\right)\rho_2\left(t + \frac{1}{\gamma} - \eta\right)d\gamma\ dwd\eta, \tag{2.8}$$

which can be bounded as

$$\Bigg\| \int\int_0^\infty \left[e^{iw\tilde{R}^t_{A,z}} - e^{iw\tilde{R}^t_{B,z}}\right]$$
$$e^{i\gamma w}\rho_1\left(t + \frac{1}{\gamma} + \eta\right)\rho_2\left(t + \frac{1}{\gamma} - \eta\right)d\gamma\ dwd\eta\Bigg\|$$
$$\leq \int \left\|\left(e^{iw\tilde{R}^t_{A,z}} - e^{iw\tilde{R}^t_{B,z}}\right)\right\|$$
$$\left|\int e^{i\gamma w}\rho_1\left(t + \frac{1}{\gamma} + \eta\right)\rho_2\left(t + \frac{1}{\gamma} - \eta\right)d\gamma\right|dwd\eta. \tag{2.9}$$

The assumption we made on the supports of ρ_1, ρ_2 implies that $-\frac{\mathbf{R}}{2} < \eta < \frac{\mathbf{R}}{2}$, and the choice $-2\mathbf{R} - 1 < t < -2\mathbf{R}$ implies $-\frac{5}{2}\mathbf{R} - 1 < t \pm \eta < -\frac{3\mathbf{R}}{2}$. This implies that

$$\left\{\gamma : \psi_{t,\eta}(\gamma) \neq 0, -\frac{5}{2}\mathbf{R} - 1 < t \pm \eta < -\frac{3\mathbf{R}}{2}\right\} \subset \left(\frac{2}{2 + 7\mathbf{R}}, \frac{2}{3\mathbf{R}}\right),$$

where $\psi_{t,\eta}(\gamma) = \rho_1\left(t + \frac{1}{\gamma} + \eta\right)\rho_2\left(t + \frac{1}{\gamma} - \eta\right)$. Thus for fixed t, η, the function $\psi_{t,\eta}(\gamma)$ is of compact support and has a τ-Hölder continuous derivative as a function of γ, for the τ stated as in the Theorem. Also, the derivative of $\psi_{t,\eta}$ is uniformly τ-Hölder continuous and the constant in the corresponding bound is uniform in t, η, which follows from the support properties of $\psi_{t,\eta}$ and the bounds on t, η. Therefore, if we denote the Fourier transform of $\psi_{t,\eta}(-\gamma)$ by $\widehat{\psi_{t,\eta}}$, then standard Fourier analysis gives the bound,

$$\left|\int e^{i\gamma w}\rho_1\left(t + \frac{1}{\gamma} + \eta\right)\rho_2\left(t + \frac{1}{\gamma} - \eta\right)d\gamma\right|$$
$$\leq \frac{C}{|w|^{1+\tau}}\left(\left\||w|^{1+\tau}\widehat{\psi_{t,\eta}}(w)\right\|_\infty\right) \leq \frac{\tilde{C}}{|w|^{1+\tau}} \quad for\ |w| \gg 1$$

for some $\tilde{C}$ depends on ρ_1, ρ_2 but not on t, η.

Again using the bounds on t, η and γ, we see that for small $|w|$, the w integral is bounded uniformly in t, η, by the L^∞ norm of ρ_1 and ρ_2 and hence $\tilde{C}$ is (t, η)-independent for all w.

On other hand using the Lemma A.2, we have

$$\left\|e^{iw\tilde{R}^t_{A,z}} - e^{iw\tilde{R}^t_{B,z}}\right\| \leq 2^{1-s}|w|^s\left\|\tilde{R}^t_{A,z} - \tilde{R}^t_{B,z}\right\|^s$$

for $0 < s < 1$. By choosing $s < \tau/2$ and using above bounds in (2.9) we have

$$\left\|\int\int_0^\infty\left(e^{iw\tilde{R}^t_{A,z}} - e^{iw\tilde{R}^t_{B,z}}\right)\right.$$
$$\left. e^{i\gamma w}\rho_1\left(t + \frac{1}{\gamma} + \eta\right)\rho_2\left(t + \frac{1}{\gamma} - \eta\right)d\gamma\, dwd\eta\right\| \tag{2.10}$$
$$\leq \hat{C}\left(1 + \int_1^\infty \frac{1}{w^{1+\tau-s}}dw\right)\int\left\|\tilde{R}^t_{A,z} - \tilde{R}^t_{B,z}\right\|^s d\eta. \tag{2.11}$$

The integral we started with is independent of t so we can integrate it with respect to the Lebesgue measure on an interval of length one. Therefore, combining the inequalities (2.7, 2.8, 2.9, 2.10) and integrating t over an interval of length 1, yields

$$\left\|\int\sqrt{F}(R(A, x_1, x_2, z) - R(B, x_1, x_2, z)\sqrt{F}\ \rho_1(x_1)\rho_2(x_2)dx_1dx_2\right\|$$
$$= \int_{-2R-1}^{-2R}\left\|\int\sqrt{F}(R(A, x_1, x_2, z) - R(B, x_1, x_2, z)\sqrt{F}\ \rho_1(x_1)\rho_2(x_2)dx_1dx_2\right\|dt$$
$$\leq C\int_{-2R-1}^{-2R}\int_{-\frac{R}{2}}^{\frac{R}{2}}\left\|\tilde{R}^t_{A,z} - \tilde{R}^t_{B,z}\right\|^s d\eta dt$$
$$\leq C\int\int\left\|\sqrt{F}\left(A + \hat{x}_1F_1 + \hat{x}_2F_2 - z\right)^{-1}\sqrt{F}\right.$$
$$\left. - \sqrt{F}\left(B + \hat{x}_1F_1 + \hat{x}_2F_2 - z\right)^{-1}\sqrt{F}\right\|^s \phi_{\mathbf{R}}(\hat{x}_1)\phi_{\mathbf{R}}(\hat{x}_2)\ d\hat{x}_1d\hat{x}_2.$$

For the last inequality we used the definition of $\tilde{R}^t_{X,z}$ changed variables $\hat{x}_1 = t+\eta$, $\hat{x}_2 = t - \eta$ along with a slight increase in the range of integration to accommodate the bump $\phi_{\mathbf{R}}$ to have their supports in $(-\frac{5}{2}\mathbf{R} - 1, -\frac{\mathbf{R}}{2})$. □

3. The Discrete Case

Let $\mathbb{G}$ denote a un-directed connected graph with a graph-metric d. Let $\{x_n\}_n$ denote an enumeration of $\mathbb{G}$ satisfying $d(\Lambda_N, x_{N+1}) = 1$ for any $N \in \mathbb{N}$, where

$$\Lambda_N = \{x_n : n \leq N\}, \quad \Lambda_\infty = \mathbb{G}, \tag{3.1}$$

and

$$\liminf_{N\to\infty} \frac{d(x_0, \mathbb{G}\backslash\Lambda_N)}{g(N)} = r_{\mathbb{G}} > 0, \tag{3.2}$$

for some increasing function g on $\mathbb{R}^+$. Typically, we will have $g(N) = N^{1/d}$ for $\mathbb{G} = \mathbb{Z}^d$ and $g(N) = \log_K(N)$ for the Bethe lattice with connectivity $K > 2$. Henceforth for indexing $\mathbb{G}$ we will say $n \in \mathbb{G}$ to mean $x_n \in \mathbb{G}$.

Let $\mathscr{H}$ be a complex separable Hilbert space equipped with a countable family $\{P_n\}_{n\in\mathbb{G}}$ of finite rank orthogonal projections such that $\sum_{n\in\mathbb{G}} P_n = Id$, with the maximum rank of P_n being finite, thus

$$\mathscr{H} = \bigoplus_{n\in\mathbb{G}} Ran(P_n).$$

Let h_0 denote a bounded self-adjoint operator on $\mathscr{H}$ and consider the random operator, we stated in Eq. (1.2),

$$h^\omega = h_0 + \sum_{n\in\Lambda_\infty} \omega_n P_n, \tag{3.3}$$

where the random variables ω_n satisfies Hypothesis 3.1 below. Given a finite subset $\Lambda \subset \mathbb{G}$, we will denote $P_\Lambda = \sum_{n\in\Lambda} P_n$, $\mathscr{H}_\Lambda = P_\Lambda \mathscr{H}$ and

$$h^\omega_\Lambda = P_\Lambda h^\omega P_\Lambda \tag{3.4}$$

denotes the restriction of h^ω to $\mathscr{H}_\Lambda$.

We abused notation to denote P for two different objects, P_n denoting projections onto sites $x_n \in \mathbb{G}$ and P_Λ to denote the sum of P_n when x_n varies in Λ, but we are sure the reader will not be confused and the meaning would be clear from the context.

We have the following assumptions on the quantities involved in the model.

Hypothesis 3.1. *We assume that the random variables ω_n are independent and distributed according to a density ρ_n which are compactly supported in* $(0, 1)$*and satisfy $\rho_n \in C^m((0, 1))$ for some $m \in \mathbb{N}$ and*

$$\mathcal{D} = \sup_n \max_{\ell\leq m} \|\rho_n^{(\ell)}\|_\infty < \infty. \tag{3.5}$$

We note that as long as $\rho_n \in C^m((a, b))$ for some $-\infty < a < b < \infty$, a scaling and translation will move its support to $(0, 1)$. So our support condition is no loss of generality.

Hypothesis 3.2. *A compact interval $J \subset\subset \mathbb{R}$ is said to be in* region of localization *for h^ω with exponent* $0 < s < 1$ *and rate of decay $\xi_s > 0$, if there exist $C > 0$ such that*

$$\sup_{\Re(z)\in J, \Im(z)>0} \mathbb{E}_\omega\left[\left\| P_n(h^\omega - z)^{-1} P_k \right\|^s\right] \leq Ce^{-\xi_s d(n,k)} \tag{3.6}$$

for any $n, k \in \mathbb{G}$. For the operators $h^{\omega}_{\Lambda_K}$ exponential localization is defined with $\Lambda, h^{\omega}_{\Lambda_K}, \xi_{s,\Lambda_K}$ replacing $\mathbb{G}, h^{\omega}, \xi_s$ respectively in the above bound.

We assume that for our models, for all Λ_K, with $K \geq N$ the inequality (3.6) holds for some $\xi_s > 0$ and $\xi_{s,\Lambda_K} \geq \xi_s$, for all Λ_K with $K \geq N$. We also assume that the constants C, ξ_s do not change if we replace the distribution ρ_n with one of its derivatives at finitely many sites n.

Remark 3.3. For large disorder models one can get explicit values for ξ_s from the papers of Aizenman and Molchanov [4] or Aizenman [2]. For example the Anderson model on $\ell^2(\mathbb{Z}^d)$ with disorder parameter $\lambda >> 1$, typically $\xi_s = -s \ln \frac{C_{s,\rho} 2d}{\lambda}$, for some constant $C_{s,\rho} < \infty$ that depends on the single-site density ρ and is independent of Λ. So $\xi_{s,\Lambda} = \xi_s > 0$ for large enough λ. Similarly for the Bethe lattice with connectivity $K + 1 > 1$, $\xi_{s,\Lambda} = \xi_s = -s \ln \frac{C_{s,\rho}(K+1)}{\lambda}$. Going through Lemma 2.1 of their paper, and tracing through the constants, we see that our assumption about changing the distribution at finitely many sites is valid.

Henceforth let $E_A(\cdot)$ denote the projection valued spectral measure of a self-adjoint operator A. Our main goal in this section is to show that

$$\mathcal{N}(E) = \mathop{\mathbb{E}}_{\omega}[tr(P_0 E_{h^{\omega}}(-\infty, E))]$$

is m times differentiable in the region of localization, if ρ has a bit more than m derivatives, which means that the density of states DOS is $m - 1$ times differentiable. Our theorem is the following, where we tacitly assume that the spectrum $\sigma(h^{\omega})$ is a constant set a.s., a fact proved by Pastur [45] for a large class of random self-adjoint operators. While it may not be widely known, it is also possible to show the constancy of spectrum for operators that do not have ergodicity but when there is independent randomness involved see for example Kirsch et al. [34]. In such non-ergodic cases when there is no limiting eigenvalue distribution, our results are still valid for the spectral measures considered.

Theorem 3.4. *Consider the random self-adjoint operators h^{ω} given in Eq. (3.3) on the Hilbert space $\mathscr{H}$ and a graph $\mathbb{G}$ satisfying the condition (3.2) with $g(N) = N^{\alpha}$, for some $\alpha > 0$. We assume that ω_n is distributed with density ρ_n satisfying the Hypothesis 3.1 and, with m as in the Hypothesis, $\rho_n^{(m)}$ is τ-Hölder continuous for some $0 < \tau < 1$. Assume that J is an interval in the region of localization for which the Hypothesis 3.2 holds for some $0 < s < \tau$. Then the function*

$$\mathcal{N}(E) = \mathop{\mathbb{E}}_{\omega}[tr(P_0 E_{h^{\omega}}(-\infty, E))] \in C^{m-1}(J) \tag{3.7}$$

and $\mathcal{N}^{(m)}(E)$ exists a.e. $E \in J$.

Remark 3.5. 1. We stated the Theorem in this generality so that it applies to multiple models, such as the Anderson models on $\mathbb{Z}^d$, other lattice or graphs, having the property that the number of points at a distance N from any fixed point grow polynomially in N. The models for which this Theorem is valid also include higher rank Anderson models, long range hopping with some restrictions, models with off-diagonal disorder to state a few. In all of these models, by including sufficiently high diagonal disorder, through a coupling constant λ on the diagonal part, we will have exponential localization for the corresponding operators via the Aizenman-Molchanov method. So this Theorem gives the Regularity of DOS in all such models. For the

Bethe lattice and other countable sets for which $g(N)$ is like $\ln(N)$, our results hold but the order of smoothness m that can be obtained is restricted by the localization length by a condition such as $\xi_s > m \ln K$. So in this work we do not consider such type of setting.

2. This Theorem also gives smoothness of DOS in the region of localization for the intermediate disorder cases considered for example by Aizenman [2] who exhibited exponential localization for such models in part of the spectrum.
3. In the case h^ω is not the Anderson model, all these results are new and it is not clear that the method of proof using super symmetry, as done for the Anderson model at high disorder, will even work for these models.
4. We note that in the proof we will take at most $m-1$ derivatives of resolvent kernels in the upper half-plane and show their boundedness, but we have a condition that the function ρ has a τ-Hölder continuous derivative. The extra $1+\tau$ 'derivatives' are needed for applying the Theorem 2.2 to obtain the inequality (3.19) from the equality (3.18).

Proof. Since the orthogonal projection P_0 is finite rank, we can write $P_0 = \sum_{i=1}^{r} |\phi_i\rangle\langle\phi_i|$ using a set $\{\phi_i\}$ of finitely many orthonormal vectors. Then we have,

$$\mathcal{N}(E) = \sum_{i=1}^{r} \mathop{\mathbb{E}}_{\omega}\left(\langle\phi_i, E_{h^\omega}((-\infty, E))\phi_i\rangle\right).$$

The densities of the measures $\langle\phi_i, E_{h^\omega}(\cdot)\phi_i\rangle$ are bounded by Lemma A.4 for each $i = 1, \ldots, r$. Hence $\mathcal{N}$ is differentiable almost everywhere and its derivative, almost everywhere, is given by the boundary values,

$$\frac{1}{\pi}\mathop{\mathbb{E}}_{\omega}\left(tr\left(P_0 \Im(h^\omega - E - i0)^{-1}\right)\right)$$

is bounded. The Theorem follows from Lemma A.1 once we show

$$\sup_{\Re(z)\in J, \Im(z)>0} \frac{d^\ell}{dz^\ell}\mathop{\mathbb{E}}_{\omega}\left[tr(P_0(h^\omega - z)^{-1})\right] < \infty, \tag{3.8}$$

for all $\ell \le m-1$, since such a bound implies that $m-1$ derivatives of η are continuous and its mth derivative exists almost everywhere, since h^ω are bounded operators. The projection P_0 is finite rank which implies that the bounded operator valued analytic functions $P_0(h^\omega - z)^{-1}$, $P_0(h^\omega_\Lambda - z)^{-1}$ are trace class for $z \in \mathbb{C}^+$. Therefore the linearity of the trace and the dominated convergence theorem together imply that

$$\mathop{\mathbb{E}}_{\omega}\left[tr(P_0(h^\omega - z)^{-1} - P_0(h^\omega_\Lambda - z)^{-1})\right] \xrightarrow{\Lambda\to\mathbb{G}} 0, \tag{3.9}$$

compact uniformly in $\mathbb{C}^+$. For the rest of the proof we set $h^\omega_K = h^\omega_{\Lambda_K}$ for ease of writing.

The convergence given in Eq. (3.9) implies that the telescoping sum,

$$\begin{aligned}\mathop{\mathbb{E}}_{\omega}\left[tr(P_0(h^\omega_M - z)^{-1})\right] = \sum_{K=N}^{M}\left(\mathop{\mathbb{E}}_{\omega}\left[tr(P_0(h^\omega_{K+1} - z)^{-1})\right] - \mathop{\mathbb{E}}_{\omega}\left[tr(P_0(h^\omega_K - z)^{-1})\right]\right)\\ + \mathop{\mathbb{E}}_{\omega}\left[tr(P_0(h^\omega_N - z)^{-1})\right]\end{aligned}$$

also converges compact uniformly, in $\mathbb{C}^+$ to

$$\mathbb{E}_{\omega}\Big[tr(P_0(h^{\omega}-z)^{-1})\Big],$$

which implies that their derivatives of all orders also converge compact uniformly in $\mathbb{C}^+$.

Therefore the inequality (3.8) follows if we prove the following uniform bound, for all $0 \le \ell \le m-1$ and N large,

$$\sum_{K=N}^{\infty} \sup_{\Re(z)\in J} \left| \frac{d^{\ell}}{dz^{\ell}} \left(\mathbb{E}_{\omega}\Big[tr(P_0(h^{\omega}_{K+1}-z)^{-1}) - \mathbb{E}_{\omega}\Big[tr(P_0(h^{\omega}_{K}-z)^{-1})\Big] \right) \Big] \right| < \infty. \tag{3.10}$$

To this end we only need to estimate

$$\left| \frac{d^{l}}{dz^{l}} \mathbb{E}_{\omega}\Big[tr(P_0(h^{\omega}_{K+1}-z)^{-1}P_0) - tr(P_0(h^{\omega}_{K}-z)^{-1}P_0)\Big] \right| \tag{3.11}$$

for $\Re(z) \in J$ where we used the trace property to get an extra P_0 on the right and set $G^{\omega}_M(z) = P_0(h^{\omega}_M - z)^{-1}P_0, \quad M \in \mathbb{N}$ for further calculations.

The function

$$f_{\epsilon}(\vec{\omega} - E\vec{1}) = tr(G^{\omega}_K(E+i\epsilon))$$

is a complex valued bounded measurable function on $\mathbb{R}^{K+1}$ for each fixed $\epsilon > 0$. Therefore we compute the derivatives in E of its expectation

$$h_{\epsilon}(E) = \mathbb{E}_{\omega}\big(f_{\epsilon}(\vec{\omega} - E\vec{1})\big) = \mathbb{E}\big(tr(G^{\omega}_K(E+i\epsilon)^{-1})\big)$$

using Theorem 2.1. This calculation gives in the notation of that Theorem,

$$\frac{d^{\ell}}{dE^{\ell}} \mathbb{E}_{\omega}\big(tr(G^{\omega}_K(E+i\epsilon))\big) = \int tr(G^{\omega}_K(E+i\epsilon))\mathbf{D}^{\ell}\Phi_K(\vec{\omega})d\vec{\omega}, \tag{3.12}$$

where we set $\Phi_K(\vec{\omega}) = \prod_{n\in\Lambda_K} \rho_n(\omega_n), \ d\vec{\omega} = \prod_{n\in\Lambda_K} d\omega_n.$

It is not hard to see that for each $0 \le \ell \le m-1$,

$$\int tr(G^{\omega}_K(E+i\epsilon))\mathbf{D}^{\ell}\Phi_K(\vec{\omega})d\vec{\omega}, = \int tr(G^{\omega}_K(E+i\epsilon))\mathbf{D}^{\ell}\Phi_{K+1}(\vec{\omega})d\vec{\omega}, \tag{3.13}$$

since the integrand $tr(G^{\omega}_K(E+i\epsilon))$ is independent of $\omega_n, n \in \Lambda_{K+1}\backslash\Lambda_K$ and ρ_n satisfies $\int \rho_n^{(j)}(x)dx = \delta_{j0}$. We set

$$R(\omega, K, E, \epsilon) = tr\left(G^{\omega}_{K+1}(E+i\epsilon) - G^{\omega}_K(E+i\epsilon)\right) \tag{3.14}$$

to simplify writing. We may write the argument ω of $R(\omega, K, E, \epsilon)$ below in terms of the vector notation $\vec{\omega}$ for uniformity as it is a function of the variables $\{\omega_n, n \in \Lambda_{K+1}\}$.

Then combining the Eqs. (3.12, 3.13) inside the absolute value of the expression in Eq. (3.11) to be estimated we have to consider the quantity, for $K \ge N$,

$$T_{K,\ell}(E,\epsilon) = \frac{d^\ell}{dE^\ell} \mathop{\mathbb{E}}_{\omega}\left[tr\left(G^\omega_{K+1}(E+i\epsilon) - G^\omega_K(E+i\epsilon)\right)\right]$$
$$= \int_{\mathbb{R}^{K+1}} R(\vec{\omega}, K, E, \epsilon)(\mathbf{D}^\ell \Phi_{K+1})(\vec{\omega})d\vec{\omega}. \tag{3.15}$$

To prove the theorem we need to show that

$$\sum_{K=N}^{\infty} \sup_{E\in J, \epsilon>0} |T_{K,\ell}(E,\epsilon)| < \infty. \tag{3.16}$$

Multinomial expansion of $\mathbf{D}^\ell = \left(\sum_{n\in\Lambda_{K+1}} \frac{\partial}{\partial\omega_n}\right)^\ell$ gives the relation

$$T_{K,\ell}(E,\epsilon) = \sum_{\substack{k_0+\cdots+k_K=\ell \\ k_n\geq 0}} \binom{\ell}{k_0,\ldots,k_K} \int_{\mathbb{R}^{K+1}} R(\vec{\omega}, K, E, \epsilon)\left(\prod_{n=0}^{K+1} \frac{\partial^{k_n}}{\partial^{k_n}\omega_n}\rho_n(\omega_n)d\omega_n\right). \tag{3.17}$$

We use Fubini to interchange the trace and an integral over ω_0 to get

$$T_{K,\ell}(E,\epsilon)$$
$$= \sum_{\substack{k_0+\cdots+k_K=\ell \\ k_n\geq 0}} \binom{\ell}{k_0,\ldots,k_K} \int_{\mathbb{R}^K} tr\left(\int \left(G^\omega_{K+1}(E+i\epsilon) - G^\omega_K(E+i\epsilon)\right)\rho_0^{(k_0)}(\omega_0)d\omega_0\right)$$
$$\times\left(\prod_{n\in\Lambda_{K+1}, n\neq 0} \rho_n^{(k_n)}(\omega_n)d\omega_n\right). \tag{3.18}$$

We take the absolute value of T and estimate the ω_0 integrals using the Theorem 2.2, displaying explicitly the dependence on the ρ or its derivatives in the constant Ξ appearing in that theorem, to get, for $0 < s < 1/2$ (the choice for s will become clear in Lemma 3.1),

$$T_{K,\ell}(E,\epsilon) \leq \sum_{\substack{k_0+\cdots+k_K=\ell \\ k_n\geq 0}} \binom{\ell}{k_0,\ldots,k_K} \Xi(\rho_0^{(k_0)}) tr(P_0)$$
$$\times \int_{\mathbb{R}^K}\left(\int \|\left(G^\omega_{K+1}(E+i\epsilon) - G^\omega_K(E+i\epsilon)\right)\|^s \phi_{\mathbf{R}}(\omega_0)d\omega_0\right)$$
$$\times\left(\prod_{n\in\Lambda_{K+1}, n\neq 0} |\rho_n^{(k_n)}(\omega_n)|d\omega_n\right). \tag{3.19}$$

We set

$$\tilde{\rho_n} = \frac{\rho_n^{(k_n)}}{\|\rho_n^{(k_n)}\|_1}, \quad n\neq 0, \quad \tilde{\rho_0} = \frac{\phi_{\mathbf{R}}}{\|\phi_{\mathbf{R}}\|_1} \tag{3.20}$$

and set, using the inequality (3.5), $C_0 = \max\{\mathcal{D}, \|\phi_{\mathbf{R}}\|_1\}$, where $\mathcal{D}$ is such that $\|\rho_n^{(k_n)}\|_1 \leq \mathcal{D}\ \forall\ n \neq 0$. We note here that at most ℓ of $\tilde{\rho}_n$ differ from ρ_n itself and that $\|\rho_n\|_1 = 1$. We then get the bound

$$|T_{K,\ell}(E,\epsilon)| \leq C_0^\ell \sup_{j\leq\ell} \Xi(\rho^{(j)}) tr(P_0) \sum_{\substack{k_0+\cdots+k_K=\ell \\ k_n\geq 0}} \binom{\ell}{k_0,\ldots,k_K} \times \int_{\mathbb{R}^{K+1}} \|\big(G^\omega_{K+1}(E+i\epsilon) - G^\omega_K(E+i\epsilon)\big)\|^s \prod_{n\in\Lambda_{K+1}} \tilde{\rho}_n(\omega_n) d\omega_n. \tag{3.21}$$

We denote the probability measure

$$d\mathbb{P}_K(\vec{\omega}) = \prod_{n\in\Lambda_K} \tilde{\rho}_n(\omega_n) d\omega_n,$$

and expectation as $\mathbb{E}_K$. We also set,

$$C_{1,m} = \sup_{0\leq\ell\leq m} \{C_0^\ell\},\ C_{2,m} = \sup_{n\in\mathbb{G},\, j\leq m} \{\Xi(\rho_n^{(j)})\}.$$

Then the inequality (3.21) becomes

$$|T_{K,\ell}(E,\epsilon)| \leq C_{1,m} C_{2,m} tr(P_0) \sum_{k_0+\cdots+k_K=\ell} \binom{\ell}{k_0,\ldots,k_K} \times \mathbb{E}_{K+1}\Big[\|\big(G^\omega_{K+1}(E+i\epsilon) - G^\omega_K(E+i\epsilon)\big)\|^s \Big]. \tag{3.22}$$

We use the estimate for the expectation $\mathbb{E}_{K+1}(\cdot)$ from Lemma 3.1 to get the following bound, for some constant C_6 independent of K,

$$\sup_{E\in J,\epsilon>0} |T_{K,\ell}(E,\epsilon)| \leq C_6 C_5 \sum_{k_0+\cdots+k_K=\ell} \binom{\ell}{k_0,\ldots,k_K} (1+2K) e^{-\xi_{2s} K^\alpha} \leq C_6 C_5 (K+1)^\ell (1+2K) e^{-\xi_{2s}|K|^\alpha}. \tag{3.23}$$

From this bound the summability stated in the inequality (3.16) follows since we assumed that $\xi_{2s} > 0$, completing the proof of the Theorem. □

We needed the exponential bound on the resolvent estimate, which is the focus of the following lemma.

Lemma 3.1. *We take the interval J stated in Theorem 3.4, then we have the bound*

$$\sup_{\Re(z)\in J, \Im(z)>0} \mathbb{E}_{K+1}\Big[\|\big(G^\omega_{K+1}(z) - G^\omega_K(z)\big)\|^s \Big] \leq C_5(m, Rank(P_0), \mathcal{D}, h_0, R, s)(2K+1) e^{-\xi_{2s} K^\alpha}.$$

Proof. We start with the resolvent identity

$$\begin{aligned} G^{\omega}_K(z) - G^{\omega}_{K+1}(z) &= P_0\Big((h^{\omega}_K - z)^{-1} - (h^{\omega}_{K+1} - z)^{-1}\Big)P_0 \\ &= P_0(h^{\omega}_K - z)^{-1}[h^{\omega}_{K+1} - h^{\omega}_K](h^{\omega}_{K+1} - z)^{-1}P_0 \\ &= P_0(h^{\omega}_{K+1} - z)^{-1}P_{\Lambda_K}h_0P_{K+1}(h^{\omega}_{K+1} - z)^{-1}P_0. \end{aligned} \tag{3.24}$$

In the above equation, the terms corresponding to the random part $\omega_{K+1}P_{K+1}$ and the part $P_{K+1}h_0P_{\Lambda_K}$ (appearing in the difference $[h^{\omega}_{K+1} - h^{\omega}_K]$) are zero, since they are multiplied by $P_0(h^{\omega}_K - z)^{-1}$ on the left and $P_0(h^{\omega}_K - z)^{-1}P_{K+1}$ being the operator $P_0(P_{\Lambda_K}h^{\omega}P_{\Lambda_K} - z)^{-1}P_{K+1}$ is obviously zero since $P_0P_{K+1} = 0$ if $K > 1$. It is to be noted that this fact is independent of how h_0 looks! We estimate the last line in the Eq. (3.24), by first by expanding $P_{\Lambda_K} = \sum_{n\in\Lambda_K}P_n$ and estimate the norms of the operators (using $\|B\sum_{i=1}^N A_i\|^s \le \|B\|^s\sum_{i=1}^N\|A_i\|^s$ for any finite collection $\{B, A_i, i = 1, \ldots, N\}$ of bounded operators and $0 < s < 1$) to get

$$\begin{aligned} \|\big(G^{\omega}_{K+1}(z) - G^{\omega}_K(z)\big)\|^s &\le \|h_0\|^s\|P_{K+1}(h^{\omega}_{K+1} - z)^{-1}P_0\|^s\|P_0(h^{\omega}_{K+1} - z)^{-1}P_{\Lambda_K}\|^s \\ &\le \|h_0\|^s\|P_{K+1}(h^{\omega}_{K+1} - z)^{-1}P_0\|^s \\ &\quad\times\sum_{n\in\Lambda_K}\|P_0(h^{\omega}_{K+1} - z)^{-1}P_n\|^s. \end{aligned} \tag{3.25}$$

We take expectation of both the sides of the above equation, then interchange the sum and the expectation on the right hand side and use Cauchy–Schwartz inequality to get the bound

$$\begin{aligned} &\mathbb{E}_{K+1}\big(\|\big(G^{\omega}_{K+1}(z) - G^{\omega}_K(z)\big)\|^s\big) \\ &\le \|h_0\|^s\sum_{n\in\Lambda_K}\big(\mathbb{E}_{K+1}\big(\|P_{K+1}(h^{\omega}_{K+1} - z)^{-1}P_0\|^{2s}\big)\big)^{\frac{1}{2}} \\ &\big(\mathbb{E}_{K+1}\big(\|P_0(h^{\omega}_{K+1} - z)^{-1}P_n\|^{2s}\big)\big)^{\frac{1}{2}}. \end{aligned} \tag{3.26}$$

We now estimate the above terms by getting an exponential decay bound for the term with operators kernels of the form $P_{K+1}[\cdot]P_0$ while the remaining factors are uniformly bounded with the bound independent of K, by using the Hypothesis 3.2.

Applying the bound on the fractional moments given in the Hypothesis 3.2, inequality 3.6 we get

$$\begin{aligned} &\mathbb{E}_{K+1}\big(\|P_n(h^{\omega}_K - z)^{-1}P_0\|^{2s}\big) \le C, \quad n \in \Lambda_K, \\ &\mathbb{E}_{K+1}\big(\|P_0(h^{\omega}_{K+1} - z)^{-1}P_n\|^{2s}\big) \le C, \quad n \in \Lambda_K \\ &\mathbb{E}_{K+1}\big(\|P_0(h^{\omega}_{K+1} - z)^{-1}P_{K+1}\|^{2s}\big) \le Ce^{-\xi_s K^{\alpha}}, \\ &\mathbb{E}_{K+1}\big(\|P_{K+1}(h^{\omega}_{K+1} - z)^{-1}P_0\|^{2s}\big) \le Ce^{-\xi_s K^{\alpha}}. \end{aligned}$$

Using these bounds in the inequality (3.26), we get the bound (after noting that the sum has $2K$ terms, so we get $(1 + 2K)$ as the only K dependence other than the exponential decay factor),

$$\le C_5(m, Rank(P_0), \mathcal{D}, h_0, R, s)(1 + 2K)e^{-\xi_{2s}K^{\alpha}},$$

which is the required estimate to complete the proof of the Lemma. □

4. The Continuous Case

In this section we show that the density of states of some Random Schrödinger operators are almost as smooth as the single site distribution. On the Hilbert space $L^2(\mathbb{R}^d)$ we consider the operator

$$H_0 = \sum_{i=1}^{d} \left(-i \frac{\partial}{\partial x_i} + A_i(x) \right)^2,$$

with the vector potential $\vec{A}(x) = (A_1(x), \cdots, A_d(x))$ assumed to have sufficient regularity so that H_0 is essentially self-adjoint on $C_0^\infty(\mathbb{R}^d)$.

The random operators considered here are given by

$$H^\omega = H_0 + \lambda \sum_{n \in \mathbb{Z}^d} \omega_n u_n, \tag{4.1}$$

where $\{\omega_n\}_{n\in\mathbb{Z}^d}$ are independent real random variables satisfying Hypothesis 3.1, u_n are operators of multiplication by the functions $u(x-n)$, for $n \in \mathbb{Z}^d$ and $\lambda > 0$ a coupling constant.

We have the following hypotheses on the operators considered above to ensure H^ω continue to be essentially self-adjoint on $C_0^\infty(\mathbb{R}^d)$ for all ω. By now it is well known in the literature (see for example the book of Carmona and Lacroix [10]) that the spectral and other functions of these operators we consider below will have the measurability properties, as functions of ω, required for the computations we perform on them and we will not comment further on measurability.

Hypothesis 4.1. *1. The random variables $\{\omega_n\}_n$ satisfy the Hypothesis 3.1.*
2. The function $0 \le u \le 1$ is a non-negative smooth function on $\mathbb{R}^d$ such that for some $0 < \epsilon_2 < \frac{1}{2}, 0 < \epsilon_1 < 1$, it satisfies

$$u(x) = \begin{cases} 0, & x \notin (-\frac{1}{2} - \epsilon_1, \frac{1}{2} + \epsilon_1)^d \\ 1, & x \in (-\frac{1}{2} + \epsilon_2, \frac{1}{2} - \epsilon_2)^d \end{cases}$$
$$\sum_{n\in\mathbb{Z}^d} u(x-n) = 1 \qquad x \in \mathbb{R}^d.$$

We need some notation before we state our results. Given a subset $\Lambda \subset \mathbb{Z}^d$, we set

$$[\Lambda] = \left\{ x \in \mathbb{R}^d : \sum_{n\in\Lambda} u(x-n) = 1 \right\} \tag{4.2}$$

and denote the restrictions of H_0, H^ω to $[\Lambda]$ respectively by $H_{0,\Lambda}$, H^ω_Λ. As an abuse of notation, whenever we talk about restricting the operator on Λ, we will mean restriction onto $[\Lambda]$. We need this distinction because $\sum_{n\in\Lambda} u(x-n) = 1$ only on $[\Lambda]$ and we need the complete covering condition. While the boundary conditions are not that important, we will work with Dirichlet boundary conditions in this section. We will also denote $u_{n,\Lambda}$ to be the restriction of u_n to $[\Lambda]$ when the need arises. We denote by $E_A(\cdot)$ the projection valued spectral measure of a self-adjoint operator A and from the context it will be clear that this symbol will not be confused with points in the spectrum denoted by E. We denote the Integrated Density of States (IDS) by

$$\mathcal{N}_\Lambda(E) = \mathop{\mathbb{E}}_{\omega}\Big[tr(u_0 E_{H^\omega_\Lambda}(-\infty, E])\Big] \qquad for\ E \in \mathbb{R}, \tag{4.3}$$

and the subscript Λ on the IDS is dropped in the case of the operator H^ω.

We start with our Hypothesis on the localization. where we set P_n to be the orthogonal projection onto $L^2(supp(u_n))$.

Hypothesis 4.2. *A compact interval $J \subset \mathbb{R}$ is said to be in* the region of localization *for H^ω with rate of decay ξ_s and exponent $0 < s < 1$, if there exists $C, \xi_s > 0$ such that*

$$\sup_{\Re(z)\in J, \Im(z)>0} \mathop{\mathbb{E}}_{\omega}\Big[\Big\| P_n(H^\omega - z)^{-1} P_k\Big\|^s\Big] \le C e^{-\xi_s \|n-k\|} \tag{4.4}$$

for any $n, k \in \mathbb{Z}^d$. For the operators H^ω_Λ exponential localization is similarly defined with Λ, H^ω_Λ, $\xi_{s,\Lambda}$ replacing $\mathbb{Z}^d$, H^ω, ξ_s respectively in the bound for the same J.

We assume that for all Λ large enough $\xi_{s,\Lambda} \ge \xi_s$ for J in the region of localization and the constants C, ξ_s do not change if we change the density ρ_n with one of its derivatives at finitely many n.

Remark 4.3. We note that the above Hypothesis holds with $\xi_s > 0$, for the models of the type we consider under a *large disorder* condition, introduced via a coupling constant. The condition $\xi_s > 0$ is sufficient for our Theorem and there is no need to specify how large it should be. Similarly the multiscale analysis which is the starting point of the fractional moment bounds, uses apriori bounds that depend on the Wegner estimate which depends on only the constant $\mathcal{D}$. So changing the distribution ρ_n with one of its derivatives at finitely many points n does not affect the constants C, ξ_s.

Our main Theorem given next, is the analogue of the Theorem 3.4. We already know from Lemma A.5, that $u_0 E_{H^\omega}(-\infty, E)$ is trace class for any $E \in \mathbb{R}$, hence we will be working with

$$\mathcal{N}(E) = \mathop{\mathbb{E}}_{\omega}[tr(u_0 E_{H^\omega}(-\infty, E))] \qquad for\ E \in \mathbb{R}. \tag{4.5}$$

The function $\mathcal{N}$ is well defined by Lemma A.5 and is known to be continuous (see [14, Theorem 1.1] for example) whenever ρ is continuous.

By the Pastur-Shubin trace formula for the IDS, the function $\mathcal{N}$ is at most a constant multiple of IDS, since $\int u_0(x)dx$ may not be equal to 1, but this discrepancy does not affect the smoothness properties, so we will refer to $\mathcal{N}$ as the IDS below.

Our main Theorem given below implies that the density of states DOS is $m-1$ times differentiable in J when ρ satisfies the conditions of the Theorem.

Theorem 4.4. *On the Hilbert space $L^2(\mathbb{R}^d)$ consider the self-adjoint operators H^ω given by (4.1), satisfying the Hypothesis 4.1. Let J be an interval in the region of localization satisfying the Hypothesis 4.2 with $\xi_s > 0$ for some $0 < s < 1/6$. Suppose the density $\rho \in C_c^m((0, \infty))$, and $\rho^{(m)}$ is τ-Hölder continuous for some $s < \tau/2$. Then $\mathcal{N} \in C^{(m-1)}(J)$ and $\mathcal{N}^{(m)}$ exists almost everywhere in J.*

Remark 4.5. A Theorem of Aizenman et al. [3, Theorem 5.2] shows that there are operators H^ω of the type we consider for which the Hypothesis 4.2 is valid for large coupling λ, where it was required that $0 < s < 1/3$. We take $0 < s < 1/6$ as we need to controls $2s$-th moment of averages of norms of resolvent kernels in our proof.

Proof. We consider the boxes $\Lambda_L = \{-L, \cdots, L\}^d$, and set $H_L^\omega = H_{\Lambda_L}^\omega$, $\mathcal{N}_L = \mathcal{N}_{\Lambda_L}^\omega$.

The strong resolvent convergence of $H_{\Lambda_L}^\omega$ to H^ω, which is easy to verify, implies that $\mathcal{N}_{\Lambda_L}$ converges to $\mathcal{N}$ point wise since $\mathcal{N}$ is known to be a continuous function for the operators we consider. Since $tr(u_0 E_{H_L^\omega}((-\infty, E]))$ is a bounded measurable complex valued function, $\mathcal{N}_L \in C^m(J)$, by Theorem 2.1. Therefore it is enough to show that $\mathcal{N}(\cdot) - \mathcal{N}_{\Lambda_N}(\cdot)$ (which is a difference of distribution functions of the σ-finite measures $tr(u_0 E_{H^\omega}(\cdot))$ and $tr(u_0 E_{H_N^\omega}(\cdot))$ appropriately normalized) is in $C^m(J)$ for some N. We will need to use the Borel–Stieltjes transforms of these measures for the rest of the proof, but these transforms are not defined because $u_0(H_N^\omega - z)^{-1}$ fails to be in trace class. Therefore we have to approximate u_0 using finite rank operators first.

To this end let Q_k be a sequence of finite rank orthogonal projections, in the range of u_0 such that they converge to the identity on this range. We then define,

$$\mathcal{N}_{L,Q_k}(E) = \mathop{\mathbb{E}}_{\omega}\Big(tr(Q_k u_0 E_{H_L^\omega}(-\infty, E])\Big). \tag{4.6}$$

Since the projections Q_k strongly converge to the identity on the range of u_0, the projections $Q_k u_0 E_{H_L^\omega}((-\infty, E))$ also converge strongly to $u_0 E_{H_L^\omega}((-\infty, E))$ point wise in E. This convergence implies that $\mathcal{N}_{L,Q_k}(E)$ converge point wise to $\mathcal{N}_L(E)$ for any fixed L. Henceforth we drop the subscript on Q_k but remember that the rank of Q is finite.

Since Q is finite rank, the measures $tr(Qu_0 E_{H_L^\omega}(\cdot))$ are finite measures. Therefore we can define the Borel–Stieltjes transform of the finite signed measure

$$\mathop{\mathbb{E}}_{\omega}\Big[tr(Qu_0 E_{H_{L+1}^\omega}(\cdot)) - tr(Qu_0 E_{H_L^\omega}(\cdot))\Big],$$

namely

$$\begin{aligned} &\mathop{\mathbb{E}}_{\omega}\Big[tr(Qu_0(H_{L+1}^\omega - z)^{-1} - tr(Qu_0(H_L^\omega - z)^{-1})\Big] \\ &\quad = \int \frac{1}{x-z}\; d\mathop{\mathbb{E}}_{\omega}\Big[tr(Qu_0 E_{H_{L+1}^\omega}(x)) - tr(Qu_0 E_{H_L^\omega}(x))\Big], \end{aligned} \tag{4.7}$$

where the signed measure has finite total variation for each Q and each L. Then the derivatives of $\mathcal{N}_{L+1,Q}(E) - \mathcal{N}_{L,Q}(E)$ are given by

$$\begin{aligned} &\lim_{\epsilon \downarrow 0} \frac{1}{\pi} \mathop{\mathbb{E}}_{\omega}\Big[tr(Qu_0 \Im(H_{L+1}^\omega - E - i\epsilon)^{-1}) - tr(Qu_0 \Im(H_L^\omega - E - i\epsilon)^{-1})\Big] \\ &\quad = \lim_{\epsilon \downarrow 0} \frac{1}{\pi} \mathop{\mathbb{E}}_{\omega}\Bigg[tr\Bigg[Qu_0\Big(\Im(H_{L+1}^\omega - E - i\epsilon)^{-1} - \Im(H_L^\omega - E - i\epsilon)^{-1}\Big)\Bigg]\Bigg]. \end{aligned} \tag{4.8}$$

Then, using the idea of a telescoping sum, as done in the previous section, we need to prove that

$$\sum_{L=N}^{\infty} \sup_{\Re(z)\in J} \left| \frac{d^\ell}{dz^\ell}\Bigg(\mathop{\mathbb{E}}_{\omega}\Big[tr(Qu_0(H_{L+1}^\omega - z)^{-1}) - \mathop{\mathbb{E}}_{\omega}\Big[tr(u_0(H_L^\omega - z)^{-1})\Big]\Bigg)\Big]\right| < \infty. \tag{4.9}$$

We set (taking $\kappa(L)$ as the volume of $\Lambda_L\backslash\{0\}$),

$$G^{\omega}_{L}(z) = Qu_0(H^{\omega}_{L} - z)^{-1}u_0, \quad S(\vec{\omega}, Q, L, z) = G^{\omega}_{L+1}(z) - G^{\omega}_{L}(z),$$
$$\Phi_{L+1}(\vec{\omega}) = \prod_{n\in\Lambda_{L+1}} \rho(\omega_n), \quad \kappa(L) = |\Lambda_L| - 1. \tag{4.10}$$

Then, following the sequence of steps leading from Eqs. (3.11) to (3.16), we need only to consider

$$T(L, \ell, Q, z) = \frac{d^{\ell}}{dE^{\ell}} \mathop{\mathbb{E}}_{\omega}\big[tr(G^{\omega}_{L+1}(E + i\epsilon) - G^{\omega}_{L}(E + i\epsilon))\big]$$
$$= \int_{\mathbb{R}^{\kappa(L+1)+1}} tr(S(\vec{\omega}, Q, L, E, \epsilon))(\mathbf{D}^{\ell}\Phi_{L+1})(\vec{\omega})d\vec{\omega}, \tag{4.11}$$

to estimate and show that

$$\sum_{L=N}^{\infty} \sup_{\substack{\Re(z)\in J, \Im(z)>0,\\ \ell\le m,\\ Q}} |T(L, \ell, Q, z)| < \infty, \tag{4.12}$$

to prove the theorem. Using the steps followed from getting Eq. (3.18) from the equality (3.17), which is an identical calculation here, to get

$$T(L, \ell, Q, z) = \sum_{\substack{\sum_{n=1}^{\Lambda_{L+1}} k_n=\ell\\ k_n\ge 0}} \binom{\ell}{k_0,\dots,k_{\kappa(L+1)+1}} \int_{\mathbb{R}^{\kappa(L+1)}} tr\bigg(\int \big(G^{\omega}_{L+1}(z) - G^{\omega}_{L}(z)\big)\rho_0^{(k_0)}(\omega_0)d\omega_0\bigg)$$
$$\cdot\bigg(\prod_{n\in\Lambda_{L+1}\backslash\{0\}} \rho_n^{(k_n)}(\omega_n)d\omega_n\bigg). \tag{4.13}$$

To proceed further, we need to get a uniform bound in the projection Q. We will show that the expression

$$\mathcal{G}(L, z, \omega) = u_0(H^{\omega}_{L+1} - z)^{-1} - u_0(H^{\omega}_{L} - z)^{-1}, \tag{4.14}$$

automatically comes with a trace class operator. This fact helps us drop the Q occurring in the expression

$$\big(G^{\omega}_{L+1}(z) - G^{\omega}_{L}(z)\big) = Q\mathcal{G}(L, z, \omega)u_0, \tag{4.15}$$

making estimates on the trace.

We need a collection of $d+2$ smooth functions $0 \le \Theta_j \le 1$, $j = 0, \dots, d+1$, where d is the dimension we are working with. Setting

$$\alpha_j = 2^{j+2}, \; j \in \{0, 1, 2, \dots, 2d + 2\}, \tag{4.16}$$

we choose the functions Θ_j from $C^{\infty}(\mathbb{R}^d)$ satisfying

$$\Theta_j(x) = \begin{cases} 1, & |x| \le \alpha_{2j}, \\ 0, & |x| > \alpha_{2j+1}, \end{cases} \qquad j = 0, \dots d + 1 \tag{4.17}$$

and note that all the derivatives of Θ_j are bounded for all j, because they are all continuous and supported in a compact set. These functions satisfy the property

$$\Theta_{j+1}\phi = \phi, \quad if\ supp(\phi) \subset supp(\Theta_j), \quad j = 0, \ldots, d, \tag{4.18}$$

in particular

$$\Theta_{j+1}\Theta_j = \Theta_j, \quad for\ \ all\ \ j = 0, \ldots, d. \tag{4.19}$$

We then take a free resolvent operator $R^0_{L,a} = (H_{0,\Lambda_L} + a)^{-1}$, with $a >> 1$. Since, H_0 is bounded below, $R^0_{L,a}$ is a bounded positive operator for any L. It is a fact that, for any smooth bump function ϕ,

$$[\phi, H_0]R^0_{L,a}, R^0_{L,a}u_j \in \mathcal{I}_p, \quad p > d. \tag{4.20}$$

See Combes et al. [14, Lemma A.1] and Simon [50, Chapter 4] for further details. Using the definition of $\mathcal{G}$ given in Eq. (4.15), the relation (4.19) and the resolvent equation we get

$$\begin{aligned}
\mathcal{G}(L, z, \omega)\Theta_0 &= u_0\Big[(H^\omega_{L+1} - z)^{-1} - (H^\omega_L - z)^{-1}\Big]\Theta_0 \\
&= u_0\Big[(H^\omega_{L+1} - z)^{-1} - (H^\omega_L - z)^{-1}\Big]\Theta_1\Theta_0 \\
&= u_0\Big[(H^\omega_{L+1} - z)^{-1}\Theta_1 - \Theta_1 R^0_{L,a} + \Theta_1 R^0_{L,a} - (H^\omega_L - z)^{-1}\Theta_1\Big]\Theta_0 \\
&= u_0\Big[\big((H^\omega_{L+1} - z)^{-1}\Theta_1 - \Theta_1 R^0_{L,a}\big) - \big((H^\omega_L - z)^{-1}\Theta_1 - \Theta_1 R^0_{L,a}\big)\Big]\Theta_0 \\
&= u_0\Big[\big((H^\omega_{L+1} - z)^{-1}\Big(\Theta_1(H_{0,L} + a) - (H^\omega_{L+1} - z)\Theta_1\Big)R^0_{L,a}\big) \\
&\quad - \big((H^\omega_L - z)^{-1}\Big((H_{0,L} + a)\Theta_1 - (H^\omega_L - z)\Theta_1\Big)R^0_{L,a}\big)\Big]\Theta_0 \\
&= u_0\Big[\big((H^\omega_{L+1} - z)^{-1}\Big(\Theta_1 H_{0,L} - H_{0,L+1}\Theta_1 + (z + a - V^\omega_{L+1})\Theta_1\Big)R^0_{L,a}\big) \\
&\quad - \big((H^\omega_L - z)^{-1}\Big(H_{0,L}\Theta_1 - H_{0,L}\Theta_1 + (z + a - V^\omega_L)\Theta_1\Big)R^0_{L,a}\big)\Big]\Theta_0 \\
&= u_0\big[(H^\omega_{L+1} - z)^{-1} - (H^\omega_L - z)^{-1}\big] \\
&\quad \cdot \Big[[\Theta_1, H_0] + \Big(z + a - \sum_{|n|\le\alpha_1} \omega_n u_n\Big)\Theta_1\Big]R^0_{L,a}\Theta_0 \\
&= \mathcal{G}(L, z, \omega)\Big[[\Theta_1, H_0] + \Big(z + a - \sum_{|n|\le\alpha_1} \omega_n u_n\Big)\Theta_1\Big]R^0_{L,a}\Theta_0 \\
&= \mathcal{G}(L, z, \omega)\Big(A_0(z, a, \alpha_1, H_0) + \sum_{|n|\le\alpha_1} \omega_n B_{0,n}(a, \alpha_1)\Big)
\end{aligned} \tag{4.21}$$

where we used the definition

$$A_0(z, a, \alpha_1, H_0) = \big([\Theta_1, H_0] + (z+a)\Theta_1\big) R^0_{L,a}\Theta_0$$
$$B_{0,n}(a, \alpha_1) = -u_n \Theta_1 R^0_{L,a}\Theta_0 \tag{4.22}$$

and in passing from equality 6 to equality 7 of the above equation, used the fact that the support of Θ_1 is far away from the boundary of Λ_L, Λ_{L+1}, so V^ω_L, V^ω_{L+1} agree on the support of Θ_1 and also the commutators of Θ_1 with $H_{0,L}$, $H_{0,L+1}$ are the same and agree with that of H_0. In the above A_0, $B_{0,n}$ are operators independent of ω, each of which is in $\mathcal{I}_p$, by Eq. (4.19). Using the definitions and properties of Θ_j, we see that

$$\Theta_2 A_0(z, a, \alpha_1, H_0) = A_0(z, a, \alpha_1, H_0), \quad \textit{and} \;\; \Theta_2 B_{0,n}(a, \alpha_1) = B_{0,n}(a, \alpha_1).$$

Therefore we can repeat this argument by defining for $j = 0, \dots d$,

$$A_j(z, a, \alpha_{2j+1}, H_0) = ([\Theta_{2j+1}, H_0] + (z+a)\Theta_{2j+1}) R^0_{L,a}\Theta_{2j}$$
$$B_{j,n}(a, \alpha_{2j+1}) = -u_n \Theta_{2j+1} R^0_{L,a}\Theta_{2j}, \;\; |n| \le \alpha_{2j+1}, \tag{4.23}$$

by using the fact that

$$\Theta_{2j} A_{j-1}(z, a, \alpha_{2j+1}, H_0) = A_{j-1}(z, a, \alpha_{2j+1}, H_0), \quad \textit{and}$$
$$\Theta_{2j} B_{j-1,n}(a, \alpha_{2j+1}) = B_{j-1,n}(a, \alpha_{2j+1}), \tag{4.24}$$

for each $j = 1, 2, \dots d$. We can then re-write the Eq. (4.21) as

$$\mathcal{G}(L, z, \omega) = \mathcal{G}(L, z, \omega) \overleftarrow{\prod_{j=0}^{d}} \Bigg(A_j(z, a, \alpha_{2j+1}, H_0) + \sum_{|n| \le \alpha_{2j+1}} \omega_n B_{j,n}(a, \alpha_{2j+1}) \Bigg), \tag{4.25}$$

where the arrow on the product indicates an ordered product with the operator sum with a lower index j coming to the right of the one with a higher index j.

Now, counting the number of terms there are in the product, we see that each sum $\sum_{|n| \le \alpha_{2j+1}}$ has a maximum of $(2\alpha_{2j+1})^d = 2^{d(2j+4)}$ terms. A simple computation shows that there are a maximum of $2^{d^2(d+4)}$ terms, if we completely expand out the product. In other words the number of terms are dependent on d but not on L.

We will now write the expression in Eq. (4.25) as

$$\mathcal{G}(L, z, \omega) = \sum_{|n| \le \alpha_{2d+2}} \mathcal{G}(L, z, \omega) u_n \Bigg(\sum_{r_1, r_2 = 0}^{d+1} \omega_n^{r_1} \omega_0^{r_2} P_{n,0}(k, r, \omega) \Bigg), \tag{4.26}$$

where $P_{n,0}(k, r)$, $r = (r_1, r_2)$ is a trace class operator valued function of ω, but independent of ω_0, ω_n for each k, r. Note that even though A_d and B_d are supported in $supp(\Theta_d)$, $\sum_{|n| \le \alpha_{2d+1}} u_n$ is not one on the support of Θ_d, so we have to take a larger sum in the above expression. We can see from the structure of the product that the trace norms satsify a bound

$$\sup_{\Re(z) \in J, 0 < \Im(z) \le 1} \|P_{n,0}(k, r)\|_1 \le C_7(d, a, J),$$

since an inspection of the product in Eq. (4.25), shows that in any product, z and $\{\omega_{\tilde{n}}, \tilde{n} \neq 0, n\}$ occurs at most to a power of $d+1$. The uniform boundedness of the trace norm as a function of z, $\omega_{\tilde{n}}$ is clear since these variables are in compact sets. As for the finiteness of the trace norm itself, we note that any product has $d+1$ factors from the set $\{A_j, B_j, j = 0, \dots, d\}$, hence by the claim in Eq. (4.20), such a product is trace class.

Using Eqs. (4.10, 4.13, 4.14, 4.15) and Eq. (4.26) in Eq. (4.13), we get, using the fact that $P_{n,0}()$ are independent of ω_0, ω_n,

$$\begin{aligned} &T(L, \ell, Q, z) \\ &= \sum_{\substack{\sum_{n=1}^{k(L+1)+1} k_n = l \\ k_n \geq 0}} \binom{\ell}{k_0, \dots, k_{\kappa(L+1)+1}} \int_{\mathbb{R}^{\kappa(L+1)-1}} \sum_{|n| \leq \alpha_{2d+2}} \sum_{r_1, r_2 = 0}^{d+1} tr\Bigg(Q\Big[\int u_0\big((H^\omega_{L+1} - z)^{-1} \\ &\quad -(H^\omega_L - z)^{-1}\big) u_n \omega_n^{r_2} \omega_0^{r_1} \rho_n^{(k_n)}(\omega_n) \rho_0^{(k_0)}(\omega_0) d\omega_n d\omega_0 \Big] P_{n,0}(k, r, \omega) \Bigg) \\ &\quad \times \prod_{m \in \Lambda_{L+1} \setminus \{0,n\}} \rho_m^{(k_m)}(\omega_m) d\omega_m. \end{aligned} \tag{4.27}$$

We now estimate the absolute value of the trace in Eq. (4.27) using the Theorem 2.2(1), taking the $\phi_{\mathbf{R}}$ that appears there, for bounding the norm of the integral with respect to ω_n, ω_0, since $2s < \tau$.

$$\begin{aligned} &|T(L, \ell, Q, z)| \\ &\leq \sum_{\substack{\sum_{n=1}^{k(L+1)+1} k_n = l \\ k_n \geq 0}} \binom{\ell}{k_0, \dots, k_{\kappa(L+1)+1}} \int_{\mathbb{R}^{\kappa(L+1)-2}} \sum_{|n| \leq \alpha_{2d+2}} \sum_{r_1, r_2 = 0}^{d+1} \|Q\| \|P_{n,0}(k, r, \omega)\|_1 \\ &\quad \Big[\int \|(u_0 + u_n)^{\frac{1}{2}} \big((H^\omega_{L+1} - z)^{-1} - (H^\omega_L - z)^{-1}\big)(u_n + u_0)^{\frac{1}{2}}\|^s \phi_{\mathbf{R}}(\omega_0) \phi_{\mathbf{R}}(\omega_n) d\omega_n d\omega_0 \Big] \\ &\quad \prod_{m \in \Lambda_{L+1} \setminus \{0,n\}} |\rho_m^{(k_m)}(\omega_m)| d\omega_m. \end{aligned} \tag{4.28}$$

In the above inequality we also used the fact that $u_0(u_0 + u_n)^{-\frac{1}{2}}, u_n(u_0 + u_n)^{-\frac{1}{2}}$ are both bounded uniformly in n and replaced u_0, u_n by $(u_0 + u_n)^{\frac{1}{2}}$ on either side of the resolvents.

We would prefer to work with probability measures in above equation, so we normalize $|\rho_m^{(k_m)}(x)|dx$ by their L^1 norm. We also do the same for $\phi_{\mathbf{R}}$. We then follow the steps involved in obtaining the inequality (3.21). We set $\eta(m, \rho) = (\sup_{n \in \mathbb{Z}^d, k_n \leq m} \|\rho_n^{k_n}\|_1 + \|\rho_n^{k_n}\|_\infty) + \|\phi_{\mathbf{R}}\|_1$ to get,

$$\begin{aligned} &|T(L, \ell, Q, z)| \\ &\leq \sum_{\substack{\sum_{n=1}^{k(L+1)+1} k_n = l \\ k_n \geq 0}} \binom{\ell}{k_0, \dots, k_{\kappa(L+1)+1}} \sum_{|n| \leq \alpha_{2d+2}} C_9(a, d, J, \eta(\rho, m)) \\ &\quad \times \mathbb{E}_{L+1}\Big[\|(u_0 + u_n)^{\frac{1}{2}} \big((H^\omega_{L+1} - z)^{-1} - (H^\omega_L - z)^{-1}\big)(u_n + u_0)^{\frac{1}{2}}\|^s \Big], \end{aligned} \tag{4.29}$$

where $\mathbb{E}_{L+1}$ is the expectation with respect to the probability density

$$\frac{\phi_{\mathbf{R}}(\omega_0)d\omega_0}{\|\phi_{\mathbf{R}}\|_1}\frac{\phi_{\mathbf{R}}(\omega_n)d\omega_n}{\|\phi_{\mathbf{R}}\|_1}\prod_{m\in\Lambda_{L+1}\setminus\{0,n\}}\frac{|\rho_m^{(k_m)}(\omega_m)|}{\|\rho_m^{(k_m)}\|_1}d\omega_m.$$

We define a smooth radial function $0\leq\Psi\leq 1$ such that

$$\Psi(x)=\begin{cases}1, & |x|\leq L/2,\\ 0, & |x|>L/2+4\end{cases}.$$

Then $\Psi_L\sqrt{u_0+u_n}=\sqrt{u_0+u_n},\ \ |n|\leq\alpha_{2d+2}$. Following the steps similar to obtaining the inequality (4.21), using the relation $(H_{0,L}+a)R^0_{L,a}=Id$, we have

$$\begin{aligned}
&(u_0+u_n)^{\frac{1}{2}}\big((H^\omega_{L+1}-z)^{-1}-(H^\omega_L-z)^{-1}\big)(u_n+u_0)^{\frac{1}{2}}\\
&\quad=(u_0+u_n)^{\frac{1}{2}}\big((H^\omega_{L+1}-z)^{-1}[\Psi_L,H_0](H^\omega_L-z)^{-1}\big)(u_n+u_0)^{\frac{1}{2}}\\
&\quad=(u_0+u_n)^{\frac{1}{2}}(H^\omega_{L+1}-z)^{-1}[\Psi_L,H_0]\big[R^0_{L,a}+(H^\omega_L-z)^{-1}-R^0_{L,a}\big](u_n+u_0)^{\frac{1}{2}}\\
&\quad=(u_0+u_n)^{\frac{1}{2}}(H^\omega_{L+1}-z)^{-1}[\Psi_L,H_0]R^0_{L,a}\big(I+(z+a-V^\omega_L)(H^\omega_L-z)^{-1}\big)(u_n+u_0)^{\frac{1}{2}}\\
&\quad=(u_0+u_n)^{\frac{1}{2}}(H^\omega_{L+1}-z)^{-1}\Big[-\sum_{i=1}^d\frac{\partial^2}{\partial x_i^2}\Psi_L+2\sum_{i=1}^d\Big(\frac{\partial}{\partial x_i}\Psi_L\Big)\Big(-i\frac{\partial}{\partial x_i}+A_i\Big)\Big]\\
&\qquad\times R^0_{L,a}\big(I+(z+a-V^\omega_L)(H^\omega_L-z)^{-1}\big)(u_n+u_0)^{\frac{1}{2}}.
\end{aligned}\tag{4.30}$$

We take a smooth bounded radial function $0\leq\Upsilon_L\leq 1$ which is 1 in a neighbourhood of $L/2\leq r\leq L/2+4$ and zero outside a neighbourhood of radial width 10. Then using the fact that

$$\begin{aligned}
\Upsilon_L\Big(\sum_{i=1}^d\frac{\partial^2}{\partial x_i^2}\Psi_L\Big)&=\Big(\sum_{i=1}^d\frac{\partial^2}{\partial x_i^2}\Psi_L\Big)\\
\Upsilon_L\Big(\frac{\partial}{\partial x_i}\Psi_L\Big)&=\Big(\frac{\partial}{\partial x_i}\Psi_L\Big),\ \ for\ \ all\ \ i=1,\dots,d
\end{aligned}\tag{4.31}$$

and (4.30), we can now bound the expectation in the inequality (4.29),by

$$\begin{aligned}
&\mathbb{E}_{L+1}\Big[\|(u_0+u_n)^{\frac{1}{2}}\big((H^\omega_{L+1}-z)^{-1}-(H^\omega_L-z)^{-1}\big)(u_n+u_0)^{\frac{1}{2}}\|^s\Big]\\
&\quad\leq\mathbb{E}_{L+1}\Big[\|(u_n+u_0)^{\frac{1}{2}}(H^\omega_{L+1}-z)^{-1}\Upsilon_L\|^s\\
&\qquad\Big\|\Big[\Big(-\sum_{i=1}^d\frac{\partial^2}{\partial x_i^2}\Psi_L\Big)+2\sum_{i=1}^d\Big(\frac{\partial}{\partial x_i}\Psi_L\Big)\Big(-i\frac{\partial}{\partial x_i}+A_i\Big)\Big]R^0_{L,a}\Big\|^s\\
&\qquad\big(1+|z|+a+\|V^\omega_L\|_\infty\big)\|\chi_{\Lambda_L}(H^\omega_L-z)^{-1}\sqrt{u_0+u_n}\|^s\Big].
\end{aligned}\tag{4.32}$$

Then using Cauchy–Schwartz inequality and Hypothesis 4.2 we get an exponential bound for the first factor, a uniform bound for the second factor after noting that $dist(supp(\Upsilon_L, \{n : |n| \leq \alpha_{2d} + 1\}) \geq L/4$, $\|\Lambda_L\| \leq (2L)^d$, we get the estimate

$$\sup_{z:\Re(z)\in J, \Im(z)\leq 1} \mathbb{E}_{\omega}\left[\|(u_0 + u_n)^{\frac{1}{2}} \left((H^{\omega}_{L+1} - z)^{-1} - (H^{\omega}_L - z)^{-1}\right)(u_n + u_0)^{\frac{1}{2}}\|^s \right] \leq C_{10}(a, J, d) L^d e^{-\xi_{2s} L}. \quad (4.33)$$

Using this inequality in (4.29) we get the bound

$$\sup_{\substack{z:\Re(z)\in J, \Im(z)\leq 1, \\ Q \\ \ell\leq m}} |T(L, \ell, Q, z)| \leq C_{11}(a, d, J, \eta(\rho, m))(L + 1)^{d(m+1} e^{-\xi_{2s} L}, \quad (4.34)$$

as the combinatorial sum

$$\sum_{\substack{\sum_{n=1}^{k(L+1)+1} k_n = l \\ k_n \geq 0}} \binom{\ell}{k_0, \ldots, k_{\kappa(L+1)}}$$

is easily seen to add up to $(L + 1)^{d\ell}$, which is still polynomial in L. This bounds shows the summability in Eq. (4.9) completing the proof. □

Acknowledgements. AM thanks ICTS-TIFR for a post doctoral fellowship, DD thanks the DST for the INSPIRE Faculty Fellowship and both of them thank Ashoka University for short visits when this work was done. We thank the referee of this paper for very detailed, numerous and critical comments that helped us improve the presentation of the results.

A. Appendix

We collect a few Lemmas in this appendix that are used in the main part of the paper. All these Theorems are well known and proved elsewhere in the literature, but we state them in the form we need and also give their proofs for the convenience of the reader.

Lemma A.1. *Consider a positive function $\rho \in L^1(\mathbb{R}, dx)$ and $J \subset \mathbb{R}$ an interval. Let $F(z) = \int \frac{1}{x-z} \rho(x)dx$. Then, for any $m \in \mathbb{N}$,*

$$\operatorname*{ess\,sup}_{x\in J} \left| \frac{d^m}{dx^m} \rho \right| (x) < \infty$$

whenever

$$\sup_{z\in\mathbb{C}^+,\ \Re(z)\in J} \left| \frac{d^m}{dx^m} \Im F \right| (z) < \infty.$$

Proof. Since $\rho(x)dx$ is a finite positive measure, F is analytic in $\mathbb{C}^+$, and the assumption on F implies that functions $\frac{d^\ell}{dz^\ell}\Im F$ are bounded harmonic functions in the strip $\{z \in \mathbb{C}^+ : \Re(z) \in J\}$, $0 \le \ell \le m$. Therefore the boundary values

$$h_\ell(E) = \lim_{\epsilon \to 0} \frac{d^\ell}{dz^\ell} \Im F(E + i\epsilon)$$

exist for Lebesgue almost every $E \in J$ and h_ℓ are essentially bounded in J, $0 \le \ell \le m$. For any $E_0 \in J$ for which $h_\ell(E_0)$ is defined for all $0 \le \ell \le m$ and we have for $0 \le \ell \le m - 1$,

$$\frac{\partial^\ell}{\partial x^\ell}(\Im F)(E+i\epsilon) - \frac{\partial^\ell}{\partial x^\ell}(\Im F)(E_0+i\epsilon) = \int_{E_0}^{E} \frac{\partial^{\ell+1}}{\partial x^{\ell+1}}(\Im F)(x+i\epsilon)\, dx, \quad E \in J. \tag{A.1}$$

Since the integrands above are Harmonic functions in the strip, their boundary values exist, they are uniformly bounded in the strip, so by the dominated convergence theorem the integral converges to

$$\int_{E_0}^{E} h_{\ell+1}(x)\, dx, \quad E \in J.$$

On the other hand the left hand side of Eq. (A.1) converges to $h_\ell(E) - h_\ell(E_0)$, showing that $h_\ell(E)$ is differentiable in J. Since, $\rho(x) = \frac{1}{\pi} h_0(x),\ x \in J$, a simple induction argument now gives the Lemma. □

Lemma A.2. *On a separable Hilbert space $\mathscr{H}$, let A and B be two bounded operators generating strongly differentiable contraction semi-groups e^{tA}, e^{tB} respectively, then for any* $0 < s < 1$,

$$\left\| e^{tA} - e^{tB} \right\| \le 2^{1-s} |t|^s \, \|A - B\|^s .$$

Proof. Since e^{tA}, e^{tB} are strongly differentiable, the fundamental Theorem of calculus gives the bound,

$$\left\| e^{tA} - e^{tB} \right\| = \left\| \int_0^t e^{(t-s)A}(A - B)e^{sB}\, ds \right\| \le |t| \, \|A - B\| .$$

Since e^{tA}, e^{tB} are contractions we have the trivial bound

$$\left\| e^{tA} - e^{tB} \right\| \le 2,$$

so the Lemma follows by interpolation. □

Lemma A.3. *Let g be a probability density with a τ-Hölder continuous derivative. Suppose A is a bounded operator on a separable Hilbert space $\mathscr{H}$ with $\Im A > 0$ and satisfies*

$$\left\| (A + \lambda I)^{-1} \right\| < C < \infty, \quad \lambda \in supp(g).$$

Then

$$\int g(\lambda)(A + \lambda I)^{-1} d\lambda = -\int_0^\infty e^{itA} \left(\int g(\lambda) e^{it\lambda} d\lambda \right) dt. \tag{A.2}$$

Proof. Since $(A + \lambda I)^{-1}$ is bounded we have, in the strong sense,

$$(A + \lambda I)^{-1} = \lim_{\epsilon \downarrow 0} (A + \epsilon + \lambda I)^{-1}.$$

Since $\Im A > 0$, the bounded operator $(A+i\epsilon)$ is the generator of a contraction semi-group, so using [59, Corollary 1, Section IX.4] we have

$$\begin{aligned}\int g(\lambda)(A + i\epsilon + \lambda I)^{-1} d\lambda &= \int g(\lambda) \int e^{it(A+i\epsilon+\lambda I)} dt d\lambda \\ &= \int \int g(\lambda) e^{(-\epsilon+\lambda)t} e^{itA} dt d\lambda. \end{aligned} \tag{A.3}$$

Since g has a τ-Hölder continuous derivative, its Fourier transform is a bounded integrable function. Therefore by Fubini we can interchange the λ and t integrals on the right hand side of the above equation to get the right hand side of Eq. (A.2). On the other hand using the fact that $\left\|(A + \epsilon + \lambda I)^{-1}\right\| < 2C$ for $0 < \epsilon < \frac{1}{2C}$ and g is a probability density, we have

$$\begin{aligned}\lim_{\epsilon \downarrow 0} \int g(\lambda)(A + i\epsilon + \lambda I)^{-1} d\lambda \qquad &= \int g(\lambda) \left[\lim_{\epsilon \downarrow 0} (A + i\epsilon + \lambda I)^{-1} \right] d\lambda \\ = \int g(\lambda)(A + \lambda I)^{-1} d\lambda. \end{aligned}$$

This set of equalities when applied to the left hand side of the Eq. (A.2) gives the Lemma after letting ϵ go to zero. □

We give the Lemma below which is a consequence of proofs of results in Stollmann [56] and Combes et al. [13]. These papers essentially prove the result, but we write it here since it does not occur in the form we need to use.

Lemma A.4. *Suppose A is a self-adjoint operator on a separable Hilbert space $\mathscr{H}$ and suppose B is a non-negative bounded operator. Consider the operators $A(t) = A + tB, \ \ t \in \mathbb{R}$, $\phi \in Range(B)$ and $\nu^{\phi}_{A(t)}$ the spectral measure of $A(t)$ associated with the vector ϕ. Suppose μ is a finite absolutely continuous measure with bounded density, then*

$$\sup_{z \in \mathbb{C}^+} \int \Im(\langle \phi, (A(t) - z)^{-1} \phi \rangle) d\mu(t) < \infty. \tag{A.4}$$

In particular the measure $\int \nu^{\phi}_{A(t)} \, dt$ has bounded density.

Proof. We set $\tilde{\nu} = \int \nu^{\phi}_{A(t)} \, dt$, then $\tilde{\nu}$ is a positive finite measure. We recall that the modulus of continuity of a measure ν is defined as

$$s(\nu, \epsilon) = \sup\{\nu([a, a + \epsilon]) : a \in \mathbb{R}\}.$$

This definition immediately implies that an absolutely continuous measure μ with bounded density ρ, satisfies $s(\mu, \epsilon) \leq \|\rho\|_\infty \epsilon$. Therefore the Theorem 3.3 of Stollman [56], implies that

$$s(\tilde{\nu}, \epsilon) \leq 6\|B\|\|\phi\| s(\mu, \epsilon) \leq C\|\rho\|_\infty \epsilon.$$

This inequality implies that the density of $\tilde{\nu}$ is bounded. Since the function

$$F(z) = \int \Im(\langle \phi, (A(t) - z)^{-1} \phi \rangle)\, d\mu(t) = \int \Im(\frac{1}{x - z})\, d\tilde{\nu}$$

is positive Harmonic in $\mathbb{C}^{+}$, by the maximum principle its supremum is attained on $\mathbb{R}$. The boundary values of F on $\mathbb{R}$ exist and equal the density of the measure $\tilde{\nu} = \int \nu_{A(t)}^{\phi}\, dt$ Lebesgue almost everywhere , by Theorem 1.4.16 of Demuth and Krishna [21], giving the result. □

Lemma A.5. *Consider the operators H^{ω}, H_{Λ}^{ω} given in Eq. (4.1) and the discussion following it. Then for any finite $E \in \mathbb{R}$, the operators $u_0 E_{H_{\Lambda}^{\omega}}((-\infty, E))$, $u_0 E_{H^{\omega}}((-\infty, E))$ are trace class for all ω. The traces of these operators are uniformly bounded in ω for fixed E.*

Proof. We will give the proof for H^{ω}, the proof for the others is similar. The hypotheses on H^{ω} imply that it is bounded below and the pair H_0, H^{ω} are relatively bounded with respect to each other, being bounded perturbations of each other, the operators $(H_0{+}a)^d E_{H^{\omega}}((-\infty, E))$ are bounded for any fixed (E, a, ω). So taking a in the resolvent set of H_0 and using the fact that $u_0(H_0 + a)^{-d}$ is trace class we see that

$$u_0 E_{H^{\omega}}(-\infty, E) = u_0(H_0 + a)^{-d}(H_0 + a)^d E_{H^{\omega}}(-\infty, E),$$

is a product of a trace class operator and a bounded operator for each fixed (ω, a, E) with a positive and large. Therefore $u_0 E_{H^{\omega}}(-\infty, E)$ is also trace class for each E, ω. The uniform boundedness statement is obvious from the assumptions on the random potential. □

References

1. Acosta, V., Klein, A.: Analyticity of the density of states in the Anderson model on the Bethe lattice. J. Stat. Phys. **69**(1–2), 277–305 (1992)
2. Aizenman, M.: Localization at weak disorder: some elementary bounds. Rev. Math. Phys. **6**(5A), 1163–1182 (1994). **(Special issue dedicated to Elliott H. Lieb)**
3. Aizenman, M., Elgart, A., Naboko, S., Schenker, J .H., Stolz, Gunter: Moment analysis for localization in random Schrödinger operators. Invent. Math. **163**(2), 343–413 (2006)
4. Aizenman, M., Molchanov, S.: Localization at large disorder and at extreme energies: an elementary derivation. Commun. Math. Phys. **157**(2), 245–278 (1993)
5. Anderson, P.W.: Absence of diffusion in certain random lattices. Phys. Rev. **109**, 1492–1505 (1958)
6. Bourgain, J., Kenig, C.E.: On localization in the continuous Anderson–Bernoulli model in higher dimension. Invent. Math. **161**(2), 389–426 (2005)
7. Bourgain, J., Klein, A.: Bounds on the density of states for Schrödinger operators. Invent. Math. **194**(1), 41–72 (2013)
8. Bovier, A., Campanino, M., Klein, A., Perez, J.F.: Smoothness of the density of states in the Anderson model at high disorder. Commun. Math. Phys. **114**(3), 439–461 (1988)
9. Campanino, M., Klein, A.: A supersymmetric transfer matrix and differentiability of the density of states in the one-dimensional Anderson model. Commun. Math. Phys. **104**(2), 227–241 (1986)
10. Carmona, R., Lacroix, J.: Spectral Theory of Random Schrödinger Operators. Probability and Its Applications. Birkhäuser Boston Inc., Boston (1990)
11. Chulaevsky, V.: Universality of smoothness of density of states in arbitrary higher-dimensional disorder under non-local interactions i. From viéte–euler identity to anderson localization (2016). arXiv:1604.08534
12. Combes, J.-M., Hislop, P.D.: Landau hamiltonians with random potentials: localization and the density of states. Commun. Math. Phys. **177**(3), 603–629 (1996)

13. Combes, J.-M., Hislop, P.D., Klopp, F.: Regularity properties for the density of states of random Schrödinger operators. In: Kuchment, P. (ed.) Waves in Periodic and Random Media (South Hadley, MA, 2002), vol. 339 of Contemp. Math., pp. 15–24. Amer. Math. Soc., Providence (2003)
14. Combes, J.M., Hislop, P.D., Klopp, F.: An optimal Wegner estimate and its application to the global continuity of the integrated density of states for random Schrödinger operators. Duke Math. J. **140**(3), 469–498 (2007)
15. Combes, J.M., Hislop, P.D., Tip, A.: Band edge localization and the density of states for acoustic and electromagnetic waves in random media. Annales de l'IHP Physique théorique **70**, 381–428 (1999)
16. Constantinescu, F., Fröhlich, J., Spencer, T.: Analyticity of the density of states and replica method for random Schrödinger operators on a lattice. J. Stat. Phys. **34**(3–4), 571–596 (1984)
17. Craig, W., Simon, B.: Log Hölder continuity of the integrated density of states for stochastic Jacobi matrices. Commun. Math. Phys. **90**(2), 207–218 (1983)
18. 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 edn. Springer, Berlin (1987)
19. de Monvel, A.B., Naboko, S., Stollmann, P., Stolz, G.: Localization near fluctuation boundaries via fractional moments and applications. Journal d'Analyse Mathématique **100**(1), 83–116 (2006)
20. Delyon, F., Lévy, Y., Souillard, B.: Anderson localization for multidimensional systems at large disorder or large energy. Commun. Math. Phys. **100**(4), 463–470 (1985)
21. Demuth, M., Krishna, M.: Determining Spectra in Quantum Theory. Progress in Mathematical Physics, vol. 44. Birkhäuser Boston Inc., Boston (2005)
22. Disertori, M.: Smoothness of the averaged density of states in a random band matrix model. Markov Process. Relat. Fields **9**(2), 311–322 (2003). **(Inhomogeneous random systems (Cergy-Pontoise, 2002))**
23. Disertori, M.: Density of states for GUE through supersymmetric approach. Rev. Math. Phys. **16**(9), 1191–1225 (2004)
24. Disertori, M., Pinson, H., Spencer, T.: Density of states for random band matrices. Commun. Math. Phys. **232**(1), 83–124 (2002)
25. Disertori, M., Lager, M.: Density of states for random band matrices in two dimensions. Ann. Henri Poincaré **18**(7), 2367–2413 (2017)
26. Fröhlich, J., Martinelli, F., Scoppola, E., Spencer, T.: Constructive proof of localization in the Anderson tight binding model. Commun. Math. Phys. **101**(1), 21–46 (1985)
27. Fröhlich, J., Spencer, T.: A rigorous approach to Anderson localization. Phys. Rep. **103**(1–4), 9–25 (1984). **(Common trends in particle and condensed matter physics (Les Houches, 1983))**
28. Germinet, F., Klein, A.: Explicit finite volume criteria for localization in random media and applications. Geom. Funct. Anal. **13**(6), 1201–1238 (2003)
29. Germinet, F., De Bievre, S.: Dynamical localization for discrete and continuous random Schrödinger operators. Commun. Math. Phys. **194**(2), 323–341 (1998)
30. Glaffig, C.: Smoothness of the integrated density of states on strips. J. Funct. Anal. **92**(2), 509–534 (1990)
31. Ja Goldsheid, I., Molčanov, S .A., Pastur, L .A.: A random homogeneous Schrödinger operator has a pure point spectrum. Funkc. Anal. i Priložen. **11**(1), 1–10, 96 (1977)
32. Kaminaga, M., Krishna, M., Nakamura, S.: A note on the analyticity of density of states. J. Stat. Phys. **149**(3), 496–504 (2012)
33. Kirsch, W.: An invitation to random Schrödinger operators. In: Remy, B. (ed.) Random Schrödinger operators. Panorama & Synthèses, vol. 25, pp. 1–119. Soc. Math. France, Paris (2008). **(With an appendix by Frédéric Klopp)**
34. Kirsch, W., Krishna, M., Obermeit, J.: Anderson model with decaying randomness—mobility edge. Math. Zeit. **235**(3), 421–433 (2000)
35. Kirsch, W., Metzger, B.: The integrated density of states for random Schrödinger operators. In: Gesztesy, F., Deift, P., Galvez, C., Perry, P., Schlag, W. (eds.) Spectral Theory and Mathematical Physics: A Festschrift in Honor of Barry Simon's 60th Birthday. Proceedings of Symposia in Pure Mathematics, vol. 76, pp. 649–696. Amer. Math. Soc., Providence (2007)
36. Kirsch, W., Krishna, M.: Regularity of the Density of States for the Cauchy distribution **(in Preparation)**
37. Klein, A., Lacroix, J., Speis, A.: Regularity of the density of states in the Anderson model on a strip for potentials with singular continuous distributions. J. Stat. Phys. **57**(1–2), 65–88 (1989)
38. Klein, A., Speis, A.: Smoothness of the density of states in the Anderson model on a one-dimensional strip. Ann. Phys. **183**(2), 352–398 (1988)
39. Klein, A., Speis, A.: Regularity of the invariant measure and of the density of states in the one-dimensional Anderson model. J. Funct. Anal. **88**(1), 211–227 (1990)
40. Klopp, F.: Localization for some continuous random Schrödinger operators. Commun. Math. Phys. **167**(3), 553–569 (1995)
41. Krishna, M.: Smoothness of density of states for random decaying interaction. Proc. Indian Acad. Sci. Math. Sci. **112**(1), 163–181 (2002). **(Spectral and inverse spectral theory (Goa, 2000))**

42. Krishna, M., Stollmann, P.: Direct integrals and spectral averaging. J. Oper. Theory **69**(1), 279–285 (2013)
43. Kunz, H., Souillard, B.: Sur le spectre des opérateurs aux différences finies aléatoires. Commun. Math. Phys. **78**(2), 201–246 (1980/81)
44. March, P., Sznitman, A.: Some connections between excursion theory and the discrete Schrödinger equation with random potentials. Probab. Theory Relat. Fields **75**(1), 11–53 (1987)
45. Pastur, L.: Spectra of random self-adjoint operators. Russ. Math. Surv. **28**(1), 1–67 (1973)
46. Pastur, L., Figotin, A.: Spectra of Random and Almost-Periodic Operators: Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 297. Springer, Berlin (1992)
47. Reed, M., Simon, B.: IV: Analysis of Operators, vol. 4. Elsevier, Amsterdam (1978)
48. Schlag, W.: On the integrated density of states for Schrödinger operators on $\mathbb{Z}^2$ with quasi periodic potential. Commun. Math. Phys. **223**(1), 47–65 (2001)
49. Simon, B.: Regularity of the density of states for stochastic Jacobi matrices: a review. In: Random Media (Minneapolis, Minn., 1985). The IMA Volumes in Mathematics and its Applications, vol. 7, pp. 245–266. Springer, New York (1987)
50. Simon, B.: Trace Ideals and Thier Applications Mathematical Surveys and Monographs, vol. 20. American Mathematical Society, Providence (2005)
51. Simon, B., Taylor, M.: Harmonic analysis on SL(2, R) and smoothness of the density of states in the one-dimensional Anderson model. Commun. Math. Phys. **101**(1), 1–19 (1985)
52. Simon, B., Wolff, T.: Singular continuous spectrum under rank one perturbations and localization for random Hamiltonians. Commun. Pure Appl. Math. **39**(1), 75–90 (1986)
53. Speis, A.: Smoothness of the density of states in the Anderson model on a one-dimensional strip. ProQuest LLC, Ann Arbor, MI. Thesis (Ph.D.)–University of California, Irvine (1987)
54. Speis, A.: Weak disorder expansions for the Anderson model on a one-dimensional strip at the center of the band. Commun. Math. Phys. **149**(3), 549–571 (1992)
55. Stollmann, P.: Caught by Disorder. Progress in Mathematical Physics, vol. 20. Birkhäuser Boston Inc., Boston (2001). **(Bound states in random media)**
56. Stollmann, P.: From uncertainty principles to Wegner estimates. Math. Phys. Anal. Geom. **13**(2), 145–157 (2010)
57. Veselić, I.: Existence and Regularity Properties of the Integrated Density of States of Random Schrödinger Operators: Lecture Notes in Mathematics, vol. 1917. Springer, Berlin (2008)
58. Wegner, F.: Bounds on the density of states in disordered systems. Z. Phys. B Condens. Matter **44**(1), 9–15 (1981)
59. Yosida, K.: Functional Analysis: Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 123, 6th edn. Springer, Berlin (1980)

Communicated by W. Schlag