Commun. Math. Phys. 157, 245–278 (1993)

Communications in Mathematical Physics

Localization at Large Disorder and at Extreme Energies: An Elementary Derivation*

Michael Aizenman[1] **and Stanislav Molchanov**[2]

[1] Departments of Physics and Mathematics, Princeton University, Jadwin Hall, P.O. Box 708, Princeton, New Jersey 08544, USA. Visiting as Fairchild Scholar at Caltech, Pasadena CA 91125, USA

[2] Department of Probability Theory, Faculty of Mathematics and Mechanics, Moscow State University, Moscow 117234, Russia, and Department of Mathematics, TRB 314, Univ. South. Cal., Los Angeles, CA 90089, USA

Received September 29, 1992

Abstract. The work presents a short proof of localization under the conditions of either strong disorder ($\lambda > \lambda_0$) or extreme energies for a wide class of self adjoint operators with random matrix elements, acting in ℓ^2 spaces. A prototypical example is the discrete Schrödinger operator $H = -\Delta + U_0(x) + \lambda V_x$ on Z^d, $d \geqq 1$, with $U_0(x)$ a specified background potential and $\{V_x\}$ generated as random variables. The general results apply to operators with $-\Delta$ replaced by a non-local self adjoint operator T whose matrix elements satisfy: $\sum_y |T_{x,y}|^s \leqq \text{Const.}$, uniformly in x, for some $s < 1$. Localization means here that within a specified energy range the spectrum of H is of the pure-point type, or equivalently – the wave functions do not spread indefinitely under the unitary time evolution generated by H. The effect is produced by strong disorder in either the potential or in the off-diagonal matrix elements $T_{x,y}$. Under rapid decay of $T_{x,y}$, the corresponding eigenfunctions are also proven to decay exponentially. The method is based on resolvent techniques. The central technical ideas include the use of low moments of the resolvent kernel, i.e., $\langle |G_E(x, y)|^s \rangle$ with s small enough (<1) to avoid the divergence caused by the distribution's Cauchy tails, and an effective use of the simple form of the dependence of $G_E(x, y)$ on the individual matrix elements of H in elucidating the implications of the fundamental equation $(H - E)G_E(x, x_0) = \delta_{x,x_0}$. This approach simplifies previous derivations of localization results, avoiding the small denominator difficulties which have been hitherto encountered in the subject. It also yields some new results which include localization under the following sets of conditions: i) potentials with an inhomogeneous non-random part $U_0(x)$, ii) the Bethe lattice, iii) operators with very slow decay in the off-diagonal terms ($T_{x,y} \approx 1/|x - y|^{(d+\varepsilon)}$), and iv) localization produced by disordered boundary conditions.

Contents

* Work supported in part by NSF Grants PHY-8912067 and PHY-9214654 (MA), and ONR Grant N00014-91-J-1526 (SM)

1. Introduction

In a variety of situations one encounters linear operators whose matrix elements have a random component, for which it is of interest to know whether the generalized eigenfunctions are extended or localized. We discuss this question here for a broad class of operators, which includes the following examples.

1) The discrete Schrödinger operator with random potential, acting in $l^2(\mathbf{Z}^d)$,

$$H = -\Delta + U(x)\,, \tag{1.1}$$

where Δ is the difference operator (the discrete Laplacian)

$$\Delta f(x) = \sum_{|n|=1} [f(x+n) - f(x)]\,, \tag{1.2}$$

and $U(\cdot)$ consists of some specified background $U_0(x)$ and a random potential λV_x:

$$U(x) = U_0(x) + \lambda V_x\,. \tag{1.3}$$

$\{V_x\}$ is a collection of random variables (independent or correlated), and λ is a parameter expressing the strength of the disorder.

2) More general matrix operators, acting in $\ell^2(\Gamma) - \Gamma$ a countable set, of the form

$$H = T + U(x) \tag{1.4}$$

with T a Hermitian matrix $T_{x,y} = \bar{T}_{y,x}$ not necessarily real (but satisfying decay conditions stated in Sect. 3), and U as above.

3) Operators with randomness in the off-diagonal terms, such as the linear operator associated with the quadratic form

$$Q(\psi) = (\bar{\psi}, H\psi) = \sum_{x,y} K_{x,y} |\psi_x - \psi_y|^2 \tag{1.5}$$

with $\{K_{x,y}\}$ random variables. E.g., Q may be an elasticity tensor, with $K_{x,y} \geqq 0$.

Example 1) appears in elementary discussions of solids, in the approximation in which the system of the valance electrons is treated as a Fermi gas with an effective one-body potential. For an ordered material $U_0(\cdot)$ is periodic, and $\{V_x\}$ may

represent the effect of impurities. At $\lambda = 0$ the spectrum is described by the Bloch/Floquet theory: the spectrum shows zone structure (with gaps in $d = 1$), but the eigenfunctions are all extended (Bloch states). For $\lambda \neq 0$, the conductivity properties depend both on the spectral gaps and on the nature of the spectrum in the vicinity of the Fermi level.

Example 3) appears in the vibrational problem described by the (hyperbolic) equation:

$$\frac{\partial^2 u}{\partial t^2} = -Hu\,, \tag{1.6}$$

whose normal modes (i.e., solutions of $u(t, x) = e^{i\omega t}\psi(x)$) are given by the eigenfunctions of H.

The spectral aspect of localization, within a specified energy range, is the existence of a complete system of square integrable (versus extended) eigenfunctions, in which case the spectrum is said to be there of the pure-point type. The direct dynamical manifestation of this feature is the non-spreading of the wave-packet. I.e., pp. spectrum for the energy range (a, b) means equivalently (by the RAGE theorem based on the Wiener criterion [RS]) that for any function ψ in the range of the spectral projection $P_{(a,b)}$ there is A_ψ such that $A_\psi(R) \to 0$, as $R \to \infty$, and

$$\int_{|x| \geqq R} |e^{-itH}\psi(x)|^2\, dx \leqq A_\psi(R) \tag{1.7}$$

uniformly in time t. With few exceptions (such as operators with only power-law decay in the off-diagonal terms) when we prove localization the analysis also shows that the eigenfunctions decay exponentially fast. For the case 1) that is more than enough for the vanishing of the electrical conductance – as given by the Kubo formula.

The localizing effect of disorder was pointed out in the context of the first example by Anderson [A], Mott and Twose [MT] and Landauer [L]. These works and related developments [T, AALR] have triggered a large body of theoretical physics and mathematical studies.

One can identify different circumstances under which localization may occur: i) localization at high disorder, ii) at the edges of the spectrum, and iii) in low dimensions, e.g., $d = 1$ where complete localization (i.e., at all energies) is known to be produced by systematic disorder of any non-zero strength.

For one dimension (first analyzed rigorously by Goldsheid, Molchanov, and Pastur [GMP]) there is now an extensive theory, based on rather direct methods (transfer operators, phase formalism, Lyapunov exponents). Description of the main results, historical reviews, and extensive bibliographies can be found in the recent monographs [CFKS, CL, FP]. In the multidimensional case there exist various open problems. Among them are questions concerning the possible coexistence of pure point spectrum and continuous spectrum above a lower critical dimension. Of particular interest is also the apparently critical case of $d = 2$ dimensions (see, e.g., [T, AALR]).

A central role for the mathematical results concerning localization in $d > 1$ dimensions has been played by the analysis of Fröhlich and Spencer [FS], which presents cluster resolvent expansion and a "multiscale method" for the dealing with the small denominators which appear in the resolvent kernels. The multiscale analysis has been somewhat simplified in the work of Spencer [Sp] and von Dreifus

[D], though it is still not quite elementary. One of the features of this important method – which is not covered by this work – is a finite volume criterion for localization.

The main results we present are elementary proofs of localization in regimes of high disorder, and/or extreme (high and low) energies, in arbitrary dimension. The argument does not require the multiscale analysis, and extends easily to i) potentials with an inhomogeneous non-random part, $U_0(x)$, ii) the Bethe lattice (which purportedly presented some difficulties to the formerly available methods), iii) operators with very slow decay in the off-diagonal terms ($\approx 1/|x-y|^{(d+\varepsilon)}$), and iv) $-\Delta$ with random boundary conditions.

The random potential may be correlated. Our analysis focuses on the derivation of exponential decay for Green's function, $\langle 0|\dfrac{1}{H-E}|x\rangle$, at specified energies. For that, the regularity conditions required of the probability distribution of the potential are quite mild. However, in order to deduce localization, we invoke the Simon–Wolff criterion, and this step requires the absolute continuity of the measure (a more restrictive assumption than that of refs. [CKM, DK1]). These regularity conditions are discussed in Sect. 3. An application of the method to off-diagonal disorder is discussed in Sect. 5.

2. Demonstration: Localization in the Anderson Tight Binding Model

We introduce the basic ideas of our approach in the context of a classical example, formulated by P.W. Anderson [1]. It is a prototype of more general constructions and results which are discussed in the following sections.

Let H be the discrete Schrödinger operator with random potential, acting in $\ell^2(\mathbf{Z}^d)$ and described by (1.1)–(1.3), with $\{V_x\}$ a collection of independent random variables *uniformly distributed* in the interval $[-1, 1]$. The background term $U_0(\cdot)$ can at this point be arbitrary. Examples of particular interest include:

i) $U_0(\cdot) \equiv 0$.
ii) $U_0(x)$ periodic in x. (For $\lambda = 0$ the spectrum is then given by the Floquet/Bloch theory, which plays a basic role in discussions of condensed matter.)
iii) $U_0(x)$ also generated by a random process (independent of $\{V_x\}$).

The latter case shows that the discussion of this chapter actually covers a broad range of random potentials ($U(\cdot)$).

Among the preliminary observations one may note that the operator H is essentially self adjoint on the domain of functions of compact support, and if $U_0(\cdot)$ is sufficiently homogeneous (e.g.: periodic, quasi-periodic, or random but generated by an ergodic process) the spectrum of H is non-random: $\sigma(H) = \sigma(-\Delta + U_0) \dot{+} [-1, 1]$ – for almost every realization of $\{V_x\}$ (and, when appropriate, for a.e. U_0), [CL, FP].

The following proposition presents the two localization statements for which our analysis is geared, in the context described above.

Theorem 2.1. *For each $0 < s < 1$ there is a constant $\kappa_s > 0$, and a non-decreasing function $\zeta_s(\cdot)$ on $\mathbb{R}$ such that*

$$\zeta_s(z) \text{ is } \begin{cases} > 0 & \text{for } z > \kappa_s \\ = -\infty & \text{for } z \leqq \kappa_s \end{cases} \tag{2.1a}$$

and

$$\lim_{z\to\infty}\frac{\zeta_s(z)}{z}=1 \tag{2.1b}$$

with which the following holds in any dimension d:

i. (*High disorder*) *If for some* $s\in(0,1)$

$$|\lambda|>\frac{(2d)^{1/s}}{\kappa_s}, \tag{2.2}$$

then for any $U_0(\cdot)$, *the operator H has, for almost every realization of* $\{V_x\}\equiv\omega$, *a complete set of orthonormal eigenfunctions,* $\{\varphi_n(x)\}$. *These eigenfunctions are exponentially localized, satisfying bounds of the form:*

$$|\varphi_n(x)|\leqq A_n(\omega)\exp[-m|x-x_n(\omega)|] \tag{2.3}$$

with a non-random m $(=\log[(|\lambda|\kappa_s)^s/2d])>0$, *and a varying collection of* $A_n(\omega)<\infty$ *and* $x_n(\omega)\in\mathbb{Z}^d$.

ii. (*Extreme energies*) *With no restriction on* λ, *in the energy range*

$$\left\{E\in\mathbb{R}\colon |E-2d|>\|U_0\|_\infty+|\lambda|\zeta_s\left(\frac{(2d)^{1/s}}{|\lambda|}\right)\right\}, \tag{2.4}$$

the operator H has, for almost every realization, only point spectrum with exponentially localized eigenfunctions (*with* $m=m(E,\lambda)$).

Remark. The second statement adds information only if there is a value of λ at which

$$|\lambda|\zeta_s\left(\frac{(2d)^{1/s}}{|\lambda|}\right)<2d+|\lambda|B,\quad \text{for } B=1\,. \tag{2.5}$$

That is based on the observation that over the spectrum of H: $|E-2d|\leqq 2d+\|U_0\|_\infty+|\lambda|B$, with B a uniform bound on V_x. We present the statement without clarifying the question of the existence of such λ, since the boundedness of V_x is not essential for the proof. A similar result is proven below for unbounded potentials, where $B=\infty$.

As in the previous derivations of such results ([FS, DLS, FMSS, SW]), our analysis focuses on the resolvent kernel $G_E(x,y)=(H-E)^{-1}(x,y)$. Two alternative regularizing cutoffs for these quantities are: an imaginary component for the energy ($G_E(x,y)$ being analytic in E outside of any strip $|\mathrm{Im}\,E|>\varepsilon$ (>0)), and, alternatively, the replacement of H by its natural restriction to a finite-volume $\Lambda\subset Z^d$. In the latter approach one deals with

$$G_E^\Lambda(x,y)=\langle x|\frac{1}{H_\Lambda-E}|y\rangle \tag{2.6}$$

which are rational function of E (with poles at $E_n\in\mathbb{R}$, $n=1,\ldots,|\Lambda|$).

The quantities $G_E^\Lambda(x,y)$ are random variables. Though their dependence on the potential U is somewhat complicated, it is possible to gain some insight into the distribution of each such term by considering its dependence on the potential at $x-$ at fixed values of the potential elsewhere. A suitable tool for this purpose is provided by the finite rank perturbation formulae. E.g., for a diagonal term we

regard H as a rank-one perturbation of the operator $\tilde{H}$ which is obtained from it by changing V_x to 0. By a rank-one perturbation formula:

$$G_E(x, x) = \frac{1}{\tilde{G}_E^{-1}(x, x) + \lambda V_x}, \tag{2.7}$$

where $\tilde{G}_E(x, x)$ is the resolvent of $\tilde{H}$, and as such it has no dependence on V_x.

The relation (2.7) shows that the event that $G_E(x, x)$ is large can be viewed as a resonance of V_x with a value determined by the potential elsewhere. With the natural estimate for the conditional probability, conditioned on $V_{\{x\}^c} \equiv \{V_y | y \in \mathbf{Z}^d, y \neq x\}$, one obtains:

$$\text{Prob}(|G_E(x, x)| \geqq t) \leqq \frac{1}{\lambda t}. \tag{2.8}$$

For the off-diagonal term $G_E(x, y)$ an argument based on the rank-two perturbation formula which is given below in Lemma 2.2, yields a similar bound (proven here in Appendix I, Theorem II.1):

$$\text{Prob}(|G_E(x, y)| \geqq t) \leqq \frac{2\sqrt{2}}{\lambda t}. \tag{2.9}$$

The study of rank-one perturbations plays a central role in Wegner's estimate of the density of states [W], which has played a fundamental role in the subject, and in the analysis of Simon and Wolff [SW]. In fact, Wegner's estimate can be derived from (2.8) and the argument given above was already employed for that purpose in [CKM], where Wegner's like estimate is derived for a large class of probability distributions which need not be absolutely continuous. (We have, however, no previous reference for the off-diagonal bound (2.9).)

The main new contribution of this work is a simple proof of the following statement, which provides a key step towards Theorem 2.1.

Lemma 2.1. *For $0 < s < 1$, there are $\kappa_s > 0$ and $\zeta_s(\cdot)$ as described in the statement of Theorem* 2.1, *such that under any of the two conditions presented there,* i.e., (2.2) *or* (2.4):

$$\langle |G_E^A(x, y)|^s \rangle \leqq D \exp(-m|x - y|) \tag{2.10}$$

with $D < \infty$, $m > 0$, uniformly in the finite volume $\Lambda \subset Z^d$. Furthermore, such a bound holds also for the fractional moments of the infinite-volume quantity $G_{E+i0}(x, y)$, at Lebesgues-almost every E in the indicated range.

Exponential bounds on the resolvent were first derived for the multidimensional case in the work of Fröhlich and Spencer [FS]. The transition from such a bound to the exponential localization was not automatic, and was proposed simultaneously in the three papers [FMSS, DLS, SW]. Ref. [FMSS] (and, based on it, [DK1]) give a direct construction of the eigenstates with exponential decay, for which, strictly speaking, the bound (2.10) is not essential. Refs. [DLS, SW] discuss a simple but very effective argument of Kotani and contain general functional – analytic arguments which guarantee exponential localization if (2.10) is satisfied, assuming some additional restrictions on the inter-dependence of the potential values. An explicit statement of this nature is provided by the work of Simon and Wolff [SW], which reduces the problem of localization to that of the

square summability of the kernel $G_{E+i0}(x, y)$. Following is a version of their generally applicable criterion. (An even more versatile version is found in Sect. 3.)

The Simon–Wolff Criterion (I). Let $H = H_0 + V_x$ be a self adjoint operator in $l^2(\Gamma)$ (Γ a countable set of sites) with H_0 a bounded operator and $\{V_x\}_{x\in\Gamma}$ a collection of random variables, having the property that for each site x the conditional probability distribution of V_x-conditioned on the values of the potential at all other sites, is absolutely continuous with respect to the Lebesgue measure (dV).

If for all $x\in\Gamma$, and Lebesgue-a.e. $E\in[a, b]$:

$$\lim_{\varepsilon\to 0}\sum_{y\in\Gamma}|G_{E+i\varepsilon}(x, y)|^2 < \infty \ , \tag{2.11}$$

for almost every realization of $\{V_x\}$, then almost surely the operator H has only pure point spectrum in the interval $[a, b]$.

Furthermore, if under the conditions stated above, $G_{E+i0}(x, y)$ (which exists for $\mathscr{L}$-a.e. E) decays exponentially, in some metric on Γ, then so do the eigenfunctions $\varphi_E(y)$, for $E\in[a, b]$.

Remarks. 1) The sum in (2.11) always has a limit (possibly ∞) since the expression equals $\langle x|\dfrac{1}{(H-E)^2+\varepsilon^2}|x\rangle$ which is monotone in ε.

2) One should appreciate that there is a fine point in the above statements, without which the condition would be of no use. And that is that the bounds on the resolvent are only required to hold separately at each energy. In fact, if the criterion applies, then for each typical potential there would be a countable set of energies at which the resolvent diverges, and (3.1) fails. The reason that this observation does not empty the stated criterion is seen in the eigenfunction expansion, which shows that (2.11) may hold if the eigenfunctions are sufficiently localized, and the eigenstates with energies increasingly resonant with any a-priori chosen energy E have rapidly decreasing weights at any given site.

We shall use the SW criterion in deducing Theorem 2.1 from the exponential bound of Lemma 2.1, to whose proof we turn now.

It may be worth pointing out that the first simplifying idea presented in this paper is to consider the moments seen in (2.10). A glance at the formula (2.15) shows that one should expect $\langle|G_E(x, y)|\rangle$ to diverge for E in the spectrum of H (the distribution of $G_E(x, y)$ has long tails, similar to that of the Cauchy distribution). However, as (2.9) and (2.8) show, $|G_E(x, y)|^s$ is integrable for any power $s < 1$, with

$$\begin{aligned}\langle|G_E^A(x, y)|^s\rangle &= \int_0^\infty \mathrm{Prob}(|G_E^A(x, y)|^s \geqq t)dt \\ &\leqq \int_0^\infty \min\left\{1, \frac{2\sqrt{2}}{\lambda t^{1/s}}\right\}dt = \frac{1}{1-s}\frac{(2\sqrt{2})^s}{|\lambda|^s} := D \ ,\end{aligned} \tag{2.12}$$

where the factor $2\sqrt{2}$ can be replaced by 1 when $x = y$. The fractional moments provide a convenient hold on the probability distribution, and, as we shall see, are particularly well suited for our purpose.

In the derivation of Lemma 2.1 we fix x_0, and denote simply $G(x) = G_E^A(x_0, x)$. Our discussion can also be applied directly to $G_{E+i0}(x_0, x)$ though in that case one

has to acknowledge that $\lim_{\varepsilon \to 0} G_{E+i\varepsilon}(x_0, x)$ is initially guaranteed to exist only for $\mathscr{L}$-a.e. E. In either case, $G(\cdot)$ is a solution of the equation:

$$[-\Delta + U(x) - E]G(x) = \delta_{x_0,x} \tag{2.13a}$$

which for $x \neq x_0$ means

$$[\lambda V_x + U_0(x) + 2d - E]G(x) = \sum_{z \in \mathbf{Z},\, |n|=1} G(x+n)\,. \tag{2.13b}$$

Since for $s < 1$: $|a_1 + \cdots + a_k|^s \leqq |a_1|^s + \cdots + |a_k|^s$ for any collection of complex numbers $\{a_i\} \subset \mathbb{C}$, Eq. (2.13b) yields:

$$|\lambda V_x + U_0(x) + 2d - E|^s |G(x)|^s \leqq \sum_{|n|=1} |G(x+n)|^s \quad (\text{for } x \neq x_0)\,. \tag{2.14}$$

Our argument proceeds by taking the expectation value of the inequality (2.14). The goal is to obtain a statement for $\langle |G(x)|^s \rangle$ as a function of x. However one first encounters the difficulty that on the left side we obtain the expectation value of a product of $|G(x)|^s$ with another quantity. This could be a serious problem. We resolve it by focusing on the conditional expectation of (2.14) – conditioned on $V_{\{x\}^c}$, and bringing up the fact that at fixed $V_{\{x\}^c}$, $G(x)$ is a simple function on V_x, namely it is of the form

$$G(x) = \frac{\alpha}{V_x - \beta} \tag{2.15}$$

with α and β some functions of $V_{\{x\}^c}$. The representation is implied by the following finite-rank perturbation formula.

Lemma 2.2 (Krein formula). *Let $\hat{H}$ be the operator obtained from H by changing V_x and V_y to 0, and let $R_{x,y}$ be the orthogonal projection on the two dimensional space (assuming $x \neq y$) spanned by the vectors $|x\rangle$ and $|y\rangle$ (i.e., $\delta_{\cdot,x}$ and $\delta_{\cdot,y}$). With some abuse of notation, we denote by $A_{2\times 2}$ the 2×2 matrix which gives the restriction of the resolvent of $\hat{H}$ to the range of $R_{x,y}$:*

$$A_{2\times 2} = R_{x,y} \frac{1}{\hat{H} - E} R_{x,y}\,. \tag{2.16}$$

Then,

$$G_E^{\Delta}(x, y) = \left[\begin{pmatrix} \lambda V_x & 0 \\ 0 & \lambda V_y \end{pmatrix} + A_{2\times 2}^{-1} \right]^{-1}_{x,y}, \tag{2.17}$$

where $[\cdot]^{-1}$ is interpreted as the inverse of a 2×2 matrix. In case $x = y$, the one dimensional analog holds, and takes the form of Eq. (2.7).

The proof of the formula (2.17) is given in the appendix. A somewhat striking aspect of it is that the complete information on the conditional dependence of $G_E(x, y)$ on V_x is contained in not more than four real numbers (three in case of real operators like our H). Applying Kramer's rule for the inverse of a matrix

$$G_E^{\Delta}(x, y) = -\frac{A^{-1}_{x,y}}{\det\left[\begin{pmatrix} \lambda V_x & 0 \\ 0 & \lambda V_y \end{pmatrix} + A_{2\times 2}^{-1} \right]}\,. \tag{2.18}$$

Since the determinant is linear in V_x, (2.18) yields the representation (2.15).

The following decoupling lemma plays now a key role. (In Sect. 3 and Appendix III we present its extensions to other measures and other collections of functions with controllable singularities.)

Lemma 2.3 (A decoupling principle). *For each $0 < s < 1$, there is an increasing, positive, function $\theta_s(\cdot)$ on $\mathbb{R}_+$ such that*

$$\lim_{\eta \to \infty} \frac{\theta_s(\eta)}{\eta} = 1 \tag{2.19}$$

with which

$$\int_{-1}^{1} \frac{dV}{2} |\eta - V|^s \frac{1}{|\beta - V|^s} \geqq [\theta_s(|\eta|)]^s \int_{-1}^{1} \frac{dV}{2} \frac{1}{|\beta - V|^s} \tag{2.20}$$

for all $\eta, \beta \in \mathbb{C}$.

Remark. The quantities mentioned in Theorem 2.1 are defined as follows:

$$\kappa_s = \min\{\theta_s(\eta) \colon \eta \in \mathbb{R}\}\,, \tag{2.21a}$$

and

$$\zeta_s(\cdot) = \theta_s^{-1}(\cdot)\,, \tag{2.21b}$$

where the inverse is in the function theoretic sense.

Proof. The bound (2.20) is obviously satisfied in the region $|\eta| \geqq 2$, with the tentative choice $\theta_s(\eta) = |\eta| - 1$. The following argument covers the regime $|\eta| \leqq 2$.

Let $\delta < 1$. Outside the disc $\{V \subset \mathbb{C} \,|\, |V - \eta| < \delta\}$ the first factor is bounded below: $|\eta - V|^s \geqq \delta^s$. Thus it suffices to show that

$$\min_{\beta,\, |\eta| \leqq 2} \frac{\int_{-1}^{1} \frac{dV}{2} I[|V - \eta| < \delta] |\beta - V|^{-s}}{\int_{-1}^{1} \frac{dV}{2} |\beta - V|^{-s}} > 0\,. \tag{2.22}$$

For $s < 1$, the quantity which is minimized is:

i) a jointly continuous function of β and η, which in all $\mathbb{C} \times \mathbb{C}$ is strictly positive,

ii) for $|\beta| > 2$ it is bounded below by $(1 - \delta)\dfrac{(|\beta| - 1)^s}{(|\beta| + 1)^s} \geqq \dfrac{1 - \delta}{3^s}$.

These two imply that the minimum over $\mathbb{C} \times \{\eta \in \mathbb{C} \colon |\eta| \leqq 2\}$ is indeed strictly positive. (For an explicit value of the constant, let us mention that the minimum occurs at $(\beta, \eta) = (0, 0)$.)

The desired function $\theta_s(\cdot)$ is constructed by taking the monotone increasing minorant of the function obtained by the combination of the bounds covering the two regimes. ■

Proof of Lemma 2.1. i. Let us return to (2.14), and average it over the randomness by first taking the conditional expectation – conditioned on $V_{\{x\}^c}$. Using the representation (2.15), the expected value of the expression on the left side can be cast in the following form:

$$\langle |\lambda V_x + U_0(x) + 2d - E|^s |G(x)|^s \rangle = \int dV_{\{x\}^c} \int_{-1}^{1} \frac{dV_x}{2} |\eta - \lambda V_x|^s \frac{\alpha}{|\beta - \lambda V_x|^s} \tag{2.23}$$

with $\{\alpha, \beta, \eta\}$ determined by $V_{\{x\}^c}$, U_0, E, and λ. The decoupling principle (Lemma 2.3) yields the lower bound:

$$\geqq |\lambda|^s \kappa_s^s \int dV_{\{x\}^c} \int_{-1}^{1} \frac{dV_x}{2} \frac{\alpha}{|\beta - \lambda V_x|^s} = (|\lambda| \kappa_s)^s \langle |G(x)|^s \rangle \,. \tag{2.24}$$

Combining (2.24) with (2.14), we find that

$$\langle |G(x)|^s \rangle \leqq \frac{\gamma}{2d} \sum_{n \in \mathbf{Z},\, |n|=1} \langle |G(x+n)|^s \rangle, \quad \text{for all } x \neq x_0 \,, \tag{2.25}$$

with

$$\gamma = \frac{2d}{(|\lambda| \kappa_s)^s} \,. \tag{2.26}$$

The condition (2.2) on λ means that $\gamma < 1$. In that case, we write

$$\gamma = e^{-m}, \quad \text{with } m > 0 \,. \tag{2.27}$$

For each x the inequality (2.25) can be safely iterated $|x - x_0|$ times without generating terms which violate the condition $x + n_1 + n_2 + \cdots \neq x_0$). Summing the resulting terms, and using the uniform bound (2.12), we obtain

$$\langle |G(x)|^s \rangle \leqq \gamma^{|x - x_0|} D = D \exp(-m|x_0 - x|) \,, \tag{2.28}$$

which is (2.10).

ii. The high energy statement is derived by the same route. The difference is that instead of passing from (2.23) to (2.24) we pick the best factor provided by (2.20), and not just κ_s. For uniformly bounded U_0 the result is:

$$\begin{aligned} \langle |\lambda V_x + U_0(x) + 2d - E|^s |G(x)|^s \rangle &\geqq |\lambda|^s \inf_x \left\{ \theta_s \left(\frac{|E - 2d - U_0(x)|}{|\lambda|} \right)^s \right\} \langle |G(x)|^s \rangle \\ &\geqq \left[|\lambda| \theta_s \left(\left| \frac{|E - 2d| - \| U_0 \|}{|\lambda|} \right|_+ \right) \right]^s \langle |G(x)|^s \rangle \,, \end{aligned} \tag{2.29}$$

where last step is by the monotonicity of $\theta(\cdot)$ and the inequality $|E - 2d - U_0(x)| \geqq |E - 2d| - \| U_0 \|$. Thus (2.25) holds and the rest of the above argument applies, with γ modified to:

$$\gamma = \frac{2d}{\left[|\lambda| \theta_s \left(\left| \frac{|E - 2d| - \| U_0 \|}{|\lambda|} \right|_+ \right) \right]^s} \,. \tag{2.30}$$

The energy condition (2.4) is equivalent to the requirement that $\gamma < 1$.

iii. Let us consider now the resolvent of the infinite-volume operator H, under the stated conditions. All the arguments used above are valid also when $G(x)$ refers to the quantity $G_{E+i\varepsilon}(x_0, x)$, with E replaced by $E + i\varepsilon$, $\varepsilon \neq 0$, and $\mathrm{Im}(E) = 0$. By general arguments (see [SW]), the limit $\varepsilon \to 0+$ exists at given $\{V_x\}$ for $\mathscr{L}$ (Lebesgue)-almost every E. Reversing this statement (with the aid of Fubini's theorem): for $\mathscr{L}$-a.e. E the limit defining $G_{E+i0}(x_0, x)$ exists for almost every

potential $\{V_x\}$. For such energies E:

$$\langle |G_{E+i0}(x_0, x)|^s \rangle \leqq \limsup_{\varepsilon \to 0+} \langle |G_{E+i\varepsilon}(x_0, x)|^s \rangle \leqq D \exp(-m|x_0 - x|) \,. \tag{2.31}$$

■

[Let us comment that one could also derive the above result without the restriction to $\mathscr{L}$-a.e. E, by constructing $G_{E+i0}(x_0, x)$ by means of weak limits, in the sence of $l^2(\mathbf{Z}^d)$, of $G_E^{\Lambda}(x_0, x)$ with $\Lambda \to Z^d$. Part of the argument is seen in the proof of Lemma 3.2 below.]

Proof of Theorem 2.1. Assume that one of the two conditions stated in Theorem 2.1 is satisfied, with some $s < 1$. Lemma 2.1 implies then the exponential decay of $\langle |G_{E+i0}(x_0, x)|^s \rangle$ with some $m > 0$ (for $\mathscr{L}$-a.e. E). For any $a < m$:

$$\left\langle \left[\sum_{x \in \mathbf{Z}} |G_{E+i0}(x_0, x)|^2 \exp\left(+ \frac{2a}{s} |x_0 - x| \right) \right]^{s/2} \right\rangle$$

$$\leqq \left\langle \sum_{x \in \mathbf{Z}} |G_{E+i0}(x_0, x)|^s \exp(+a|x_0 - x|) \right\rangle < \infty \,, \tag{2.32}$$

where the last statement is by (2.10). It follows that the sum in the bracket $[\,\cdot\,]$ is finite for $\mathscr{L} \times \rho(dV)$ almost every $\{E, V\}$, which means that the Simon Wolff criterion for localization, in fact exponential localization, is satisfied. That proves the claim made in Theorem 2.1. ■

Let us add some words to clarify the phenomenon discussed here. Writing Eq. (2.13b) in the form

$$G(x) = \frac{1}{[\lambda V_x + U_0(x) + 2d - E]} \sum_{n \in \mathbf{Z}, |n| = 1} G(x + n) \,, \tag{2.33}$$

one sees that for large λ, and at large E, $G(x)$ is strictly subharmonic in x – except at the rare sites (which however occur with positive density) where the denominator is close to zero. Intuitively, away from such sites the local spectrum of H does not include E. The existence of the exceptional sites poses what is referred to as the small denominator problem, which in ref. [FS] is addressed by means of the multiscale analysis (of which a simplified form is found in [Sp, D, DK1]). As that analysis and/or the one presented here prove, under suitable conditions the exceptional sites do not affect the picture that much.

It is easy to see from the example presented in this section that our new approach to localization is applicable more generally. The basic elements of this approach are:

1) Estimation of low moments of the resolvent kernel $G_E(x_0, x)$. (That bypasses the difficulty caused by the Cauchy tails of the distribution of this quantity.)

2) Study the dependence of $G_E(x, y)$ on any single potential by means of the suitable finite rank perturbation formula. This basic tool was employed in the analysis of Wegner [W] and Simon and Wolff [SW]. For us, such formulae yield uniform bounds on the fractional moments, and imply that $G_E(x, y)$ is a simple rational function of any single potential.

3) The use of the conditional expectation, in which one averages the equation characterizing $G_E(\cdot, \cdot)$ over a local potential at fixed values of the potential elsewhere.

4) The decoupling lemma, which shows that terms like $|V_x - \eta|$ and $G_E(x, y)$ "are not too dependent". The key to it is control of the possible singularities, and zeros, of these quantities as functions of the single variable V_x.

5) Derivation of exponential decay from the strict subharmonicity of $G_E(x_0, \cdot)$.

One may note that in Step 4 (which of course is enabled only by the other ideas) we in effect resolve the small denominator problem mentioned below (2.33). Somewhat similar difficulties appear in a variety of other situations, e.g. statistical mechanical models with random couplings, and it is quite possible that the ideas presented here may be of even broader use.

3. Generalizations: Operators with Diagonal Disorder

We now turn to various extensions and generalizations of the localization results discussed in the previous section.

a. *The Regularity Assumptions on the Distribution of* $\{V_x\}$. The variables $\{V_x\}$ need not be independent. All that is required of their joint distribution, $\rho(dV)$, is a certain regularity of $\rho(dV_x \mid V_{\{x\}^c})$ – the conditional distributions of V_x conditioned on the values of the potential elsewhere. (In Sect. 5 we further relax the condition presented below by allowing the values of the potential to be rigidly correlated within finite size blocks.)

In order to apply the Simon–Wolff criterion (see Sect. 2) the probability measure is required to be *conditionally absolutely continuous* with respect to the corresponding Lebesgue measure, i.e.: $\rho(dV_x \mid V_{\{x\}^c}) \ll dV_x$. However, the exponential decay of the resolvent, which is a key partial result, is derived here under just the regularity condition which is seen in the following definition.

Definition. 1) *A probability measure* $\rho(du)$, *on the real line, is said (here) to be* τ*-regular (with* $0 < \tau \leqq 1$*) if, with some* $v > 0$ *and* $C < \infty$,

$$\rho([z - \delta, z + \delta]) \leqq C|\delta|^\tau \rho([z - v, z + v]) \tag{3.1}$$

for all $0 < \delta < 1$ *and* $z \in R$.

2) *A joint probability measure* $\rho(dV)$ *of a collection of variables* $\{V_x\}$ *is conditionally* τ*-regular if the conditional distributions* $\rho(dV_x \mid V_{\{x\}^c})$ *satisfy the regularity conditions* (3.1) *with uniform values for all the constants.*

3) *If – in addition: for some* $\varepsilon > 0$, *the conditional expectations of* $|V_x|^\varepsilon$ *are uniformly bounded:*

$$\rho(|V_x|^\varepsilon \mid V_{\{x\}^c}) \leqq B \tag{3.2}$$

then the joint probability measure $\rho(dV)$ *is said to be conditionally* (τ, ε)*-regular.*

Remarks. The regularity condition (3.1) allows $\rho(\cdot)$ to be supported on Cantor sets of any non-zero dimension (τ), though it requires slightly more than Hölder continuity of the measure. We note that an even weaker condition is required in references [CKM, DK1], where $\rho(\cdot)$ need only have a Hölder continuous component, given some estimates on the density of states. Extension of our method in this direction are not to be ruled out, but are not discussed here.

For *absolutely continuous* measures, with $\rho(dx) = \tilde{\rho}(x)dx$, a sufficient condition is that $\tilde{\rho}(x)$ is a piecewise-continuous density function (allowed only isolated discontinuities) which for $|x|$ large enough is monotone decreasing as $|x| \to \infty$.

Following is a generalization of the decoupling principle expressed by Lemma 2.3.

Lemma 3.1. *Let $\rho(du)$ be a τ-regular measure on $\mathbb{R}$. Then for each $s < \tau$ and a pair of integers (n, k), with $s \cdot k < \tau$, there is $\kappa_s^{n,k} > 0$ such that*

$$\int |u - \eta|^s |\mathscr{R}(u)|^s \rho(du) \geqq (\kappa_s^{n,k})^s \int |\mathscr{R}(u)|^s \rho(du) \,. \tag{3.3}$$

for all rational functions:

$$\mathscr{R}(u) = \frac{\mathscr{P}(u)}{\mathscr{Q}(u)} \tag{3.4}$$

with

$$\deg \mathscr{P} \leqq n, \quad \text{and} \quad \deg \mathscr{Q} \leqq k \; (< \tau/s) \,, \tag{3.5}$$

and all η.

If, furthermore, for some $\varepsilon > s$

$$\int |u|^\varepsilon \rho(du) < \infty \tag{3.6}$$

then the inequality (3.3) *can be strengthened as follows:*

$$\int |u - \eta|^s |\mathscr{R}(u)|^s \rho(du) \geqq [\xi_s^{n,k}(|\eta|)]^s \int |\mathscr{R}(u)|^s \rho(du) \,, \tag{3.7}$$

with $\xi_s^{n,k}(\cdot)$ an increasing positive function on $\mathbb{R}_+$, satisfying

$$\lim_{u \to \infty} \frac{\xi_s^{n,k}(u)}{u} = 1 \,. \tag{3.8}$$

The proof is given at the end of Appendix III. With no loss of generality we can assume that $\inf\{\xi_s^{n,k}(\cdot)\} = \kappa_s^{n,k}$, and for consistency with Lemma 2.3 we denote

$$\xi_s^{0,1}(\cdot) = \theta_s(\cdot), \quad \kappa_s^{n,k} \equiv \inf\{\xi_s^{n,k}(\cdot)\} = \kappa_s \,. \tag{3.9}$$

All these quantities depend on the measure $\rho(du)$ – though only through the few quantities which appear explicitly in the regularity conditions (3.1), (3.2) and (3.6).

b. *The General Setup.* The results of Sect. 2 for the Schrödinger operator are included in a more general statement, which is formulated for operators of the form

$$H = T + U_0(x) + \lambda V_x \tag{3.10}$$

on $\ell^2(\Gamma)$, with: i) Γ a countable set, ii) the potential as in (1.3), and iii) T a bounded self adjoint operator whose matrix elements (in the natural basis of $\ell^2(\Gamma)$ are denoted as

$$T_{x,y} = \bar{T}_{y,x} \,. \tag{3.11}$$

With no loss we assume $T_{x,x} \equiv 0$. (When they cannot be absorbed by an energy shift, the diagonal terms can be viewed as incorporated in U.)

Following are two general results. Their derivation is given after the demonstration of some of their implications.

Theorem 3.1 (High disorder). *If*:

i) *$\rho(dV)$ is conditionally absolutely continuous and conditionally τ-regular, with some* $(1 \geqq) \; \tau > 0$, *and*

ii) *T is a bounded self-adjoint operator which for some $s < \tau$ satisfies*

$$\sum_{y \in \Gamma} |T_{x,y}|^s \exp(+a\delta(x, y)) \leqq B_{s,a} \quad \textit{(uniformly in } x \in \Gamma) \tag{3.12}$$

with some pseudo-metric $\delta(\cdot, \cdot)$ on Γ and $a \geqq 0$, then for each $m \leqq a$ there is some $\lambda_0(m) < \infty$ such that, with any $U_0(\cdot)$, for $\lambda > \lambda_0$ the operator H has almost surely a complete set of orthonormal eigenfunctions, which decay exponentially in δ:

$$|\varphi_n(x)| \leqq A_n(\omega) \exp[-m\delta(x, x_n(\omega))] \,. \tag{3.13}$$

We note that the value $a = 0$ is admissible here. I.e., the basic requirement of T for localization (p.p. spectrum) at high disorder is that for some $s < \tau$,

$$\sum_{y \in \Gamma} |T_{x,y}|^s \leqq B_{s,0} < \infty \,, \tag{3.14}$$

uniformly in $x \in \Gamma$.

Theorem 3.2 (Extreme energies). *If:*

i) *$\rho(dV)$ is conditionally absolutely continuous and conditionally (τ, ε)-regular and*

ii) *T is a bounded self adjoint operator satisfying* (3.12) *with $s < \min\{\tau, \varepsilon\}$,*

then there is function $E_0(\lambda)$, with $\sup\{E_0 \colon \lambda \in \mathbb{R}\} = \hat{E}_0 < \infty$, such that for each bounded $U_0(\cdot)$ the operator H has almost surely only pure point spectrum in the energy range $|E| > E_0(\lambda) + \|U_0\|_\infty$, and the corresponding eigenfunctions obey (3.13).

The above results obviously apply to Schrödinger operators on graphs for which the number of neighbors of $x \in \Gamma$ is uniformly bounded – such as the lattices Z^d, and Bethe lattices. Other interesting examples on these lattices/graphs are operators with T not necessarily real (i.e., without the time-reflection invariance) having $\{T_{x,y}\}$ uniformly bounded and of finite range. The natural choice for the metric $\delta(x, y)$ is the Euclidean distance on Z^d, or the length of a minimal connecting path. However, as we describe next, the general results can also be applied with some less obvious choices of the metric $\delta(\cdot, \cdot)$.

c. *Operators with Long Range Hopping Terms.* The condition (3.14) allows the operator T to have slowly decaying off diagonal terms. In case of the d-dimensional lattice $\Gamma = Z^d$, it suffices for T to have the power-law decay:

$$|T_{x,y}| \leqq \frac{\text{const.}}{|x - y|^{d+\varepsilon}} \,, \tag{3.15}$$

with some $\varepsilon' > 0$. We note that for translation invariant $T_{x,y}$ $(= T_{x-y})$, the last condition assures also the boundedness of the operator T (as is easily seen in the Fourier-transform representation).

For T with slowly decaying off-diagonal elements the eigenfunctions discussed in Theorems 3.1 and 3.2 do not decay exponentially, and indeed the condition (3.12) fails for the standard choice: $\delta(x, y) = |x - y|$ ($|\cdot|$ – the Euclidean norm). However, (3.12) does hold (with some $a > 0$) for another choice of the metric, namely:

$$\delta(x, y) = \log(1 + |x - y|) \,, \tag{3.16}$$

with which the condition becomes:

$$\sum_{y\in\Gamma} |T_{x,y}|^s |x-y|^a \leqq B_{s,a} \,. \tag{3.17}$$

With this choice for $\delta(\cdot,\cdot)$, our analysis yields power-law upper bounds on the decay of the Green function and of the eigenfunctions.

The use of the norm (3.16) for an automatic extension of results on exponential decay to power-law decay was pointed out, in a different context, in an earlier work of L. Gross [G]. The statement we obtain goes beyond the previously proven results for operators with long-range off diagonal terms. These include the proof of singular spectrum under the condition

$$d=1, \quad T_{x,y} \leqq \frac{\text{Const.}}{|x-y|^{4+\varepsilon'}} \tag{3.18}$$

([SiSp]), and of pure point spectrum if $T_{x,y} \leqq C/|x-y|^{6+\varepsilon'}$ ([MS]). Although in these works, which refer only to $d=1$, λ can be arbitrary.

d. *Derivation of the General Statement.* As in the proof of Theorem 2.1, our analysis focuses on the fractional moments of the resolvent. First, in lieu of (2.9) we have the following estimate, which follows from (II.3) of Appendix II and the conditional τ-regularity of the measure:

$$\text{Prob}(|G_E^\Lambda(x,y)| \geqq t| \leqq \frac{\text{Const.}}{t^\tau} \,. \tag{3.19}$$

Hence (2.10) can be extended as follows

$$\langle |G_E^\Lambda(x,y)|^s \rangle \leqq D_s \,, \tag{3.20}$$

with $D_s < \infty$ for $s < \tau$.

As in the discussion of Sect. 2, our main new contribution is the elementary derivation of the following extension of Lemma 2.1.

Lemma 3.2. *Under the conditions described in Theorem* 3.1 *and Theorem* 3.2, *the resolvents of the finite volume restrictions* H_Λ ($= R_\Lambda H R_\Lambda$ *acting in* $\ell^2(\Lambda)$) *satisfy*:

$$\sum_{y\in\Lambda} \langle |G_E^\Lambda(x,y)|^s \rangle \exp[+m\delta(x,y)] = \hat{D} \,, \tag{3.21}$$

with a finite constant $\hat{D}$ *which is uniform in* $x\in\Lambda$, *and in* $\Lambda\subset\Gamma$. *Furthermore, the same bound applies to the moments of the infinite-volume resolvent* $G_{E+i\varepsilon}(x,y)$, *with* $\varepsilon>0$, *and at* $\mathscr{L}$-a.e. *energy* E, *within the indicated range, it also holds for* $G_{E+i0}(x,y)$.

Lemma 3.2 is derived by following the steps taken in the proof of Lemma 2.1, with only few significant differences.

Proof. 1) $\Lambda\subset\Gamma$ finite. With E fixed, we denote:

$$G(x) = G_E^\Lambda(x_0,x) \equiv \langle x_0| \frac{1}{H-E} |x\rangle \,. \tag{3.22}$$

For $x \neq x_0$ it satisfies the equation:

$$[\lambda V_x + U_0(x) - E\,]G(x) = \sum_{y \in \Gamma} T_{x,y} G(y)\,, \tag{3.23}$$

which is a generalization of (2.13b).

From (3.24) we deduce (as in (2.14)), that for $0 < s < 1$, and $x \neq x_0$;

$$\langle |\lambda V_x + U_0(x) - E|^s |G(x)|^s \rangle \leqq \sum_{y \in \Gamma} \langle |T_{x,y}|^s |G(y)|^s \rangle\,, \tag{3.24}$$

where the bracket represents the average over $\{V_x\}$.

The Krein formula, and the implied representation (2.15) do apply in the more general setup. Decoupling the expectation of the product by means of Lemma 3.1, which is used in the same fashion as Lemma 2.3 in the derivation of (2.24) and (2.29), we obtain

$$\langle |\lambda V_x + U_0(x) - E|^s |G(x)|^s \rangle \geqq [|\lambda|\beta_s]^s \langle |G(x)|^s \rangle \tag{3.25}$$

with two possible choices for β_s:

i. For $\rho(dV)$ which is only known to be conditionally τ-regular, at $\tau > s$, we have

$$\beta_s = \kappa_s\,. \tag{3.26}$$

ii. Under the additional assumption that $\rho(dV)$ is conditionally (τ, ε)-regular, and $s < \min\{\tau, \varepsilon\}$, we also have (3.25) with

$$\beta_s = \inf_x \left\{ \theta_s \left(\frac{|E - U_0(x)|}{|\lambda|} \right) \right\} \geqq \theta_s \left(\left| \frac{|E| - \|U_0\|}{|\lambda|} \right|_+ \right). \tag{3.27}$$

The two relations (3.24) and (3.25) yield:

$$\langle |G(x)|^s \rangle \leqq (|\lambda|\beta_s)^{-s} \sum_{y \in \Gamma} |T_{x,y}|^s \langle |G(y)|^s \rangle, \quad x \neq x_0\,. \tag{3.28}$$

Notice that under the stated conditions $(\lambda\beta_s) \to \infty$ when either $\lambda \to \infty$ or, in case ii, $E \to \infty$.

In the general case (with $T_{x,y}$ possibly of unbounded range) we cannot derive the exponential decay by just iteration of (3.28), the way (2.24) was handled. A convenient method to bypass this difficulty employs the weighted sum:

$$\Xi(s, m) = \sum_{x \in \Lambda} \langle |G(x)|^s \rangle \exp(+m\delta(x_0, x))\,, \tag{3.29}$$

which for $s < \tau$ is finite, by (3.20). Let us apply (3.29) to the terms with $x \neq x_0$. Making a judicious use of the triangle inequality, in the form:

$$\exp(+m\delta(x_0, x)) \leqq \exp(+m\delta(x_0, y))\exp(+m\delta(x, x)) \tag{3.30}$$

one finds that

$$\Xi(s, m) \leqq D_s + \frac{B_{s,m}}{(|\lambda|\beta_s)^s}\, \Xi(s, m)\,. \tag{3.31}$$

We now let $s\ (<\tau)$ take a value with which (3.12) is assumed to be satisfied for some $a > 0$. By the dominated convergence theorem, $B_{s,m}$ is continuous in m, for

$0 \leqq m \leqq a$. Letting either λ or E to infinity, under the conditions spelled above one reaches the region of values for which $(|\lambda|\beta_s)^s > B_{s,0}$, e.g., (see (3.27)) where

$$|\lambda|\theta_s\left(\left|\frac{|E| - \|U_0\|}{|\lambda|}\right|_+\right) > B_{s,0} \,. \tag{3.32}$$

By continuity, for each such value of (λ, E), there is some $m > 0$ for which also:

$$\frac{B_{s,m}}{(|\lambda|\beta_s)^s} < 1 \,. \tag{3.33}$$

If (3.12) holds only with $a = 0$, then we still have (3.33), but with $m = 0$.

Under the condition (3.33), the inequality (3.31) yields

$$\Xi(s, m) \leqq \frac{D_s}{1 - B_{s,m}/(|\lambda|\beta_s)^s} \,. \tag{3.34}$$

The resulting uniform bound on $\Xi(s, m)$ directly implies the claim made in (3.21) for the resolvents of the finite volume operators.

2) *The Infinite Volume Case.* The extension of the above analysis to the resolvents of the full (infinite volume) operator requires an additional argument, since our deduction of (3.34) by way of (3.32) and (3.33) hinges on the finiteness of $\Xi(s, m)$. For the full operator that condition is not initially obvious, even for energies off the real axis.

In the infinite volume there are typically many solutions to the equation

$$\sum_{y \in \Gamma} (H - E)_{z,y} g(y) = \delta_{z,x} \,, \tag{3.35}$$

though there is not always a square summable one. If there is such a solution, then either the solution is non-unique in $\ell^2(\Gamma)$, in which case E is a proper eigenvalue, or (when unique) $(H - E)^{-1}|x\rangle$ is well defined and given by $g(\cdot)$, and then

$$g(y) = G_{E+i0}(x, y) = G_{E-i0}(x, y) \,. \tag{3.36}$$

Since the probability that a given energy is a proper eigenvalue can be non-zero for only a countable collection of values of E, we find that for $\mathscr{L}$-a.e. E the former case may be discounted. Thus, our claim will follow if it is proven that under the assumptions of Lemma 3.1, for every E within the relevant range, Eq. (3.35) admits, ρ-almost surely, a square summable solution, and these solutions (which are unique for $\mathscr{L} \times \rho(dV)$ a.e. $(E, \{V\})$) satisfy

$$\left\langle \sum_{x \in \Gamma} |g(x)|^s \exp(+m\delta(x_0, x)) \right\rangle \leqq \frac{D_s}{1 - B_{s,m}/(|\lambda|\beta_s)^s} \,. \tag{3.37}$$

We construct the above solutions of (3.35) from the finite volume resolvents. Let $\Lambda_n \to \Gamma$ be a sequence of finite subsets. For convenience we regard $G_E^{\Lambda_n}(x_0, \cdot)$ as functions on Γ, which vanish outside of Λ_n. Using Fatou's lemma, and the just

derived (3.34):

$$\left\langle \liminf_{n\to\infty} \sum_{x\in\Gamma} |G_E^{\Lambda_n}(x_0, x)|^s \exp(+m\delta(x_0, x)) \right\rangle$$

$$\leqq \liminf_{n\to\infty} \left\langle \sum_{x\in\Gamma} |G_E^{\Lambda_n}(x_0, x)|^s \exp(+m\delta(x_0, x)) \right\rangle$$

$$\leqq \frac{D_s}{1 - B_{s,m}/|\lambda|\beta_s} . \tag{3.38}$$

Hence for ρ almost every realization of $\{V_x\}$, the "$\liminf_{n\to\infty}$" of the above sum is finite. For such $\{V_x\}$ one may pick a subsequence for which the above $\liminf_{n\to\infty}$ is attained as a limit. Along this subsequence, the functions $G_E^{\Lambda_n}(x_0, \cdot)$ are of uniformly bounded $\ell^2(\Gamma)$ norms (by the Hölder inequality, seen in (2.32)). By the weak compactness of the unit ball, there is then a further subsequence for which $G_E^{\Lambda_n}(x_0, \cdot)$ converge weakly in $\ell^2(\Gamma)$ (and hence also pointwise on Γ):

$$G_E^{\Lambda_n}(x_0, \cdot) \xrightarrow[n\to\infty]{} \hat{G}(\cdot) \quad \text{(weakly in } \ell^2(\Gamma)) . \tag{3.39}$$

It is easy to see that the limit obeys Eq. (3.35), is in $\ell^2(\Gamma)$, and satisfies the bound

$$\sum_{x\in\Gamma} |\hat{G}(x)|^s \exp(+m\delta(x_0, x)) \leqq \liminf_{n\to\infty} \sum_{x\in\Gamma} |G_E^{\Lambda_n}(x_0, x)|^s \exp(+m\delta(x_0, x)) . \tag{3.40}$$

By our previous discussion, that proves the last part of Lemma 3.2. ■

Proofs of Theorems 3.1 and 3.2. The last statement in Lemma 3.2 directly implies that the Simon–Wolff criterion for localization is satisfied under the assumptions made in either of the two theorems. ■

4. Some Further Examples

In this section we mention examples indicating other situations in which the method presented above can be applied. We abstain here from stating the most general theorems, and abbreviate the proofs, which are based on the arguments seen above.

a. *Δ with Random Boundary Conditions.* Consider the Laplace operator, $H = -\Delta$, in the upper half space $\mathbf{Z}^{d-1} \times [-1, \infty)$ with the boundary conditions

$$\psi(x, -1) = -\lambda V(x, \omega)\psi(x, 0) \quad (x \in \mathbf{Z}^{d-1}) . \tag{4.1}$$

The corresponding spectral problem can alternatively be described in terms of the operator H acting on functions defined over the half space $\mathbf{Z}^d_+ = \{(x, z) \in \mathbf{Z}^d : x \in \mathbf{Z}^{d-1}, z \geqq 0\}$ as:

$$H\psi(x, z) = -\Delta\psi(x, z) \quad \text{for points } (x, z), \text{ with } x \in \mathbf{Z}^{d-1}, z > 0 ,$$

$$\text{and} \quad H\psi(x, 0) = -\left[\psi(x, 1) + \sum_{|x'-x|=1} \psi(x', 0)\right] + [2d + \lambda V(x, \omega)]\psi(x, 0) . \tag{4.2}$$

Our discussion of the effect of the random boundary conditions can be easily generalized to operators of the form

$$-\Delta + U_0(x, z) + \lambda\xi(x, \omega)\delta_{z,0}\ , \quad (x, z) \in \mathbf{Z}^d_+\ , \tag{4.3}$$

with $H_0 = -\Delta + U_0(x, z)$ bounded. Such models appear naturally in considerations of the effects of surface disorder in somewhat idealized models on solid state physics, and were studied in ref. [KP] (concerning the density of states and some problems connected with Lifshitz tails).

In the following we focus on the prototypical case (4.2). Regardless of $V(x, \omega)$:

$$\sigma(H) \supset \sigma_{\text{ess}}(H) \supset \sigma(-\Delta) = [0, 4d]\ . \tag{4.4}$$

By standard arguments (analogous to those found in [CL, FP]): if $V(x, \omega)$ is a homogeneous ergodic process on $\mathbf{Z}^{d-1}$ then the spectrum $\sigma(H)$ is non-random, and if $V(x, \omega)$ are IID random variables then almost surely:

$$\sigma(H) = \sigma_{\text{ess}}(H) = ([1, 4d - 1] \dot{+} \operatorname{supp} \operatorname{dist} V) \cup [0, 4d]\ . \tag{4.5}$$

In particular, if $V(x, \omega)$ are IID with distribution which is supported on all $\mathbb{R}$ then $\sigma(H) = \mathbb{R}$ (a.s.).

Theorem 4.1. *If the joint distribution of* $\{V(x)\colon x \in \mathbf{Z}^{d-1}\}$ *is conditionally absolutely continuous and conditionally* (τ, ε)*-regular then there exists a function* $\delta(\lambda)$ $(< \infty)$ *such that within the energy range*

$$\{E\colon \operatorname{dist}\{E, \sigma(-\Delta)\} > \delta(\lambda)\} \equiv \{E\colon |E - 2d| > 2d + \delta(\lambda)\}\ ; \tag{4.6}$$

the Laplacian $-\Delta$ *with the random boundary conditions* (4.1) *has only pure point spectrum, with exponentially localized eigenfunctions. Furthermore,* $\lim_{\lambda \to \infty} \delta(\lambda) = 0$.

The proof can be presented in two ways, and it is interesting to see it from both perspectives. We employ in it the original, local, version of the Simon–Wolff [SW] criterion which is more general than the one needed in Sect. 2. This version is used also in other extensions presented below, in Sects. 4.b and 5.

The Simon–Wolff Criterion (II). Let H be a self adjoint in $\ell^2(\Gamma)$, Γ a countable set of sites, which includes a rank-one term with a random coefficient, i.e.: $H = H_0 + u|\alpha\rangle\langle\alpha|$ with u a random variable, whose probability distribution is absolutely continuous $(\rho(du)) \ll \mathscr{L}(du))$. If for $\mathscr{L}$-a.e. $E \in [a, b]$:

$$\sum_{y \in \Gamma} \left| \langle \alpha | \frac{1}{H - E - i0} | y \rangle \right|^2 < \infty\ , \tag{4.7}$$

for ρ-a.e. u, then ρ-almost surely the spectral measure of H associated with $|\alpha\rangle$ includes only pure point spectrum in the interval $[a, b]$. I.e., within the subspace for which $|\alpha\rangle$ is a cyclic vector – $\sigma(H) \cap [a, b] = \sigma_{\text{pp}}(H)$.

Furthermore, if for some metric $\delta(\cdot, \cdot)$ on Γ:

$$\left| \langle \alpha | \frac{1}{H - E - i0} | y \rangle \right| \leqq C(u, E) \exp[-m(E)\delta(x_0, y)] \tag{4.8}$$

with $C(u, E) < \infty$ for $\mathscr{L} \times \mathscr{L}$-a.e. pair (E, u), and $m(E)$ which is strictly positive on $[a, b]$, then the spectral projections associated with the eigenvalues $(\{E_n\})$ of H satisfy

$$|\langle \alpha | P_{\{E_n\}} | y \rangle| \leqq A(u, E_n) \exp[-m(E)\delta(x_0, y)] \tag{4.9}$$

with a function $A(\cdot, \cdot)$ which is finite at $\rho \times \mathscr{L}$-almost every (u, E).

Proof of Theorem 4.1. One way to derive the statement is to note that in the bulk (i.e., $z > 0$) and on the boundary ($z = 0$) a suitable version of inequality (2.25) is satisfied, though for different reasons:

i) for (x, z) with $z > 0$, (2.25) holds with

$$\gamma = \frac{2d}{|E - 2d|^s}\,; \tag{4.10}$$

ii) at (x, z) with $z = 0$, the relevant version of (2.25) is with $2d$ replaced by $(2d - 1)$ and

$$\gamma = \frac{2d - 1}{(|\lambda|\kappa_s)^s}\,. \tag{4.11}$$

The argument seen in the proof of Lemma 2.1 yields therefore exponential decay of the resolvent in the energy range $\{E\colon |E - 2d| > 2d + \delta\}$, provided (δ, λ) satisfy

$$\max\left\{\frac{2d}{(2d + \delta)^s}, \frac{2d - 1}{(|\lambda|\kappa_s)^s}\right\} < 1\,. \tag{4.12}$$

Thus we have an analog of Lemma 2.1, with the conditions described in Theorem 4.1.

Since the potential fluctuates only along the boundary of the half space, $\mathbf{Z}^{d-1} \times \{0\}$, we can use only the local version of the SW criterion (i.e., the above version II). The conclusion it yields directly is that within the specified energy range the spectral measures of H associated with functions supported on the boundary of the half space (i.e. on $\mathbf{Z}^{d-1} \times \{0\}$) are of the pure point type, with exponentially localized eigenfunctions. That leaves room for other spectral behavior of H within the specified range, but only within the subspace of functions ψ with the property that for all $k \geqq 0$: $H^k\psi(x, 0) \equiv 0$, i.e., $\Delta^k\psi(x, 0) \equiv 0$. It is not difficult to see that this subspace consists of only the null function 0, however it is even easier to see that the spectrum of H within this subspace does not leave $\sigma(-\Delta)$. Either way, we have a proof of the claim. ■

An alternative derivation of Theorem 4.1 can be based on the reduction of the d-dimensional spectral problem to a $(d - 1)$-dimensional spectral problem with a non-local operator. More specifically, for energies $E \notin \sigma(-\Delta)$, the resolvent $G(x, z) \equiv G_E((x, z), (0, 0))$ can be determined from its boundary values as the suitable solution of the Dirichlet problem

$$-\Delta\psi(x, z) = E\Psi(x, z), \quad (x, z) \in \mathbf{Z}^d_+, \quad z > 0\,,$$

and

$$\Psi(x, 0) = G_E((x, 0), (0, 0))\,. \tag{4.13}$$

For $E \notin \sigma(-\Delta)$, the problem is well posed, and in particular one can express:

$$G(x, 1) = \sum_{y \in \mathbf{Z}^{d-1}} D_E((x, 1), (y, 0))G(y, 0)\,, \tag{4.14}$$

where $D_E((x, z), (x', z'))$ is the resolvent kernel of $-\Delta$ in the half space ($z > 0$), which in the specified range of energies decays exponentially. Equation (4.14) permits to reduce the d-dimensional resolvent equation to a $(d - 1)$ dimensional problem concerning an operator of the form

$$\hat{H} = -\Delta_{(d-1)} + V_x + T_E \tag{4.15}$$

with T given by

$$T_E(x, y) = D_E((x, 1), (y, 1)) . \tag{4.16}$$

For E in the energy range (4.6), $T_E(x, y) \leqq \text{Const.} \exp(-m|x - y|)$, and hence the discussion of operators with long range hopping terms (Sect. 3) applies, and yields an alternative proof of Theorem 4.1. One may note that the projected spectral problem has a double dependence on E. That however does not affect our analysis, since the results of Sect. 3 are sufficiently uniform.

Remark. When $d = 2$, the above reduction yields a random potential problem in one dimension. The known 1D results suggest that in this case for each $\lambda \neq 0$ the spectrum is localized at all energies $E \notin \sigma(-\Delta)$ (i.e., $E \notin [0, 8]$). Unfortunately, we cannot use for this purpose the results of [MS], since its estimates are not sufficiently uniform. (The problem is related to the above mentioned double dependence on E.)

b. *Strongly Inhomogeneous Randomness.* The following result concerns another situation with inhomogeneous randomness.

Theorem 4.2. *Let H be a discrete Schrödinger operator of the form*

$$H = -\Delta + U_0(x) + D(x)\xi(x) \tag{4.17}$$

on $\ell^2(\mathbf{Z}^d)$, with $D(\cdot)$ a fixed function on $\mathbf{Z}^d$ and $\{\xi(x)\}$ I.I.D. *random variables whose distribtion is τ-regular, with $\tau > 0$. If*

$$\lim_{|x| \to \infty} D(x) = \infty , \tag{4.18}$$

then $\sigma(H) = \sigma_{\text{pp}}(\text{H})$ almost surely, and the eigenfunctions of H decrease exponentially (in fact, super-exponentially).

Proof. The proof requires only minor adjustments in the analysis which led to Theorem 2.1. The key difference is that the inequality (2.25) [which corresponds to (3.29)] is valid with $\gamma < 1$ only where $D(x) > \lambda_0$. That condition however is satisfied in the complement of a finite set, i.e., in $\{x \in Z^d: |x| \geqq R\}$ with R a finite distance beyond which $D(x) > \lambda_0$.

Even with the above proviso, for each pair of sites $\{x, y\}$, the bound on $\langle |G(x, y)|^s \rangle$ can be iterated, from either one end or the other, the total of at least $(|x - x_0| - 2R)$ times. Hence, the following modified form of (2.28) holds:

$$\langle |G(x, y)|^s \rangle \leqq D\gamma^{|x-y|-2R} . \tag{4.19}$$

The rest of the analysis presented in Sect. 2, with the Simon–Wolff criterion (I) replaced by the local version (II), implies that H has the spectral properties claimed in Theorem 4.2 within the subspace spanned by the functions obtained by repeatedly applying H to functions with support in $\{x \in \mathbf{Z}^d: |x| \geqq R\}$. The orthogonal complement of this subspace includes only functions supported in $\{x \in \mathbf{Z}^d: |x| \leqq R\}$, and hence any spectrum left out by the above analysis consists of only eigenfunctions of compact support. (In fact, by a separate argument, that orthogonal complement consists of only the 0 element). ■

This example can be somewhat generalized: $\{\xi(x)\}$ need not be identically distributed, e.g., it suffices for their distributions to be uniformly regular, and $D(\cdot)$ may just satisfy

$$\liminf_{|x|\to\infty} D(x) > \lambda_0 \,, \tag{4.20}$$

with λ_0 of Theorem 3.1. A more complete analysis of the spectral problem for the potential

$$V(x) = |x|^{\alpha}\xi(x) \tag{4.21}$$

(which shows spectral bifurcations with respect to the exponent α) with $\xi(x)$ I.I.D. random variables uniformly distributed on $[0, 1]$ or on $[-1, 1]$ was done in [GJMS], using a quite different method.

Continuing with the setup described in (4.17), if $U_0 \equiv 0$, and

$$\liminf_{|x|\to\infty} D(x) > 0 \,, \tag{4.22}$$

then in the energy range $|E| > E_0$ (with E_0 as in Theorem 3.2) H has only p.p. spectrum, with exponentially localized eigenfunctions.

c. *Correlated Randomness.* The property of a measure being conditionally regular is closely related to the decomposability of the potential as a sum:

$$V_x = V_x^{(1)}(\omega) + V_x^{(2)}(\omega) \tag{4.23}$$

(ω representing the randomness) where $V_{(\cdot)}^{(1)}$ and $V_{(\cdot)}^{(2)}$ are independent, and $V_x^{(2)}$ form I.I.D. random variables whose distribution satisfies the suitable regularity condition. In particular, if $V_x^{(2)}$ is (τ, ε)-regular then so is V_x, and the results of Sect. 3 apply.

For example, such a decomposition is easily evident for a broad class of homogeneous gaussian fields on $\mathbf{Z}^d$. Such fields admit the spectral representation:

$$V(x, \omega) = \int_{[-\pi,\pi]^d} e^{i(k\cdot x)}\mu(dk, \omega) \tag{4.24}$$

with $\mu(dk, \omega)$ a random (gaussian) measure with uncorrelated values, and

$$\langle |\mu(dk, \omega)|^2 \rangle = F(dk) \,. \tag{4.25}$$

$F(dk)$ is a measure on $[-\pi, \pi]^d$, in terms of which:

$$B_x \equiv \langle V_y\, V_{y+x} \rangle = \int_{[-\pi,\pi]^d} e^{i(k\cdot x)}F(dk, \omega) \,. \tag{4.26}$$

If $F(dk)$ dominates the Lebesgue measure dk, on $[-\pi, \pi]^d$, i.e.,

$$F(dk) \geqq \text{const. } dk \quad \text{(as measures on } [-\pi, \pi]^d) \,, \tag{4.27}$$

then V_x admits the decomposition (4.23) with V_x^2 forming IID gaussian random variables.

Consequently, our analysis applies to a wide class of stochastically homogeneous potentials (not necessarily gaussian) for which the correlations, or mixing coefficients (B_x), can exhibit arbitrarily slow decay. Interestingly, even in such cases the method proves exponential decay for the Green's function, and thus for the eigenfunctions (in the appropriate regimes).

The role of the indeterminacy, or the lack of deterministic interpolation, of the joint distribution has been emphasized and elucidated in the 1D case by Kotani [K] (with other previous results found in [Si, P]). Beyond one dimension, localization for correlated potentials was derived in [DK2] (using the multiscale analysis), though the results found there yield only power-law decay.

d. *Potentials with Block Structure.* We can also consider random Schrödinger operators for which the values of the potential V_x in some regions are strongly dependent, even restricted to be equal. For example, let $\mathcal{Q} = \{Q_\alpha\}$ be a partition of $\mathbf{Z}^d$ into non-intersecting cubes, of side lengty L, and let

$$V_x = \mathscr{V}_\alpha, \quad \text{if } x \in Q_\alpha \,, \tag{4.24}$$

with $\{\mathscr{V}_\alpha\}$ independent identically distributed random variables.

We will omit here the adaptation, or the bypass, of the Simon–Wolff analysis suitable for this case. However, well within the scope of the methods discussed here explicitly is the following result.

Theorem 4.3. *Let $\{\mathscr{V}_\alpha\}$ be conditionally absolutely continuous (among the block variables) and conditionally (τ, ε) regular, for some $\tau, \varepsilon > 0$, and let*

$$H = -\Delta + U_0(x) + \lambda V_x \tag{4.25}$$

on $\ell^2(Z^d)$, with $\{V_x\}$ given by (4.24). Then there is a range of values of (λ, E) in which the resolvents of the finite volume restrictions satisfy

$$\langle |G_E^\Lambda(x, y)|^s \rangle \leqq \hat{D} \exp(-m|x - y|) \tag{4.26}$$

with some $s, m(\lambda, E) > 0$ and $\hat{D} < \infty$, uniformly in $x \in \mathbf{Z}^d$ and in $\Lambda \subset \mathbf{Z}^d$. Furthermore, for some λ_0 and E_0 which are defined independently of U_0, the range of values of (λ, E) for which (4.26) holds includes the two regimes:

i) $$\lambda > \lambda_0 \quad (\textit{and all } E) \textit{ and} \tag{4.27}$$

ii) $$|E| \geqq E_0 + \|U_0\|_\infty \quad (\textit{and all } \lambda) \,. \tag{4.14}$$

The main difference between this case, and the one discussed in Lemma 3.2, is in the dependence of $G_E(x_0, x)$ on the potential at x. The representation (2.15) is not useful since V_x is determined by the values of $V_{\{x\}^c}$. However, V_x has a conditionally regular distribution when conditioned on the values of the other block variables. Let α be the index of the block containing x. In lieu of (2.15) we have the following representation:

$$G_E(x_0, x) = \frac{P_n(\mathscr{V}_\alpha)}{\tilde{P}_n(\mathscr{V}_\alpha)} \,, \tag{4.15}$$

where $P_n(\cdot)$ and $\tilde{P}_n(\cdot)$ are polynomials of degree $\leqq L^d = |Q_\alpha|$, with coefficient determined by E, λ, and the other $\mathscr{V}$ variables (other than $\mathscr{V}_\alpha$).

The representation (4.14) reflects the fact that $\mathscr{V}_\alpha$ appears in H as the coupling coefficient for an operator of only the finite rank L^d, and is a direct consequence of the formula (I.4) (in Appendix I).

Despite the difference between (4.14) and (2.15), our decoupling Lemma 3.2 does apply, provided s is chosen so that

$$s \cdot L^d < \min\{\tau, \varepsilon\} \,. \tag{4.16}$$

Thus, with only minimal adjustments the derivation of Lemma 2.1 yields Theorem 4.3.

5. Localization Due to Off-Diagonal Disorder

The method presented above can be easily adapted to handle also operators as H in Eq. (1.4)/(3.10) with randomness in the off-diagonal terms $\{T_{x,y}\}$. Off-diagonal disorder has previously been studied by Faris [F], using the multiscale analysis.

We note that our method does not apply to the large off-diagonal disorder in the way it does to the diagonal disorder. For example, for operators H with dual disorder, $H = \lambda_1 T + \lambda_2 V$, with V the diagonal part (as in (1.4)), our previous analysis on localization at high disorder can be extended to the regime where the *ratio* λ_2/λ_1 is large. Nevertheless, the analysis of localization at high energies is similar for the two types of disorder.

Of special interest is the case of non-negative quadratic forms:

$$Q(f) \equiv (f, Hf) = \frac{1}{2} \sum_{\langle x, y\rangle \subset \Gamma} K_{x,y} |f(x) - f(y)|^2 \tag{5.1}$$

with random $K_{x,y} \geqq 0$, indexed by the unordered pairs $\langle x, y\rangle$. Such forms appear, for example, as elasticity tensors of various structures.

In a prototypical example, $\Gamma = \mathbf{Z}^d$, and the sum in (5.1) is over only the nearest neighbor pairs, for which the coefficients $K_{x,y}$ are independent and identically distributed, with probability measure whose density $\tilde{\rho}_{x,y}(K)$ decays at infinity by at least some power law. For such cases, we prove that above certain energies (with a calculable bound) the spectrum can only be of the pure point type, with exponentially localized eigenstates. It follows that when the support of $\tilde{\rho}_{x,y}(K)$ is unbounded the p.p. spectrum is not empty. (It would be of interest to elucidate the nature of the low energy states, which in sufficiently high dimensions may be continuous.)

Somewhat more generally:

Theorem 5.1. *Let H be an operator on $\ell^2(\Gamma)$ of the form* (5.1) *with*

$$K_{x,y} = t_{x,y}\zeta_{x,y}\,, \tag{5.2}$$

where $t_{x,y}$ are non-random coefficients and $\zeta_{x,y}$ random variables (not necessarily of definite sign), which satisfy the following three conditions.

i) *$\zeta_{x,y}$ are conditionally absolutely continuous and uniformly conditionally (τ, ε)-regular* (i.e., *with conditional distributions satisfying* (3.1) *and* (3.2) *with uniform value for the constants).*

ii) *The graph defined over Γ by the set of bonds $\{\langle x, y\rangle : t_{x,y} \neq 0\}$ has only infinite connected components (possibly only one).*

iii) *For some metric $\delta(\cdot,\cdot)$ on Γ, $a > 0$ and $s < \min\{\tau, \varepsilon/2\}$, $\{t_{x,y}\}$ satisfy*

$$\sum_{y \in \Gamma} |\tau_{x,y}|^s \exp(+a\delta(x, y)) \leqq B_{s,a} \; (< \infty) \quad \textit{(uniformly in } x \in \Gamma)\,. \tag{5.3}$$

Then, there is $E_0 < \infty$ such that for almost every realization of the couplings $\{\zeta_{x,y}\}$, H has only pure point spectrum in the energy range $E > E_0$ and the corresponding eigenfunctions obey (3.13), i.e., *are exponentially localized.*

Remark. A comparison of (5.3) with (3.12) shows a similarity in the required decay of the off-diagonal coefficients. In particular, by the discussion found in Sect. 3.c, Theorem 5.1 applies to operators with $t_{x,y}$ decaying by only a power law. For $\Gamma = \mathbf{Z}^d$, and $\zeta_{x,y}$ having an optimally regular distribution (i.e., $\tau = 1$, and $\varepsilon = \infty$

[or just $\varepsilon \geqq 2$]), the relevant condition is the exact analog of (3.15), with $T_{x,y}$ replaced by $t_{x,y}$.

Proof. First let us note that each coefficient $K_{x,y}$ multiplies a rank one term, which can be written as $|\alpha\rangle\langle\alpha|$ with $|\alpha\rangle = |x\rangle - |y\rangle$. Condition ii) assures that the linear span of the collection of vectors $\{(|x\rangle - |y\rangle) \in \ell^2(\Gamma)\colon x, y \in \Gamma\ t_{x,y} \neq 0\}$ is dense in $\ell^2(\Gamma)$. Hence the Simon–Wolff criterion (II), in the version stated in Sect. 4, applies and shows that a sufficient condition for p.p. spectrum and exponential localization, in some energy range, is the almost sure exponential decay of the resolvent. This is of course what our method is geared for, although the setup is now somewhat different from what is seen above.

By arguments spelled out in the proof of Theorem 2.1, it suffices to handle the finite volume problem, provided that is done with uniform estimates. That is what we shall now look at, suppressing the finite volume cutoff in our notation.

For the operator H defined by (5.1), the equation satisfied by the resolvent, $G(x) \equiv G_E(x_0, x)$, is

$$EG(x) = \sum_{y \in \Gamma \setminus \{x\}} K_{x,y}[G(x) - G(y)] - \delta_{x_0,x}\,. \tag{5.4}$$

One way to proceed is to break the sum and estimate the contribution of each term to $\langle |G(x)|^s \rangle$. However, one gets better estimates by paying attention to the structure present here.

As we noted above, $K_{x,y}$ is the coefficient of a rank one term, $|\alpha\rangle\langle\alpha|$, in the operator H. The difference $G(x) - G(y)$ can be written as a matrix element involving this vector:

$$G(x) - G(y) = \langle \alpha | \frac{1}{H - E} | x_0 \rangle \tag{5.5}$$

with $|\alpha\rangle = |x\rangle - |y\rangle$. This representation carries the implication that $[G(x) - G(y)]$ as a function of $K_{x,y}$ takes the simple form which we have encountered in (2.15):

$$G(x) - G(y) = \frac{\alpha}{K_{x,y} - \beta} \tag{5.6}$$

with α and β determined by the other parameters. (The proof is by the argument used for (2.15). For an alternative derivation, one can look at $\partial[G(x) - G(y)]/\partial K_{x,y}$).)

We now proceed by following our basic stratagem. From (5.4) we get:

$$|E|^s \langle |G(x)|^s \rangle \leqq \sum_{y \in \Gamma \setminus \{x\}} \langle |K_{x,y}|^s |G(x) - G(y)|^s \rangle \tag{5.7}$$

for any $x \neq x_0$, and $s < 1$. By (5.6), the product seen in the expectation on the right side involves only explicitly controllable singularities. The relevant decoupling principle, which is slightly different from those encountered in earlier sections, is stated below as Lemma 5.1. Using it, in much the same way (though in a different direction) as Lemma 2.3 was used in the derivation of (2.24), we obtain the following inequality,

$$|E|^s \langle |G(x)|^s \rangle \leqq \mu \sum_{y \in \Gamma \setminus \{x\}} |t_{x,y}|^s \langle |G(x) - G(y)|^s \rangle\,, \tag{5.8}$$

with $\mu < \infty$ provided $s < \min\{\tau, 2\varepsilon\}$. Since $s < 1$:

$$\langle |G(x) - G(y)|^s \rangle \leqq [\langle |G(x)|^s \rangle + \langle |G(y)|^s \rangle] \,. \tag{5.9}$$

Gathering the terms, we obtain

$$\left(|E|^s - \mu_s \sum_{y \in \Gamma \setminus \{x\}} |t_{x,y}|^s \right) \langle |G(x)|^s \rangle \leqq \mu_s \sum_{y \in \Gamma \setminus \{x\}} |t_{x,y}|^s \langle |G(y)|^s \rangle \,. \tag{5.10}$$

That inequality is very similar to (3.29). Following the argument given there, one easily obtains exponential decay of $\langle |G(x)|^s \rangle$ for each $E > E_0$, with

$$E_0 = \left[2\mu_s \sum_{y \in \Gamma \setminus \{x\}} |t_{x,y}|^s \right]^{1/s} . \quad \blacksquare \tag{5.11}$$

The decoupling tool used in the above derivation is:

Lemma 5.1. *Let $\rho(du)$ be a (τ, ε)-regular measure on $\mathbb{R}$. Then for any $0 < s < \min\{\tau, \varepsilon/2\}$ there is $\mu_s > 0$ such that*

$$\int \frac{|\eta|^s}{|\eta - \alpha|^s} \rho(d\eta) \leqq \mu_s \int \frac{1}{|\eta - \alpha|^s} \rho(d\eta) \tag{5.12}$$

for all $\alpha \in \mathbb{C}$.

This lemma is an immediate corollary of Theorem III.2, of Appendix III, and is explained there.

Appendix I. A Finite-Rank Perturbation Formula

Following is the proof of Lemma 2.2 (the Krein formula), which states that

$$G_E^A(x, y) = \left[\begin{pmatrix} \lambda V_x & 0 \\ 0 & \lambda V_y \end{pmatrix} + A_{2\times 2}^{-1} \right]^{-1}_{x,y} , \tag{I.1}$$

where $A_{2\times 2}$ is the restriction of $\dfrac{1}{\hat{H} - E}$ to the two dimensional space spanned by $\delta_{x,\cdot}$ and $\delta_{y,\cdot}$, $\hat{H}$ being the operator obtained from H by setting $V_x = V_y = 0$. Equation (I.1) is a special case of the following statement.

Theorem I.1. *Let H be a self adjoint operator H, acting in some Hilbert space $\mathscr{H}$. If*

$$H = \hat{H} + W \tag{I.2}$$

with W (a finite rank term) satisfying

$$W = RWR \tag{I.3}$$

for some finite dimensional orthogonal projection R, then, for E with $\mathrm{Im}(E) \neq 0$:

$$\left[R \frac{1}{H - E} R \right] = \left[W + \left[R \frac{1}{\hat{H} - E} R \right]^{-1} \right]^{-1} , \tag{I.4}$$

where $[\,\cdot\,]$ denotes a finite dimensional matrix $(\mathrm{rank}(R) \times \mathrm{rank}(R))$ describing the restriction of an operator to the range of R, and $[\,\cdot\,]^{-1}$ represents the matrix inverse.

This formula is of use in analyzing the dependence of the resolvent $G_E(x, y)$ on any finite collection of parameters of H. The restriction to energies off the real axis may be dropped for finite dimensional $\mathscr{H}$, in which case the above are meromorphic functions of E.

Proof. The resolvent formula (for E is off the real axis, where all our algebraic manipulations are fully justified)

$$\frac{1}{\hat{H}-E}=\frac{1}{H-E}+\frac{1}{\hat{H}-E}\,W\,\frac{1}{H-E}\,, \tag{I.5}$$

and Eq. (I.3) imply the matrix relation:

$$\left[R\ \frac{1}{\hat{H}-E}\ R\right]=\left[R\ \frac{1}{H-E}\ R\right]+\left[R\ \frac{1}{\hat{H}-E}\ R\right][R\ W\ R]\left[R\ \frac{1}{H-E}\ R\right]. \tag{I.6}$$

Multiplying Eq. (I.6) from left and right by the corresponding inverses, one gets

$$[R(H-E)^{-1}R]^{-1}=[R(\hat{H}-E)^{-1}R]^{-1}+[W]\ , \tag{I.7}$$

from which the claim (I.4) readily follows. ■

Appendix II. A Bound on $G_E(x, y)$

In this appendix we derive general "a-priori" bounds on the probability distribution of the resolvent kernel. These bounds are closely related to Wegner's estimate of the density of states.

Let $H = T + U_0(x) + \lambda V_x$, as in Eq. (3.10), and assume that the conditional probability $\rho(dV_x \mid V_{\{x\}^c})$ is uniformly bounded for intervals of length t by some function $g(t)$. I.e.:

$$\sup_u \rho(V_x \in [u, u+t] \mid V_{\{x\}^c}) \leqq g(t)\ . \tag{II.1}$$

Theorem II.1. *Assuming (II.1), for any finite volume Λ and any pair of sites $x, y \in \Lambda$,*

$$\mathrm{Prob}(|G_E^\Lambda(x, x)| \geqq t) \leqq g\left(\frac{2}{\lambda t}\right), \tag{II.2}$$

$$\mathrm{Prob}(|G_E^\Lambda(x, y)| \geqq t) \leqq 2g\left(\frac{2\sqrt{2}}{\lambda t}\right). \tag{II.3}$$

Proof. For simplicity let us denote $G(x, y) \equiv G_E^\Lambda(x, y)$.

i) By the rank one case of Eq. (I.4),

$$G(x, x)=\left[\lambda V_x+\frac{1}{\hat{G}(x, x)^{-1}}\right]^{-1}. \tag{II.4}$$

Thus in order for $G(x, x)$ to be large, λV_x has to be within the distance $1/t$ from a value which is determined by the potential elsewhere. By our regularity assumption, the conditional probability for that is bounded above by $g\left(\frac{2}{\lambda t}\right)$.

ii) For the off-diagonal terms, we deduce from (I.1) that

$$|G(x, y)| \leqq \left\| \begin{pmatrix} \lambda V_x & 0 \\ 0 & \lambda V_y \end{pmatrix} + A_{2\times 2}^{-1} \right]^{-1} \Bigg\| = \frac{1}{\min\{|\eta_1|, |\eta_2|\}} \tag{II.5}$$

with $\|\cdot\|$ denoting the matrix operator-norm, and $\{\eta_i\}_{i=1,2}$ the two eigenvalues of the (total) matrix. Let S be the subset of $\mathbb{R}\times\mathbb{R}$ consisting of the joint values of (V_x, V_y) for which the above matrix has a zero eigenvalue (i.e., $\det[\cdot] = 0$). Then (II.5) easily implies that:

a. S is a hyperbola:

$$S = \{(V_x, V_y)\in \mathbb{R}\times\mathbb{R}: (V_x - u_x)(V_y - u_y) = \alpha^2\} \tag{II.6}$$

with the parameters $\{u_x, u_y, \alpha = |A_{x,y}^{-1}|/\sqrt{|\lambda|}\}$ determined by E, λ, and the values of $V_{\{x,y\}^c} \equiv \{V_z: \{x, y\}^c z \neq x, y\}$.

b. $$|G(x, y)| \leqq \frac{1}{\lambda d_1((V_x, V_y), S)}, \tag{II.7}$$

where $d_1((V_x, V_y), S)$ is the distance of $(V_x, V_y)\in\mathbb{R}\times\mathbb{R}$ from the set S, along the diagonal of slope 1.

The subset of $\mathbb{R}\times\mathbb{R}$ on which $d_1((V_x, V_y), S) \leqq \frac{1}{\lambda t}$ consists of a pair of strips of width $\frac{2}{\lambda t}$, in the (1, 1) direction, centered on the two branches of S. Elementary geometric considerations show that there is a choice of functions $S_i(\cdot)$, $i = 1, 2$, with which this set is covered by the union $B_1 \cup B_2$, with

$$B_1 = \left\{(x, y)\in\mathbb{R}\times\mathbb{R}: |y - S_1(x)| \leqq \frac{\sqrt{2}}{\lambda t}\right\}$$

and

$$B_2 = \left\{(x, y)\in\mathbb{R}\times\mathbb{R}: |x - S_2(y)| \leqq \frac{\sqrt{2}}{\lambda t}\right\}. \tag{II.8}$$

Instead of presenting the algebraic expression for $S_i(\cdot)$, let us describe these functions by sketching their graphs. Let $\mathscr{a}$ be the symmetry axis of S which intersects the two branches of its graph. As x varies from $-\infty$ to $+\infty$, the graph of $S_1(x)$ follows the "lower branch" of S until it meets the symmetry line $\mathscr{a}$, it then crosses along $\mathscr{a}$ to the upper branch of S, and then continues along it as $x \to \infty$. The graph of $S_2(y)$ is obtained by switching the roles of the two coordinates.

With the above choice,

$$\left\{(V_x, V_y)\in\mathbb{R}\times\mathbb{R}: d_1((V_x, V_y), S) \leqq \frac{1}{\lambda t}\right\} \subset B_1 \cup B_2\,. \tag{II.9}$$

(The detailed demonstration is based on elementary arguments which will be omitted here.) The conditional probability of B_1, conditioned on $V_{\{x,y\}^c}$, can be

estimated by first integrating V_y at fixed V_x. For B_2 the order is reversed. The net result is

$$\text{Prob}\left(\left\{d_1((V_x, V_y), S) \leqq \frac{1}{\lambda t}\right\}\right) \leqq 2g\left(\frac{2\sqrt{2}}{\lambda t}\right), \tag{II.10}$$

where Prob$(\cdot)$ refers to the conditional probability $\rho(dV | V_{\{x,y\}^c})$. Together, (II.7) and (II.10) imply the stated bound (II.3). ■

Remark. The inequality (II.2) is closely related with Wegner's bound [W] on the density of states (which can also be derived by this route). We have no prior reference for the off-diagonal bound (II.3).

Appendix III. Decoupling Estimates

In this appendix we present the analytical tool which we have used for decoupling the expectations of products of random variables. The measures referred to here are all on the real line, and the considerations involved are simple real-analysis arguments. Our main goal is to derive the decoupling lemmas whose proofs appear at the end.

First, let us repeat the definition given in Sect. 3.

Definition. *A measure $\rho(du)$, on the real line, is said to be τ-regular (with $0 < \tau \leqq 1$) if, with some $v \geqq 1$ and $C_1 < \infty$,*

$$\rho([z - \delta, z + \delta]) \leqq C|\delta|^\tau \rho([z - v, z + v]) \tag{III.1}$$

for all $0 < \delta < 1$ and $z \in R$.

This naturally leads to a natural of uniform regularity.

Definition. *A collection of probability measures on the real line, is said to be uniformly τ-regular if the measure satisfies the regularity condition* (III.1) *with common values for the constants appearing there.*

Following is a fundamental result concerning this notion.

Theorem III.1. *Let $\rho(dx)$ be a τ-regular measure on $\mathbb{R}$. Then for each pair of integers (n, k), with*

$$k < \tau/s\,, \tag{III.2}$$

the family of measures of the form

$$\hat{\rho}(du) = \text{Const.} \left|\frac{\mathscr{P}(u)}{\mathscr{Q}(u)}\right|^s \rho(du)\,. \tag{III.3}$$

With $\mathscr{P}(\cdot)$ and $\mathscr{Q}(\cdot)$ polynomials of bounded degrees:

$$\det \mathscr{P} \leqq n, \quad \textit{and} \quad \deg \mathscr{Q} \leqq k\,, \tag{III.4}$$

is uniformly $(\tau - sk)$-regular.

Remark. We shall later normalize $\hat{\rho}(du)$ to be probability measures. However, the regularity condition (III.1) is homogeneous in $\rho(\cdot)$, and thus the normalization is irrelevant for this condition.

By the factorization of polynomials, this theorem is immediately implied by the following two lemmas.

Lemma III.1. *Let $\rho(du)$ be a τ-regular measure on $\mathbb{R}$. Then for any $s < \tau$, the measures*

$$\hat{\rho}(du) = \text{Const.} \frac{1}{|u-\alpha|^s} \rho(du) \tag{III.5}$$

with $\alpha \in \mathbb{C}$ are uniformly $(\tau - s)$-regular, satisfying (III.1) *with the modified constants $\hat{\nu}$ and $\hat{C}$:*

$$\hat{\nu} = \nu + 1, \quad \hat{C} = C\left[1 + \frac{s}{\tau - s}\right](3\nu + 1)^s . \tag{III.6}$$

Lemma III.2. *Let $\rho(du)$ be a τ-regular measure on $\mathbb{R}$. Then for any $s > 0$ the measures*

$$\hat{\rho}(du) = \text{Const.} |u-\beta|^s \rho(du) \tag{III.7}$$

are also τ-regular, satisfying (III.1) *with constants which can be chosen uniformly in $\beta \in \mathbb{C}$.*

Proof of Lemma III.1. For given $z, \delta \in \mathbb{R}$, we split the discussion into two cases, according to whether $|z-\alpha| \geqq 2\nu$, or not.

i) If $|z-\alpha| \geqq 2\nu$, then within the interval $[z-\nu, z+\nu]$ (which contains $[z-\delta, z+\delta]$), $\dfrac{1}{|x-\alpha|^s}$ does not vary by a factor exceeding 3^s. Hence (III.1) implies

$$\hat{\rho}([z-\delta, z+\delta])/\hat{\rho}([z-\nu, z+\nu]) \leqq 3^s C\delta^\tau \tag{III.8}$$

for all $0 < \delta \leqq 1$, which is a stronger bound than the claimed one.

ii) For the other case, when the singularity is close to z, we express $\hat{\rho}([z-\delta, z+\delta])$ as:

$$\begin{aligned}\hat{\rho}([z-\delta, z+\delta]) &= \int_{|u-z|\leqq\delta} \rho(du) \frac{1}{|u-\alpha|^s} \\ &= \int_0^\infty dt \int \rho(du) I[|u-z| \leqq \delta] I\left[\frac{1}{|u-\alpha|^s} \geqq t\right] .\end{aligned} \tag{III.9}$$

Splitting the first integral at the point $t = 1/\delta^s$, and applying (III.1) we get the following estimate, which holds irrespectively of the condition $|z-\alpha| \leqq 2\nu$,

$$\begin{aligned}\hat{\rho}([z-\delta, z+\delta]) &\leqq C\delta^{\tau-s}\rho([z-\nu, z+\nu]) \\ &\quad + C\int_{1/\delta^s}^\infty dt\, t^{-\tau/s} \rho([z-(\nu+1), z+(\nu+1)]) \\ &\leqq \delta^{\tau-s}\left[1 + \frac{s}{\tau-s}\right] C_1 \rho([z-(\nu+1), z+(\nu+1)]) .\end{aligned} \tag{III.10}$$

This is to be compared to the measure $\hat{\rho}$ of the interval on the right. Since $|z - \alpha| \leqq 2v$,

$$\hat{\rho}([z-(v+1), z+(v+1)]) \geqq \rho([z-(v+1), z+(v+1)]) \frac{1}{(3v+1)^s}, \quad \text{(III.11)}$$

and the claimed inequality readily follows. ■

Proof of Lemma III.2. As in the first part of the above proof, if $|z - \beta| \geqq 2v$ then the analog of (III.8) holds, with the constants seen there.

In case $|z - \beta| \leqq 2v$, then (for $\delta \leqq v$)

$$\hat{\rho}([z-\delta, z+\delta]) \leqq (3v)^s \rho([z-\delta, z+\delta]) \leqq (3v)^s \delta^\tau C \rho([z-v, z+v]) . \quad \text{(III.12)}$$

To estimate $\hat{\rho}(\cdot)$ of the surrounding interval we note that the regularity condition (III.1) implies that an interval of length $b = 2 \cdot (3C)^{-\tau}$ has no more than 1/3 of the $\rho(\cdot)$ measure of the concentric interval of length $2 \cdot v$. Since $|z - \beta| \leqq 2v$, the interval of length $2v$ centered at β is included in $[z - 3v, z + 3v]$. Hence

$$\begin{aligned} \hat{\rho}([z-3v, z+3v]) &= \int_{|u-z| \leqq 3v} \rho(du) |u-\beta|^s \geqq b \int_{|u-z| \leqq 3v} \rho(du) I[|u-\beta| \geqq b] \\ &\geqq \frac{2}{3} b^s \rho([z-3v, z+3v]) \geqq \frac{2}{3} b^s \rho([z-v, z+v]) . \quad \text{(III.13)} \end{aligned}$$

The comparison of (III.12) with (III.13) shows that $\hat{\rho}(\cdot)$ is τ-regular, uniformly in β. ■

Taken together, the last two lemmas are equivalent to Theorem III.1, with some specific information on the constants (III.6)). In addition, the estimates seen in the proofs imply the following useful statement.

Corollary III.1. *Under the conditions of Theorem* III.1, *for any interval B,*

$$\int_B \frac{\prod_{i=1}^n |u-\alpha_i|^s}{\prod_{j=1}^k |u-\beta_j|^s} \rho(du) \leqq \text{Const.} \int_{B'} \frac{\prod_{i=1}^n (1+|u-\alpha_i|)^s}{\prod_{j=1}^k (1+|u-\beta_j|)^s} \rho(du) \quad \text{(III.14)}$$

with $B' = B + [-2v, 2v]$ and the constant independent of $\{\alpha_i, \beta_j\}$. Furthermore, with B and B' exchanged and a different value for the constant, the reversed inequality also holds.

Theorem III.1 plays a key role in the decoupling estimates needed for the high disorder regime. For results concerning the regimes of extreme energies, independent of the strength of the coupling, we use

Theorem III.2. *Let $\rho(du)$ be a τ-regular probability meausure on $\mathbb{R}$ with a finite fractional moment:*

$$\rho(|u|^\varepsilon) \equiv \int |u|^\varepsilon \rho(du) < \infty , \quad \text{(III.15)}$$

at some $\kappa > 0$. Then for any pair of integers (n, k) satisfying:

$$s \cdot (n+k) < \varepsilon, \quad \text{and} \quad s \cdot k < \tau \quad \text{(III.16)}$$

the collection of probability measures given by (III.3), *suitably normalized, and* (III.4) *is tight, in the sense that the measures satisfy uniform decay conditions. More*

specifically, these measures obey

$$\hat{\rho}(|u|^{\varepsilon - s(n+k)}) \leqq \text{Const.}\, \rho(|u|^{\varepsilon}) \,, \tag{III.17}$$

with a uniform constant (depending on (n, k) and the constants with which ρ obeys (III.1)).

Proof. Let the polynomials $\mathscr{P}$ and $\mathscr{Q}$ have the factorization seen in (III.14). Then, using (III.14), in its two forms, the corresponding measure is seen to satisfy:

$$\hat{\rho}(\{u \in \mathbb{R}: |u| > t\}) \leqq \text{Const.} \int\limits_{|u| > t - 2\nu} \frac{\prod_{i=1}^{n} (1 + |u - \alpha_i|)^s}{\prod_{j=1}^{k} (1 + |u - \beta_j|)^s} \rho(du) \Bigg/ \int\limits_{|u| \leqq u_0} \frac{\prod_{i=1}^{n} (1 + |u - \alpha_i|)^s}{\prod_{j=1}^{k} (1 + |u - \beta_j|)^s} \rho(du) \,, \tag{III.18}$$

with a constant uniform in $\{\alpha_i, \beta_j\}$, and any u_0. (We use now the symbol Const. to represent such quantities, which are not necessarily all equal.)

The following is a helpful inequality for polynomials (easily seen to be true in its logarithmic version):

$$\frac{\prod_{i=1}^{n} (1 + |u - \alpha_i|)^s}{\prod_{j=1}^{k} (1 + |u - \beta_j|)^s} \leqq (1 + |u - u'|)^{s(k+n)} \frac{\prod_{i=1}^{n} (1 + |u' - \alpha_i|)^s}{\prod_{j=1}^{k} (1 + |u' - \beta_j|)^s} \,. \tag{III.19}$$

Choosing u_0 so that $\rho(\{u: |u| \leqq u_0\}) > 0$, and applying (III.19) in (III.18), we obtain a bound of the form

$$\hat{\rho}(\{u \in \mathbb{R}: |u| > t\}) \leqq \text{Const.} \int\limits_{|u| > t - 2\nu} (1 + |u|)^{s(k+n)} \rho(du) \,. \tag{III.20}$$

The claim follows now by the Tchebychev inequality. ■

We shall now apply these results for the derivation of the decoupling lemmas used in this work.

Proof of Lemma 5.1. Lemma 5.1 is an immediate corollary of Theorem III.2. To see that, all one has to do is rewrite the condition (5.12) in terms of the probability measures $\hat{\rho}(du) = |u - \alpha|^{-s} \rho(du)/\text{Norm}$. The corresponding values of (n, k) are $(0, 1)$ and the claimed bound takes the form:

$$\hat{\rho}(|u|^s) \leqq \mu_s \quad \text{(uniformly in } \alpha) \,. \tag{III.21}$$

The boundedness of μ_s follows from (III.17), provided $s < \varepsilon - s$, i.e.: $s < \varepsilon/2$. ■

Next we turn to the proof of Lemma 3.1. For convenience, we restate it here.

Lemma 3.1. *Let $\rho(du)$ be a τ-regular measure on $\mathbb{R}$. Then:*

i) *for each $s < \tau$ and a pair of integers (n, k), with $s \cdot k < \tau$, there is $\kappa_s^{n,k} > 0$ such that for all rational functions, $\mathscr{R}(u) = \mathscr{P}(u)/\mathscr{Q}(u)$ with:*

$$\deg \mathscr{P} \leqq n, \quad \textit{and} \quad \deg \mathscr{Q} \leqq k \; (< \tau/s) \,, \tag{III.22}$$

and all η:

$$\int |u - \eta|^s |\mathscr{R}(u)|^s \rho(du) \geqq (\kappa_s^{n,k})^s \int |\mathscr{R}(u)|^s \rho(du) \,. \tag{III.23}$$

ii) *If, furthermore,*

$$\int |u|^s \rho(du) < \infty \ , \tag{III.24}$$

then the inequality (3.3*b*) *can be strengthened as follows*:

$$\int |u-\eta|^s |\mathscr{R}(u)|^s \rho(du) \geqq [\xi_s^{n,k}(|\eta|)]^s \int |\mathscr{R}(u)|^s \rho(du) \ , \tag{III.25}$$

with $\xi_s^{n,k}(\cdot)$ *an increasing positive function on* $\mathbb{R}_+$, *satisfying*

$$\lim_{u \to \infty} \frac{\xi_s^{n,k}(u)}{u} = 1 \ . \tag{III.26}$$

Proof. i) With $\hat{\rho}(du)$ defined by (III.3), normalized so as to form a probability measure,

$$\int |u-\eta|^s |\mathscr{R}(u)|^s \rho(du) / \int |\mathscr{R}(u)|^s \rho(du) = \int |u-\eta|^s \hat{\rho}(du) \geqq |a|^s \hat{\rho}(\{u: |u-\eta| \geqq a\}) \ . \tag{III.27}$$

According to Theorem III.1, $\hat{\rho}(\cdot)$ belongs to a class of uniformly $(\tau - sk)$-regular measures. Denoting by $\hat{C}$ the constant with which (III.1) is satisfied within this class, we have

$$\hat{\rho}(\{u: |u-\eta| \geqq a\}) \geqq \frac{1}{2} \hat{\rho}(\mathbb{R}) = \frac{1}{2} \tag{III.28}$$

with $a = 2 \cdot (2\hat{C})^{-\tau}$. Hence (III.23).

ii) The second part of the assertion is implied by the bound:

$$\int |u-\eta|^s \hat{\rho}(du) \geqq |\eta|^s (1-\varepsilon)^s [1 - \hat{\rho}(\{u: |u| \leqq \varepsilon |\eta|\})] \ , \tag{III.29}$$

and the tightness statement expressed in Theorem III.2, by which the probability seen on the right side tends uniformly to zero, as $|\eta| \to \infty$. ■

The above results demonstrate the general utility and versatility of the simple analytic tools presented in this appendix, which play a key role in our discussion of localization.

Acknowledgements. We thank B. Simon, T. Spencer, and T. Wolff for a number of stimulating discussions on the subject of this work. M.A. wishes to also express his gratitude for the hospitality accorded him at the Institute for Advanced Studies at an early phase of this work, and at California Institute of Technology, on a visit supported by the Fairchild Foundation.

References

[A] Anderson, P.W.: Absence of diffusion in certain random lattices. Phys. Rev. **109**, 1492–1505 (1958)

[AALR] 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)

[CFKS] Cycon, H., Froese, R., Kirsh, W., Simon, B.: Topics in the Theory of Schrödinger Operators. Berlin, Heidelberg, New York: Springer-Verlag 1987

[CL] Carmona, R., Lacroix, J.: Spectral theory of random Schrödinger operators. Boston: Birkhäuser 1990

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

[D] von Dreifus, H.: On the effects of randomness in ferromagnetic models and Schrödinger operators. NYU Ph.D.. Thesis (1987)

[DK1] von Dreifus, H., Klein, A.: A new proof of localization in the Anderson tight binding model. Commun. Math. Phys. **124**, 285–299 (1989)

[DK2] von Dreifus, H., Klein, A.: Localization for random Schrödinger operators with correlated potentials. Commun. Math. Phys. **140**, 133–147 (1991)
[DLS1] Delyon, F., Levy, Y., Souillard, B.: Anderson localization for multidimensional systems at large disorder or low energy. Commun. Math. Phys. **100**, 463–470 (1985)
[DLS2] Delyon, F., Levy, Y., Souillard, B.: Anderson localization for one and quasi one-dimensional systems. J. Stat. Phys. **41**, 375 (1985)
[F] Faris, W.: Localization estimates for off-diagonal disorder. Localization estimates for off-diagonal disorder. In Mathematics of Random Media, Lecture Notes in Appl. Mat. **27**, W.E. Kohler, B.J. White (eds.), pp. 391–406, Providence, R.I.: AMS 1991
[FMSS] Fröhlich, J., Martinelli, F., Scoppola, E., Spencer, T.: A constructive proof of localization in Anderson tight binding model. Commun. Math. Phys. **101**, 21–46 (1985)
[FP] Figotin, A., Pastur, L.: Spectral properties of disordered systems in the one-body approximation. Berlin, Heidelberg, New York: Springer-Verlag 1991
[FS] 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)
[G] Gross, L.: Decay of Correlations at High Temperature. Commun. Math. Phys. **68**, 9–27 (1979)
[GJMS] Gordon, A., Jaksiĉ, V., Molchanov, S., Simon, B.: Spectral properties of random Schrödinger operators with unbounded potential. Commun. Math. Phys. ■
[GMP] Goldsheid, I., Molchanov, S., Pastur, L.: A pure point spectrum of the one-dimensional Schrödinger operator. Funct. Anal. Appl. **11**, 1–10 (1977)
[K] Kotani, S.: Lyapunov exponent and spectra for one-dimensional random Schrödinger operators. In: Proc. conf. on random matrices and their applic. Contemp. Math. **50**. Providence RI, 1986
[KP] Kirsh, W., Pastur, L.: Bochum preprint (1990)
[KS] Kunz, H., Souillard, B.: The localization transition on the Bethe lattice. J. Phys. (Paris) Lett. **44**, 411–414 (1983)
[L] Landauer, R.: Phil. Mag. **21**, 863 (1970)
[MS] Molchanov, S., Simon B.: Localization theorem for non-local Schrödinger operator in one-dimensional lattice case. Caltech preprint
[MT] Mott, N., Twose, W.: The theory of impurity conduction. Adv. Phys. **10**, 107–163 (1961)
[P] Pastur, L.A.: Random and almost periodic operators: New examples of spectral behaviour. In: IXth International Congress on Mathematical Physics, Simon, B., Truman, A., Davis, I.M. (eds.) Bristol, UK: Adam Hilger 1989
[RS] Reed, M., Simon, B.: Methods of modern mathematical physics, vol. **II** (Fourier analysis, self-adjointness). New York: Acad. Press 1975
[Si] Simon, B.: Localization in general one dimensional systems, I. Jacobi matrices. Commun. Math. Phys. **102**, 327–336 (1985)
[Sp] Spencer, T.: Localization for random and quasiperiodic potentials. J. Stat. Phys. **51**, 1009–1019 (1988)
[SiSp] Simon, B., Spencer, T.: Trance class perturbation and the absence of absolutely continuous spectrum. Commun. Math. Phys. **125**, 113–125 (1989)
[SW] Simon, B., Wolff, T.: Singular continuous spectrum under rank one perturbations and localization for random Hamiltonians. Commun. Pure Appl. Math. **39**, 75–90 (1986)
[T] Thouless, D.: Electrons in disordered systems and the theory of localization. Phys. Rev. **13**, 93–106 (1974)
[W] Wegner, F.: Bounds of the density of states in disordered systems. Z. Phys. B. Condensed Matter **44**, 9–15 (1981)

Note added in proof. The method presented here can be extended also to localization at weak disorder and moderate energies. Such supplementary results, and further developments, are found in:

[A2] Aizenman, M.: Localization at weak disorder: some elementary bounds. To appear in Rev. Math. Phys.

Communicated by T. Spencer