LOCALIZATION AND UNIVERSALITY OF POISSON STATISTICS NEAR ANDERSON TRANSITION

Wei-Min Wang

UMR 8628 du CNRS
Département de mathématiques
Université Paris-Sud
Batiment 425
91405 Orsay CEDEX, FRANCE

Abstract. By further developing the supersymmetric formalism initiated in [W1] (see also [SjW]), we obtain Anderson localization with the mean-field Lyapunov exponent and Poisson statistics for eigenvalue spacing at weak disorder near the Anderson transition.

1991 *Mathematics Subject Classification.* 35P, 60K, 81V.

Key words and phrases. supersymmetry, localization, Poisson statistics.

We are indebted to M. Aizenman for explaining to us the results in [AFHS]. We thank the organizers of the AMS workshop on "wave phenomenon in complex media" in June 1999, for a stimulating meeting, where we first learned of [AFHS]. We also thank J. Sjöstrand for some helpful comments on a previous version of the introduction.

I. Introduction

Let

$$H = \Delta + \gamma V, \text{ on } \ell^2(\mathbf{Z}^d), \tag{1.1}$$

where

$$\begin{aligned} \Delta_{ij} &= 1 \quad |i-j|_{\ell_1} = 1 \\ &= 0 \quad \text{otherwise}; \end{aligned} \tag{1.2}$$

γ is a positive parameter, the potential function V is a diagonal matrix: $V = \text{diag}(v_j)$, $j \in \mathbf{Z}^d$, where $\{v_j\}$ is a family of independently identically distributed (iid) real random variables with distribution dg. (dg is only assumed to be a measure.) From now on, we write $|\,|$ for $|\,|_{\ell_1}$. ℓ_2 norms will be denoted by $\|\,\|$. The probability measure is taken to be $P = \prod_{j\in\mathbf{Z}^d} dg(v_j)$. We use the notation $\langle\,\rangle$ to denote the expectation value with respect to P.

For hypercubes Λ: $\Lambda = [-L, L]^d$ ($L \in \mathbf{N}$), we define

$$H_\Lambda = \Delta_\Lambda + \gamma V, \text{ on } \ell^2(\Lambda), \tag{1.3}$$

where the matrix elements of Δ_Λ are defined as in (1.2) for $i, j \in \Lambda$ and 0 otherwise.

As is well known, the spectrum $\sigma(\Delta)$ is $[-2d, 2d]$. Let supp dg be the support of dg, then we have the well established fact (See e.g., [CFKS,PF].) that

$$\sigma(H) = [-2d, 2d] + \gamma \,\text{supp}\, dg \tag{1.4}$$

almost surely. For all Λ, we also have

$$\sigma(H_\Lambda) \subset [-2d, 2d] + \gamma \,\text{supp}\, dg \tag{1.5}$$

In this paper we study spectral properties of H and eigenvalue statistics of H_Λ for large Λ in the weak disorder limit: $\gamma << 1$. We use the Green's function formulation. For $E \notin \sigma(H)$, we define $G(E) = (E - H)^{-1}$ and its kernel

$$G(E, i, j) = (\delta_i, (E - H)^{-1}\delta_j), \qquad i, j \in \mathbf{Z}^d, \tag{1.6}$$

to be the Green's function of H at energy E. Similarly we define

$$G_\Lambda(E, i, j) = (E - H_\Lambda)^{-1}(i, j), \qquad E \notin \sigma(H_\Lambda), \tag{1.7}$$

to be the finite volume Green's function.

In [W1], we assumed that $E \notin \sigma(H)$, dg has semi-bounded support and satisfy some regular variation conditions. Using supersymmetry and Witten Laplacians, we computed the asymptotics of Lyapunov exponents:

$$\beta_E(\hat{j}) = \beta_E\Big(\frac{j}{\|j\|}\Big) = \lim_{j\to\infty} -\frac{\log\langle G(E, 0, j)\rangle}{\|j\|}, \qquad E \notin \sigma(H), \tag{1.8}$$

for $\gamma << 1$. (The limit in (1.8) exists by using general arguments, see [Sz].) This result is new in $d \geq 2$. However we were unable to let $E \to \inf \sigma(H)^-$ or $E \to \sup \sigma(H)^+$ due to the loss of convexity and the occurrence of degeneracy at infinity. Convexity and non-degeneracy are crucial for the Witten Laplacian formulation. (For more details, see sect. III of [W1]. See also [BJS,Jo,Sj1,2].)

In this paper, we further develop the supersymmetry formalism. Under assumptions (H1-3) to be stated below, we show that for $\gamma << 1$ the spectrum of H is pure point in intervals near the infimum or supremum of the spectrum. The lengths of the intervals go to $\mathbf{O}(1)\gamma$ as $\gamma \to 0$, which is the physically expected answer. We obtain the mean field lower bound: $a\sqrt{\gamma}$, ($a > 0$, $\gamma << 1$) on the Lyapunov exponents (defined slightly differently from (1.8), see (1.10), see also (1.9)) for E in these intervals. (The corresponding upper bound seem to be out of reach for the moment.) We also obtain Poisson statistics for eigenvalue spacing. These results are new in the weak disorder regime $\gamma << 1$ for $d \geq 2$.

In a separate paper ([W2]), we specialize to the case $E \to \inf \sigma(H)^-$ or $E \to \sup \sigma(H)^+$. Similar to (1.8), one defines the almost sure Lyapunov exponents α_E, (see [Sz,Z] [W2]):

$$(1.9) \qquad \alpha_E(\hat{j}) = \alpha_E(\frac{j}{||j||}) = \lim_{j\to\infty} -\frac{\log G(E,0,j)}{||j||}, \ a.s. \qquad E \notin \sigma(H),$$

We show that the bound $\mathbf{O}(1)\sqrt{\gamma}$ on α_E is optimal here, i.e., $\alpha_E \sim \sqrt{\gamma}$ as $\gamma \to 0$. This lends support to the belief that the bound in (1.10) is optimal.

Remark. Previously known lower bounds in [Sz,Z], which are valid for all γ, give only $\mathbf{O}(1)\gamma$ for $\gamma << 1$.

In order to state our results, we make three assumptions on dg:

(H1) supp $dg \subseteq (-\infty, 0]$ or supp $dg \subseteq [0, +\infty)$;

(H2) dg has finite first and second moments, i.e.,

$$|\int v dg(v)| < \infty, \qquad \int v^2 dg(v) < \infty;$$

(H3) dg is absolutely continuous with a bounded density:

$$dg(v) = \tilde{g}(v)dv \quad \text{with} \, ||\tilde{g}||_\infty < \infty.$$

(H1) implies that

$$\sigma(H) = (-\infty, 2d] \quad or \quad \sigma(H) = [-2d, +\infty).$$

Remark. To carry through our argument in sect. II, which relies on some positivity properties we need to assume that supp dg is semi-bounded, since we have the parameter E at our disposal, we could assume (H1) without loss of generality. For Theorem 4.3, to be stated below, on spectral properties of H, (H3) could be weakened to requiring only some Hölder regularity. (See (3.7).) For the eigenvalue statistics to be stated in Theorem 5.2, (H3) seems to be needed. (See [Mi].)

Results on localization at weak disorder: $\gamma << 1$.

Our first result is

Theorem 4.3. *Assume dg satisfies (H1-3). For all $\delta > 0$, let $I_\delta = [2d - \gamma^{1+\delta}, 2d]$ if $\operatorname{supp} dg \subseteq (-\infty, 0]$ and $I_\delta = [-2d, -2d + \gamma^{1+\delta}]$ if $\operatorname{supp} dg \subseteq [0, \infty)$. There exist $a > 0$, $\gamma_0 > 0$ (a independent of γ_0), such that for all $\gamma \in (0, \gamma_0)$, with probability one $\sigma(H) \cap I_\delta$ is pure point, and the eigenfunctions ψ corresponding to eigenvalues in I_δ satisfy*

$$\limsup_{\|i\|\to\infty} \frac{\log|\psi(i)|}{\|i\|} \leq -a\sqrt{\gamma}. \tag{1.10}$$

Remarks. 1. The rate of decay $a\sqrt{\gamma}$ has the correct dependence on γ for $\gamma << 1$. This is precisely the "mean field" rate mentioned earlier. The constant a however is not optimal. For $E \in \sigma(H)$, we take the LHS of (1.10) to be the definition of the almost sure Lyapunov exponent α_E.

2. When $\gamma = 0$, $\sigma(H) = \sigma(\Delta) = [-2d, 2d]$ and is absolutely continuous. The eigenfunctions of Δ are extended (plane waves), $\alpha_E = 0$ for all $E \in [-2d, 2d]$. So Theorem 4.3 is *not* a perturbation result about Δ in the conventional sense. Note also that the result in (1.10) is subtler than the corresponding localization results in the large disorder regime $\gamma >> 1$. The decay rate for $\gamma << 1$, $a\sqrt{\gamma} \to 0$ as $\gamma \to 0$; while the decay rate for $\gamma >> 1$ remains to be at least $\mathbf{O}(1)$.

3. Theorem 4.3 is a generalization of the well known Furstenberg's theorem [Fu] to all dimensions $d \geq 1$ for almost all $E \in I_\delta$. Furstenberg's theorem is valid only in one dimension but for almost all $E \in \mathbf{R}$. Moreover Theorem 4.3 is *quantitative*, it shows that the Lyapunov exponent $\alpha_E(\gamma)$ does not go to 0 faster than $\sqrt{\gamma}$ as $\gamma \to 0$. To the best of our knowledge, Furstenberg's theorem only indicates that $\alpha_E(\gamma) > 0$ for all $\gamma > 0$, but it does not give information on how $\alpha_E(\gamma) \to 0$ as $\gamma \to 0$.

4. The spectral phenomenon exhibited in Theorem 4.3 is called Anderson localization, after the physicist P. W. Anderson [An]. Theorem 4.3 is interesting because of the conjecture of spectral transition for $d \geq 3$, which we detail later.

A sketch of the proof. The proof of Theorem 4.3 relies on two inputs:

(i) For some sufficiently large but *fixed* $\Lambda = [-L, L]^d$, and $E \in I_\delta$,

$$|G_\Lambda(E, i, j)| = |(E - H_\Lambda)^{-1}(i, j)| \leq Ce^{-a_L\sqrt{\gamma}|i-j|}, \quad (a_L > 0,\ |i-j| >> 1/\sqrt{\gamma},\ \gamma << 1),$$

with high (relative to the scale L) probability,

(ii) Some apriori regularity of eigenvalue spacing.

Since it is well known that (H3) implies (ii), our main task is to prove (i). The proof of (i) is the main novelty of the paper. It is deduced from the estimate on $\langle (E_0 - \Delta_\Lambda - \gamma V)^{-1}_{ij} \rangle$, where $E_0 = 2d$ or $-2d$ pending on $\operatorname{supp} dg$, by using Tchebychev inequality and perturbation theory. Tchebychev inequality is applicable here, because for all fixed i, $j \in \Lambda$, $(E_0 - \Delta_\Lambda - \gamma V)^{-1}_{ij}$ has a definite sign, which is independent of V. (See the very beginning of sect. II and Lemma 2.5, Corollary 2.6. below.)

To estimate $\langle (E_0 - \Delta_\Lambda - \gamma V)^{-1}_{ij} \rangle$ for $E_0 = 2d$ or $-2d$, we need to assume (H1,2). We use the supersymmetric representation as in [W1,2] (see also [SjW]), which expresses $\langle (E_0 - \Delta_\Lambda - \gamma V)^{-1}_{ij} \rangle$ as a correlation function of a (real positive) measure as in statistical mechanics.

The crux of the proof is the use of some Jacobian-free "integration by parts" special to supersymmetry and some determinant estimates of the free Laplacian $E - \Delta_\Lambda$ ($E \geq 2d$ or $E \leq -2d$). This enables us to gain extra powers in γ in the resolvent expansion of $\langle (E_0 - \Delta_\Lambda - \gamma V)^{-1}_{ij} \rangle$ about $(E_0 - \Delta_\Lambda - \gamma \langle v \rangle)^{-1}_{ij}$, which is necessary, as pointwise in ω, all the terms in the resolvent series are of the same order. This way we are able to control the series. The fact that for fixed i, $j \in \Lambda$, $(E_0 - \Delta_\Lambda - \gamma V)^{-1}_{ij}$ has a definite sign plays again a crucial role here. (Note the difference in the positioning of $\langle \, \rangle$. For more details, see Proposition 2.4.)

Estimating directly $\langle (E_0 - \Delta_\Lambda - \gamma V)^{-1}_{ij} \rangle$ is a way of controling the probability measure of the configurations of $\{v_i\}_{i \in \Lambda}$ such that $H_\Lambda = \Delta_\Lambda + \gamma V$ has eigenvalues close to $E_0 = 2d$ or $-2d$. This problem clearly has a combinatorial aspect. Our proof of Proposition 2.3 is however purely analytical upon using the supersymmetric structure naturally present in the problem.

Choosing an appropriate $\Lambda = [-L, L]^d$, we obtain (i) with the mean field rate of decay (relative to the scale L): $a_L \sqrt{\gamma}$ ($a_L > 0$, $\gamma << 1$). Using (i,ii) in the established mechanisms for localization in [FS,FMSS,vDK] stated as Theorem 4.3 in sect. IV or [AFHS], Theorem 4.2, which proves in essense that as $L \to \infty$, $a_L \to a > 0$, i.e., the mean field rate of decay is preserved, we obtain Theorem 4.3.

Previously known results and conjectures. Prior to Theorem 4.3, there have been essentially two cases, in which (i) was shown to hold in any dimension d:

(a) Large disorder: $\gamma >> 1$, for all $E \in \sigma(H)$;

(b) $\gamma = 1$, $E \to \inf \sigma(H)^-$ or $\sup \sigma(H)^+$ or more generally $E \to$ spectral edges. These are the intervals where the density of states (for a definition, see (1.11) and (1.12)) is small. They are the so called Lifshitz tails.

Using [FS,FMSS,vDK] or [AFHS] in case (a), one proves that $\sigma(H)$ is pure point almost surely for $\gamma >> 1$, hence excluding the interesting possibility of a spectral transition, called Anderson transition in this context. Similarly in case (b), one proves pure point spectrum in the Lifshitz regions.

In one dimension, (i) holds under general conditions for all $\gamma > 0$, all E and it is first proven in [GMP] that the spectrum is pure point.

However, it is conjectured, in fact strongly believed by physicists, that for $\gamma << 1$ and in dimensions greater or equal to three, there is Anderson transition from point spectrum to continuous spectrum. The region of possible spectral transition is called the mobility edge. For $\gamma << 1$, it is therefore important to locate the mobility edge.

Physicists believe that the transition occurs immediately to the left of $I_0 = (2d - \gamma\langle v \rangle, 2d]$ if $\operatorname{supp} dg \subseteq (-\infty, 0]$ (or to the right of $I_0 = [-2d, -2d + \gamma \langle v \rangle)$ if $\operatorname{supp} dg \subseteq [0, \infty)$) and that the spectrum in I_0 is pure point. Theorem 4.3 proves that the spectrum in $I_\delta = [2d - \gamma^{1+\delta}, 2d]$, if $\operatorname{supp} dg \subseteq (-\infty, 0]$ (or $I_\delta = [-2d, -2d + \gamma^{1+\delta}]$, if $\operatorname{supp} dg \subseteq [0, \infty)$) is pure point almost surely for all $\delta > 0$. So in some sense, Theorem 4.3 is "close".

To the best of our knowledge, for weak disorder $\gamma << 1$, there were no previous results of this precision in the published mathematics literature. This is the closest yet to the mobility ' edge, where localization is proven. In [Ai], Aizenman proved localization at weak disorder for energies E further away from the mobility edge. More precisely, he provd localization for E's such that $\mathrm{dist}(E, \sigma(\Delta)) \geq \gamma^{\frac{1}{1+d+\epsilon}} >> \mathbf{O}(\gamma)$ $(\epsilon > 0)$ for $\gamma << 1$. If $\mathrm{supp}\, dg \subseteq (-\infty, 0]$ and is unbounded, then for $E < 0$, this compliments Theorem 4.3. But for $E > 0$, these energies are *outside* of the spectrum H. Similarly for $\mathrm{supp}\, dg \subseteq [0, \infty)$.

A mathematical proof of Anderson transition remains notoriously difficult. Theorem 4.3 is very far from accomplishing that. Nonetheless, it does make one feel a small step closer to the region of possible spectral transition.

Comments. A more natural way to obtain (i) (or more precisely, Proposition 3.3 or 3.4) would have been to introduce the correponding complex measure similar to [SjW]. Unfortunately, this seemingly more natural and direct approach leads us to the very difficult problem of treating a *complex* and degenerate measure in large dimensions, which we do not know how to attack frontally. Instead we avoid this fundamental difficulty here by studying a close by real measure (even though still non-convex and degenerate) and subsequently using perturbation theory.

We believe that a good way of understanding the role of supersymmetry in random Schrödinger operators is by drawing the analogy with the Ising model. Recall that by using the supersymmetric representation, $\langle (E_0 - \Delta_\Lambda - \gamma V)^{-1}_{ij} \rangle$ is expressed as an integral on $\mathbf{R}^{2|\Lambda|}$. Supersymmetry refers to certain "symmetry" of a special class of differential forms on $\mathbf{R}^{2|\Lambda|}$. (For more details, see sect. III.)

In the Ising model, the partition function is calculated exactly by using also the dual of $\mathbf{Z}^2$: $(\mathbf{Z}^2)^*$. (See e.g., [McW] or the original article of Onsager [On].) Analogously, to compute $\langle (E_0 - \Delta_\Lambda - \gamma V)^{-1}_{ij} \rangle$ we use not only the original space $\mathbf{R}^{2|\Lambda|}$, but also its dual $(\mathbf{R}^{2|\Lambda|})^*$, the differential 1-forms. Once one is used to (or comes to terms with) this slightly unfamiliar way of doing analysis with supersymmetry, the proof of Proposition 3.3 is simple and very tractable.

It is well known in the physics literature, see e.g., [ID] that the quantum mechanical theory of disordered systems can be mapped to a field theory with constraints. The constraints here are expressed in terms of supersymmetry. (See sect. II, in particular (2.30).) This point of view was first incorporated in a mathematically rigorous manner by Klein etc. in [BCKP,KS], see also [K], where they obtained results in the large disorder regime: $\gamma >> 1$. This paper together with [W1,2] does justice to this approach by developing it mathematically and demonstrating its prowess in the weak disorder regime: $\gamma << 1$.

After we finished this paper, we noticed that (i) or more precisely Proposition 3.3 in fact coincides with some statements made by Spencer in [Sp], where he sketched the use of a cluster expansion (a type of combinatorial estimates) to prove (3.10). The proof of Proposition 3.3 in this paper uses entirely different methods. It develops further the supersymmetric method introduced in [W1] (see also [SjW]), to the non-convex, degenerate case. In fact we are not aware of any other precursor (in this field) to the method employed here to prove Proposition 3.3.

The result on statistics of eigenvalue spacing.

We now turn to the direction of eigenvalue statistics. We first introduce the few necessary notions in order to state our second theorem. For more details on these notions, consult sect. V.

For $\Lambda = [-L, L]^d \subset \mathbf{Z}^d$, let $E_1(\Lambda) \leq \cdots \leq E_{|\Lambda|}(\Lambda)$ be the eigenvalues of H_Λ defined in (1.3). Since $V = \{v_i\}$ is an ergodic random field. The limit

$$N(E) = \lim_{\Lambda \nearrow \mathbf{Z}^d} \frac{1}{|\Lambda|} \sharp\{j | E_j(\Lambda) \leq E\} \tag{1.11}$$

exists for almost all realizations of V. $N(E) = \langle N(E) \rangle$ almost surely and is called the integrated density of states. Since $N(E)$ is a monotone increasing function of E, the derivative:

$$\frac{dN(E)}{dE} = n(E) \tag{1.12}$$

exists a.e. in E and is called the density of states.

For a given $E \in \mathbf{R}$, define the rescaled eigenvalues ξ_j by

$$\xi_j(\Lambda, E) = |\Lambda|(E_j(\Lambda) - E), \qquad j = 1, \cdots, |\Lambda|, \tag{1.13}$$

and the point process $\xi(\Lambda, E)$ by

$$\begin{aligned} \xi(\Lambda, E)(dx) &= \sum_j \delta_{\xi_j}(dx) \\ &= \sum_j \delta_{|\Lambda|(E_j(\Lambda) - E)}(dx), \end{aligned} \tag{1.14}$$

where δ_{ξ_j} is the Dirac delta function concentrated at ξ_j. Note that (1.13) is the natural scaling as there are $|\Lambda|$ eigenvalues for H_Λ.

For any point process ξ^ω, define $\mu(dx)$ to be its intensity measure if

$$\mu(A) = \mathbf{E}[\xi^\omega(A)] \tag{1.15}$$

for each Borel set A. The Poisson point process ξ^ω with intensity $\mu(dx)$ is the process which satisfies

(Pi) For each bounded Borel set A, $\xi^\omega(A)$ obeys the Poisson distribution with parameter $\mu(A)$, i.e.,

$$\text{Prob}\,(\xi^\omega(A) = k) = e^{-\mu(A)} \frac{\mu(A)^k}{k!} \qquad k \geq 0, \tag{1.16}$$

(Pii) If $A_1 \cdots A_n$ are disjoint, then $\xi(A_1) \cdots \xi(A_n)$ are independent random variables.

Theorem 5.2. *Assume dg satisfies (H1-3). For all $\delta > 0$, let $I_\delta = [2d - \gamma^{1+\delta}, 2d]$ if $\operatorname{supp} dg \subseteq (-\infty, 0]$ and $I_\delta = [-2d, -2d + \gamma^{1+\delta}]$ if $\operatorname{supp} dg \subseteq [0, \infty)$. There exists $\gamma_0 > 0$, such that for all $\gamma \in (0, \gamma_0)$, Lebesgue almost all $E \in \sigma(H) \cap I_\delta$, as $L \to \infty$, the point process $\xi(\Lambda, E)$ defined in (1.14) converges weakly to the Poisson point process ξ defined in (Pi,Pii) with intensity measure $n(E)dx$.*

Poisson statistics for eigenvalue spacing was first proved in one dimension: $d = 1$ by Molchanov in [Mo2]. It is based on some precise analysis of localization on a large finite interval in [Mo1]. For $d = 1$, Poisson statistics holds for the whole spectrum. Molchanov's approach was further developed by Minami in [Mi], where he showed that appropriate exponential estimates on G_Λ for all finite cubes Λ suffice to prove Poisson statistics in *any* dimension d. Minami's result is quoted in Theorem 5.1 of sect. V.

Using Proposition 3.4 in Theorem 4.2, we see that the condition for applying Theorem 5.1 is satisfied. We henceforth obtain Poisson statistics in the interval I_δ in Theorem 5.2.

Let A be an interval in $\mathbf{R}$, $|A|$ be its length. For Lebesgue a.e. E in I_δ, denote by $\tilde{A}$, the corresponding rescaled interval with length $|A|/n(E) < \infty$, where $n(E)$ is the density of states. Theorem 5.2 implies in particular that for almost all $E \in I_\delta$, as $\Lambda \nearrow \mathbf{Z}^d$,

$$\begin{aligned} \operatorname{Prob}(\xi^\omega(\Lambda, E)(\tilde{A}) = k) &\to \operatorname{Prob}(\xi^\omega(\tilde{A}) = k) \\ &= e^{-|A|}\frac{|A|^k}{k!}, \qquad (k \in \mathbf{N}). \end{aligned} \tag{1.17}$$

We see therefore that after rescaling by the density of states $n(E)$, which is a model specific quantity, we have *universal* behavior for eigenvalue statistics in I_δ.

For "large random matrices", (random matrices with full entries), which have many symmetry properties, universality of eigenvalue statistics is widely expected and has been proven under increasing generality, see e.g., [Me]. The operator H defined in (1.1) is in a completely different category, it has few symmetry. As a consequence, the proof of Theorem 5.2 is very different from the corresponding ones for large random matrices. Due to the symmetry properties, large random matrices have strong connections to integrable systems, see e.g., [Me]. The only symmetry here that one uses to study H, if one could borrow that term is supersymmetry. (See sect. II.)

The universal eigenvalue statistics that one obtains is also very different. In particular, from (1.17) there is no level repulsion, i.e., the probability of two eigenvalues coincide is not zero. Level repulsion is one of the dominant features of large random matrices. It is not surprising that level repulsion does not occur here, as Theorem 4.3 states that almost surely eigenfunctions of H decay exponentially fast for $E \in I_\delta$. There is an effective decoupling at distances of the order $\mathbf{O}(\gamma^{-1/2})$. (For more details, see [Mi].) However it is conjectured and supported by physics arguments [Ef] that in the conjectured continuous part of the spectrum of H, the eigenvalue statistics is the same as that of the so called Dyson's circular ensemble [Dy], which is a type of large random matrices. It would indeed be most interesting to study this cross over behavior.

II. A general remainder estimate

For definitiveness, we assume $\operatorname{supp} dg \subseteq (-\infty, 0]$. The other case clearly works in the same way. This is because the replacement of V by $-V$, followed by the unitary map $u(n) \to (-1)^{|n|} u(n)$ takes H into $-H$. Let $E = 2d$. Let $\Lambda = [-L, L]^d$, $L \in \mathbf{N}$. We define H_Λ, G_Λ as in (1.3), (1.7). It is well known that for all cubes Λ in $\mathbf{Z}^d$,

$$E - \Delta_\Lambda = 2d - \Delta_\Lambda > 0 \tag{2.1}$$

as an operator. Hence

$$E - \Delta_\Lambda - \gamma V > 0 \tag{2.2}$$

when $\operatorname{supp} dg \subseteq (-\infty, 0]$. The finite volume Green's function

$$G_\Lambda = (E - \Delta_\Lambda - \gamma V)^{-1} \tag{2.3}$$

is well defined at $E = 2d$.

Our aim in this section is to estimate $\langle (E - \Delta_\Lambda - \gamma V)^{-1}_{\mu\nu} \rangle$, μ, $\nu \in \Lambda$, at $E = 2d$ for some *fixed* but sufficiently large (depending on γ) Λ. Recall from [W1] that for $E > 2d$, we know how to *compute* the asymptotics of $\langle (E - \Delta_\Lambda - \gamma V)^{-1}_{\mu\nu} \rangle$, as $\Lambda \nearrow \mathbf{Z}^d$, $\gamma \searrow 0$ and for all μ, ν. We use the supersymmetric representation as in [W1,2]. Our main result is Proposition 2.3, whose proof relies on the crucial Proposition 2.4. The relevance of Proposition 2.3 is manefested in sect. III, where we derive Proposition 3.3 from it. Proposition 3.3 furnishes the necessary first step toward proofs of localization at *weak* disorder: $\gamma << 1$.

The supersymmetric representation.

For readers' convenience and for the sake of independence, we present below a brief introduction to the supersymmetric representation. For a general proper formulation, see e.g., [Be,V]. This is essentially a reorganization of sect. II of [W1]. The raison d'etre of this interlude is the formulae in Proposition 2.2.

For any symmetric $n \times n$ matrix A with $\Re A > 0$, we recall the following well-known formula

$$A^{-1}_{\mu\nu} = \int x_\mu \cdot x_\nu \det A e^{-\sum_{i,j=1}^n A_{ij} x_i \cdot x_j} \prod_{i=1}^n \frac{d^2 x_i}{\pi}, \tag{2.4}$$

where $x_i \in \mathbf{R}^2$ and $x_i \cdot x_j$ is the usual scalar product in $\mathbf{R}^2$. Substituting $E - H_\Lambda$ for A in the above formula, and drop the subscript Λ for simplicity, we then have

$$G(E, \mu, \nu) = \int x_\mu \cdot x_\nu \det(E - H) e^{-\sum_{i,j \in \Lambda} (E-H)_{ij} x_i \cdot x_j} \prod_{i \in \Lambda} \frac{d^2 x_i}{\pi}. \tag{2.5}$$

(In this section, since we only work in Λ, if it is convenient, we will continue to drop the subscript Λ without further notice.)

Let $|\Lambda|$ be the number of points in Λ. We use the Grassmann algebra of $2|\Lambda|$ generators to express $\det[E-H]$. This algebra is generated by $2|\Lambda|$ anticommuting variables ξ_i, η_i, $i \in \Lambda$ satisfying the relations:

$$
\begin{aligned}
[\xi_i, \eta_j] &= \xi_i\eta_j + \eta_j\xi_i = 0, \\
[\xi_i, \xi_j] &= \xi_i\xi_j + \xi_j\xi_i = 0, \\
[\eta_i, \eta_j] &= \eta_i\eta_j + \eta_j\eta_i = 0,
\end{aligned}
\tag{2.6}
$$

where we write $[a,b] = ab+ba$ for the anti-commutator. It is denoted by $\Lambda[\xi_1, \eta_1, .., \xi_{|\Lambda|}, \eta_{|\Lambda|}]$ (if we identify Λ with $\{1, .., |\Lambda|\}$). From (2.4), we see in particular that $\xi_i^2 = \eta_i^2 = 0$. We write these anticommutative variables collectively as ξ, η.

To make connections with better known objects, ξ_j, η_j can be thought of as the two 1-forms on the j^{th} copy of $\mathbf{R}^2$. "C^∞ functions" $F(\xi_i, \eta_j)$ of these anticommuting variables are defined by Taylor's formula at $(0,0)$ which contains a finite number of terms because of nilpotency. In this way $F(\xi, \eta)$ becomes an element of the Grassmann algebra. For example if

$$
F(\xi, \eta) = e^{A_{ij}\xi_i\eta_j},
\tag{2.7}
$$

then

$$
F(\xi, \eta) = 1 + A_{ij}\xi_i\eta_j.
\tag{2.8}
$$

This is the function that we need in writing the determinant. We also need to define the notions of differentiation and integration. Define:

$$
\begin{aligned}
\frac{\partial}{\partial \xi_i}(\xi_i) &= 1, \\
\frac{\partial}{\partial \eta_i}(\eta_i) &= 1.
\end{aligned}
\tag{2.9}
$$

We require further that these differentiations be linear operators and that Leibnitz' rule hold. We can then define integrals (with respect to ∂) as follows:

$$
\int 1 d\xi_i = 0, \ \int \xi_i d\xi_i = 1, \ \int 1 d\eta_i = 0, \ \int \eta_i d\eta_i = 1.
\tag{2.10}
$$

A multiple integral is defined to be a repeated integral. For example,

$$
\int \xi_i \eta_j d\xi_i d\eta_j = -\int \eta_j \xi_i d\xi_i d\eta_j = -\int \eta_j d\eta_j = -1.
\tag{2.11}
$$

Using (2.7)–(2.11), we obtain

$$
\det(E - H) = \int e^{-\sum_{j,k\in\Lambda}(E-H)_{jk}\eta_j\xi_k} \prod_{j\in\Lambda}(d\eta_j d\xi_j),
\tag{2.12}
$$

$$
(E-H)^{-1}_{\mu\nu} \det(E-H) = \int \xi_\mu \eta_\nu e^{-\sum_{j,k\in\Lambda}(E-H)_{jk}\eta_j\xi_k} \prod_{j\in\Lambda}(d\eta_j d\xi_j),
\tag{2.13}
$$

We illustrate (2.12) in the case where $(E-H)=M$ is a 2×2 matrix. The integrand in the RHS of (2.12) is

$$e^{-\sum_{j,k=1,2} M_{jk}\eta_j\xi_k} = \prod_{j,k=1,2}(1-M_{jk}\eta_j\xi_k), \tag{2.14}$$

where we used the commutativity of $e^{-M_{jk}\eta_j\xi_k}$ and (2.15, 2.16). Doing the integral in the RHS of (2.14) upon using (2.10) and (2.11), we obtain

$$\text{RHS} = M_{11}M_{22} - M_{12}M_{21} = \det M, \tag{2.15}$$

as expected. Similarly, we obtain for the RHS of (2.13):

$$\text{RHS} = (-1)^{|\mu|+|\nu|}\det(M^{\mu\nu}) = M^{-1}_{\mu\nu}\det M, \qquad \mu,\, \nu = 1,\, 2, \tag{2.16}$$

where $M^{\mu\nu}$ is the co-matrix with the μth row and νth column removed. (In the 2×2 case it is just the one matrix element that is left.)

Combining (2.5) with (2.12), we obtain the following expression:

$$G(E,\mu,\nu) = \int x_\mu \cdot x_\nu e^{-\sum_{j,k\in\Lambda}(E-H)_{jk} X_j\cdot X_k} \prod_{j\in\Lambda} d^2X_j, \tag{2.17}$$

where

$$\begin{aligned} x_j \in& \mathbf{R}^2, \\ X_j \overset{\text{def}}{=}& (x_j, \xi_j, \eta_j), \\ X_j\cdot X_k \overset{\text{def}}{=}& x_j\cdot x_k + \frac{1}{2}(\eta_j\xi_k + \eta_k\xi_j), \\ d^2X_j \overset{\text{def}}{=}& \frac{d^2x_j}{\pi} d\eta_j d\xi_j, \end{aligned} \tag{2.18}$$

and we used that $(E-H)$ is a symmetric matrix.

Let $\hat{g}$ be the Laplace transform of dg:

$$\hat{g}(t) = \int e^{tv} dg(v), \qquad t\geq 0 \tag{2.19}$$

Write

$$\hat{g}(t) = e^{k(t)}, \tag{2.20}$$

as $\hat{g}>0$. For example, for the Bernoulli measure:

$$dg(v) = \frac{\delta(v+1)+\delta(v)}{2}, \tag{2.21}$$

$\hat{g}(t) = e^{-t/2}\cosh(t/2)$. Using (2.17) and taking the expectation value with respect to dg, we obtain

(2.22)

$$\begin{aligned}\langle G(E,\mu,\nu)\rangle &= \int x_\mu \cdot x_\nu e^{-(\sum_{j\in\Lambda} EX_j\cdot X_j - \sum_{j,k\in\Lambda,|j-k|_1=1} X_j\cdot X_k - \gamma\sum_{j\in\Lambda} v_j X_j\cdot X_j)} \\ &\qquad \prod_{j\in\Lambda} dg(v_j) \prod_{j\in\Lambda} d^2X_j \\ &= \int x_\mu \cdot x_\nu e^{-(\sum_{j\in\Lambda} EX_j\cdot X_j - \sum_{j,k\in\Lambda,|j-k|_1=1} X_j\cdot X_k - \sum_{j\in\Lambda} k(\gamma X_j\cdot X_j))} \prod_{j\in\Lambda} d^2X_j \\ &\stackrel{\text{def}}{=} \int x_\mu \cdot x_\nu e^{-2\Phi(X)} \prod_{j\in\Lambda} d^2X_j,\end{aligned}$$

where

$$\begin{aligned}k(\gamma X_j\cdot X_j) &= k(\gamma x_j\cdot x_j + \gamma\eta_j\xi_j) \\ &= k(\gamma x_j\cdot x_j) + \gamma k'(x_j\cdot x_j)\eta_j\xi_j,\end{aligned} \tag{2.23}$$

by using (2.8) to expand $e^{\gamma v_j\xi_j\eta_j}$ and direct computations, and

$$2\Phi(X) \stackrel{\text{def}}{=} \sum_{j\in\Lambda} EX_j\cdot X_j - \sum_{i,j\in\Lambda,|i-j|_1=1} X_i\cdot X_j - \sum_{j\in\Lambda} k(\gamma X_j\cdot X_j). \tag{2.24}$$

We notice in (2.22), (2.24) that apart from the factor $x_\mu\cdot x_\nu$, the integrand is only a "function" of $X_j\cdot X_k$. More precisely, we consider the algebra $C^\infty(\mathbf{R}^{2|\Lambda|})\otimes \Lambda[\xi_1,\eta_1,..,\xi_{|\Lambda|},\eta_{|\Lambda|}]$. We call elements of this algebra, (super)functions. Clearly,

$$\begin{aligned}D(X) = X_i\cdot X_j &= x_i\cdot x_j + \frac{1}{2}(\eta_i\xi_j + \eta_j\xi_i) \\ &\in C^\infty(\mathbf{R}^{2|\Lambda|})\otimes \Lambda[\xi_1,\eta_1,..,\xi_{|\Lambda|},\eta_{|\Lambda|}],\end{aligned}$$

and is hence a (super)function. We call the scalars (which are functions in $C^\infty(\mathbf{R}^{2|\Lambda|})$) in front of the Grassmann generators ξ, η, coefficients. In the case of $X_i\cdot X_j$, the coefficients are $x_i\cdot x_j$ and $1/2$. Those (super)functions which are only functions of $X_i\cdot X_j$ $(i,\ j\in\Lambda)$ are called supersymmetric functions. More precisely, supersymmetries are defined to be the set of transformations (in the alegebraic sense) that leave the dot products $X_i\cdot X_j$ $(i,\ j\in\Lambda)$ invariant. Two obvious transformations that leave the dot product invariant are the usual rotations O in $\mathbf{R}^2$,

$$x_i = x_i'O \qquad (i=1,\cdots,n)$$

and the transformations $A\in Sp(2)$ acting on ξ_i, η_i $(i=1,\cdots,n)$ such that

$$(\xi_i,\eta_i) = (\xi_i',\eta_i')A,$$

where $\{x_i',\xi_i',\eta_i'\}_{i=1}^n$ is another set of coordinates, x_i' being the even ones and ξ_i', η_i' the odd ones. We put $X_i' = (x_i',\xi_i',\eta_i')$ $\quad (i=1,\cdots,n)$. Aside from these two linear transformations, supersymmetries also include transformations generated by (super)vector fields of the type:

$$
\begin{aligned}
V &= \sum_i V_i \\
&= \sum_i (\xi_i a + \eta_i b)\frac{\partial}{\partial x_i} + 2(b\cdot x_i)\frac{\partial}{\partial \xi_i} - 2(a\cdot x_i)\frac{\partial}{\partial \eta_i},
\end{aligned}
$$

where a, $b \in \mathbf{R}^2$, and

$$
\begin{aligned}
a\frac{\partial}{\partial x_i} &\stackrel{\text{def}}{=} a^{(1)}\frac{\partial}{\partial x_i^{(1)}} + a^{(2)}\frac{\partial}{\partial x_i^{(2)}}, \\
b\frac{\partial}{\partial x_i} &\stackrel{\text{def}}{=} b^{(1)}\frac{\partial}{\partial x_i^{(1)}} + b^{(2)}\frac{\partial}{\partial x_i^{(2)}}.
\end{aligned}
$$

(Note that it is the same transformation in all the X_i.) The above transformation is to be understood in the algebraic sense. We check that $VD(X_i, X_j) = 0$. Analogous to the usual change of variables, the (super) change of variables also generates a Jacobian, which is called a Berezinian. We will not go into the details, except mentioning that the Berezinian here is 1.

Let τ be a supersymmetric transformation. Let $X_i = \tau X_i'$ $(i = 1, \cdots, n)$.

Definition. *A superfunction F is supersymmetric if it is invariant under all supersymmetries:*

$$F(X_1, \cdots, X_n) = F(X_1', \cdots, X_n'),$$

for all τ supersymmetric transformations.

Clearly, supersymmetric functions belong to a rather restricted class of functions. For example, in $\mathbf{R}^2$, with two generators ξ, η, F is supersymmetric if and only if there exists f: $[0, \infty) \mapsto \mathbf{R}$ of class C^∞, such that

$$F(X) = f(X\cdot X) = f(x\cdot x) + f'(x\cdot x)\eta\xi. \tag{2.25}$$

For the general case, see *e.g.*, [KS].

Define $dX_i = (d^2x_i/\pi)d\eta_i d\xi_i$ $(i = 1, \cdots, n)$. One of the most useful properties of the supersymmetric functions, which we rely on for the proof of Proposition 2.4, is the following:

Proposition 2.1. *If F is supersymmetric with all of its coefficients in $\mathcal{S}(\mathbf{R}^{2n})$, then*

$$\int F(X_1, \cdots, X_n)dX_1 \cdots dX_n = F(0, \cdots 0). \tag{2.26}$$

For a proof, see e.g., [K,KS]. Proposition 2.1 is related to the so called "localization formula" in geometry, see e.g., [BGV].

Proposition 2.2.

$$\begin{aligned}\langle G(E,\mu,\nu)\rangle &= \langle (E-\Delta-\gamma V)^{-1}_{\mu\nu}\rangle \\ &= \int \xi_\mu \eta_\nu e^{-2\Phi} \prod_{j\in\Lambda} d^2 X_j \\ &= \int (E-\Delta-\gamma\,\mathrm{diag}\,k')^{-1}_{\mu\nu} e^{-2\phi} \prod_{j\in\Lambda} \frac{d^2 x_j}{\pi},\end{aligned} \tag{2.27}$$

where

(2.28)

$$\phi(x) = \frac{1}{2}[\sum_{i,j}(E-\Delta)_{ij} x_i\cdot x_j - \sum_{j\in\Lambda} k(\gamma x_j\cdot x_j) - \log\det(E-\Delta-\gamma\,\mathrm{diag}\,k'(\gamma x_j\cdot x_j))],$$

$$x_i \in \mathbf{R}^2,$$

k is as defined in (2.20), and $\mathrm{diag}\,k'$ *denotes the diagonal matrix with entries* $(\mathrm{diag}\,k')jj = k'(\gamma x_j\cdot x_j)$.

Remark. It is instructive to compare the first and the third equalities in (2.27).

Proof. Let $F(X) = X_\mu\cdot X_\nu e^{-2\Phi}$. Using Proposition 2.1, (2.23), $F(0)=0$, (2.18), (2.24) and integrating over the anti-commuting variables as in (2.13), we have:

$$\begin{aligned}\langle G(E,\mu,\nu)\rangle &= \int x_\mu\cdot x_\nu e^{-2\Phi} \prod_{j\in\Lambda} d^2 X_j \\ &= \int \xi_\mu \eta_\nu e^{-2\Phi} \prod_{j\in\Lambda} d^2 X_j \\ &= \int (E-\Delta-\gamma\,\mathrm{diag}\,k')^{-1}_{\mu\nu} e^{-2\phi} \prod_{j\in\Lambda} \frac{d^2 x_j}{\pi}.\end{aligned} \tag{2.29}$$

□

We end this subsection by remarking that the measure $e^{-2\phi}$ with ϕ defined in (2.28) is normalized:

$$\int e^{-2\phi} = 1 \qquad \text{for all}\,\Lambda, \tag{2.30}$$

as $\prod_{j\in\Lambda} dg(v_j)$ is normalized. This can also be verified by using (2.24,2.26) and Proposition 2.1

A general remainder estimate.

Using the representation in Proposition 2.2, we arrive at our main estimate of this section:

Proposition 2.3. *Assume dg satisfies (H1,2). There exists $\gamma_0 > 0$, such that for all $0 < \gamma < \gamma_0$, all $n \in \mathbf{N}$, all cubes $\Lambda = [-L, L]^d$ in $\mathbf{Z}^d$, all μ, $\nu \in \Lambda$,*

$$\begin{aligned}\langle (E-\Delta_\Lambda-\gamma V)^{-1}_{\mu\nu}\rangle =&(E-\Delta_\Lambda-\gamma\langle v\rangle)^{-1}_{\mu\nu}+\mathbf{O}(\gamma\tilde\gamma)(E-\Delta_\Lambda-\gamma\langle v\rangle)^{-2}_{\mu\nu}\\ &+\mathbf{O}(\gamma\tilde\gamma)^2(E-\Delta_\Lambda-\gamma\langle v\rangle)^{-3}_{\mu\nu}+\cdots\\ &+\mathbf{O}_n(\tilde\gamma^n)L^2, \qquad\qquad (E\geq 2d),\end{aligned}\tag{2.31}$$

where

$$\tilde\gamma=\begin{cases}\gamma L & d=1\\ \gamma\log L & d=2\\ \gamma & d\geq 3.\end{cases}\tag{2.32}$$

Remark. We note that the error term is independent of $|\mu-\nu|$. (Compare with Theorem 7.2 of [W1].) (2.31) is therefore useful only for fixed Λ and μ, ν. The dependence of $\tilde\gamma$ on the spatial dimension d reflects the well known dependence of the free Green's function $(E-\Delta_\Lambda)^{-1}_{ij}$ $(E\geq 2d)$, on d. See Lemma 2.5 below.

To prove Proposition 2.3, we use the third equality in (2.27). Using the resolvent equation, we expand the integrand about the constant potential $k'(0)=\langle v\rangle$:

$$\begin{aligned}&(E-\Delta_\Lambda-\gamma\,\mathrm{diag}\,k')^{-1}_{\mu\nu}\\ &=(E-\Delta_\Lambda-\gamma k'(0))^{-1}_{\mu\nu}\\ &\quad+\gamma\sum_{j_1\in\Lambda}(E-\Delta_\Lambda-\gamma k'(0))^{-1}_{\mu j_1}(k'(\gamma x_{j_1}\cdot x_{j_1})-k'(0))(E-\Delta_\Lambda-\gamma k'(0))^{-1}_{j_1\nu}\\ &\quad+\gamma^2\sum_{j_1,j_2\in\Lambda}(E-\Delta_\Lambda-\gamma k'(0))^{-1}_{\mu j_1}(k'(\gamma x_{j_1}\cdot x_{j_1})-k'(0))(E-\Delta_\Lambda-\gamma k'(0))^{-1}_{j_1j_2}\\ &\qquad\qquad(k'(\gamma x_{j_2}\cdot x_{j_2})-k'(0))(E-\Delta_\Lambda-\gamma k'(0))^{-1}_{j_2\nu}\\ &\quad+\cdots\\ &\quad+\gamma^n\sum_{j_1\cdots j_n\in\Lambda}(E-\Delta_\Lambda-\gamma k'(0))^{-1}_{\mu j_1}(k'(\gamma x_{j_1}\cdot x_{j_1})-k'(0))\cdots\\ &\qquad\qquad(E-\Delta_\Lambda-\gamma k'(0))^{-1}_{j_{n-1}j_n}(k'(\gamma x_{j_n}\cdot x_{j_n})-k'(0))(E-\Delta_\Lambda-\gamma\,\mathrm{diag}\,k')^{-1}_{j_n\nu}.\end{aligned}\tag{2.33}$$

Note that the last factor in (2.33) is $(E-\Delta_\Lambda-\gamma\,\mathrm{diag}\,k')^{-1}$ itself. In view of (2.33), the proof of the Proposition 2.3 depends on the following:

Proposition 2.4. *For all $m\in\mathbf{N}$, if dg satisfies (H1,2), then*

$$\begin{aligned}&\int\prod_{i=1}^m(k'(\gamma x_{\ell_i}\cdot x_{\ell_i})-k'(0))e^{-2\phi}\prod_{j\in\Lambda}\frac{d^2x_j}{\pi}\\ &=\int\prod_{i=1}^m|(k'(\gamma x_{\ell_i}\cdot x_{\ell_i})-k'(0))|e^{-2\phi}\prod_{j\in\Lambda}\frac{d^2x_j}{\pi}\\ &\overset{\text{def}}{=}\mathbf{O}_m(1)\begin{cases}\gamma L & d=1\\ \gamma\log L & d=2\\ \gamma & d\geq 3.\end{cases}\\ &\overset{\text{def}}{=}\mathbf{O}_m(1)\tilde\gamma^m.\end{aligned}\tag{2.34}$$

The proof of Proposition 2.4 uses Proposition 2.1, induction and the following standard estimate on the free Green's function, which we state without proof. See e.g., [La]. Let $\Lambda = [-L, L]^d$ as before.

Lemma 2.5. *For $E \geq 2d$,*

$$0 < (E - \Delta_\Lambda)^{-1}_{ij} \leq \begin{cases} \mathbf{O}(L), & d = 1 \\ \mathbf{O}(\log(\frac{L}{|i-j|+1})) & d = 2 \\ \frac{\mathbf{O}_d(1)}{(|i-j|+1)^{d-2}} & d \geq 3, \end{cases} \tag{2.35}$$

for all i, $j \in \Lambda$ and $\mathbf{O}_d(1)$ is uniform in L.

Corollary 2.6. *(maximum principle) Let $S \subseteq \Lambda$, be any subset of $\Lambda = [-L, L]^d$. Define*

$$\tilde{\Delta}_\Lambda(i,j) = \begin{cases} \Delta_\Lambda(i,j), & i, j \in \Lambda \backslash S, \\ 0 & \text{otherwise}; \end{cases} \tag{2.36}$$

$$\tilde{M} = E - \tilde{\Delta}_\Lambda - \gamma \operatorname{diag} k' \qquad (E \geq 2d).$$

Then

$$0 < \tilde{M}^{-1}_{ij} \leq \begin{cases} \mathbf{O}(L), & d = 1 \\ \mathbf{O}(\log(\frac{L}{|i-j|+1})) & d = 2 \\ \frac{\mathbf{O}_d(1)}{(|i-j|+1)^{d-2}} & d \geq 3, \end{cases} \tag{2.37}$$

for all i, $j \in \Lambda \backslash S$ and $\mathbf{O}_d(1)$ is uniform in L.

Proof. By the definition (2.20),

$$\hat{g}(t) = \int e^{tv} dg(v) = e^{k(t)} \qquad (t \geq 0, \operatorname{supp} dg \subseteq (-\infty, 0]). \tag{2.38}$$

So

$$k'(t) = \frac{\int v e^{tv} dg(v)}{\int e^{tv} dg(v)} \stackrel{\text{def}}{=} \langle v \rangle_t. \tag{2.39}$$

From (H2),

$$-\infty < \langle v \rangle_0 = \langle v \rangle < 0, \tag{2.40}$$

and since

$$\frac{d\langle v \rangle_t}{dt} = \langle v^2 \rangle_t - \langle v \rangle_t^2 \geq 0, \tag{2.41}$$

we have that

$$-\infty < k'(t) < 0 \tag{2.42}$$

for all $t \geq 0$ by using (2.39), (2.41).

We first prove the lower bound. Since

$$\frac{\partial}{\partial E}\tilde{M}^{-1}(E)_{ij} = \frac{\partial}{\partial E}(E - \tilde{\Delta}_\Lambda - \gamma \operatorname{diag} k')^{-1}_{ij} = -(E - \tilde{\Delta}_\Lambda - \gamma \operatorname{diag} k')^{-2}_{ij}, \tag{2.43}$$

in order to prove $\tilde{M}(E)^{-1}_{ij} > 0$, for all $E \geq 2d$, it suffices to prove $\tilde{M}(E)^{-1}_{ij} > 0$ for some sufficiently large E. For such large E, we obtain by using the resolvent equation:

$$\begin{aligned} &(E - \tilde{\Delta}_\Lambda - \gamma \operatorname{diag} k')^{-1} \\ =&(E - \gamma \operatorname{diag} k')^{-1} + (E - \gamma \operatorname{diag} k')^{-1}\tilde{\Delta}_\Lambda(E - \gamma \operatorname{diag} k')^{-1} \\ &+ \cdots + [(E - \gamma \operatorname{diag} k')^{-1}\tilde{\Delta}_\Lambda]^{-n}(E - \gamma \operatorname{diag} k')^{-1} \\ &+ \cdots \end{aligned} \tag{2.44}$$

The above series is convergent for E sufficiently large. Since $\tilde{\Delta}_\Lambda$ has non-negative entries, we obtain that

$$(E - \tilde{\Delta}_\Lambda - \gamma \operatorname{diag} k')^{-1}_{ij} > 0 \tag{2.45}$$

for large E. Using (2.43), we then obtain the lower bound in (2.37) for $E \geq 2d$.

Set $\delta_S = \Delta_\Lambda - \tilde{\Delta}_\Lambda$. δ_S has non-negative entries. Using the resolvent equation again, we have

$$\begin{aligned} (E - \tilde{\Delta}_\Lambda - \gamma \operatorname{diag} k')^{-1} =&(E - \Delta_\Lambda + \delta_S - \gamma \operatorname{diag} k')^{-1} \\ =&(E - \Delta_\Lambda)^{-1} \\ &- (E - \Delta_\Lambda)^{-1}(\delta_S - \gamma \operatorname{diag} k')(E - \tilde{\Delta}_\Lambda - \gamma \operatorname{diag} k')^{-1}. \end{aligned} \tag{2.46}$$

Since δ_S has non-negative entries and $-\infty < k'(t) < 0$ from (2.42), we obtain

$$(E - \tilde{\Delta}_\Lambda - \gamma \operatorname{diag} k')^{-1}_{ij} < (E - \Delta_\Lambda)^{-1}_{ij}. \tag{2.47}$$

Using Lemma 2.5, we obtain the upper bound in (2.37). □

Proof of Proposition 2.4. We first show that under (H1,2)

$$0 \leq k''(t) < \infty \qquad (t \geq 0), \tag{2.48}$$

which implies that $k'(t) - k'(0) \geq 0$ $(t \geq 0)$ and hence prove the first equality in (2.34). Since

$$\begin{aligned} \hat{g}(t) &= \int e^{tv} dg(v) = e^{k(t)}, \\ \hat{g}'(t) &= \int v e^{tv} dg(v) = k'(t)e^{k(t)}, \\ \hat{g}''(t) &= \int v^2 e^{tv} dg(v) = ({k'}^2(t) + k''(t))e^{k(t)}, \qquad (t \geq 0,\, v \leq 0) \end{aligned} \tag{2.49}$$

$$k''(t) = \frac{\int v^2 e^{tv} dg(v)}{\int e^{tv} dg(v)} - \left(\frac{\int v e^{tv} dg(v)}{\int e^{tv} dg(v)}\right)^2 \tag{2.50}$$
$$= \langle v^2 \rangle_t - \langle v \rangle_t^2$$
$$\geq 0.$$

We therefore obtain the lower bound in (2.48).

From (H2), $\langle v^2 \rangle_0 = \langle v^2 \rangle < \infty$. Since

$$\frac{d}{dt}\langle v^2 \rangle_t = \langle v^3 \rangle_t - \langle v^2 \rangle_t \langle v \rangle_t \leq 0 \tag{2.51}$$

by a trivial application of the FKG inequality [FKG,Li], as v^2, v are monotone in v, we obtain that k'' is bounded.

We now use induction to prove the estimate in (2.34). For $m = 1$, write ℓ for ℓ_1. Recall from (2.25) that

$$k'(\gamma X_\ell \cdot X_\ell) = k'(\gamma x_\ell \cdot x_\ell) + \gamma k''(\gamma x_\ell \cdot x_\ell)\eta_\ell \xi_\ell \tag{2.52}$$

and that $k'(\gamma X_\ell \cdot X_\ell)e^{-2\Phi}$ is a supersymmetric function. Using Proposition 2.1, we therefore have

$$\begin{aligned} &\int (k'(\gamma x_\ell \cdot x_\ell) - k'(0))e^{-2\phi} \prod_{j\in\Lambda} \frac{d^2 x_j}{\pi} \\ =&\int (k'(\gamma x_\ell \cdot x_\ell) - k'(\gamma X_\ell \cdot X_\ell))e^{-2\Phi} \prod_{j\in\Lambda} \frac{d^2 X_j}{\pi} \\ =&\int \gamma k''(\gamma x_\ell \cdot x_\ell)\xi_\ell \eta_\ell e^{-2\Phi} \prod_{j\in\Lambda} \frac{d^2 X_j}{\pi}, \end{aligned} \tag{2.53}$$

where Φ is as in (2.24), $d^2 X_j$ is as defined in (2.18) and we used (2.52). Note the change of ϕ, $d^2 x_j$ to Φ, $d^2 X_j$ in the second and third lines of (2.53).

Integrating over the anticommuting variables ξ, η in (2.53), we obtain

$$\begin{aligned} &\int (k'(\gamma x_\ell \cdot x_\ell) - k'(0))e^{-2\phi} \prod_{j\in\Lambda} \frac{d^2 x_j}{\pi} \\ =&\gamma \int k''(\gamma x_\ell \cdot x_\ell)(E - \Delta_\Lambda - \gamma \operatorname{diag} k')^{-1}_{\ell\ell} e^{-2\phi} \prod_{j\in\Lambda} \frac{d^2 X_j}{\pi} \\ =&\mathbf{O}(1) \begin{cases} \gamma L & d = 1 \\ \gamma \log L & d = 2 \\ \gamma & d \geq 3. \end{cases} \\ =&\mathbf{O}(\tilde{\gamma}), \end{aligned} \tag{2.54}$$

where we used (2.48) and Corollary 2.6.

To illustrate the mechanism, we also work out the case $m = 2$ explicitly. For simplicity of notation, we write ℓ for ℓ_1, n for ℓ_2. From Proposition 2.1, we have

$$\begin{aligned}
&\int (k'(\gamma X_\ell \cdot X_\ell) - k'(0))(k'(\gamma X_n \cdot X_n) - k'(0))e^{-2\Phi} \prod_{j\in\Lambda} \frac{d^2 X_j}{\pi} \\
=&\int (k'(\gamma x_\ell \cdot x_\ell) - k'(0) + \gamma k''(\gamma x_\ell \cdot x_\ell)\eta_\ell \xi_\ell) \\
&\qquad (k'(\gamma x_n \cdot x_n) - k'(0) + \gamma k''(\gamma x_n \cdot x_n)\eta_\ell \xi_\ell)e^{-2\Phi} \prod_{j\in\Lambda} \frac{d^2 X_j}{\pi} \\
=&0.
\end{aligned} \tag{2.55}$$

Hence

$$\begin{aligned}
&\int (k'(\gamma x_\ell \cdot x_\ell) - k'(0))(k'(\gamma x_n \cdot x_n) - k'(0))e^{-2\Phi} \prod_{j\in\Lambda} \frac{d^2 X_j}{\pi} \\
=&\int (k'(\gamma x_\ell \cdot x_\ell) - k'(0))(k'(\gamma x_n \cdot x_n) - k'(0))e^{-2\phi} \prod_{j\in\Lambda} \frac{d^2 x_j}{\pi} \\
=&\gamma \int (k'(\gamma x_\ell \cdot x_\ell) - k'(0))k''(\gamma x_n \cdot x_n)\xi_n \eta_n e^{-2\Phi} \prod_{j\in\Lambda} \frac{d^2 X_j}{\pi} \\
&+ \gamma \int (k'(\gamma x_n \cdot x_n) - k'(0))k''(\gamma x_\ell \cdot x_\ell)\xi_\ell \eta_\ell e^{-2\Phi} \prod_{j\in\Lambda} \frac{d^2 X_j}{\pi} \\
&- \gamma^2 \int k''(\gamma x_\ell \cdot x_\ell)k''(\gamma x_n \cdot x_n)\xi_\ell \eta_\ell \xi_n \eta_n e^{-2\Phi} \prod_{j\in\Lambda} \frac{d^2 X_j}{\pi}
\end{aligned} \tag{2.56}$$

Writing Δ for Δ_Λ, we perform the Grassmann integrations:

$$\begin{aligned}
&\int \xi_\ell \eta_\ell e^{-2\Phi} \prod_{j\in\Lambda} d\eta_j d\xi_j \\
=&\det(E - \Delta - \gamma \operatorname{diag} k')^{\ell\ell} e^{-2\phi_0} \\
=&\frac{\det(E - \Delta - \gamma \operatorname{diag} k')^{\ell\ell}}{\det(E - \Delta - \gamma \operatorname{diag} k')} e^{-2\phi_0} \det(E - \Delta - \gamma \operatorname{diag} k') \\
=&(E - \Delta - \gamma \operatorname{diag} k')^{-1}_{\ell\ell} e^{-2\phi},
\end{aligned} \tag{2.57}$$

where $(E-\Delta-\gamma \operatorname{diag} k')^{\ell\ell}$ is the minor obtained from $(E-\Delta-\gamma \operatorname{diag} k')$ by deleting the ℓ^{th} "row" and the ℓ^{th} "column",

$$\phi_0 = \frac{1}{2}[\sum_{i,j} (E - \Delta)_{ij} x_i \cdot x_j - \sum_{j\in\Lambda} k(\gamma x_j \cdot x_j)], \tag{2.58}$$

and we used (2.28) to reach the last equality in (2.57); and

(2.59)

$$\begin{aligned}
&\int \xi_\ell \eta_\ell \xi_n \eta_n e^{-2\Phi} \prod_{j\in\Lambda} d\eta_j d\xi_j \\
=&(1-\delta_{n\ell}) \det(E-\Delta-\gamma \operatorname{diag} k')^{nn,\ell\ell} e^{-2\phi_0} \\
=&(1-\delta_{n\ell}) \frac{\det(E-\Delta-\gamma \operatorname{diag} k')^{nn,\ell\ell}}{\det(E-\Delta-\gamma \operatorname{diag} k')^{nn}} \frac{\det(E-\Delta-\gamma \operatorname{diag} k')^{nn}}{\det(E-\Delta-\gamma \operatorname{diag} k')} e^{-2\phi_0} \det(E-\Delta-\gamma \operatorname{diag} k') \\
=&(1-\delta_{n\ell})(E-\Delta^{(n)}-\gamma \operatorname{diag} k')^{-1}_{\ell\ell}(E-\Delta-\gamma \operatorname{diag} k')^{-1}_{nn} e^{-2\phi},
\end{aligned}$$

where $(E-\Delta-\gamma \operatorname{diag} k')^{nn,\ell\ell}$ is the minor obtained from $(E-\Delta-\gamma \operatorname{diag} k')$ by deleting the n^{th} "row", n^{th} "column" and the ℓ^{th} "row", ℓ^{th} "column", (The order is irrelevant.) and

$$\Delta^{(n)}_{ij} = \begin{cases} \Delta_{ij} & i,\, j \in \Lambda\backslash\{n\}, \\ 0 & \text{otherwise}, \end{cases} \tag{2.60}$$

The Kronecker δ function originates from anti-commutativity.

Sustituting (2.57) (2.59) in (2.56), we obtain

(2.61)

$$\begin{aligned}
&\int (k'(\gamma x_\ell \cdot x_\ell) - k'(0))(k'(\gamma x_n \cdot x_n) - k'(0)) e^{-2\Phi} \prod_{j\in\Lambda} \frac{d^2 X_j}{\pi} \\
=&\gamma \int (k'(\gamma x_\ell \cdot x_\ell) - k'(0)) k''(\gamma x_n \cdot x_n)(E-\Delta-\gamma \operatorname{diag} k')^{-1}_{nn} e^{-2\phi} \prod_{j\in\Lambda} \frac{d^2 x_j}{\pi} \\
&+ \gamma \int (k'(\gamma x_n \cdot x_n) - k'(0)) k''(\gamma x_\ell \cdot x_\ell)(E-\Delta-\gamma \operatorname{diag} k')^{-1}_{\ell\ell} e^{-2\phi} \prod_{j\in\Lambda} \frac{d^2 x_j}{\pi} \\
&- \gamma^2(1-\delta_{n\ell}) \int k''(\gamma x_\ell \cdot x_\ell) k''(\gamma x_n \cdot x_n)(E-\Delta^{(n)}-\gamma \operatorname{diag} k')^{-1}_{\ell\ell} \\
&\qquad\qquad (E-\Delta-\gamma \operatorname{diag} k')^{-1}_{nn} e^{-2\phi} \prod_{j\in\Lambda} \frac{d^2 x_j}{\pi} \\
=&\mathbf{O}(\tilde\gamma) \int (k'(\gamma x_\ell \cdot x_\ell) - k'(0)) e^{-2\phi} \prod_{j\in\Lambda} \frac{d^2 x_j}{\pi} \\
&+ \mathbf{O}(\tilde\gamma) \int (k'(\gamma x_n \cdot x_n) - k'(0)) e^{-2\phi} \prod_{j\in\Lambda} \frac{d^2 x_j}{\pi} + \mathbf{O}(\tilde\gamma^2) \\
=&\mathbf{O}_2(\tilde\gamma^2)
\end{aligned}$$

where we used (2.48), (2.37) and (2.54) to reach the last estimate.

Assume (2.34) holds up to m, for $m+1$ we have from Proposition 2.1

$$\begin{aligned}
&\int \prod_{i=1}^{m+1} (k'(\gamma X_{\ell_i} \cdot X_{\ell_i}) - k'(0)) e^{-2\Phi} \prod_{j\in\Lambda} \frac{d^2 X_j}{\pi} \\
=&\int \prod_{i=1}^{m+1} (k'(\gamma x_{\ell_i} \cdot x_{\ell_i}) - k'(0) + \gamma k''(\gamma x_{\ell_i} \cdot x_{\ell_i}) \eta_{\ell_i} \xi_{\ell_i}) e^{-2\Phi} \prod_{j\in\Lambda} \frac{d^2 X_j}{\pi} \\
=&0.
\end{aligned} \tag{2.62}$$

So

$$
\begin{aligned}
(2.63)\qquad & -\int \prod_{i=1}^{m+1} (k'(\gamma x_{\ell_i} \cdot x_{\ell_i}) - k'(0)) e^{-2\Phi} \prod_{j\in\Lambda} \frac{d^2 X_j}{\pi} \\
&= -\int \prod_{i=1}^{m+1} (k'(\gamma x_{\ell_i} \cdot x_{\ell_i}) - k'(0)) e^{-2\phi} \prod_{j\in\Lambda} \frac{d^2 x_j}{\pi}
\end{aligned}
$$

$$
\begin{aligned}
=&\gamma \sum_{\alpha=1}^{m} \int \prod_{i\neq\alpha} (k'(\gamma x_{\ell_i} \cdot x_{\ell_i}) - k'(0)) k''(\gamma x_{\ell_\alpha} \cdot x_{\ell_\alpha}) \eta_{\ell_\alpha} \xi_{\ell_\alpha} e^{-2\Phi} \prod_{j\in\Lambda} \frac{d^2 X_j}{\pi} \\
&+ \gamma^2 \sum_{\alpha=1}^{m} \sum_{\beta=1,\, \beta\neq\alpha}^{m} \int \prod_{i\neq\alpha,\, i\neq\beta} (k'(\gamma x_{\ell_i} \cdot x_{\ell_i}) - k'(0)) k''(\gamma x_{\ell_\alpha} \cdot x_{\ell_\alpha}) \eta_{\ell_\alpha} \xi_{\ell_\alpha} \\
&\qquad k''(\gamma x_{\ell_\beta} \cdot x_{\ell_\beta}) \eta_{\ell_\beta} \xi_{\ell_\beta} e^{-2\Phi} \prod_{j\in\Lambda} \frac{d^2 X_j}{\pi} \\
&+ \cdots + \gamma^{m+1} \int \prod_{i=1}^{m+1} k''(\gamma x_{\ell_i} \cdot x_{\ell_i}) \eta_{\ell_i} \xi_{\ell_i} e^{-2\Phi} \prod_{j\in\Lambda} \frac{d^2 X_j}{\pi}.
\end{aligned}
$$

We make the important observation that in the last equality of (2.63), there are *at most* m factors of $k'(\gamma x_{\ell_i} \cdot x_{\ell_i}) - k'(0)$ under the product sign $\prod$. Combining with the estimate

(2.64)

$$
\begin{aligned}
&\int \xi_{\ell_1} \eta_{\ell_1} \cdots \xi_{\ell_p} \eta_{\ell_p} e^{-2\Phi} \prod_{j\in\Lambda} d\eta_j d\xi_j \\
=&\det(E - \Delta - \gamma \operatorname{diag} k')^{\ell_1\ell_1, \cdots \ell_p\ell_p} e^{-2\phi_0} \\
=&\frac{\det(E - \Delta - \gamma \operatorname{diag} k')^{\ell_1\ell_1, \cdots \ell_p\ell_p}}{\det(E - \Delta - \gamma \operatorname{diag} k')^{\ell_1\ell_1, \cdots \ell_{p-1}\ell_{p-1}}} \frac{\det(E - \Delta - \gamma \operatorname{diag} k')^{\ell_1\ell_1, \cdots \ell_{p-1}\ell_{p-1}}}{\det(E - \Delta - \gamma \operatorname{diag} k')^{\ell_1\ell_1, \cdots \ell_{p-2}\ell_{p-2}}} \\
&\cdots \frac{\det(E - \Delta - \gamma \operatorname{diag} k')^{\ell_1\ell_1}}{\det(E - \Delta - \gamma \operatorname{diag} k')} \det(E - \Delta - \gamma \operatorname{diag} k') e^{-2\phi_0} \\
=&(E - \Delta^{(\ell_1 \cdots \ell_{p-1})} - \gamma \operatorname{diag} k')^{-1}_{\ell_p\ell_p} \cdots (E - \Delta - \gamma \operatorname{diag} k')^{-1}_{\ell_1\ell_1} e^{-2\phi} \\
\leq&(E - \Delta - \gamma \operatorname{diag} k')^{-1}_{\ell_p\ell_p} \cdots (E - \Delta - \gamma \operatorname{diag} k')^{-1}_{\ell_1\ell_1} e^{-2\phi} \\
=&\mathbf{O}(\tilde{\gamma}^p),
\end{aligned}
$$

where $p \leq |\Lambda|$, $\ell_1 \neq \ell_2 \cdots \neq \ell_p$ (otherwise the integral is identically 0), the various notations invoving the superscrips $\ell_1, \cdots, \ell_p$ are just the natural extensions of the corresponding notations used in (2.57), (2.59). (2.64) is the p^{th} repeated version of (2.57), (2.59). We obtain (2.34), by (2.63), (2.64) and induction. □

Proof of Proposition 2.3. Since

$$
(2.65)\qquad \langle v\rangle = \int v dg(v) = k'(0),
$$

the first n terms follow by using (2.34). To estimate the last term, we use the bound

$$
(2.66)\qquad 0 < (E - \Delta - \gamma \operatorname{diag} k')^{-1}_{j_n\nu} < (E - \Delta)^{-1}_{j_n\nu},
$$

and Proposition 2.4, we have

(2.67)

$$\gamma^n \sum_{j_1\cdots j_n\in\Lambda}\int (E-\Delta_\Lambda-\gamma k'(0))^{-1}_{\mu j_1}(k'(\gamma x_{j_1}\cdot x_{j_1})-k'(0))\cdots$$
$$(E-\Delta_\Lambda-\gamma k'(0))^{-1}_{j_{n-1}j_n}(k'(\gamma x_{j_n}\cdot x_{j_n})-k'(0))(E-\Delta_\Lambda-\gamma\,\mathrm{diag}\,k')^{-1}_{j_n\nu}$$
$$e^{-2\phi}\prod_{\ell\in\Lambda}\frac{d^2X_\ell}{\pi}$$
$$\leq\gamma^n[(E-\Delta_\Lambda-\gamma k'(0))^{-n}(E-\Delta_\Lambda)^{-1}](\mu,\nu)\cdot\mathbf{O}_n(\tilde\gamma^n)$$
$$\leq\gamma^n\mathbf{O}_n(\tilde\gamma^n)\|(E-\Delta_\Lambda-\gamma k'(0))^{-n}\|_{\ell^2}\|(E-\Delta_\Lambda)^{-1}\|_{\ell^2}$$
$$=\mathbf{O}_n(\tilde\gamma^n)L^2,$$

where we used

$$\|(E-\Delta_\Lambda-\gamma k'(0))^{-n}\|_{\ell^2}=\mathbf{O}(\gamma^{-1}) \tag{2.68}$$

and

$$\|(E-\Delta_\Lambda)^{-1}\|_{\ell^2}=\mathbf{O}(L^2). \tag{2.69}$$

□

III. The initial estimates for localization

For a fixed Λ, we now transform the decay estimate on $\langle(E-\Delta_\Lambda-\gamma V)^{-1}_{\mu\nu}\rangle$ valid for $E\geq 2d$ in (2.31) into probabilistic decay estimates on $|(E-\Delta_\Lambda-\gamma V)^{-1}_{\mu\nu}|$ valid for E in the interval $I_\delta=[2d-\gamma^{1+\delta},2d]\subset\sigma(H)$, $(\delta>0)$. Note the appearance of the absolute value sign as for $E\in\sigma(H)$, $(E-\Delta_\Lambda-\gamma V)^{-1}_{\mu\nu}$ no longer has a definite sign. The length of the interval $|I_\delta|\to\mathbf{O}(1)\gamma$ as $\delta\to 0$. This is the content of Proposition 3.3, which is the necessary initial estimate in order to use Theorem 4.1 in sect. IV to prove localization. Using Proposition 3.3, we in turn derive in Proposition 3.4 an estimate on $\langle|(E-\Delta_\Lambda-\gamma V)^{-1}_{\mu\nu}|^s\rangle$ $(0<s<1)$ valid in the same interval I_δ. (Note the restriction on s in order to control the poles. See (3.26).) Proposition 3.4 is the necessary initial estimate for Theorem 4.2, in sect. IV, the alternative mechanism to prove localization.

Let

$$L=-[\frac{\log\gamma^p}{\sqrt{\gamma}}],\qquad(\gamma<<1,\,p>0), \tag{3.1}$$

where $[\,]$ denotes the integer part. The RHS of (2.31) can be further resummed by using the resolvent equation, we obtain more compactly,

(3.2)

$$\langle(E-\Delta_\Lambda-\gamma V)^{-1}_{\mu\nu}\rangle=(E-\Delta_\Lambda-\gamma\langle v\rangle-\mathbf{O}_n(\gamma\tilde\gamma))^{-1}_{\mu\nu}+\mathbf{O}_n(\tilde\gamma^n),\qquad E\geq 2d,\,n\in\mathbf{N}.$$

Assume μ, ν are such that $|\mu-\nu|\geq L/m$ for some $m>0$ to be determined later. For a given p, adjust n so that for all such μ, ν, the remainder term is of the same or smaller order as the principal term. Since $\langle v\rangle<0$ for $\operatorname{supp} dg\subseteq(-\infty,0]$,

$$\begin{aligned}&(E-\Delta_\Lambda-\gamma\langle v\rangle-\mathbf{O}_n(\gamma\tilde{\gamma}))^{-1}_{\mu\nu}\\ \leq&Ce^{-c\sqrt{\gamma}|\mu-\nu|}\\ \stackrel{\text{def}}{=}&Ce^{-d(\mu,\nu)},\qquad |\mu-\nu|\geq L/m,\quad E\geq 2d,\end{aligned}\tag{3.3}$$

where C, $c>0$, by using weighted estimates. Therefore n is restricted to be $n>\text{const.}p$ for some const.>0. For such n's, we have from (2.31)

$$\langle(E-\Delta_\Lambda-\gamma V)^{-1}_{\mu\nu}\rangle\leq Ce^{-d(\mu,\nu)},\qquad |\mu-\nu|\geq L/m,\quad E\geq 2d,\tag{3.4}$$

with a slightly larger C.

For fixed μ, ν, since $(E-\Delta_\Lambda-\gamma V)^{-1}_{\mu\nu}>0$ $(E\geq 2d)$, using Tchebychev, we have for all $q>0$

$$(E-\Delta_\Lambda-\gamma V)^{-1}_{\mu\nu}\leq C|\mu-\nu|^q e^{-d(\mu,\nu)},\qquad |\mu-\nu|\geq L/m,\quad E\geq 2d,\tag{3.5}$$

with probability $1-|\mu-\nu|^{-q}$. Since in Λ there are at most $(2L+1)^{2d}$ numbers of pairs of indices, we obtain

Lemma 3.1. *Assume dg satisfies (H1,2). With probability $1-C_d m^q/L^{q-2d}$ $(q>2d)$*

$$\begin{aligned}(E-\Delta_\Lambda-\gamma V)^{-1}_{\mu\nu}\leq&C|\mu-\nu|^q e^{-d(\mu,\nu)}\\ \stackrel{\text{def}}{=}&Ce^{-\tilde{d}(\mu,\nu)}\qquad (E\geq 2d)\end{aligned}\tag{3.6}$$

holds for all μ, $\nu\in\Lambda=[-L,L]^d$, such that $|\mu-\nu|\geq L/m$, where L is as defined in (3.1).

In order for Lemma 3.1 to be useful, we need to choose q satisfying $q<\text{const.}p$. Hence we need $2d<q<\text{const.}p$. Using Lemma 3.1 to obtain similar estimates on $(E-\Delta_\Lambda-\gamma V)^{-1}_{\mu\nu}$ for $E\in\sigma(H)$, we need some apriori estimate on the regularity of eigenvalue spacing. This is usually provided by assuming further regularity on dg. We assume (H3), even though for the results in Theorem 4.3 in sect. IV, this could be weakened to only requiring some Hölder regularity:

$$g[a,b]=\int_a^b dg\leq C|b-a|^\rho\tag{3.7}$$

for some $0<\rho\leq 1$, $C>0$ and all $a<b\in\mathbf{R}$. (See e.g.,[vDK].) However, in order to have Theoorem 5.2 on eigenvalue statistics, which is one of the main applications of the present work, (H3) seems to be needed. See [Mi].

Lemma 3.2. *(Wegner's Lemma) Assume dg satisfies (H3). Then for any finite set $\Lambda\subset\mathbf{Z}^d$, $E\in\mathbf{R}$ and $\epsilon>0$*

$$\mathit{Prob}\{\mathit{dist}(E,\sigma(H_\Lambda))\leq\epsilon\}\leq 2\frac{\sup\tilde{g}}{\gamma}|\Lambda|\epsilon.\tag{3.8}$$

This is a well known lemma, for a proof see [Sp,vDK]. Using Lemma 3.1, the resolvent expansion and Lemma 3.2, we obtain our first initial estimate:

Proposition 3.3. *Assume dg satisfies (H1-3). Let*

$$\Lambda = [-L, L]^d = [\frac{\log \gamma^p}{\sqrt{\gamma}}, -\frac{\log \gamma^p}{\sqrt{\gamma}}] \cap \mathbf{Z}^d \qquad p > 0. \tag{3.9}$$

Fix $\delta > 0$, let $I_\delta = [2d - \gamma^{1+\delta}, 2d]$. There exist C_1, C_2, γ_0, $p_0 > 0$, (C_1, C_2 independent of γ, p), such that for all $0 < \gamma < \gamma_0$, all $p > p_0$, with probability $\geq 1 - L^{-C_1\sqrt{p\delta}}$,

$$|(E - \Delta_\Lambda - \gamma V)^{-1}_{\mu\nu}| \leq \gamma^{C_2\sqrt{p\delta}} \tag{3.10}$$

for all $E \in I_\delta$, all μ, ν such that $|\mu - \nu| \geq L$.

Proof. Set $E_0 = 2d$. For all $E \in I_\delta$, using the resolvent equation we have
(3.11)

$$\begin{aligned}(E - \Delta_\Lambda - \gamma V)^{-1}_{\mu\nu} =&(E_0 - \Delta_\Lambda - \gamma V)^{-1}_{\mu\nu} + (E_0 - E)(E_0 - \Delta_\Lambda - \gamma V)^{-2}_{\mu\nu} \\ &+ (E_0 - E)^2(E_0 - \Delta_\Lambda - \gamma V)^{-3}_{\mu\nu} + \cdots (E_0 - E)^{m-2}(E_0 - \Delta_\Lambda - \gamma V)^{-m+1}_{\mu\nu} \\ &+ (E_0 - E)^{m-1}\sum_{j\in\Lambda}(E_0 - \Delta_\Lambda - \gamma V)^{-m+1}_{\mu j}(E - \Delta_\Lambda - \gamma V)^{-1}_{j\nu},\end{aligned}$$

where m is the same as the m in Lemma 3.1. Writing out in detail, the ℓ^{th} ($1 \leq \ell \leq m-1$) term is
(3.12)

$$\begin{aligned}&(E_0 - E)^{\ell-1}(E_0 - \Delta_\Lambda - \gamma V)^{-\ell}_{\mu\nu} \\ =&(E_0 - E)^{\ell-1}\sum_{j_1\cdots j_\ell\in\Lambda}(E_0 - \Delta_\Lambda - \gamma V)^{-1}_{\mu j_1}(E_0 - \Delta_\Lambda - \gamma V)^{-1}_{\mu j_1 j_2}\cdots(E_0 - \Delta_\Lambda - \gamma V)^{-1}_{j_\ell\nu}.\end{aligned}$$

Combining Corollary 2.6 with Lemma 3.1 and letting
(3.13)

$$\bar{d}(j_1, j_2) \stackrel{\text{def}}{=} \begin{cases} 0 & |j_1 - j_2| < L/m \\ \tilde{d}(j_1, j_2) - (d-2)\log(|j_1 - j_2| + 1), & d \geq 3, \quad |j_1 - j_2| \geq L/m\ , \\ \tilde{d}(j_1, j_2), & d < 3, \quad |j_1 - j_2| \geq L/m, \end{cases}$$

where $\tilde{d}$ is as defined in (3.6). We have the following estimate: with probability $\geq 1 - \frac{C_d m^q}{L^{q-2d}}$ ($2d < q < \text{const.}p$)

$$(E_0 - \Delta_\Lambda - \gamma V)^{-1}_{j_1 j_2} \leq e^{-\bar{d}(j_1,j_2)} \begin{cases} \mathbf{O}(L) & d = 1 \\ \mathbf{O}(\log \frac{L}{|j_1 - j_2| + 1}) & d = 2 \\ \frac{\mathbf{O}_d(1)}{(|j_1 - j_2| + 1)^{d-2}} & d \geq 3, \end{cases} \tag{3.14}$$

where $E_0 = 2d$, holds for *all* j_1, $j_2 \in \Lambda$.

Since $|\mu - \nu| \geq L$ and $1 \leq \ell < m - 1$, inspecting the RHS of (3.12), we see that for all possible configurations of $\{j_1, \cdots j_\ell\} \in \Lambda^\ell$, there always exists s, ($0 \leq s \leq \ell$), such that $|j_s - j_{s+1}| \geq L/m$, where we define $j_0 = \mu$, $j_{\ell+1} = \nu$. Using (3.6) in the RHS of (3.12) and summing over $j_1 \cdots j_\ell \in \Lambda$, we obtain that with probability $\geq 1 - \frac{C_d m^q}{L^{q-2d}}$,

$$
\begin{aligned}
&(E_0 - E)^{\ell-1}(E_0 - \Delta_\Lambda - \gamma V)^{-\ell}_{\mu\nu} \\
\leq&\text{const.}(\gamma^{1+\delta})^{\ell-1}(L^2 \log L)^\ell e^{-\bar{d}(0,|\mu-\nu|/m)} \qquad (3.15)\\
=&\text{const.}\gamma^{\frac{\text{const.}p}{m} - \frac{q}{2} - \frac{d}{2} + 1} \qquad (\gamma << 1),
\end{aligned}
$$

where we used (3.14).

To estimate the remainder in the RHS of (3.11), we use Lemma 3.2 and take $\epsilon = |\Lambda|^{-2\beta-3}$ ($\beta > 0$), we have

$$
\begin{aligned}
&(E_0 - E)^{m-1}|[(E_0 - \Delta_\Lambda - \gamma V)^{-m+1}(E - \Delta_\Lambda - \gamma V)^{-1}]_{\mu\nu}| \\
\leq&(\gamma^{1+\delta})^{m-1}\|(E_0 - \Delta_\Lambda - \gamma V)^{-1}\|^{m-1}_{\ell^2}\|(E - \Delta_\Lambda - \gamma V)^{-1}\|_{\ell^2} \\
\leq&\text{const.}(\gamma^{1+\delta})^{m-1}L^{2(m-1)}|\Lambda|^{2\beta+3} \qquad (3.16)\\
=&\text{const.}\gamma^{\delta(m-1)-(\beta+3/2)d}(|\log \gamma^p|)^{2(m-1)+2\beta d} \\
\leq&\text{const.}\gamma^{\tilde{\delta}(m-1)-(\beta+3/2)d},
\end{aligned}
$$

for γ sufficiently small, some $\tilde{\delta} \in (0, \delta)$, with probability $\geq 1 - \frac{C_d m^q}{L^{q-2d}} - \frac{2 \sup \tilde{g}}{L^{2\beta d}}$ ($\beta > 0$). Combining (3.15) with (3.16), taking p large enough (depending on d), and adjusting m, β and q we obtain the Proposition. □

From Proposition 3.3, we further deduce our second initial estimate:

Proposition 3.4. *Assume dg satisfies (H1-3). Let*

$$
\Lambda = [-L, L]^d = [\frac{\log \gamma^p}{\sqrt{\gamma}}, -\frac{\log \gamma^p}{\sqrt{\gamma}}] \cap \mathbf{Z}^d \qquad p > 0.
$$

Fix $\delta > 0$, let $I_\delta = [2d - \gamma^{1+\delta}, 2d]$. There exist γ_0, $p_0 > 0$, such that for all $0 < \gamma < \gamma_0$, all $p > p_0$, all $s \in (0, 1)$,

$$
\langle |(E - \Delta_\Lambda - \gamma V)^{-1}_{\mu\nu}|^s \rangle \leq \gamma^{C_s\sqrt{p\delta}} \qquad (3.17)
$$

for Lebesgue almost all $E \in I_\delta$, all μ, ν such that $|\mu - \nu| \geq L$.

Proof. We use finite rank perturbation theory as in [AM] to express $(E-\Delta_\Lambda-\gamma V)^{-1}_{\mu\nu}$ in a convenient way, in order to control $(E - \Delta_\Lambda - \gamma V)^{-1}_{\mu\nu}$ on the set where (3.10) is not valid.

Let Π (finite rank) be an orthogonal projection and W (finite rank) be such that

$$
\begin{aligned}
H_\Lambda &= \widetilde{H}_\Lambda + W, \qquad (3.18)\\
W &= \Pi W \Pi,
\end{aligned}
$$

where $H_\Lambda = \Delta_\Lambda + \gamma V$ as defined earlier in (1.3). From the resolvent equation, we have

$$
\begin{aligned}
\widetilde{G}_\Lambda(E) &= (E - \widetilde{H}_\Lambda)^{-1} \qquad (3.19)\\
&= (E - H_\Lambda)^{-1} - (E - \widetilde{H}_\Lambda)^{-1}W(E - H_\Lambda)^{-1}.
\end{aligned}
$$

(3.18) and (3.19) imply

$$\Pi\widetilde{G}_\Lambda(E)\Pi = \Pi G_\Lambda(E)\Pi - [\Pi\widetilde{G}_\Lambda(E)\Pi][\Pi W\Pi][\Pi G_\Lambda(E)\Pi]. \tag{3.20}$$

Multiplying (3.20) from the left and the right by the corresponding inverse, one obtains

$$(\Pi G_\Lambda(E)\Pi)^{-1} = (\Pi\widetilde{G}_\Lambda(E)\Pi)^{-1} - W. \tag{3.21}$$

So

$$(\Pi G_\Lambda\Pi) = \left[(\Pi\widetilde{G}_\Lambda\Pi - W)^{-1}\right]^{-1}. \tag{3.22}$$

Let

$$W = \begin{pmatrix} \gamma v_\mu & 0 \\ 0 & \gamma v_\nu \end{pmatrix} \qquad \mu \neq \nu \tag{3.23}$$

in (3.21), $A = \Pi\widetilde{G}_\Lambda\Pi$, which is a 2×2 matrix. We have the following representation

$$\begin{aligned} G_\Lambda(E,\mu,\nu) &= \left[-\begin{pmatrix} \gamma v_\mu & 0 \\ 0 & \gamma v_\nu \end{pmatrix} + A^{-1}\right]^{-1}_{\mu\nu} \\ &= -\frac{A^{-1}_{\mu\nu}}{\det\left[-\begin{pmatrix} \gamma v_\mu & 0 \\ 0 & \gamma v_\nu \end{pmatrix} + A^{-1}\right]}. \end{aligned} \tag{3.24}$$

Similarly, when $\mu = \nu$,

$$G_\Lambda(E,\nu,\nu) = \frac{1}{[\widetilde{G}_\Lambda(E,\nu,\nu)]^{-1} - \gamma v_\nu}. \tag{3.25}$$

Since we only aim for an estimate in $|\mu-\nu|$, we fix $\mu = 0$. (3.24) can then be written in the form

$$G_\Lambda(E,0,\nu) = \frac{\alpha}{\gamma v_\nu - \beta}, \tag{3.26}$$

where α, β are some functions of $V_{\{\nu\}^C}$ only.

(3.25), (3.26) together with (H3) imply that

$$\mathrm{Prob}(|G_\Lambda(E,\mu,\nu)| \geq \epsilon) \leq \frac{\text{const.}}{\epsilon} \qquad (\text{for all}\, \epsilon > 0). \tag{3.27}$$

(See also [AM].) Note that (3.27) is consistent with Lemma 3.2. It follows from (3.27) that

$$\langle |G_\Lambda(E,\mu,\nu)|^s\rangle \leq C_s\gamma^{-1} \tag{3.28}$$

for all $s \in (0,1)$. Let I be the set on which (3.10) holds and χ_I the characteristic function of I. We then have

$$\begin{aligned} \langle |(E-\Delta_\Lambda-\gamma V)^{-1}_{\mu\nu}|^s\rangle =& \langle |G_\Lambda(E,\mu,\nu)|^s\chi_I\rangle + \langle |G_\Lambda(E,\mu,\nu)|^s\chi_{I^C}\rangle \\ \leq& \gamma^{C_2\sqrt{p\delta}} + \langle |G_\Lambda(E,\mu,\nu)|^{st}\rangle^{1/t}\langle (\chi_{I^C})^{t'}\rangle^{1/t'} \\ \leq& \gamma^{C_s\sqrt{p\delta}}, \end{aligned} \tag{3.29}$$

(assuming large enough $p\delta$), where $1/t + 1/t' = 1$ and we used Proposition 3.3 and (3.28). $\square$

IV. Localization theorem at weak disorder

In this section, we use Lemma 3.2, Proposition 3.3 or alternatively Proposition 3.4 to deduce the nature of the spectrum of $H = \Delta + \gamma V$ at weak disorder: $\gamma << 1$, in the interval $I_\delta = [2d - \gamma^{1+\delta}, 2d]$ ($\delta > 0$) and the exponential decay of eigenfunctions. Our result is summarized in Theorem 4.3. We use the well established Theorem 4.1 or alternatively Theorem 4.2 to prove Theorem 4.3.

To state Theorem 4.1, which is for $H = \Delta + V$ with γ set to be 1, we first define the following notion. Let $m > 0$, $E \in \mathbf{R}$, $\Lambda_L(i) = [-L, L]^d + i$, $i \in \mathbf{Z}^d$, $L \in \mathbf{N}$, a cube centered at i and $\partial\Lambda_L(i) = \{j \,|\, |j - i|_\infty = L,\ j \in \Lambda\}$.

Definition. $\Lambda_L(i)$ is (m, E)-regular (for a fixed potential) if $E \notin \sigma(H_{\Lambda_L(i)})$ and

$$\sum_{j \in \partial\Lambda_L(i)} |G_{\Lambda_L(i)}(E, i, j)| \le e^{-mL}. \tag{4.1}$$

Theorem 4.1. *(*[FS,FMSS,vDK]*) Let I be an interval in $\mathbf{R}$ and $m_0 > 0$. Let d be the spatial dimension. Suppose that for some $L_0 > 0$ we have for all $E \in I$*

(Li) $\mathrm{Prob}\{\Lambda_{L_0(0)}$ *is* (m_0, E) *regular*$\} \ge 1 - 1/(L_0)^p$ *for some $p > d$;*

(Lii) $\mathrm{Prob}\{\mathrm{dist}(E, \sigma(H_{\Lambda_L(0)}) < e^{-L^\beta}\} \le 1/L^q$ *for some $\beta \in (0,1)$, all $L \ge L_0$ and $q > 4p + 6d$.*

For all $m \in (0, m_0)$, there exists $B > 0$ such that if $L_0 > B$, then almost surely $\sigma(H) \cap I$ is pure point and if ψ is a generalized eigenfunction of H associated to $E \in I$, one has

$$\limsup_{\|i\| \to \infty} \frac{\log |\phi(i)|}{\|i\|} \le -m. \tag{4.2}$$

The above theorem was fist proven by Fröhlich, Martinelli, Scoppolla and Spencer in [FMSS], using the crucial estimates supplied by [FS]. Its proof was subsequently simplified by von Dreifus and Klein in [vDK]. Theorem 4.1 is a re-rendition of Theorem 2.1 in [vDK]. Its proof relies crucially on a *multi*-scale analysis, which we do not state here, see Theorem 2.2 of [vDK].

In [AM], Aizenman and Molchanov gave a new proof of the statement of Theorem 4.1, using the conditions (Li,Lii) in a different way than the previous proofs in [FS,FMSS,vDK]. It is simpler by virtue of the fact that it is a *single* scale proof. However that proof is only valid in $\mathbf{Z}^d$ and at high disorder $\gamma >> 1$ or large energy, where the bloc size is effectively one single lattice site ($L_0 = 0$ in Theorem 4.1), while Theorem 4.1 remains valid in $\mathbf{R}^d$ and for any disorder. (It is only the search for (Li), which is treacherous.)

Recently Aizenman, Friedrich, Hundertmark and Shankar in [AFHS] found an important generalization of [AM], which is valid in $\mathbf{Z}^d$ at any disorder. Writing Λ_L for $\Lambda_L(0)$, their result for $H = \Delta + V$ is the following

Theorem 4.2. *(*[AFHS]*) Assume dg has Hölder regularity as defined in (3.7) with exponent $\rho \in (0, 1]$. Let $E \in \mathbf{R}$, $m_0 > 0$, $s \in (0, \rho)$. Suppose that for some $L_0 > 0$*

$$\sum_{j \in \partial\Lambda_{L_0}} \langle |G_{\Lambda_{L_0}}(E, 0, j)|^s \rangle \le e^{-m_0 L_0}. \tag{4.3}$$

Then for all $m \in (0, m_0)$, there exists $D > 0$ so that if $L_0 > D$, then there exists $A \in (0, \infty)$, such that for all $L \geq L_0$,

$$\tag{4.4} \begin{aligned} \langle |G_{\Lambda_L}(E, i, j)|^s \rangle &\leq A e^{-m|i-j|}; \qquad && \textit{for all } i, j \in \Lambda_L \\ \langle |G(E + i\epsilon, i, j)|^s \rangle &\leq A e^{-m|i-j|}, \qquad && \textit{for all } \epsilon > 0, \ \textit{all } i, j \in \mathbf{Z}^d. \end{aligned}$$

The above theorem relies on a crucial decoupling lemma for $E \in \sigma(H)$. For more details, see [AFHS]. See also Lemma 4.1, Theorem 4.2 of [W2] for a related proof in the simpler case $E = 2d$ or $-2d$ and $s = 1$. Using Lemma 3.2, Proposition 3.3 in Theorem 4.1 or alternatively using Proposition 3.4 in Theorem 4.2, we obtain

Theorem 4.3. *Assume dg satisfies (H1-3). For all $\delta > 0$, let $I_\delta = [2d - \gamma^{1+\delta}, 2d]$ if $\operatorname{supp} dg \subseteq (-\infty, 0]$ and $I_\delta = [-2d, -2d + \gamma^{1+\delta}]$ if $\operatorname{supp} dg \subseteq [0, \infty)$. There exist $a > 0$, $\gamma_0 > 0$ (a independent of γ), such that for all $\gamma \in (0, \gamma_0)$, with probability one $\sigma(H) \cap I_\delta$ is pure point, and the eigenfunctions ψ corresponding to eigenvalues in I_δ satisfy*

$$\tag{4.5} \limsup_{\|i\| \to \infty} \frac{\log |\psi(i)|}{\|i\|} \leq -a\sqrt{\gamma}.$$

Proof ♯ 1. Using Proposition 3.3, (Li) of Theorem 4.1 is satisfied for sufficiently small γ, using Lemma 3.2 (Lii) is satisfied, Theorem 4.3 then follows from Theorem 4.1. (The proof of Theorem 4.1 goes through with minor modifications due to the presence of the small parameter γ, see [vDK].) □

Proof ♯ 2. Using Proposition 3.4, we see that (4.3) is satisfied with $m_0 = \mathbf{O}(1)\sqrt{\gamma}$. So (4.4) holds for $E \in I_\delta$. (Similar to the proof of Theorem 4.1, the proof of Theorem 4.2 goes through with the presence of the small parameter γ, see [AFHS].) Since for Lebesgue almost every E in I_δ

$$\tag{4.6} \lim_{\epsilon \to 0} G(E + i\epsilon, i, j) = G(E + i0, i, j)$$

exists for almost every realization of $\{V_i\}_{i \in \mathbf{Z}^d}$, (4.4) implies that for all $m' \in (0, m)$,

$$\tag{4.7} \begin{aligned} &\langle \Big[\sum_{j \in \mathbf{Z}^d} |G(E + i0, 0, j)|^2 \exp\big(\frac{2m'}{s}|j|\big)\Big]^{s/2} \rangle \\ \leq &\langle \sum_{j \in \mathbf{Z}^d} |G(E + i0, 0, j)|^2 \exp(m'|j|) \rangle \\ \leq &\infty. \end{aligned}$$

So the Simon-Wolff criterion, Theorems 8, 9 of [SiW] is satisfied, which implies almost surely pure point spectrum and exponential decay of eigenfunctions. □

V. Poisson statistics for eigenvalue spacing

We now use Proposition 3.4 and Theorem 4.2 to deduce eigenvalue statistics in the interval I_δ. Recall from Theorem 4.3 that with probability one the spectrum of H in I_δ is pure point with exponentially decaying eigenfunctions. We use the theorem proved by Minami in [Mi], which is valid in any dimemsion d. [Mi] builds upon earlier ideas of Molchanov in [Mo2]. The proof in [Mo2] is however only valid in one dimension. In order to state the theorem in [Mi], we recall briefly the notion of point processes and their weak convergence.

Let $\mathcal{M}(\mathbf{R})$ be the space of all non-negative Radon measures on $\mathbf{R}$. We say that a sequence $\{m_n\}$ in $\mathcal{M}(\mathbf{R})$ converges to $m \in \mathcal{M}(\mathbf{R})$ vaguely if

$$\lim_{n\to\infty} \int \phi(x) m_n(dx) = \int \phi(x) m(dx) \tag{5.1}$$

holds for all ϕ in $\mathbf{C}_c^+(\mathbf{R})$, the space of all non-negative continuous functions with compact support.

This defines the vague topology on $\mathcal{M}(\mathbf{R})$. Let further $\mathcal{M}_p(\mathbf{R})$ be the subspace of $\mathcal{M}(\mathbf{R})$ consisting of point measures on $\mathbf{R}$. Each $\xi \in \mathcal{M}_p(\mathbf{R})$ can be written as

$$\xi(dx) = \sum_j c_j \delta_{\xi_j}(dx), \tag{5.2}$$

with a sequence $\{\xi_j\}$ having no finite accumulation points and δ_{ξ_j} is the Dirac measure concentrated at ξ_j. $\mathcal{M}_p(\mathbf{R})$ is closed in $\mathcal{M}(\mathbf{R})$ with respect to the vague topology. (See e.g., Chap. IV of [RS, Vol I].)

A random variable ξ^ω taking values in $\mathcal{M}_p(\mathbf{R})$ is called a point process. We say $\mu(dx)$ is its intensity measure if

$$\mu(A) = \mathbf{E}[\xi^\omega(A)] \tag{5.3}$$

for each Borel set A.

Now consider a sequence $\{\xi_n^\omega\}$ of point processes defined on a probability space $(\Omega, \mathcal{F}, P)$. This sequence converges weakly to a point process $\{\xi^{\bar\omega}\}$ defined on a (possibly larger) probability space $(\bar\Omega, \bar{\mathcal{F}}, \bar P)$ if and only if, by definition, for any bounded continuous function $\Phi(\xi)$ on $\mathcal{M}_p(\mathbf{R})$, one has

$$\lim_{n\to\infty} \int \Phi(\xi_n^\omega) P(d\omega) = \int \Phi(\xi_n^{\bar\omega}) P(d\bar\omega). \tag{5.4}$$

Let $H_\Lambda = \Delta_\Lambda + V$ ($\gamma = 1$). Let $n(E)$ be the density of states defined in (1.12) for Lebesgue a.e. in E. Define the rescaled eigenvalues $\xi_j(\Lambda, E)$ as in (1.13) and the point process $\xi(\Lambda, E)$ as in (1.14). Recall also from sect.I, the definition of a Poisson process in (Pi,Pii). Minami proved

Theorem 5.1. *([Mi]) Assume dg satisfies (H3) and that the density of states $n(E)$ exists at E and is positive. If there exist $s \in (0,1)$, $C > 0$, $m > 0$ and $r > 0$ such that*

$$\langle |G_\Lambda(z,i,j)|^s \rangle \le C e^{-m|i-j|}, \tag{5.5}$$

holds for all hypercubes $\Lambda = [-L, L]^d \subset \mathbf{Z}^d$, $i \in \Lambda$, $j \in \partial\Lambda$ and z such that $\operatorname{Im} z > 0$, $|z - E| < r$, then as $L \to \infty$, the point process $\xi(\Lambda, E)$ defined in (1.14) converges weakly to the Poisson point process ξ with intensity measure $n(E)dx$.

Using Theorem 5.1, we immediately arrive at

Theorem 5.2. *Assume dg satisfies (H1-3). For all $\delta > 0$, let $I_\delta = [2d - \gamma^{1+\delta}, 2d]$ if $\operatorname{supp} dg \subseteq (-\infty, 0]$ and $I_\delta = [-2d, -2d + \gamma^{1+\delta}]$ if $\operatorname{supp} dg \subseteq [0, \infty)$. There exists $\gamma_0 > 0$, such that for all $\gamma \in (0, \gamma_0)$, Lebesgue almost all $E \in \sigma(H) \cap I_\delta$, as $L \to \infty$, the point process $\xi(\Lambda, E)$ defined in (1.14) converges weakly to the Poisson point process ξ defined in (Pi,Pii) with intensity measure $n(E)dx$.*

Proof. Using Proposition 3.4 we see that (4.3) of Theorem 4.2 is satisfied for γ sufficiently small and Lebesgue almost all E in I_δ. So Theorem 4.2 implies that (5.5) of Theorem 5.1 holds. To conclude the proof, it suffices to recall from (1.4) that

$$\sigma(H) = [-2d, 2d] + \gamma \operatorname{supp}\ dg = [-2d, 2d] + \gamma \operatorname{supp}\ \tilde{g}, \tag{5.6}$$

and that the density of states $n(E) > 0$ for Lebesgue almost all $E \in \sigma(H)$. (See e.g., Theorem 9.8 of [CFKS].) □

References

[**Ai**] M. Aizenman, *Localization at weak disorder: some elememtary bounds*, Rev. Math. Phys. **6** (1994), 1163.

[**An**] P. Anderson, *Absence of diffusion in certain random lattices*, Phys.Rev. **109** (1958), 1492.

[**AFHS**] M. Aizenman, R. Friedrich, D. Hundertmark, S. Shankar, *Finite-volume criteria for Anderson localization*, Math-Ph/9910022 v2 (Preprint) (1999).

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

[**BJS**] V. Bach, T. Jecko and J. Sjöstrand, *Correlation asymptotics of classical lattice spin systems with nonconvex Hamilton function at low temperature*, Preprint (1998).

[**Be**] F. A. Berezin, *The Method of Second Quantization*, Academic Press, New York, 1966.

[**BGV**] N. Berline, E. Getzler and M. Vergne, *Heat kernels and Dirac operators*, Springer-Verlag, 1992.

[**BCKP**] A. Bovier, M. Campanino, A. Klein, and F. Perez, *Smoothness of the density of states in the Anderson model at high disorder*, Commun. Math. Phys. **114** (1988), 439-461.

[**CFKS**] H. L. Cycon, R. G. Froese, W. Kirsch and B. Simon, *Schrödinger Operators*, Springer-Verlag, 1987.

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

[**Dy**] F. Dyson, *Statistical theory of energy levels of complex systems, I, II, III*, J. Math. Phys. **3** (1962), 140-175.

[**Ef**] K. B. Effetov, *Supersymmetry and the theory of disordered metals*, Adv. Phys. **32** (1983), 53-127.

[**FKG**] C. M. Fortuin, P. W. Kasteleyn, J. Ginibre, *Correlation inequalities on some partially ordered sets*, Commun. Math. Phys. **22** (1971), 89-103.

[**FMSS**] J. Fröhlich, F. Martinelli, E. Scoppola and T.Spencer, *Constructive proof of localization in Anderson tight binding model*, Commun. Math. Phys. **101** (1985), 21-46.

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

[**Fu**] H. Furstenberg, *Non commuting random products*, Trans. Amer. Soc. **108** (1963), 377-428.

[**GMP**] Ya. Gol'dsheid, S. Molchanov and L. Pastur, *Pure point spectrum of stochastic one dimensional Schrödinger operators*, Func. Anal. Appl **11, 1** (1977).

[**ID**] C. Itzykson and J-M. Drouffe, *Statistical Field Theory, vol 2*, Cambridge University Press, Cambridge, 1989.

[**Jo**] J. Johnsen, *On spectral properties of Witten-Laplacians, their range projections and Brascamp-Lieb inequality*, Aalborg University preprint (1998).

[**K**] A. Klein, *The supersymmetric replica trick and smoothness of the density of states for the random Schrödinger operators*, Proceedings of Symposium in Pure Mathematics **51** (1990).

[**KS**] A. Klein, A. Spies, *Smmothness of the density of states in the Anderson model on a one-dimensional strip*, Ann. of Phys. **183** (1988), 352-398.

[**La**] G. F. Lawler, *Intersection of Random Walks*, Birkhäuser, 1991.

[**Li**] T. M. Liggett, *Interacting Particle Systems*, Springer-Verlag, 1985.

[**McW**] B. McCoy and T. T. Wu, *The Two-Dimensional Ising Model*, Harvard University Press, Cambridge, MA, 1973.

[**Me**] M. L. Mehta, *Random Matrices, 2d ed.*, Academic Press, Boston, 1991.

[**Mi**] N. Minami, *Local fluctuation of the spectrum of a multidimensional Anderson tight binding model*, Commun. Math. Phys. **177** (1996), 709-725.

[**Mo1**] S. A. Molchanov, *The structure of eigenfunctions of one-dimensional unordered structure*, Math. USSR,Izv. **12** (1978), 69-101.

[**Mo2**] S. A. Molchanov, *The local structure of the spectrum of the one-dimensional Schrödinger operator*, Commun. Math. Phys. **78** (1981), 429-446.

[**On**] L. Onsager, *Crystal statistics. I. A two-dimensional model with an order-disorder transition.*, Phys. Rev. (2) **65** (1944), 117-149.

[**PF**] L. Pastur and A. Figotin, *Spectra of Random and Almost Periodic Operators*, Springer, 1992.

[**RS**] M. Reed, B. Simon, *Methods of Modern Mathematical Physics I: Functional Analysis*, Academic Press, 1980.

[**SiW**] B. Simon, T. Wolff, *Singular continuous spectrum under rank one perturbation and localization for random Hamiltoniams*, Commun. Pure. Appl. Math **XXXIX** (1986), 75-90.

[**Sj1**] J. Sjöstrand, *Correlation asymptotics and Witten Laplacians*, Algebra and Analysis **8** (1996).

[**Sj2**] J. Sjöstrand, *Complete asymptotics for correlations of Laplace integrals in the semiclassical limit*, Ecole Polytechnique Preprint (1999).

[**SjW**] J. Sjöstrand and W. M. Wang, *Supersymmetric measures and maximum priciples in the complex domain–decay of Green's functions*, Ann. Scient. Éc. Norm. Sup. **32** (1999).

[**Sp1**] T. Spencer, *The Schrödinger equation with a random potential–a mathematical review*, Les Houches XLIII, K. Osterwalder, R. Stora (eds.) (1984).

[**Sp2**] T. Spencer, *Lifshitz tails and localization*, (unpublished notes) (1993).

[**Sz**] A. S. Sznitman, *Brownian motion, Obstacles and Random media*, Springer Monograph in Mathematics, 1998.

[**V**] T. Voronov, *Geometric Integration Theory on Supermanifolds*, Mathematical Physics Review, USSR Academy of Sciences, Moscow, 1993.

[**W1**] W. M. Wang, *Supersymmetry, Witten complex and asymptotics for directional Lyapunov exponents in* $\mathbf{Z}^d$, Orsay Preprint (1999).

[**W2**] W. M. Wang, *Mean field bounds on Lyapunov exponents in* $\mathbf{Z}^d$ *at the critical energy*, Orsay Preprint (1999).

[**Z**] M. P. W. Zerner, *Directional decay of the Green's function for a random nonnegative potential on* $\mathbf{Z}^d$, The Annals of Applied Probability (1998), 246-280.