Eigenvalue counting inequalities,
with applications to Schrödinger operators

Alexander Elgart[1] and Daniel Schmidt[1]

Abstract. We derive a sufficient condition for a Hermitian $N \times N$ matrix A to have at least m eigenvalues (counting multiplicities) in the interval $(-\epsilon, \epsilon)$. This condition is expressed in terms of the existence of a principal $(N - 2m) \times (N - 2m)$ submatrix of A whose Schur complement in A has at least m eigenvalues in the interval $(-K\epsilon, K\epsilon)$, with an explicit constant K.

We apply this result to a random Schrödinger operator H_ω, obtaining a criterion that allows us to control the probability of having m closely lying eigenvalues for H_ω – a result known as an m-level Wegner estimate. We demonstrate its usefulness by verifying the input condition of our criterion for some physical models. These include the Anderson model and random block operators that arise in the Bogoliubov–de Gennes theory of dirty superconductors.

Contents

1. Introduction

1.1. Small eigenvalues and the Green function. Let A be an invertible Hermitian $N \times N$ matrix with inverse A^{-1}, and let I_N be the $N \times N$ identity matrix. Let $G(x, y)$ denote the matrix element in the (x, y) position of A^{-1}, also known as the Green function of A. Our first objective in this work is to relate information about the small eigenvalues of A to the behavior of $G(x, y)$. Let us denote by $\mathcal{C}_\epsilon(A)$ the number of eigenvalues (counting multiplicities) of A in the interval $I_\epsilon := (-\epsilon, \epsilon)$. As a first step, let us ask the most basic question: does A has at least one eigenvalue in the interval I_ϵ? A well known result in the matrix analysis says that

$$\mathcal{C}_\epsilon(A) > 0 \iff \|A^{-1}\| > \frac{1}{\epsilon}.$$

Since $\|B\|_{\max} \leq \|B\| \leq N\|B\|_{\max}$ for any $N \times N$ matrix B with

$$\|B\|_{\max} = \max_{x,y} |B(x, y)|,$$

we obtain the relations

$$\mathcal{C}_\epsilon(A) > 0 \implies \text{ there exists a pair } \{x, y\} \text{ such that } |G(x, y)| > \frac{1}{N\epsilon}; \quad (1.1a)$$

and

$$|G(x, y)| > \frac{1}{\epsilon} \text{ for some pair } \{x, y\} \implies \mathcal{C}_\epsilon(A) > 0. \quad (1.1b)$$

It is natural to try to quantify these relations further, viz. to detect whether the matrix A has at least m small eigenvalues from the behavior of $G(x, y)$. To this end, we prove the following result.

Theorem 1.1. *Let $A = A^*$ be an $N \times N$ invertible matrix. Let $A[\alpha, \beta]$ denote the submatrix of A with rows indexed by index subset α and columns indexed by index subset β. Consider the following two assertions:*

  I. $\mathcal{C}_\epsilon(A) \geq m$;

  II. *there exist index subsets*

$$\alpha_m = \{i_1, \ldots, i_m\} \quad \text{and} \quad \beta_m = \{j_1, \ldots, j_m\}$$

  of $\{1, \ldots, N\}$, such that

$$A^{-1}[\alpha_m, \beta_m]A^{-1}[\beta_m, \alpha_m] > \frac{K^2}{\epsilon^2} I_N[\alpha_m, \alpha_m] \quad \text{for some } K > 0, \quad (1.2)$$

  where I_N is the $N \times N$ identity matrix.

Then (I) *implies* (II) *with*

$$K = \frac{C_m}{N} \quad and \quad C_m = \frac{1}{m!\, 2^{m-1}} \tag{1.3}$$

Conversely, (II) *with* $K = 1$ *implies* (I).

The constant C_m in (1.3) is not sharp for $m > 1$. However, the dependence on N is optimal (and we will be interested in small m, large N behavior in the application below).

It is often convenient to work with principal submatrices $A^{-1}[\gamma]$ of A^{-1}. One can tailor Theorem 1.1 somewhat differently to accommodate this requirement, at the cost of increasing the cardinality of the corresponding index subsets α_m, β_m. Namely, we have the following result.

Corollary 1.2. *Let* $A = A^*$ *be an* $N \times N$ *invertible matrix. Consider the following two assertions*:

  I. $\mathcal{C}_\epsilon(A) \geq m$;

 II. *there exists an index subset* $\gamma_m = \{i_1, \ldots, i_{2m}\}$ *of* $\{1, \ldots, N\}$ *such that for any subset* $\gamma \supset \gamma_m$ *for which the matrix* $A^{-1}[\gamma]$ *is invertible*

$$\mathcal{C}_{\epsilon/K}((A^{-1}[\gamma])^{-1}) \geq m \quad for\ some\ K > 0. \tag{1.4}$$

Then (I) *implies* (II) *with*

$$K = \frac{C_m}{N} \quad and \quad C_m = \frac{1}{m!\, 2^{m-1}}.$$

Conversely, (II) *with* $K = 1$ *implies* (I).

Remark. The matrix $(A^{-1}[\gamma])^{-1}$ coincides with the Schur complement of $A[\gamma^c]$ in A, see (2.1) below for details. Here $\gamma^c = \{1, 2, \ldots, N\} \setminus \gamma$.

1.2. Application to random Schrödinger operators. In quantum physics, the tight-binding approximation is often used as the prototypical model for the study of electron propagation in solids. In this model, the evolution of the wave function ψ on the d-dimensional lattice $\mathbb{Z}^d$ is given by the Schrödinger equation

$$i\hbar \dot{\psi}_t = H\psi_t; \quad \psi(0) = \psi_0, \tag{1.5}$$

where the self-adjoint Hamiltonian H is a sum of the hopping term H_0 and the potential V, of the form

$$(H\psi)(x) = (H_0\psi)(x) + V(x)\psi(x), \quad x \in \mathbb{Z}^d.$$

In this work we consider the random operators that have this functional form. Let us list few of these.

Anderson model H_A. One of the best-studied models for disordered solids was introduced by P. H. Anderson in [1]. In this model the Hilbert space is $\ell^2(\mathbb{Z}^d)$, the hopping term H_0 is the discrete Laplacian Δ, and the potential V in H above is of the form $V(x) = g v(x)$, where the potentials $v(y)$ are independent random variables. The real parameter g is a coupling constant which describes the strength of the disorder.

Alloy-type Anderson model H_{alloy}. Here the Hilbert space is also $\ell^2(\mathbb{Z}^d)$, the hopping term H_0 is a short range ergodic operator. The value of the potential $V(x)$ at a site $x \in \mathbb{Z}^d$ is generated from independent random variables $\{u(y)\}$ via the transformation

$$V(x) = g \sum_{y \in \mathbf{\Gamma}} a_{x-y} u(y) \,,$$

where the index y takes values in some sub-lattice $\mathbf{\Gamma}$ of $\mathbb{Z}^d$. The Hamiltonian H_A coincides with H_{alloy} provided $H_0 = \Delta$, $\mathbf{\Gamma} = \mathbb{Z}^d$; $a_z = \delta_{|z|}$, where δ_x is Kronecker delta function: $\delta_0 = 1$; $\delta_x = 0$ for $x \neq 0$. In general, the random potential at sites x, y is correlated for this model. As its name suggest, H_{alloy} is used to describe (random) alloys in the tight binding approximation.

Random block operator H_{block}. The Hilbert space is $\ell^2(\mathbb{Z}^d; \mathbb{C}^k) \cong (\ell^2(\mathbb{Z}^d))^k$ (the space of square-summable functions $\psi \colon \mathbb{Z}^d \to \mathbb{C}^k$). The kernel $H_0(x, y)$ of the hopping term is a deterministic, translation invariant $k \times k$ matrix. The random potential $V(x)$ at each site is an independently drawn random $k \times k$ Hermitian matrix multiplied by g.

1.2.1. Previous results. Anderson [1] argued that in the $g \gg 1$ regime, the solution of the initial value problem (1.5) for H_A stays localized in space for all times almost surely if the initial wave packet ψ_0 is localized. Mathematical study of Anderson localization is an active field; we refer readers to the recent reviews [14, 20] on the subject for the detailed bibliography. In this work, we focus our attention on a single aspect of Anderson localization – the so-called m-level Wegner estimate.

Let $|S|$ denote the cardinality of the set S. Let H_A^Λ be a restriction of the operator H_A to a finite box Λ. Then the m-level Wegner estimate is an upper bound on the probability of n eigenvalues being in the same energy interval $I_\epsilon := (E - \epsilon, E + \epsilon)$:

$$\mathbb{P}(\mathcal{C}_\epsilon(H_A^\Lambda - E) \geq m) \leq C_m (|\Lambda| \epsilon)^m,$$

for random variables $v(x)$ with a bounded density. As such, it gives some measure of the correlation between multiple eigenvalues. We will refer to the 1-level bound simply as the Wegner estimate (first established by F. J. Wegner in [22]). It plays the instrumental role in the proof of Anderson localization.

If localization occurs in some energy interval $I \subset \mathbb{R}$, the entire spectrum of H_A in I is pure point. It is then natural to study the distribution of the eigenvalues for H_A^{Λ} in this interval. Physicists expect that there is no energy level repulsion for states in the localized regime: that is, the eigenvalues should be distributed independently on the interval I. The first rigorous result in this direction, namely that the point process associated with the (rescaled) eigenvalues converges to a Poisson process, was obtained by Molchanov [19] in the setting of a one-dimensional continuum.

Minami [18] established the analogous result for H_A under the assumption that the distribution μ of every $v(n)$ has a bounded density. The key component in [18] is the 2-level Wegner estimate, which is consequently known as the Minami estimate.

By now the localization phenomenon for the original Anderson model H_A is well understood. In particular, the general m-level Wegner estimate is known to hold for essentially all distributions μ of the random potential $v(x)$; see [2, 11]. We refer the reader to [5] for the state of the art results concerning eigenvalue counting inequalities for H_A. However, the current understanding of many (in fact almost all) other random models of interest remains partial at best.

The Wegner estimate for a special class of alloy-type Anderson model H_{alloy} was first established by Kirsch in [13]. By now it is known to hold in fair generality (albeit not universally). See the recent preprint [17] for the extensive bibliography on the subject. The Wegner estimate for a random block operator H_{block} – with $V(x) = g v(x) A$ where $v(x)$ are independent random variables and A is a fixed invertible Hermitian matrix – holds in perturbative regimes. That is, it holds near the edges of the spectrum, [4] and in the strong disorder regime $1 \ll g$; see [8]. A weaker bound (weaker in terms of the volume dependence) near the edges of the spectrum was established for Fröhlich model, where the matrix-valued potential is given by $V(x) = g U(x)^* A U(x)$, where A is a fixed self-adjoint $k \times k$ matrix, and the $U(x)$ are independently chosen according to the Haar measure on $\text{SU}(k)$; see [3].

On the other hand, not much progress has been made on extensions of the multi-level Wegner estimate, besides allowing for more general background operators H_0 than the discrete Laplacian. In particular, apart from two special examples below, all previous works require a non-correlated random potential. In [16],

this limitation was partially removed in the continuum one-dimensional setting, allowing for positively correlated randomness. In [6], the authors announced the establishment of the Minami estimate and subsequently Poisson statistics for a general class of positively correlated random potentials. Unfortunately, although [6] contains a new elegant and efficient proof of Minami's estimate for H_A, its extension to the generalized setting has a significant gap, which so far has not been removed. Finally, let us mention the recent result [21], which established the Minami estimate for a special class of weakly correlated randomness for which one can transform the problem to the uncorrelated one.

The reader may wonder about the glaring disparity between the wealth of results on the 1-level Wegner estimate and the scarcity of results for its many-level counterparts. The reason can be traced to the direct (and frequently exploited) link between the former and the underlying Green function given by (1.1). The amenable nature of the Green function then allows one to establish a robust 1-level Wegner estimate in many situations of interest. In the present work, we harness the connection between the many-level Wegner estimate and the Green function given by Theorem 1.1 to establish an m-level Wegner estimate for a certain class of models with correlated randomness. Roughly speaking, our method works if the randomness in the system is sufficiently rich. (We will quantify this statement in the sequel.)

Although in most known applications (such as localization, simplicity of the spectrum, and Poisson statistics of eigenvalues) one is interested in the 1- and 2-level Wegner estimates, it is nonetheless natural from a mathematical perspective to investigate the general many-level case. From a practical perspective, it can yield some insight on the nonlinear Anderson model via multi-state resonance phenomenon, [9].

The blessing and the curse of the existing methods employed in proof of the Minami estimate (with the single exception of [16]) is that the nature of the background operator H_0 plays little if any role in the proofs. It is however clear that in the case of the correlated random potential in H_A one cannot hope to get the Minami estimate without exploiting the structure of H_0. Indeed, consider the one dimensional operator H_{alloy} with $H_0 = 0$, and the random potential at odd sites being i.i.d. random variables, while $v(2n) = v(2n - 1)$. Its spectrum consists of (the closure of) the set of eigenvalues $\{\lambda_n\} = \{v(n)\}$, each one being degenerate. Consequently, even though H_{alloy} in this setting is perfectly localized, the probability of finding two closely lying eigenvalues is equal to 1.

1.2.2. The m-level Wegner estimate for the random block model H_{block}. We will consider the class of random block models H_{block} introduced earlier.

Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a graph with degree at most κ, such that the set of vertices (sites) $\mathcal{V}$ is finite with cardinality N. The main example of this model is the restriction of the lattice $\mathbb{Z}^d$ to the box, but the greater generality does not require additional effort here.

Let $(\Omega, \mathbb{P})$ be a probability space. Let

$$\{V(x) = gA_\omega(x) \colon x \in \mathcal{V}, \quad \omega \in \Omega\}$$

be a collection of independent, identically distributed random $k \times k$ Hermitian matrices.

Basic assumption.   We now state the main technical condition that we will use as an input for our application theorem below.

(A) *For an integer n, let S be a given set of $2nk$ district integers. Then the matrix $A_\omega(x) - a$ is invertible for all $a \in S$ and all $\omega \in \Omega$. Moreover, there exists an $\alpha > 0$ such that, for any integer $a \in S$, any $\epsilon \in [0, 1]$ and arbitrary Hermitian $k \times k$ matrix J the bound*

$$\mathbb{P}(|\det((A_\omega(x) - a)^{-1} + (J + a)^{-1})| \leq \epsilon) \leq K\epsilon^\alpha \tag{1.6}$$

holds.

It guarantees that the randomness in the system is rich enough to imply the result below (Theorem 1.3). At the first glance, a more natural condition should concern the properties of the matrix $A_\omega(x) + J$ as it is the correct functional form of the corresponding Schur complement of H_ω (see Section 3 for details). However, this turns out to be an unsuitable choice because of the absence of an *a priori* bound on the norm of the background operator J (which encodes the information about the environment of the x-block in H_ω). On the other hand, for a sufficiently large set of numbers $\{a_i\}$ one can ensure that regardless of the norm of J, one of the matrices $\{(J + a_i)^{-1}\}$ is bounded in norm by 1 (see Proposition 3.2). We then exploit the fact that matrices $(A-a)^{-1}+(J+a)^{-1}$, which appears in (1.6), and $A + J$ are related:

$$(A + J)^{-1} = (A - a)^{-1} - (A - a)^{-1}((A - a)^{-1} + (J + a)^{-1})^{-1}(A - a)^{-1}, \tag{1.7}$$

provided that $A-a$ and $J+a$ are invertible. One can readily verify (1.7), which is in fact a particular case of Woodbury's matrix identity, by multiplying both sides by $A + J$.

We now introduce our single-particle Hamiltonian. Namely, let $H_\omega(g)$ be a random block operator H_{block} acting on $\ell^2(\mathcal{V}; \mathbb{C}^k)$ (the space of square-summable functions $\psi : \mathcal{V} \to \mathbb{C}^k$) as

$$(H_\omega(g)\psi)(x) = (H_0\psi)(x) + gA_\omega(x)\psi(x), \qquad (1.8)$$

where $g > 0$ is a coupling constant, H_0 is an arbitrary deterministic self-adjoint operator on $\ell^2(\mathcal{V}; \mathbb{C}^k)$, and $A_\omega(x)$ is an independently drawn random $k \times k$ Hermitian matrix as above. We use the notation $H_\omega(g)$ instead of H_{block} to stress the random nature of this operator as well as the dependence on the parameter g.

Theorem 1.3. *Assume* (**A**). *Then*

  I. *for any $E \in \mathbb{R}$ the operator $H_\omega(g) - E$ is almost surely invertible;*

 II. *Moreover, there exist $\epsilon_0 > 0$ and $C > 0$, which depend only on k, m, α, for which we have*

$$\mathbb{P}(\mathcal{C}_\epsilon(H_\omega(g) - E) \geq m) \leq C\,|\ln(N\epsilon/g)(N\epsilon/g)^\alpha|^m \qquad (1.9)$$

for any $E \in \mathbb{R}$, for any $\epsilon \in [0, \epsilon_0]$ and for all $m \leq n$. In the $m = 1$ case we can improve the above bound to

$$\mathbb{P}(\mathcal{C}_\epsilon(H_g - E) \geq 1) \leq C(N\epsilon/g)^\alpha. \qquad (1.10)$$

1.2.3. Examples

● **Anderson model H_A.** As we mentioned earlier, the nontrivial Minami estimate is well understood only for the original Anderson model among all alloy-type models. It is therefore a litmus test to verify Assumption (**A**) for H_A.

Theorem 1.4. *Suppose that the distribution μ of the $v(x)$ variables in H_A is compactly supported on the interval $I = [-b, b]$ for some $b > 0$ and is β-regular, i.e. for any Lebesgue measurable $S \subset I$ we have*

$$\mu(S) \leq C|S|^\beta.$$

Then Assumption (**A**) *holds with $\alpha = \beta$.*

Our approach to the Minami estimate is also meaningful for the Γ-trimmed Anderson model introduced in [7], near the edges of the spectrum, in the sense that Assumption (**A**) can be verified for it.

● **Fröhlich model and alloy type Anderson model H_{alloy}.**   Assumption (**A**) is either not satisfied for a single site x or is satisfied with a power α which is too small to make the result meaningful. However, the close inspection of the proof of Theorem 1.3 shows that the matrix J that appears in Assumption (**A**) is not required to be completely arbitrary. In fact, the relevant matrices J carry the structure of the Schrödinger operator (with arbitrary boundary conditions). It seems plausible (and is on our to-do list) that Assumption (**A**) can be verified for such J and sets of sites that include x and its neighbors.

●   The third model arises from the study of dirty superconductors via the Bogoliubov–de Gennes equation. After a suitable change of the coordinate basis, the Bogoliubov–de Gennes (BdG) model above can be described in terms of the operator defined in (1.8), with $A_\omega(x) = \sigma_\omega(x) \in M_{2\times 2}$, where $\sigma_\omega(x)$ is a random Pauli matrix of the form

$$\sigma_\omega(x) = \begin{bmatrix} u_x & v_x \\ v_x & -u_x \end{bmatrix}, \qquad \text{where } u_x \text{ and } v_x \text{ are random variables.} \qquad (1.11)$$

The m-Wegner estimate for these models has only been established for $m = 1$ case, for a restricted class of joint distributions of u, v variables (absolutely continuous and with support in a half plane), and for a specific background operator H_0 in [15, 10]. We establish the robust Wegner and Minami estimates for this model.

Theorem 1.5. *Let each $A_\omega(x)$ be given by* (1.11). *Suppose that the joint distribution μ of the u, v variables is supported on a unit disc O and is β-regular, i.e. for any Lebesgue - measurable $S \subset O$ we have*

$$\mu(S) \le C |S|^\beta.$$

Then Assumption (**A**) *holds with $\alpha = \beta$.*

1.3. Paper's organization.   We prove our main abstract result, Theorem 1.1, along with its corollary, in Section 2. We prove our result on eigenvalue estimates, Theorem 1.3, in Section 3. We consider the implication of the latter result for the random block operators in Section 4. These proofs depend on a number of auxiliary results, which we prove in Section 5.

2. Proof of Theorem 1.1 and Corollary 1.2

2.1. Notation. Let n be a positive integer, and let α and β be index sets, i.e. subsets of $\{1, 2, \ldots, n\}$. We denote the cardinality of an index set by $|\alpha|$ and its complement by $\alpha^c = \{1, 2, \ldots, n\} \setminus \alpha$. For an $n \times n$ matrix A, let $A[\alpha, \beta]$ denote the submatrix of A with rows indexed by α and columns indexed by β, both of which are thought of as increasing, ordered sequences, so that the rows and columns of the submatrix appear in their natural order. We will write $A[\alpha]$ for $A[\alpha, \alpha]$. If $|\alpha| = |\beta|$ and if $A[\alpha, \beta]$ is nonsingular, we denote by $A/A[\alpha, \beta]$ the Schur complement of $A[\alpha, \beta]$ in A, [23]:

$$A/A[\alpha, \beta] = A[\alpha^c, \beta^c] - A[\alpha^c, \beta](A[\alpha, \beta])^{-1}A[\alpha, \beta^c]. \qquad (2.1)$$

We will frequently use Schur's complementation and its consequences in this work; we refer the reader to the comprehensive book [23] on this topic.

For a Hermitian matrix A and a positive number a we will write

$$\mathcal{B}_a(A) := |\sigma(A) \cap [a, \infty)|.$$

Let $P_\epsilon(A)$ denote the spectral projection of the Hermitian matrix A onto the interval $(-\epsilon, \epsilon)$ for $\epsilon > 0$.

2.2. Proof of Theorem 1.1. Suppose that (I) holds. We will use the following assertion.

Proposition 2.1. *Let A be an $N \times N$ positive definite matrix, and suppose that $\mathcal{B}_a(A) = k$ for some $a > 0$. Then there exists an index subset $\alpha_k = \{i_1, i_2, \ldots, i_k\}$ of $\{1, \ldots, N\}$ such that*

$$A[\alpha_k] \geq \frac{a}{k!\, 2^{k-1} N} I_N[\alpha_k].$$

By Proposition 2.1, there exists α_m such that

$$P_\epsilon(A)[\alpha_m] \geq \frac{C_m}{N} I_N[\alpha_m].$$

with

$$C_m = \frac{1}{k!\, 2^{k-1}}.$$

Combining this bound with

$$A^{-2} > \frac{1}{\epsilon^2} P_\epsilon(A),$$

we obtain

$$A^{-2}[\alpha_m] > \frac{C_m}{N\epsilon^2} I_N[\alpha_m].$$

Since $\sigma(TT^*) \setminus \{0\} = \sigma(T^*T) \setminus \{0\}$ for any operator T, we deduce from the previous equation (with $T = I_N[\alpha_m, \alpha_N]A^{-1}$) that there exists an orthogonal projection Q of rank m such that

$$A^{-1}I_N[\alpha_m]A^{-1} > \frac{C_m}{N\epsilon^2} Q.$$

Applying now Proposition 2.1 once again, we conclude that there exists β_m such that (1.2) holds with K given by (1.3).

Conversely, suppose that (1.2) holds with $K = 1$. Since

$$A^{-2}[\alpha_m] \geq A^{-1}[\alpha_m, \beta_m]A^{-1}[\beta_m, \alpha_m],$$

the assertion follows from the Cauchy interlacing theorem for the Hermitian matrix A^{-2} and its principal submatrix $A^{-2}[\alpha_m]$. $\qquad\square$

Proof of Proposition 2.1. The proof will proceed by induction in k. If $k = 1$, the result follows from the fact that A is positive, so tr $A = \sum_{\lambda \in \sigma(A)} \lambda \geq a$. Since the trace is at least a, there exists a diagonal entry which is greater than or equal to $\frac{a}{N}$.

Suppose we have established the induction hypothesis for $k = K$. We want to verify the induction step, i.e. the case $k = K + 1$. To this end, choose the index i_1 so that $A_{i_1 i_1} \geq A_{ii}$ for all i. Without loss of generality, let us assume that $i_1 = 1$. Then A is of the block form

$$A = \begin{bmatrix} A_{11} & u \\ u^* & B \end{bmatrix}.$$

Consider now the matrix $D = A/A_{11}$. It is positive definite by the Schur complement condition for positive definiteness (as A is positive definite). Also, the matrix

$$F = \begin{bmatrix} A_{11} & u \\ u^* & B - D \end{bmatrix}$$

is rank one (since $F/A_{11} = 0$), so by the rank one perturbation theory,

$$\mathcal{B}_a(A - F) \geq K.$$

But $\mathcal{B}_a(A - F) = \mathcal{B}_a(D)$. Using the induction hypothesis, we conclude that there exists an index set $\alpha_K = \{i_2, \ldots, i_K\}$ with $1 \notin \alpha_K$ such that

$$D[\alpha_K] \geq \frac{a}{K! \, 2^{K-1} N} I_N[\alpha_K].$$

The induction step (with $\alpha_{K+1} = \alpha_K \cup \{1\}$) now follows from the following assertion. □

Lemma 2.2. *Let A be an $l \times l$ positive definite matrix of the block form*

$$A = \begin{bmatrix} A_{11} & u \\ u^* & B \end{bmatrix}. \tag{2.2}$$

Suppose that, in addition, $A_{11} \geq A_{ii}$ for all $i \in \{1, \ldots, l\}$, and $A/A_{11} \geq a$ for some $a > 0$. Then $A \geq \frac{a}{2l}$.

Proof of Lemma 2.2. To show that $A - \frac{a}{2l} \geq 0$ it suffices to check (by the Schur complement condition for positive definiteness) that

$$A_{11} - \frac{a}{2l} \geq 0; \quad \left(A - \frac{a}{2l}\right)/\left(A_{11} - \frac{a}{2l}\right) \geq 0. \tag{2.3}$$

Since $A/A_{11} \geq a$, we have $A_{ii} \geq a$ for all $i \geq 2$ as a is positive, so by assumption of the lemma $A_{11} \geq a$ as well (and hence we have established the first bound in (2.3)). Next we write

$$
\begin{aligned}
\left(A - \frac{a}{2l}\right)/\left(A_{11} - \frac{a}{2l}\right) &= B - \frac{a}{2l} - \frac{u^*u}{A_{11} - \frac{a}{2l}} \\
&= \left(B - \frac{u^*u}{A_{11}}\right) - \frac{a}{2l} - \frac{au^*u}{2lA_{11}(A_{11} - a/2l)} \\
&\geq a - \frac{a}{2l} - a\frac{u^*u}{l(A_{11})^2}.
\end{aligned}
\tag{2.4}
$$

Now observe that since A is positive, the contraction $A[\{1, i + 1\}]$ is also positive for all i, and in particular $\det A[\{1, i + 1\}] = A_{11}B_{ii} - |u_i|^2 \geq 0$. But $A_{11} \geq B_{ii}$ for all i, hence $|u_i|^2/(A_{11})^2 \leq 1$. We therefore can estimate

$$\|u^*u\| = \|u\|^2 \leq (l - 1)(A_{11})^2.$$

Substitution of this estimate into (2.4) yields the second bound in (2.3). □

2.3. Proof of Corollary 1.2. We first observe that if sets of indices α, β satisfy $\alpha \subset \beta$, then $A^{-1}[\alpha]$ is a principal submatrix of $A^{-1}[\beta]$, and we have

$$\mathcal{C}_\epsilon((A^{-1}[\beta])^{-1}) \geq \mathcal{C}_\epsilon((A^{-1}[\alpha])^{-1}) \tag{2.5}$$

by the Cauchy interlacing theorem, provided the matrices $A^{-1}[\alpha], A^{-1}[\beta]$ are invertible. Therefore, it suffices to establish the corollary for the smallest set $\gamma_{\min}$

that contains γ_m and for which $A^{-1}[\gamma_{\min}]$ is invertible. Without loss of generality, we will assume that $\gamma_{\min} = \gamma_m$.

Suppose that (I) holds. Then Assertion (II) of Theorem 1.1 holds with K given by (1.3). Construct now the set $\gamma_m = \alpha_m \cup \beta_m$ with α_m, β_m from Assertion (II) of Theorem 1.1. Let us consider the matrix

$$B := (A^{-1}[\gamma_m])^{-1}.$$

Since $A^{-1}[\alpha_m, \beta_m]$ is a submatrix of $A^{-1}[\gamma_m]$ we see that the condition (II) of Theorem 1.1 is fulfilled for B. Hence we can apply Theorem 1.1 to B to conclude that $\mathcal{C}_{\epsilon/K}(B) \geq m$.

Conversely, suppose that (1.4) holds with $K = 1$. Then (I) holds as well, as follows from (2.5) with $\alpha = \gamma_m$, $\beta = \{1, \ldots, N\}$. $\qquad\square$

3. Proof of Theorem 1.3

We first observe that by scaling it suffices to prove the result for the $g = 1$ case. We will use the shorthand notation H_ω instead of $H_\omega(1)$ in the sequel.

Next, we prove the first assertion of the theorem, using induction in N. To initiate the induction, we consider the case $N = 1$, so that $H_\omega = A_\omega(x) + K$, where K is a deterministic Hermitian matrix. It follows from Assumption (A) and (1.7) that $H_\omega - E$ is invertible almost surely.

Suppose now that the induction hypothesis holds, i.e. the matrix $H_\omega - E$ is almost surely invertible for $N \leq M$ and all E. We want to establish the induction step ($N = M + 1$ case). To this end, let $\widehat{\mathcal{V}}$ be any subset of $\mathcal{V}$ of cardinality M, and let $\widehat{H}_\omega$ be a restriction of H_ω to $\widehat{\mathcal{V}}$. By the induction hypothesis, $\widehat{H}_\omega - E$ is invertible almost surely for all E. Let us consider some configuration ω for which $\widehat{H}_\omega - E$ is invertible. Then $H_\omega - E$ is invertible if and only if the Schur complement of $\widehat{H}_\omega - E$ in $H_\omega - E$, i.e. $(H_\omega - E)/(\widehat{H}_\omega - E)$, is invertible, [23]. But $(H_\omega - E)/(\widehat{H}_\omega - E)$ is a Hermitian $k \times k$ matrix of the form $A_\omega(x) + J$, where $\{x\} = \mathcal{V} \setminus \widehat{\mathcal{V}}$, and J is a matrix independent of the randomness in $A_\omega(x)$. It follows by the same argument as in the $N = 1$ case that $(H_\omega - E)/(\widehat{H}_\omega - E)$ is invertible for almost all values of the randomness in $A_\omega(x)$.

We now prove the second assertion of the theorem. We will only consider configurations ω in Ω such that $H_\omega - E$ is invertible (for the remaining set of configurations has measure zero by the first assertion).

For the random operator T_ω, let $\mathcal{E}_\epsilon(T_\omega)$ be the event $\{\omega\colon \mathcal{C}_\epsilon(T_\omega) \geq m\}$. With this notation, we wish to estimate the size of the set $\mathcal{E}_\epsilon(H_\omega - E)$. If we enumerate the vertices $v \in \mathcal{V}$, we can think of H_ω as a $kN \times kN$ Hermitian matrix with a block form, i.e. the indices $\{lk - k + 1, \ldots, lk\}$ correspond to the vertex l in $\mathcal{V}$, with $l = 1, \ldots, N$.

Size reduction.   We first reduce the dimensionality of the original problem using Corollary 1.2. This assertion gives us the existence of the index subset γ_m with $|\gamma_m| = 2m$ such that inclusion

$$\mathcal{E}_\epsilon(H_\omega - E) \subset \mathcal{E}_{\epsilon/K}(((H_\omega - E)^{-1}[\gamma])^{-1}) \tag{3.1}$$

holds for any index set $\gamma \supset \gamma_m$ for which $(H_\omega - E)^{-1}[\gamma]$ is invertible, with K given by (1.3). (To be precise, the matrix size N in that corollary gets replaced by kN.)

In general, the submatrix $(H_\omega - E)^{-1}[\gamma]$ can be a complicated object, so it is not immediately clear that such a reduction is helpful. However, if the set γ happens to consist of the indices that agree with the block structure of H_ω, something interesting happens. More precisely, suppose that $i \in \gamma \implies (j \in \gamma$ for any j with $\lfloor j/k \rfloor = \lfloor i/k \rfloor)$, where $\lfloor \cdot \rfloor$ is the floor function. In this case we can associate γ with a subset $\mathcal{V}'$ of the original vertex set $\mathcal{V}$. Then the submatrix $((H_\omega - E)^{-1}[\gamma])^{-1}$ retains the same block form as H_ω, in the following sense: if we go back to the vertex representation for $((H_\omega - E)^{-1}[\gamma])^{-1}$ (which is possible due to the special form of the set γ), then for any $\psi \in \ell^2(\mathcal{V}'; \mathbb{C}^k)$ and any $x \in \mathcal{V}'$ we have

$$(((H_\omega - E)^{-1}[\gamma])^{-1}\psi)(x) = (T_0\psi)(x) + A_\omega(x)\psi(x). \tag{3.2}$$

This can be seen from the fact that the matrix $((H_\omega - E)^{-1}[\gamma])^{-1}$ coincides with the Schur complement of $(H_\omega - E)[\gamma^c]$ in $H_\omega - E$,

$$((H_\omega - E)^{-1}[\gamma])^{-1} = (H_\omega - E)/(H_\omega - E)[\gamma^c].$$

It is important to note that the operator T_0 in (3.2) is independent of the randomness associated with matrices $\{A_\omega(x)\}_{x \in \mathcal{V}'}$ (though it does depend on the other random variables). We also note that the matrix $((H_\omega - E)^{-1}[\gamma])$ is almost surely invertible (as follows from the first part of the theorem).

Combining these observations, we conclude that it is beneficial (and sufficient) to consider the sets γ in (3.1) that respect the block structure of H_ω and therefore contain up to $2km$ indices (i.e. up to $2m$ vertices in $\mathcal{V}'$, as in Corollary 1.2, and exactly k indices per vertex, to preserve blocks). Thus, we have obtained the following intermediate result

Lemma 3.1. *Suppose that the second assertion of Theorem* 1.3 *holds for all* $N \leq 2m$. *Then it holds for any* N.

Norm reduction.   The deterministic part of H_ω – namely the operator H_0 – can be arbitrarily large in norm (even if $\|H_0\| \leq C$ for the original H_ω, the size reduction process indicated above creates a new background operator T_0 with uncontrollable norm). Our next step in the proof will require that the background operator is bounded in norm by a constant, say by $1/2$. We achieve this by means of the following transformation.

Proposition 3.2. *Let* $B_{1,2}$ *be a pair of Hermitian* $L \times L$ *matrices with* $\|B_1\| \leq 1$. *Consider the matrices*

$$B = B_1 + B_2, \quad \hat{B} = (B_1 - aI_L)^{-1} + (B_2 + aI_L)^{-1}$$

where $a \in \mathbb{R}$. *Then there exists an integer* $a \in [-L - 3, -3] \cup [3, L + 3]$ (*which depends on* B_2 *but not on* B_1) *and* $\epsilon_0 > 0$ (*which depends only on* L) *such that, for any* $\epsilon < \epsilon_0$,

$$\max(\|(B_1 - aI_L)^{-1}\|, \|(B_2 + aI_L)^{-1}\|) \leq \frac{1}{2} \tag{3.3}$$

and

$$\mathcal{C}_{\epsilon/(225L^4)}(\hat{B}) \leq \mathcal{C}_\epsilon(B) \leq \mathcal{C}_{7L^2\epsilon}(\hat{B}). \tag{3.4}$$

We will apply this proposition to the operator $H_\omega - E$ by choosing $B_2 = H_0 - E$, $B_1 = H_\omega - H_0$. By the hypothesis of Theorem 1.3, the assumptions of Proposition 3.2 are satisfied, with $L = kN$. Combining this observation with the size reduction, we obtain the second intermediate result.

Lemma 3.3. *Assume* (A). *Let* $\hat{H}_\omega$ *be an operator acting on* $\ell^2(\mathcal{V}; \mathbb{C}^k)$ *as*

$$(\hat{H}_\omega \psi)(x) = (H_0 \psi)(x) + (A_\omega(x) - a)^{-1} \psi(x). \tag{3.5}$$

Suppose that $\|H_0\| \leq 1/2$. *If there exist* $\epsilon_0 > 0$ *and* $b > 0$ (*which depend on* k, m, α) *so that for all integers* $a \in [-km - 3, -3] \cup [3, km + 3]$ *and all* $N \leq 2m$ *the bound*

$$\mathbb{P}(\mathcal{C}_{7k^2m^2\epsilon}(\hat{H}_\omega) \geq m) \leq b|\ln \epsilon|^m \epsilon^\alpha$$

holds uniformly in H_0 *and* $\epsilon < \epsilon_0$, *then the second assertion of Theorem* 1.3 *holds.*

Reduction to the determinant.   Suppose that $\mathcal{C}_{7k^2m^2\epsilon}(\hat{H}_\omega) \geq m$. Since by construction $\|\hat{H}_\omega\| \leq 1$, the operator $\hat{H}_\omega$ can have no more than $kN - m$ large eigenvalues, and each of these can have an absolute value no larger than 1. As a result, we obtain the bound:

$$|\det \hat{H}_\omega| \leq (7k^2m^2\epsilon)^m. \tag{3.6}$$

We may now employ the following lemma to calculate the probability of the aforementioned bound on the determinant. Its proof can be found in Section 5.

Lemma 3.4. *Assume* (A). *Let $\hat{H}_\omega$ be as in* (3.5), *and let $\hat{\mathcal{E}}_\delta$ be the event*

$$\hat{\mathcal{E}}_\delta = \{\omega \in \Omega \colon |\det \hat{H}_\omega| \leq \delta\}.$$

Let

$$\delta_0 := \exp(2K\alpha^{1+1/N}).$$

Then for any $\delta \in [0, \delta_0]$ we have

$$\mathbb{P}(\hat{\mathcal{E}}_\delta) \leq (2K\alpha)^N \ln^N(\delta^{-1})\delta^\alpha. \tag{3.7}$$

Using this result in conjunction with (3.6) we obtain that there exist $\epsilon_0 > 0$ and $b > 0$ that depend on k, m, α so that

$$\mathbb{P}(\mathcal{C}_{7k^2m^2\epsilon}(\hat{H}_\omega) \geq m) \leq b|\ln \epsilon|^m \epsilon^{m\alpha}, \tag{3.8}$$

for $N \leq m$ and for $\epsilon < \epsilon_0$. The combination of Lemma 3.3 and (3.8) yields (1.9).

Improvement on the Wegner bound.   We want to improve on this bound for the special case that $m = 1$. In this case we need to verify the (improved) input for Lemma 3.1 for $N = 1$ and $N = 2$. In the former case, the bound (1.10) follows from Assumption (A) and Proposition 3.2 (where we choose $B_2 = H_0-E$, $B_1 = H_\omega - H_0$). So for the rest of the argument we will assume that $N = 2$.

Let $\mathcal{E}_\epsilon$, $\mathcal{S}_\epsilon$ be the events

$$\mathcal{E}_\epsilon = \{\omega \colon \mathcal{C}_\epsilon(H_\omega - E) \geq 1\}$$

and

$$\mathcal{S}_\epsilon = \{\omega \colon \mathcal{C}_{\epsilon^{2/3}}(H_\omega - E) \geq 2\}.$$

We first observe that it follows from (1.9) (which we already established earlier) that

$$\mathbb{P}(\mathcal{E}_\epsilon \cap \mathcal{S}_\epsilon) \leq \mathbb{P}(\mathcal{S}_\epsilon) \leq C|\ln(\epsilon^{2\alpha/3})\epsilon^{2\alpha/3}|^2 \leq C\epsilon^\alpha$$

for ϵ sufficiently small. Therefore, to get (1.10) it suffices to show that $\mathbb{P}(\mathcal{E}_\epsilon \smallsetminus \mathcal{S}_\epsilon) \leq C\epsilon^\alpha$. To this end, suppose that $\omega \in \mathcal{E}_\epsilon \smallsetminus \mathcal{S}_\epsilon$. Then

$$(H_\omega - E + \epsilon)^{-1} + 2\epsilon^{-2/3} > 0; \quad \|(H_\omega - E + \epsilon)^{-1} + 2\epsilon^{-2/3}\| \geq \frac{1}{2\epsilon}. \quad (3.9)$$

If $\mathcal{V} = \{x, y\}$, let us denote by P_x (P_y) the rank k projection onto the site x (accordingly y). The positivity of the left-hand side can be exploited by means of Lemma 3.5 below with choices $P_1 = P_x$, $P_2 = P_y$.

Lemma 3.5. *Let $A > 0$, and let $P_{1,2}$ be orthogonal projections that satisfy $P_1 P_2 = 0$. Let $P = P_1 + P_2$. Then we have*

$$\|PAP\| \leq 2\max(\|P_1 A P_1\|, \|P_2 A P_2\|). \quad (3.10)$$

Using (3.9) and (3.10), we infer that $\omega \in \mathcal{R}_\epsilon$ (and thus $\mathcal{E}_\epsilon \smallsetminus \mathcal{S}_\epsilon \subset \mathcal{R}_\epsilon$), where

$$\mathcal{R}_\epsilon = \Big\{ \omega : \max_{i=x,y} \big(\|P_i (H_\omega - E + \epsilon)^{-1} P_i\| + 2\epsilon^{-2/3} \big) \geq \frac{1}{4\epsilon} \Big\}.$$

But

$$P_x (H_\omega - E + \epsilon)^{-1} P_x = ((H_\omega - E + \epsilon)/P_x (H_\omega - E + \epsilon) P_x)^{-1}$$
$$= (A_\omega(x) + J)^{-1}$$

by the block inversion formula. Here the operator J depends on $A_\omega(y)$ but not on $A_\omega(x)$. Hence we can deduce from Assumption (**A**) and Proposition 3.2 that

$$\mathbb{P}\Big(\|P_x (H_\omega - E + \epsilon)^{-1} P_x\| \geq \frac{1}{5\epsilon} \Big) \leq \tilde{C}\epsilon,$$

with $\tilde{C}$ that depends on k, α but not on ϵ. The same bound holds with P_x replaced by P_y. Hence we infer that for ϵ small enough

$$\mathbb{P}(\mathcal{R}_\epsilon) \leq C\epsilon^\alpha,$$

and the result follows. $\qquad\square$

Proof of Lemma 3.5. Let $A_1 = P_1 A P_1$, $A_2 = P_2 A P_2$, and $A_{12} = P_1 A P_2$. Then by Schur complement condition for positive definiteness

$$A_2 > 0 \quad \text{and} \quad A_1 \geq A_{12} A_2^{-1} A_{21}.$$

Since A_2 is positive, $A_2^{-1} \geq 1/\|A_2\|$, hence

$$A_1 \geq A_{12} A_{21}/\|A_2\|,$$

and so

$$\|A_1\|\|A_2\| \geq \|A_{12}\|^2,$$

where in the last step we have used $A_{12} = A_{12}^*$. Since

$$\|PAP\| \leq \max(\|A_1\| + \|A_{12}\|, \|A_2\| + \|A_{21}\|),$$

the result follows.   □

4. Proof of Theorems 1.4 and 1.5

4.1. Proof of Theorem 1.4. Let a be an integer that satisfies $a - b \geq 2$, and let j be arbitrary fixed real number. Then if $\epsilon \in [0, 1/(2a)]$, the inequality

$$\left| \frac{1}{v_x - a} + \frac{1}{j + a} \right| < \epsilon \tag{4.1}$$

for v_x has solutions in I only if $0 < j + a < 1$. Since equation

$$\left| \frac{1}{v_x - a} + \frac{1}{j + a} \right| = \epsilon$$

defines the pair of points

$$v_x = a + \frac{1}{(j + a)^{-1} \pm \epsilon},$$

the set of v_x for which (4.1) holds is the interval $\widehat{I}$ of length

$$|\widehat{I}| = \frac{1}{(j + a)^{-1} - \epsilon} - \frac{1}{(j + a)^{-1} + \epsilon} = \frac{2\epsilon}{(j + a)^{-2} - \epsilon^2} < 4\epsilon,$$

where in the last step we used $0 < a + j < 1, \epsilon < 1/(2a) < 1/4$. Hence

$$\mathbb{P}\left(\left| \frac{1}{v_x - a} + \frac{1}{j + a} \right| < \epsilon \right) \leq C\epsilon^\beta$$

by the hypothesis of the theorem.   □

4.2. Proof of Theorem 1.5. We first establish the bounds

$$\mathbb{P}(|\det(\sigma_\omega(x) + J)| \leq \epsilon) \leq C\epsilon^\beta \tag{4.2}$$

and

$$\mathbb{P}(\mathcal{C}_\epsilon(\sigma_\omega(x) + J) \neq 0) \leq C\epsilon^\beta \tag{4.3}$$

for $\epsilon \in [0, 1]$. Indeed, note that $\det(\sigma_\omega(x) + J) = c^2 - (u_x - a)^2 - (v_x - b)^2$ with some constants a, b, c originating from J. Therefore the set $|\det(\sigma_\omega(x) + J)| \leq \epsilon$ is an intersection I of the disc O with the annulus centered at a, b and with radii $R_- = \sqrt{\max(c^2 - \epsilon, 0)}$, $R_+ = \sqrt{c^2 + \epsilon}$. The area of this set therefore cannot exceed $\pi(R_+^2 - R_-^2) \leq 2\pi\epsilon$ and (4.2) follows. To establish (4.3), we note that

$$\mathbb{P}(\mathcal{C}_\epsilon(\sigma_\omega(x) + J) \neq 0) = \mathbb{P}(\|(\sigma_\omega(x) + J)^{-1}\| \geq 1/\epsilon).$$

The value of $\|(\sigma_\omega(x) + J)^{-1}\|$ however can be evaluated explicitly and is given by

$$\|(\sigma_\omega(x) + J)^{-1}\| = \left||c| - \sqrt{(u_x - a)^2 + (v_x - b)^2}\right|^{-1},$$

with the same constants a, b, c as before. Hence the set of the points in O that satisfy $\|(\sigma_\omega(x) + J)^{-1}\| \geq 1/\epsilon$ is an intersection $\hat{I}$ of the disc O with the annulus centered at a, b and with radii $R_- = ||c| - \epsilon|$, $R_+ = |c| + \epsilon$. The area of $\hat{I}$ cannot exceed the circumference of the unit circle times the maximal thickness 2ϵ of the annulus, so $|\hat{I}| \leq 4\pi\epsilon$, and (4.3) follows.

The assertion of the theorem follows now from Lemma 4.1 below (whose proof can be found in Section 5) and bounds (4.2) and (4.3).

$\square$

Lemma 4.1. *Let A and J be Hermitian $k \times k$ matrices, and let a be some real number that satisfies $|a| \geq 2$. If $\|A\| \leq 1$ and*

$$|\det((A - a)^{-1} + (J + a)^{-1})| \leq \frac{1}{16|a|} \left\{ \frac{|a| - 1}{2(|a| + 1)^2} \right\}^{k-1}$$

then we have

$$(2(|a| + 1)^2)^{-k} |\det(A + J)| \leq |\det((A - a)^{-1} + (J + a)^{-1})|$$

or

$$\frac{1}{16|a|} \{2(|a| + 1)^2 \|(A + J)^{-1}\|\}^{1-k} \leq |\det((A - a)^{-1} + (J + a)^{-1})|.$$

5. Proofs

Proof of Proposition 3.2. If $|a| \geq 3$, then since $\|B_1\| \leq 1$, we have

$$\nu := \|(B_1 - a)^{-1}\|, \quad (L + 4)^{-1} \leq \nu \leq 1/2. \tag{5.1}$$

Since B_2 is $L \times L$, it has at most L distinct eigenvalues. On the other hand, for every set S of real numbers with $|S| = L$ there exists an integer $a \in [-L-3, -3] \cup [3, L+3]$ so that $\text{dist}(S, -a) \geq 2$, hence we can choose a that satisfies (3.3).

With this choice of a, consider the block matrix W of the form

$$W = \begin{bmatrix} (B_2 + aI_n)^{-1} & (B_1 - aI_n)^{-1} \\ (B_1 - aI_n)^{-1} & -(B_1 - aI_n)^{-1} \end{bmatrix}.$$

Note that the Schur complements to the upper and lower diagonal blocks are

$$-W/W_{11} = (B_1 - a)^{-1} + (B_1 - a)^{-1}(B_2 + a)(B_1 - a)^{-1};$$

and

$$W/W_{22} = (B_2 + a)^{-1} + (B_1 - a)^{-1}.$$

Let

$$T = (B_1 - a)^{-1}/\nu, \tag{5.2}$$

where ν is given by (5.1).

In what follows, we will need two lemmas.

Lemma 5.1.

(1) *Let $D = D^*, \tilde{D} = \tilde{D}^* \in M_{n,n}$. Then*

$$\mathcal{C}_\epsilon(D) \leq \mathcal{C}_{2\epsilon}(\tilde{D}),$$

provided $\|D - \tilde{D}\| \leq \epsilon$.

(2) *$\mathcal{C}_\epsilon(A) \leq \mathcal{C}_\epsilon(BAB)$ whenever*

$$A = A^*, \quad B = B^*, \quad \|B\| \leq 1. \tag{5.3}$$

Lemma 5.2. *Suppose $D = D^* \in M_{n,n}$ is of the form*

$$D = \begin{bmatrix} A & V \\ V^* & B \end{bmatrix},$$

with $A \in M_{k,k}$, $B \in M_{m,m}$, $\|V\| \leq 1/2$, and

$$\mathcal{C}_{2\epsilon}(B) = 0. \tag{5.4}$$

Then

$$\mathcal{C}_\epsilon(D/B) \leq \mathcal{C}_\epsilon(D); \tag{5.5}$$

and

$$\mathcal{C}_\epsilon(D) \leq \mathcal{C}_{\beta\epsilon}(D/B), \tag{5.6}$$

with $\beta = 2(\|B^{-1}\| + 1)^2$.

Armed with these results, we can infer that

$$
\begin{aligned}
\mathcal{C}_\epsilon(B) &= \mathcal{C}_{\nu^2\epsilon}(\nu^2 B) \\
&\le \mathcal{C}_{\nu^2\epsilon}(\nu^2 T B T) \\
&= \mathcal{C}_{\nu^2\epsilon}(W/W_{11}) \\
&\le \mathcal{C}_{\nu^2\epsilon}(W) \\
&\le \mathcal{C}_{\beta\nu^2\epsilon}(W/W_{22}) \\
&= \mathcal{C}_{\beta\nu^2\epsilon}(\widehat{B}),
\end{aligned}
\tag{5.7}
$$

where in the second step we have used Lemma 5.1 and in the remaining steps we have used Lemma 5.2. Here

$$
\beta = 2(\|(W_{22})^{-1}\| + 1)^2 \le 2(L+5)^2 \le 25L^2
$$

for $L \ge 2$. (It is straightforward to check that the relation $\mathcal{C}_\epsilon(B) \le \mathcal{C}_{\beta\nu^2\epsilon}(\widehat{B})$ holds for $L = 1$.) Plugging in the upper bounds for ν, β we get the second inequality in (3.4):

$$
\mathcal{C}_\epsilon(B) \le \mathcal{C}_{7L^2\epsilon}(\widehat{B}).
$$

On the other hand, let

$$
U = \kappa \begin{bmatrix} B_2 + aI_n & B_1 - aI_n \\ B_1 - aI_n & -B_1 + aI_n \end{bmatrix}; \quad \kappa = \frac{1}{2\|B_1 - a\|}.
$$

Then

$$
U/U_{22} = \kappa B;
$$

and

$$
-U/U_{11} = \kappa(B_1 - a + (B_1 - a)(B_2 + a)^{-1}(B_1 - a)).
$$

Similarly to (5.7), we obtain

$$
\begin{aligned}
\mathcal{C}_{\nu^2\epsilon/(25L^2)}(\widehat{B}) &= \mathcal{C}_{\epsilon/(25L^2)}(\kappa^{-1}T(U/U_{11})T) \\
&\le \mathcal{C}_{\kappa\epsilon/(25L^2)}(U/U_{11}) \\
&\le \mathcal{C}_{\kappa\epsilon/(25L^2)}(U) \\
&\le \mathcal{C}_{\kappa\beta\epsilon/(25L^2)}(U/U_{22}) \\
&\le \mathcal{C}_\epsilon(B),
\end{aligned}
$$

with T given by (5.2) and

$$\beta = 2(\|(U_{22})^{-1}\| + 1)^2 \le 2(\|B_1 - a\| + 1)^2 \le 25L^2.$$

Since

$$\frac{\nu^2 \epsilon}{25L^2} \ge \frac{\epsilon}{25L^2(L+4)^2} \ge \frac{\epsilon}{225L^4},$$

the first inequality in (3.4) follows.   □

Proof of Lemma 5.1. For part (1), we use the Weyl's theorem, cf. Theorem 4.3.1 in [12], which states that if

$$\sigma(A) = \{\lambda_i(A)\}_{i=1}^n,$$

$$\sigma(B) = \{\lambda_i(B)\}_{i=1}^n,$$

$$\sigma(A + B) = \{\lambda_i(A + B)\}_{i=1}^n,$$

for Hermitian A, B, with the eigenvalues arranged in increasing order, then

$$\lambda_k(A) + \lambda_1(B) \le \lambda_k(A + B) \le \lambda_k(A) + \lambda_n(B), \quad k = 1, ..., n.$$

Therefore, every number $\lambda_k(A + B)$ which lies in the interval $[-\epsilon, \epsilon]$ can be matched with $\lambda_k(A) \in [-2\epsilon, 2\epsilon]$, provided that $\|B\| \le \epsilon$.

For part (2), observe that there exists a Hermitian matrix $\widehat{A}$ such that

(1) $\|\widehat{A} - A\| \le \epsilon$,

(2) $\mathrm{nul}(\widehat{A}) = \mathcal{C}_\epsilon(A)$, and

(3) $\widehat{A}$ has no non-zero eigenvalues in the interval $(-\epsilon, \epsilon)$.

Then Sylvester's law of inertia implies that $\mathrm{nul}(\widehat{A}) \le \mathrm{nul}(B\widehat{A}B)$ (with equality in the case of nonsingular B). Since $\|BAB - B\widehat{A}B\| \le \epsilon$ we can use Weyl's theorem again to conclude that

$$\mathcal{C}_\epsilon(BAB) \ge \mathrm{nul}(B\widehat{A}B) \ge \mathrm{nul}(\widehat{A}) = \mathcal{C}_\epsilon(A).$$ □

Proof of Lemma 5.2. The relation (5.5) follows from the interlacing theorem for inverses of Hermitian matrices – see Lemma 2.3 in [23], which is itself a simple consequence of the the Schur complement formula and Cauchy interlacing theorem for Hermitian matrices.

To prove (5.6) note that there exists a matrix

$$\widehat{D} := \begin{bmatrix} \widehat{A} & \widehat{V} \\ \widehat{V}^* & \widehat{B} \end{bmatrix}$$

such that

(1) $\|\widehat{D} - D\| \le \epsilon$,

(2) nul $\widehat{D} = \mathcal{C}_\epsilon(D)$, and

(3) $\widehat{D}$ has no non-zero eigenvalues in the interval $[-\epsilon, \epsilon]$,

where nul $\widehat{D}$ is the multiplicity of the zero eigenvalue of $\widehat{D}$, and equals zero if this eigenvalue is absent. One can readily prove the existence of $\widehat{D}$ by diagonalizing D and replacing all eigenvalues less than or equal to ϵ with zeros. Using the Haynsworth inertia additivity formula, we get

$$\text{nul } \widehat{D} = \text{nul } \widehat{B} + \text{nul}(\widehat{D}/\widehat{B}).$$

Observe that the condition (1) above implies $\|\widehat{B} - B\| \le \epsilon$. We can therefore infer from (5.4) and Lemma 5.1 that

$$\mathcal{C}_\epsilon(\widehat{B}) = 0. \tag{5.8}$$

As a result we obtain the equality

$$\mathcal{C}_\epsilon(D) = \text{nul } \widehat{D} = \text{nul}(\widehat{D}/\widehat{B}). \tag{5.9}$$

Note now that

$$\begin{aligned}
&\|\widehat{V}\widehat{B}^{-1}\widehat{V}^* - VB^{-1}V^*\| \\
&\le \|(\widehat{V} - V)\| \cdot \|\widehat{B}^{-1}\widehat{V}^*\| + \|V\| \cdot \|\widehat{B}^{-1} - B^{-1}\| \cdot \|\widehat{V}^*\| \\
&\qquad + \|V\widehat{B}^{-1}\| \cdot \|(\widehat{V}^* - V^*)\| \\
&\le \epsilon\|\widehat{B}^{-1}\|\left(\frac{1}{2} + \epsilon\right) + \frac{1}{2}\|\widehat{B}^{-1} - B^{-1}\|\left(\frac{1}{2} + \epsilon\right) + \frac{1}{2}\|\widehat{B}^{-1}\|\epsilon \\
&= (\epsilon + \epsilon^2)\|\widehat{B}^{-1}\| + \left(\frac{1}{2}\epsilon + \frac{1}{4}\right)\|\widehat{B}^{-1} - B^{-1}\| \\
&\le \frac{3}{2}\epsilon\|\widehat{B}^{-1}\| + \frac{1}{2}\|\widehat{B}^{-1} - B^{-1}\|.
\end{aligned}$$

(We have assumed that $\epsilon \le \frac{1}{2}$.) Using the first resolvent identity, we get the bound

$$\|\widehat{B}^{-1} - B^{-1}\| = \|\widehat{B}^{-1}(\widehat{B} - B)B^{-1}\| \le \epsilon\|B^{-1}\| \cdot \|\widehat{B}^{-1}\|, \tag{5.10}$$

which in turn implies the estimate

$$\|\widehat{B}^{-1}\| \le \|B^{-1}\| + \epsilon\|B^{-1}\| \cdot \|\widehat{B}^{-1}\| \le 2\|B^{-1}\|,$$

where we have used (5.8) in the last step. Inserting the last inequality into the right-hand side of (5.10), we finally obtain

$$\|\widehat{B}^{-1} - B^{-1}\| \le 2\epsilon\|B^{-1}\|^2.$$

As a result, we arrive at

$$\|\widehat{V}\widehat{B}^{-1}\widehat{V}^* - VB^{-1}V^*\| \le \frac{3}{2}\epsilon\|B^{-1}\| + \epsilon\|B^{-1}\|^2,$$

hence

$$\|\widehat{D}/\widehat{B} - D/B\| \le \epsilon\Big(1 + \frac{3}{2}\|B^{-1}\| + \|B^{-1}\|^2\Big) < \epsilon(\|B^{-1}\| + 1)^2 =: \frac{\epsilon\beta}{2}.$$

Consequently, we get

$$\mathcal{C}_\epsilon(D) = \mathrm{nul}(\widehat{D}/\widehat{B}) \le \mathcal{C}_{\frac{\beta\epsilon}{2}}(\widehat{D}/\widehat{B}) \le \mathcal{C}_{\beta\epsilon}(D/B),$$

where we have used (5.9) in the first step and Lemma 5.1 in the last one. □

Proof of Lemma 3.4. We use induction in N. For $N = 1$ the result follows from Assumption (**A**). Suppose that (3.7) holds for $|\mathcal{V}| = N$. We wish to establish the induction step, i.e. (3.7) for $|\mathcal{V}| = N + 1$. We can evaluate $\det \widehat{H}_\omega$ using the Schur determinant formula. Namely, for $x \in \mathcal{V}$ let us denote by $\widehat{H}_\omega^{(x)}$ the restriction of $\widehat{H}_\omega$ to the site x. Then

$$\det \widehat{H}_\omega = \det \widehat{H}_\omega^{(x)} \det(\widehat{H}_\omega/\widehat{H}_\omega^{(x)})$$

by Schur's determinant formula. Both determinants on the right-hand side are random, but the first one depends only on randomness associated with $A_\omega(x)$, a fact which we will exploit momentarily. We note now that the Schur complement $\widehat{H}_\omega/\widehat{H}_\omega^{(x)}$ is by itself also of the form (3.5) (with $\mathcal{V}$ replaced by $\mathcal{V} \setminus \{x\}$). Note that the H_0 term in $\widehat{H}_\omega/\widehat{H}_\omega^{(x)}$ might depend on $A_\omega(x)$, but not on the other random variables $\{A_\omega(y)\}$. By the induction hypothesis, we have

$$\mathbb{P}(|\det(\widehat{H}_\omega/\widehat{H}_\omega^{(x)})| \le r) \le (2K\alpha)^N \ln^N(r^{-1})r^\alpha, \quad r \in [0, 1]. \tag{5.11}$$

Let $S := \{\omega : |\det \widehat{H}_\omega| \le \epsilon\}$, and let

$$F_\omega = |\det \widehat{H}_\omega^{(x)}|, \quad G_\omega = |\det(\widehat{H}_\omega/\widehat{H}_\omega^{(x)})|.$$

We set $Q := \{\omega\colon \min(F_\omega, G_\omega) \le \epsilon\}$, then by Assumption (**A**) and the induction hypothesis

$$\mathbb{P}(Q) \le (K + (2K\alpha)^N \ln^N(\epsilon^{-1}))\epsilon^\alpha. \tag{5.12}$$

On the other hand, we have

$$\chi(S \smallsetminus Q) = \int_\epsilon^1 \chi(sG_\omega \le \epsilon)\delta(F_\omega - s)ds.$$

Taking expectations on both sides and using (5.11), we obtain

$$\mathbb{E}\chi(S \smallsetminus Q) \le \mathbb{E}\int_\epsilon^1 ds\, \delta(F_\omega - s)\mathbb{E}(\chi(sG_\omega \le \epsilon) \mid A_\omega(x))$$

$$\le (2K\alpha)^N \epsilon^\alpha \mathbb{E}\int_\epsilon^1 \frac{\ln^N\left(\frac{s}{\epsilon}\right)\delta(F_\omega - s)}{s^\alpha}ds$$

$$= (2K\alpha)^N \epsilon^\alpha \mathbb{E}\frac{\ln^N(\epsilon^{-1}F_\omega)\chi(1 > F_\omega > \epsilon)}{(F_\omega)^\alpha} \tag{5.13}$$

$$\le (2K\alpha)^N \epsilon^\alpha \ln^N(\epsilon^{-1})\mathbb{E}\frac{\chi(1 > F_\omega > \epsilon)}{(F_\omega)^\alpha}.$$

Using now Assumption (**A**) and the layer cake representation, we get

$$\mathbb{E}\frac{\chi(1 > F_\omega > \epsilon)}{(F_\omega)^\alpha} = \int_1^{\epsilon^{-\alpha}} \mathbb{P}((F_\omega)^{-\alpha} \ge t)dt$$

$$\le K\int_1^{\epsilon^{-\alpha}} \frac{1}{t}dt \tag{5.14}$$

$$= K\alpha \ln(\epsilon^{-1}).$$

Combination of (5.12), (5.13), and (5.14) yields the induction step. $\qquad\square$

Proof of Lemma 4.1. We have

$$|\det(A + J)^{-1}|$$

$$= |\det(A - a)^{-1}||\det(J + a)^{-1}||\det((A - a)^{-1} + (J + a)^{-1})^{-1}|.$$

Suppose first that $\mathcal{C}_{16|a|}(((A - a)^{-1} + (J + a)^{-1})^{-1}) = 0$. According to (5.3)

$$\mathcal{C}_{16/|a|}((A - a)^{-1}((A - a)^{-1} + (J + a)^{-1})^{-1}(A - a)^{-1})$$

$$\le \mathcal{C}_{16/|a|}((|a|/2)^{-2}((A - a)^{-1} + (J + a)^{-1})^{-1})$$

$$= \mathcal{C}_{4|a|}(((A - a)^{-1} + (J + a)^{-1})^{-1})$$

$$= 0,$$

where we have used $\|A - a\| \geq |a|/2$. Since

$$\|(A - a)^{-1}\| \leq (|a| - 1)^{-1},$$

we can use (1.7) to decompose

$$\det(A + J)^{-1}$$
$$= \det((A - a)^{-1}((A - a)^{-1} + (J + a)^{-1})^{-1}(A - a)^{-1}) \det(H - I),$$

with

$$H = ((A - a)^{-1} + (J + a)^{-1})(A - a).$$

It follows that $\|H\| \leq 1/2$, and consequently $|\det(H - I)| \geq 2^k$. On the other hand, $|\det(A - a)| \leq (|a| + 1)^k$, and we can conclude that

$$|\det(A + J)^{-1}| \geq (2(|a| + 1)^2)^{-k}|\det((A - a)^{-1} + (J + a)^{-1})^{-1}|, \quad (5.15)$$

whenever $\mathcal{C}_{16|a|}(((A - a)^{-1} + (J + a)^{-1})^{-1}) = 0$.

On the other hand, if $\mathcal{C}_{16|a|}(((A - a)^{-1} + (J + a)^{-1})^{-1}) \neq 0$, then

$$\|((A - a)^{-1} + (J + a)^{-1})^{-1}\|^{k-1} \geq \frac{1}{16|a|}|\det((A - a)^{-1} + (J + a)^{-1})^{-1}|.$$

Hence

$$\|(A - a)^{-1}((A - a)^{-1} + (J + a)^{-1})^{-1}(A - a)^{-1}\|$$
$$\geq (|a| + 1)^{-2}\|((A - a)^{-1} + (J + a)^{-1})^{-1}\|$$
$$\geq (|a| + 1)^{-2}\left\{\frac{1}{16|a|}|\det((A - a)^{-1} + (J + a)^{-1})^{-1}|\right\}^{1/(k-1)}.$$

Using (1.7), we conclude that

$$\|(A + J)^{-1}\|$$
$$\geq (|a| + 1)^{-2}\left\{\frac{1}{16|a|}|\det((A - a)^{-1} + (J + a)^{-1})^{-1}|\right\}^{1/(k-1)} - (|a| - 1)^{-1}$$
$$\geq \frac{1}{2}(|a| + 1)^{-2}\left\{\frac{1}{16|a|}|\det((A - a)^{-1} + (J + a)^{-1})^{-1}|\right\}^{1/(k-1)},$$

$$(5.16)$$

if $\mathcal{C}_{16|a|}(((A - a)^{-1} + (J + a)^{-1})^{-1}) \neq 0$, whenever (2.2) holds.

Combining (5.15) and (5.16) we establish the assertion. □

References

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

[2] J. Bellissard, P. Hislop, and G. Stolz, Correlation estimates in the Anderson model. *J. Stat. Phys.* **129** (2007), no. 4, 649–662. MR 2360226 Zbl 1131.82016

[3] J. Bourgain, An approach to Wegner's estimate using subharmonicity. *J. Stat. Phys.* **134** (2009), no. 5-6, 969–978. MR 2518977 Zbl 1193.82018

[4] Z. Cao and A. Elgart, The weak localization for the alloy-type Anderson model on a cubic lattice. *J. Stat. Phys.* **148** (2012), no. 6, 1006–1039. MR 2975520 Zbl 1260.82045

[5] J.-M. Combes, F. Germinet, and A. Klein, Generalized eigenvalue-counting estimates for the Anderson model. *J. Stat. Phys.* **135** (2009), 201–216. MR 2505733 Zbl 1168.82016

[6] J.-M. Combes, F. Germinet, and A. Klein, Poisson statistics for eigenvalues of continuum random Schrödinger operators. *Anal. PDE* **3** (2010), no. 1, 49–80. MR 2663411 Zbl 1227.82034

[7] A. Elgart and A. Klein, Ground state energy of trimmed discrete Schrödinger operators and localization for trimmed Anderson models. *J. Spectr. Theory* **4** (2014), no. 2, 391–413. MR 3232816 Zbl 1304.82039

[8] A. Elgart, M. Shamis and S. Sodin, Localisation for non-monotone Schrödinger operators. *J. Eur. Math. Soc.* (*JEMS*) **16** (2014), no. 5, 909–924. MR 3210957 Zbl 1296.82031

[9] S. Fishman, Y. Krivolapov, and A. Soffer, Perturbation theory for the nonlinear Schrödinger equation with a random potential. *Nonlinearity 22* (*2009), no. 12, 2861–2887.* MR 2557451 Zbl 1204.37074

[10] M. Gebert and P. Müller, Localization for random block operators. In M Demuth and W. Kirsch (eds.), *Mathematical physics, spectral theory and stochastic analysis.* Operator Theory: Advances and Applications, 232. Advances in Partial Differential Equations (Basel). Birkhäuser/Springer Basel AG, Basel, 2013, 229–246. MR 3077279 MR 3087286 (collection) Zbl 1282.47060 Zbl 1264.00036 (collection)

[11] G. M. Graf and A. Vaghi, A remark on the estimate of a determinant by Minami. *Lett. Math. Phys.* **79** (2007), no. 1, 17–22. MR 2290333

[12] R. A. Horn and C. R. Johnson, *Matrix analysis,* Cambridge University Press, Cambridge, 1985. MR 0832183 Zbl 0576.15001

[13] W. Kirsch, Wegner estimates and Anderson localization for alloy-type potentials. *Math. Z.* **221** (1996), no. 3, 507–512. MR 1381595 Zbl 0863.35069

[14] W. Kirsch, An invitation to random Schrödinger operators. In M. Disertori, W. Kirsch, A. Klein, F. Klopp, Frédéric, and V. Rivasseau, *Random Schrödinger operators.* Panoramas et Synthèses, 25. Société Mathématique de France, Paris, 2008, 1–119. MR 2509110 MR 2516524 (collection) Zbl 1162.82004 Zbl 1151.35002 (collection)

[15] W. Kirsch, B. Metzger, and P. Müller, Random block operators. *J. Stat. Phys.* **143** (2011), no. 6, 1035–1054. MR 2813784 Zbl 1229.82099

[16] F. Klopp, Inverse tunneling estimates and applications to the study of spectral statistics of random operators on the real line. *J. Reine Angew. Math.* **690** (2014), 79–113. MR 3200336 Zbl 1293.35337

[17] K. Leonhardt, N. Peyerimhoff, M. Tautenhahn, and I. Veselić, Wegner estimate and localization for alloy-type models with sign-changing exponentially decaying single-site potentials. Preprint 2013. arXiv:1309.0109 [math.AP]

[18] N. Minami, Local fluctuation of the spectrum of a multidimensional Anderson tight binding model. *Comm. Math. Phys.* **177** (1996), no. 3, 709–725. MR 1385082 Zbl 0851.60100

[19] S. Molchanov, The local structure of the spectrum of the one-dimensional Schrödinger operator. *Comm. Math. Phys.* **78** (1980/81), no. 3, 429–446. MR 0603503 Zbl 0449.60049

[20] G. Stolz, An introduction to the mathematics of Anderson localization. In R. Sims and D. Ueltschi (eds.), *Entropy and the quantum* II. Lecture notes from the 2nd Arizona School of Analysis with Applications held at the University of Arizona, Tucson, AZ, March 15–19, 2010. Contemporary Mathematics, 552. American Mathematical Society, Providence, RI, 2011,71–108. MR 2868042 MR 2868059 (collection) Zbl 1244.82031 Zbl 1223.82001 (collection)

[21] M. Tautenhahn and I. Veselić, Minami's Estimate: beyond rank one perturbation and monotonicity. *Ann. Henri Poincaré* **15** (2014), no. 4, 737–754. MR 3177919 Zbl 1287.81049

[22] F. Wegner, Bounds on the density of states in disordered systems. *Z. Phys. B* **44** (1981), no. 1-2, 9–15. MR 0639135

[23] F. Zhang (ed.), *The Schur complement and its applications.* Numerical Methods and Algorithms, 4. Springer, New York, 2005. MR 2160825 Zbl 1075.15002

Alexander Elgart, Department of Mathematics, Virginia Tech, Blacksburg, VA, 24061, U.S.A.

e-mail: aelgart@vt.edu

Daniel Schmidt Department of Mathematics, Virginia Tech, Blacksburg, VA, 24061, U.S.A.

e-mail: dfs@vt.edu