

# Eigenvalue order statistics for random Schrödinger operators with doubly-exponential tails

Marek Biskup<sup>1,2</sup>, Wolfgang König<sup>3,4\*</sup>

<sup>1</sup> Department of Mathematics, UCLA, Los Angeles, California, USA

<sup>2</sup> School of Economics, University of South Bohemia, České Budějovice, Czech Republic

<sup>3</sup> Weierstraß-Institut für Angewandte Analysis und Stochastik, Berlin, Germany

<sup>4</sup> Institut für Mathematik, Technische Universität Berlin, Berlin, Germany

Version: April 21, 2015

**Abstract:** We consider random Schrödinger operators of the form  $\Delta + \xi$ , where  $\Delta$  is the lattice Laplacian on  $\mathbb{Z}^d$  and  $\xi$  is an i.i.d. random field, and study the extreme order statistics of the Dirichlet eigenvalues for this operator restricted to large but finite subsets of  $\mathbb{Z}^d$ . We show that, for  $\xi$  with a doubly-exponential type of upper tail, the upper extreme order statistics of the eigenvalues falls into the Gumbel max-order class, and the corresponding eigenfunctions are exponentially localized in regions where  $\xi$  takes large, and properly arranged, values. The picture we prove is thus closely connected with the phenomenon of Anderson localization at the spectral edge. Notwithstanding, our approach is largely independent of existing methods for proofs of Anderson localization and it is based on studying individual eigenvalue/eigenfunction pairs and characterizing the regions where the leading eigenfunctions put most of their mass.

## 1. Introduction and Results

Random Schrödinger operators have been a focus of interest among mathematicians and mathematical physicists for several decades. A good representative is the Anderson Hamiltonian  $H_\xi$  that acts on  $f: \mathbb{Z}^d \rightarrow \mathbb{C}$  as

$$(H_\xi f)(x) := \sum_{y: |y-x|=1} [f(y) - f(x)] + \xi(x)f(x), \quad x \in \mathbb{Z}^d, \quad (1.1)$$

with the potentials  $\{\xi(x)\}_{x \in \mathbb{Z}^d}$  sampled independently from a common law on  $\mathbb{R}$ . The first term on the right is the lattice Laplacian so we may also write  $H_\xi = \Delta + \xi$ .

Much is known (and unknown) about the spectral properties of  $H_\xi$ ; see Section 3 and the introduction to Section 2 for some pointers to the literature and further connections. Our principal goal here is a description of the spectral *extreme order statistics* for  $H_\xi$

---

\* © 2015 by M. Biskup and W. König. Reproduction, by any means, of the entire article for non-commercial purposes is permitted without charge.

over large finite subsets of  $\mathbb{Z}^d$ , i.e., the limiting joint distribution of its largest eigenvalues. More precisely, for a finite set  $D \subset \mathbb{Z}^d$ , let  $H_{D,\xi}$  denote the operator  $H_\xi$  restricted to functions with Dirichlet boundary condition outside  $D$ . This is a self-adjoint operator (a matrix) with real eigenvalues that we will label in decreasing order as

$$\lambda_D^{(1)}(\xi) \geq \lambda_D^{(2)}(\xi) \geq \cdots \geq \lambda_D^{(|D|)}(\xi). \quad (1.2)$$

As is common in the theory of extreme-value statistics (see, e.g., de Haan and Ferreira [23] for a comprehensive account on this theory), given a sequence  $D_L$  of finite subsets of  $\mathbb{Z}^d$  with  $D_L \uparrow \mathbb{Z}^d$ , we wish to identify sequences  $a_L$  and  $b_L$  so that, as  $L \rightarrow \infty$ , the maximal eigenvalue  $\lambda_{D_L}^{(1)}(\xi)$  behaves in law as  $a_L + b_L G$ , where  $G$  is a non-degenerate random variable. A more ambitious goal would be to show that the set of points

$$\left\{ \frac{1}{b_L} (\lambda_{D_L}^{(k)}(\xi) - a_L) : k = 1, \dots, |D_L| \right\} \quad (1.3)$$

tends in law to a non-degenerate point process on  $\mathbb{R}$ . (This still concerns only a finite number of top rescaled eigenvalues; i.e., the behavior at the top edge of the spectrum.)

Of course we cannot hope to do this just for any sequence of domains  $D_L$  so we will content ourselves with domains that arise as scaled-up lattice versions,

$$D_L := \{x \in \mathbb{Z}^d : x/L \in D\} = (LD) \cap \mathbb{Z}^d, \quad (1.4)$$

of bounded open sets  $D \subset \mathbb{R}^d$  with a piece-wise smooth (and thus rectifiable) boundary  $\partial D$ . We will use  $\mathfrak{D}$  to denote the collection of all such sets. For reasons to be explained later, we will also limit ourselves to potentials whose upper tails are close to the doubly-exponential distribution,

$$\text{Prob}(\xi(0) > r) = \exp\{-e^{r/\rho}\}, \quad (1.5)$$

where  $\rho \in (0, \infty)$ . The specific class of potentials we will consider is determined by:

**Assumption 1.1.** *Suppose  $\text{esssup } \xi(0) = \infty$  and let*

$$F(r) := \log \log(\mathbb{P}(\xi(0) > r)^{-1}), \quad r > \text{essinf } \xi(0). \quad (1.6)$$

*We assume that  $F$  is continuously differentiable on its domain and there is  $\rho \in (0, \infty)$  such that*

$$\lim_{r \rightarrow \infty} F'(r) = \frac{1}{\rho}. \quad (1.7)$$

Our results will address not only the eigenvalues but also the associated eigenfunctions. For this, let  $\{\psi_{D,\xi}^{(k)} : k = 1, \dots, |D|\}$  denote an orthonormal basis of eigenfunctions of  $H_{D,\xi}$  in a (finite) set  $D \subset \mathbb{Z}^d$  that are labeled such that

$$H_{D,\xi} \psi_{D,\xi}^{(k)} = \lambda_D^{(k)}(\xi) \psi_{D,\xi}^{(k)}, \quad k = 1, \dots, |D|. \quad (1.8)$$

Although the eigenfunctions are not uniquely determined when two or more of the eigenvalues coincide, we can and will take these to be real-valued measurable functions of  $\xi$ . The principal eigenfunction  $\psi_{D,\xi}^{(1)}$  is known to be of one sign. We then define  $X_k = X_k(\xi)$  by

$$|\psi_{D,\xi}^{(k)}(X_k)| = \max_{x \in D} |\psi_{D,\xi}^{(k)}(x)|, \quad k = 1, \dots, |D|, \quad (1.9)$$

resolving ties using the lexicographic order on  $\mathbb{Z}^d$ . Again, this can be done so that each  $X_k$  is a measurable function of  $\xi$ .

The family  $\{(\lambda_{D_L}^{(k)}(\xi), \psi_{D_L, \xi}^{(k)}): k = 1, \dots, |D_L|\}$  thus identifies a random sequence of points  $\mathbf{a}_1, \dots, \mathbf{a}_{|D_L|} \in \mathbb{R}^{d+1}$  defined by

$$\mathbf{a}_k := \left( \frac{X_k(\xi)}{L}, \frac{1}{b_L} (\lambda_{D_L}^{(k)}(\xi) - a_L) \right), \quad k = 1, \dots, |D_L|. \quad (1.10)$$

These are ordered by their last coordinate. This sequence induces a random measure  $\mathcal{X}_L$  on  $\mathbb{R}^{d+1}$  by setting, for any Borel set  $B \subset \mathbb{R}^{d+1}$ ,

$$\mathcal{X}_L(B) := \sum_{k=1}^{|D_L|} \mathbf{1}_{\{\mathbf{a}_k \in B\}}. \quad (1.11)$$

Note that  $\mathcal{X}_L(B)$  takes values in  $\mathbb{N}_0 := \{0, 1, 2, \dots\}$ , and  $\mathcal{X}_L$  is supported on  $D \times \mathbb{R}$ .

We will interpret  $\mathcal{X}_L$  as a point process, so let us take a moment to recall some basic facts about point processes. First, by a point process  $\mathcal{X}$  on  $\mathbb{R}^{d+1}$  we will generally mean a random  $\mathbb{N}_0 \cup \{\infty\}$ -valued measure on Borel sets in  $\mathbb{R}^{d+1}$  such that  $\mathcal{X}(C) < \infty$  a.s. for any compact  $C \subset \mathbb{R}^{d+1}$ . The space of such measures, endowed with the topology of vague convergence, is a Polish space so convergence in law can be defined accordingly. Explicitly,  $\mathcal{X}_L$  converges in law to  $\mathcal{X}$  if (and only if) for any continuous and compactly-supported function  $f: \mathbb{R}^{d+1} \rightarrow \mathbb{R}$ , the integral of  $f$  against  $\mathcal{X}_L$  converges in law to the integral of  $f$  against  $\mathcal{X}$ .

A process  $\mathcal{X}$  is called a Poisson point process with intensity measure  $\mu$  if  $\mathcal{X}(B)$  is, for any Borel set  $B$ , a Poisson random variable with parameter  $\mu(B)$  and if the random variables  $\mathcal{X}(B_1), \dots, \mathcal{X}(B_n)$  are independent for any pairwise disjoint Borel sets  $B_1, \dots, B_n$ . The principal result of the present paper is then:

**Theorem 1.2 (Poisson convergence; eigenfunction localization).** *Fix  $d \geq 1$  and let  $(\xi(x))_{x \in \mathbb{Z}^d}$  be i.i.d. random variables satisfying Assumption 1.1 with some  $\rho \in (0, \infty)$ . Then there is a sequence  $a_L$  with asymptotic growth*

$$a_L = (\rho + o(1)) \log \log L, \quad L \rightarrow \infty, \quad (1.12)$$

such that, for any  $D \in \mathfrak{D}$  with scaled lattice version  $D_L$  and any choice of the normalized eigenfunctions  $\{\psi_{D_L, \xi}^{(k)}\}$  as above, we have:

(1) (Eigenfunction localization) For any  $k \in \mathbb{N}$  and any  $r_L \rightarrow \infty$ ,

$$\sum_{z: |z - X_k(\xi)| \leq r_L} |\psi_{D_L, \xi}^{(k)}(z)|^2 \xrightarrow{L \rightarrow \infty} 1 \quad \text{in } \mathbb{P}\text{-probability.} \quad (1.13)$$

(2) (Poisson convergence) The process  $\mathcal{X}_L$ , defined via (1.11) by the points

$$\left\{ \left( \frac{X_k(\xi)}{L}, \frac{1}{\rho} (\lambda_{D_L}^{(k)}(\xi) - a_L) \log |D_L| \right) : k = 1, \dots, |D_L| \right\}, \quad (1.14)$$

converges in law to the Poisson point process on  $D \times \mathbb{R}$  with intensity measure  $dx \otimes e^{-\lambda} d\lambda$ .

Restricting the points in  $\mathcal{X}_L$  to the last coordinate shows that the set (1.3) converges in law to a Poisson process provided  $a_L$  is as above and  $b_L \sim \log L$ . The limit law can be described concisely and explicitly as follows:

**Corollary 1.3 (Eigenvalue order-statistics).** *Assume the setting of Theorem 1.2 and let  $D_L$  be as in (1.4). Then the (upper) order statistics of the eigenvalues lies in the domain of attraction of the Gumbel universality class. In particular, for each  $k \in \mathbb{N}$ ,*

$$\left( e^{-\frac{1}{\rho}(\lambda_{D_L}^{(1)}(\xi) - a_L) \log |D_L|}, \dots, e^{-\frac{1}{\rho}(\lambda_{D_L}^{(k)}(\xi) - a_L) \log |D_L|} \right) \xrightarrow[L \rightarrow \infty]{\text{law}} (Z_1, Z_1 + Z_2, \dots, Z_1 + \dots + Z_k), \quad (1.15)$$

where  $Z_1, Z_2, \dots$  are i.i.d. exponential with parameter one. Equivalently, the vector on the left tends in law to the first  $k$  points of a Poisson point process on  $[0, \infty)$  with intensity one.

We remind the reader that the Gumbel universality class is one out of three possible non-degenerate limit distributions for order statistics of i.i.d. random variables; see e.g. de Haan and Ferreira [23]. Gumbel is also the extreme-order class associated with the doubly exponential tails (see Section 6.1). However, as many values of the field need to “cooperate” to create conditions for an extremal eigenvalue, the shift  $a_L$  required for the  $\xi$ ’s differs from the one above by a non-vanishing amount. Explicitly, under Assumption 1.1 and for any  $D \in \mathfrak{D}$ ,

$$\max_{x \in D_L} \xi(x) - \lambda_{D_L}^{(1)}(\xi) \xrightarrow[L \rightarrow \infty]{} \chi \quad \text{in } \mathbb{P}\text{-probability}, \quad (1.16)$$

where  $\chi = \chi(\rho, d)$  is the quantity in  $(0, 2d]$  given by

$$\chi := -\sup \{ \lambda^{(1)}(\varphi) : \varphi \in \mathbb{R}^{\mathbb{Z}^d}, \mathcal{L}(\varphi) \leq 1 \}. \quad (1.17)$$

with the quantities entering this expression given as follows:  $\mathcal{L} : \mathbb{R}^{\mathbb{Z}^d} \rightarrow (0, \infty]$  is the function defined by

$$\mathcal{L}(\varphi) := \sum_{x \in \mathbb{Z}^d} e^{\varphi(x)/\rho} \quad (1.18)$$

while  $\lambda^{(1)}(\varphi)$  is the supremum of the spectrum of  $\Delta + \varphi$  on  $\mathbb{Z}^d$ ,

$$\lambda^{(1)}(\varphi) := \sup \left\{ \langle f, (\Delta + \varphi)f \rangle : f \in \ell^2(\mathbb{Z}^d), \|f\|_2 = 1 \right\}, \quad (1.19)$$

where  $\langle \cdot, \cdot \rangle$ , resp.,  $\|\cdot\|_2$  is the canonical inner product, resp., norm in  $\ell^2(\mathbb{Z}^d)$ . We note that  $\lambda^{(1)}(\varphi)$  is an isolated simple eigenvalue for all potentials  $\varphi$  with  $\mathcal{L}(\varphi) < \infty$ .

The limit (1.16) is actually well known (albeit under different assumptions) from earlier studies of  $H_\xi$  with  $\xi$  having doubly-exponential tails; see our discussion in Section 3. The function  $\varphi \mapsto \mathcal{L}(\varphi)$  is encountered in these studies as well; it is the large-deviation rate function for the field  $\xi$  and so it plays an important role in estimating the probability that  $\xi$  exceeds a given function in a given domain; see (2.12) for an explicit statement in this vein.

The proof of the limiting Poisson statistics (1.15) is based on constructing a coupling to i.i.d. random variables; cf Theorem 2.3 for a precise formulation. This coupling applies to a whole non-degenerate interval at the top of the spectrum. The statement (1.13) implies exponential localization of leading eigenfunctions at the lattice scale. We state a quantitative decay bound that concerns the leading eigenfunctions:

**Theorem 1.4.** *For each  $k \geq 1$  and  $X_k(\xi)$  as in Theorem 1.2, the following holds with probability tending to one as  $L \rightarrow \infty$ : There exist (deterministic) constants  $c_1, c_2 > 0$  such that*

$$|\psi_{D_L, \xi}^{(k)}(z)| \leq c_1 e^{-c_2 |z - X_k(\xi)|}, \quad z \in D_L. \quad (1.20)$$

Moreover, for larger separations from  $X_k(\xi)$  we in fact get

$$|\psi_{D_L, \xi}^{(k)}(z)| \leq c'_1 e^{-c'_2 (\log \log L) |z - X_k(\xi)|}, \quad |z - X_k(\xi)| \geq \log L. \quad (1.21)$$

for some non-random constants  $c'_1, c'_2 > 0$ .

The methods of the present paper are largely independent of the existing techniques for proving Anderson localization. In particular, our approach permits a rather explicit characterization of the location, size and shape of the potential and the corresponding eigenfunction for eigenvalues at and near the top of the spectrum. However, we make no attempt to prove Anderson localization in the usual sense, i.e., for the operator  $\Delta + \xi$  on the entire space  $\mathbb{Z}^d$ .

We note that our paper serves as an important input to a forthcoming article by Biskup, König and dos Santos [10] where it used to prove “one-island” localization in the parabolic version of the above (elliptic) problem.

## 2. Road map to proofs

We proceed to discuss the key ideas of the proofs. We break the main argument into a sequence of stand-alone steps which, we believe, are of independent interest. But before we do so, we will make a brief link to the mathematical theory of Anderson localization. This will give us a chance to compare our approach with the methods used there. Further links and connections to literature will be given in Section 3.

In 1958 Anderson [3] observed that conducting materials may become insulators when impurities are doped in at a sufficiently high density. Anderson’s non-rigorous derivations were later picked up by mathematicians and the subject has been intensely studied in over five decades. In the context of the model (1.1), the word “localization” thus nowadays refers to the fact that, when the fields  $\xi$  are non-degenerate i.i.d. random variables, the spectrum of the Hamiltonian  $H_\xi$  in (1.1) will contain a band of proper eigenvalues with exponentially localized eigenfunctions. This is in contrast with the situation when  $\xi$  is periodic, where the spectrum has a band structure but remains continuous, by the classic Bloch theory. (As already stated, in our class of models we do not address “localization” in this strong sense but only give a version that holds asymptotically in large-enough boxes.)

The texts of Pastur and Figotin [32], Stollmann [35], Carmona and Lacroix [11] and Hundertmark [26] can be consulted for most known details and explanations about Anderson localization. The important technical aspect is that all existing proofs seem to be based on controlling the Green function — or, more precisely, the kernel of the Green operator — associated with  $H_{D, \xi}$ ,

$$G_{D, \xi}(x, y; z) := \langle \delta_x, (z - H_{D, \xi})^{-1} \delta_y \rangle, \quad \Im z > 0, \quad (2.1)$$

in the limit as  $z$  “radially” approaches the real line from the upper half plane in  $\mathbb{C}$ . The aim is to show that  $G_{D, \xi}(x, y; z)$  exhibits exponential decay in  $|x - y|$  in the said limit;

a key challenge is to avoid  $z$  hitting an eigenvalue where the constant in front of the exponentially decaying term becomes divergent.

Various averaging methods have been developed with this purpose in mind. For instance, the fractional-moment method of Aizenman and Molchanov [1] generally yields bounds of the form

$$\mathbb{E}(|G_{D,\xi}(x,y;z)|^s) \leq c_1 e^{-c_2|x-y|} \quad (2.2)$$

uniformly in  $D$  and  $\Im m(z) > 0$ , where  $c_1, c_2 \in (0, \infty)$  are constants while the exponent  $s \in (0, 1)$  is tied to Hölder continuity of the probability density of  $\xi(0)$ . Arguments from spectral theory for *infinite-volume* operators (perfected into the so called Simon-Wolff criteria [34]) then permit one to infer from (2.2) the existence of eigenvalues with exponentially decaying eigenfunctions. Approaches based on “finite-volume criteria” also exist (Aizenman, Schenker, Friedrich and Hundertmark [2]), but they are still versed in the language of Green functions.

Our approach is different from the above in a number of aspects. We work directly with individual eigenvalues in a finite volume and control their dependence on the configuration of random fields. This permits us to characterize geometrically the regions where the eigenfunctions are localized. On the technical side, we manage to avoid working with complex weights and the Green function. Large-deviation theory naturally lurks in the background although, for the most part, we proceed by direct estimates. Although our method is, in its present form, tailored to the study of the upper edge of the spectrum for operators  $H_{D,\xi}$  in finite  $D$  with unbounded i.i.d. random fields, we believe that an extension for bounded fields is possible.

We note that ours is not the first paper marking a shift away from the Green function in the study of random Schrödinger operators. For instance, the work of Germinet and Klopp [19] analyzes eigenvalue distribution functions, including information about localization centers, in the bulk of the spectrum by way of spatial decompositions akin to those used in the present paper. Such spatial decompositions (and coarse-graining ideas in general) of course go back to the early proofs of Anderson localization based on multiscale analysis (Fröhlich and Spencer [13]).

We will now proceed to describe the main steps of our approach formulating the key parts thereof as separate theorems.

*STEP 1: Domain truncation and component trimming.* A good deal of our proof of Theorem 1.2 focuses on individual eigenvalues. The starting observation is that the field configuration  $\xi$  in regions where  $\xi$  is smaller than an eigenvalue  $\lambda$  is of little relevance for  $\lambda$ . For  $A > 0$  and  $R \in \mathbb{N}$ , consider the set

$$D_{R,A}(\xi) := \bigcup_{z \in D: \xi(z) \geq \lambda_D^{(1)}(\xi) - 2A} B(z, R) \cap D, \quad (2.3)$$

where here and henceforth

$$B(z, r) := \{x \in \mathbb{Z}^d: |z-x|_1 < r\} \quad (2.4)$$

denotes the open ball in  $\ell^1$ -metric on  $\mathbb{Z}^d$ . (A key point is that every  $x \notin D_{R,A}(\xi)$  will be at least  $R$  steps of a simple-random walk path away from all vertices where  $\xi$  exceeds  $\lambda_D^{(1)}(\xi) - 2A$ ; note that this applies to even to  $R = 1$ .) We will occasionally refer to the field values in  $D_{R,A}$  as “large” while those not in this set as “small.” Deterministic arguments show:

**Theorem 2.1 (Domain truncation).** *Let  $A > 0$  and  $R \in \mathbb{N}$  obey  $2d(1 + \frac{A}{2d})^{1-2R} \leq \frac{A}{2}$ . Then for any  $\xi$ , any  $U$  with  $D_{R,A}(\xi) \subset U \subset D$  and any  $k \in \{1, 2, \dots, |U|\}$  such that*

$$\lambda_D^{(k)}(\xi) \geq \lambda_D^{(1)}(\xi) - \frac{A}{2} \quad (2.5)$$

we have

$$|\lambda_D^{(k)}(\xi) - \lambda_U^{(k)}(\xi)| \leq 2d \left(1 + \frac{A}{2d}\right)^{1-2R}. \quad (2.6)$$

The vehicle that brings us to this conclusion is, not surprisingly, analysis of rank-one perturbations of  $H_{D,\xi}$ . However, unlike for the corresponding arguments in the proofs of, e.g., (2.2), where the perturbation occurs at a single vertex, here we address single eigenvalues (rather than the Green function) and we perturb the configuration in all of  $D \setminus U$ .

In the specific context of doubly-exponential tails, we will use the conclusion in (2.6) for  $D$  replaced by  $D_L$ , the cutoff  $A$  fixed to a small number (less than  $\chi$  from (1.17)) and  $R$  growing slowly to infinity with  $L$ . Under such conditions, the components of  $U := D_{R,A}$  become very sparse and their geometry can be analyzed by straightforward probabilistic estimates. In particular, due to Dirichlet boundary conditions, the spectrum of  $H_{U,\xi}$  is the union of the spectra in the connected components of  $U$ .

If  $C$  is such a component and  $\lambda_C^{(1)}(\xi) < \lambda_D^{(1)}(\xi) - A$ , then  $C$  cannot contribute to the set of eigenvalues covered by (2.6). We can thus remove  $C$  from  $U$  and still maintain the control provided by Theorem 2.1. This permits systematic component “trimming” that helps significantly reduce the number of connected components of concern.

*STEP 2: Reduction to one eigenvalue per component.* The removal of irrelevant components of  $D_{R,A}$  reduces, quite considerably, the geometric complexity of the underlying field configuration. An issue that comes up next is what part of the spectrum in each component needs to be taken into account. Here we will observe that in components that have a chance to contribute, all but the leading eigenvalue can safely be disregarded. This effectively couples the top of the spectrum in  $D$  to the set of principal eigenvalues in the connected components of  $D_{R,A}$ .

As a starter, for  $C \subset \mathbb{Z}^d$  finite, consider the finite-volume version of (1.18),

$$\mathcal{L}_C(\varphi) := \sum_{x \in C} e^{\varphi(x)/\rho}. \quad (2.7)$$

Similarly, we introduce the finite-volume analogue of (1.17),

$$\chi_C := -\sup\{\lambda_C^{(1)}(\xi) : \xi \in \mathbb{R}^{\mathbb{Z}^d}, \mathcal{L}_C(\xi) \leq 1\}. \quad (2.8)$$

Then we have the following deterministic estimate:

**Proposition 2.2 (Spectral gap).** *Let  $C \subset \mathbb{Z}^d$  be finite. If for some  $K \geq 0$ ,*

$$\lambda_C^{(1)}(\xi) - \lambda_C^{(2)}(\xi) \leq K, \quad (2.9)$$

then also

$$\lambda_C^{(1)}(\xi) - \rho \log \mathcal{L}_C(\xi) \leq -\chi_C + K - \rho \log 2. \quad (2.10)$$

Our use of this proposition requires observations from large-deviation theory of double-exponential i.i.d. random fields: First, for potentials satisfying Assumption 1.1,  $\mathcal{L}_C$  acts as a large-deviation rate function for finding a specific potential profile in  $C$ . More explicitly, in Section 6.1 we will show that, for some  $\widehat{a}_L \rightarrow \infty$ ,

$$\mathbb{P}(\xi(0) \geq \widehat{a}_L + s) = \frac{1}{L^{d\theta(1+o(1))}} \quad \text{where} \quad \theta := e^{s/\rho}, \quad (2.11)$$

with  $o(1) \rightarrow 0$  as  $L \rightarrow \infty$  uniformly in compact sets of  $s$ . Thus, for any given  $\varphi: C \rightarrow \mathbb{R}$ ,

$$\mathbb{P}(\xi \geq \widehat{a}_L + \varphi \text{ in } C) = L^{-d\mathcal{L}_C(\varphi)[1+o(1)]}, \quad L \rightarrow \infty. \quad (2.12)$$

The set inclusion

$$\{\xi: \lambda_C^{(1)}(\xi) \geq a\} \subseteq \bigcup_{\varphi: \mathcal{L}_C(\varphi) \geq 1} \{\xi: \xi \geq a + \chi_C + \varphi \text{ on } C\}, \quad (2.13)$$

which is derived readily from the alternative formula for  $\chi_C$ ,

$$\chi_C = -\sup_{\xi} [\lambda_C^{(1)}(\xi) - \rho \log \mathcal{L}_C(\xi)], \quad (2.14)$$

and a union bound then show that large eigenvalues in  $D_L$  will thus come only with potential profiles for which  $\varphi := \xi - \widehat{a}_L - \chi_C$  obeys  $\mathcal{L}_C(\varphi) \approx 1$ .

Returning to the role of Proposition 2.2 in our proofs, we note that its main conclusion can be used to derive

$$\begin{aligned} \lambda_C^{(1)}(\xi) \geq a' \quad \text{AND} \quad \lambda_C^{(1)}(\xi) - \lambda_C^{(2)}(\xi) &\leq \frac{1}{2}\rho \log 2 \\ \implies \xi &\geq \varphi + a + \chi_C \text{ in } C \text{ for some } \varphi \text{ satisfying } \mathcal{L}_C(\varphi) \geq u, \end{aligned} \quad (2.15)$$

where  $u$  is defined by

$$\log u := \frac{a' - a}{\rho} + \frac{1}{2} \log 2. \quad (2.16)$$

Since  $u > 0$  for  $a' > \widehat{a}_L - \chi - \frac{\rho}{2} \log 2$ , the event that a connected set  $C$  satisfies the conditions on the left of (2.15) has probability  $L^{-d-\varepsilon}$  for some  $\varepsilon > 0$ ; a union bound shows that for  $R = L^{o(1)}$  no connected component of  $D_{R,A}$  will have this property once  $L$  is sufficiently large. It follows that, whenever  $\lambda_C^{(1)}(\xi)$  for some component  $C$  in  $D_{R,A}$  is close to its optimal value,  $\lambda_C^{(2)}(\xi)$  is at least a deterministic constant below  $\lambda_C^{(1)}(\xi)$ . In short, only the top eigenvalue in each connected component of  $D_{R,A}$  need be considered.

*STEP 3: Coupling to i.i.d. variables.* Our previous observations permit us to design a coupling between the eigenvalues in a small interval near the top the spectrum of  $H_{D,\xi}$  and a family of i.i.d. random variables. There is a number of ways how such a coupling can be formulated; we will present one that is based on a regular partition of  $\mathbb{Z}^d$  into square boxes.

As we will work, from now on, with  $L$ -dependent objects, we need to fix two sequences that determine the main scales of the problem: a sequence of positive integers  $R_L$  satisfying

$$\frac{R_L}{\log \log L} \xrightarrow{L \rightarrow \infty} \infty \quad \text{but} \quad R_L = (\log L)^{o(1)}, \quad (2.17)$$



which will govern the size of the connected components and spatial range of the perturbation arguments described above, and a sequence of integers  $N_L$  such that

$$\frac{N_L}{R_L} \xrightarrow{L \rightarrow \infty} \infty \quad \text{and} \quad \limsup_{L \rightarrow \infty} \frac{\log N_L}{\log L} < 1 \quad (2.18)$$

which determines the size of the boxes in the partition.

Using a natural partition of  $\mathbb{Z}^d$  into square boxes of side  $N_L + 1$ , for each such box  $B_{N_L+1}$ , consider the sub-box of side  $N_L$  induced by the natural embedding  $B_{N_L} \subset B_{N_L+1}$ . We will call these  $N_L$ -boxes. Note that distinct  $N_L$ -boxes are thus disjoint while “neighboring”  $N_L$ -boxes share part of their external boundary. For  $D \in \mathfrak{D}$ , let  $D_L$  be its scaled lattice version (1.4) and let  $B_{N_L}^{(i)}$ ,  $i = 1, \dots, m_L$ , denote the collection of those  $N_L$ -boxes that are entirely contained in  $D_L$ . From the fact that  $\partial D$  is smooth we get  $|\partial D_L| = o(|D_L|)$  and so

$$|D_L| = m_L N_L^d (|D| + o(1)), \quad L \rightarrow \infty, \quad (2.19)$$

where  $|D|$  denotes the Lebesgue volume of  $D$ . By (2.18),  $m_L \rightarrow \infty$  as  $L \rightarrow \infty$ .

Given a configuration  $\xi$ , let  $\lambda_i(\xi)$  be a shorthand for the principal eigenvalue in  $B_{N_L}^{(i)}$ ,

$$\lambda_i(\xi) := \lambda_{B_{N_L}^{(i)}}^{(1)}(\xi), \quad i = 1, \dots, m_L. \quad (2.20)$$

Since these eigenvalues depend on disjoint parts of the  $\xi$  field (the Dirichlet boundary condition plays a role here), the random variables  $\{\lambda_i(\xi) : i = 1, \dots, m_L\}$  are i.i.d.

**Theorem 2.3 (Coupling to i.i.d. process).** *Let  $D \in \mathfrak{D}$ . For sequences  $(R_L)$ ,  $(N_L)$  and  $D_L$  as above, let  $\hat{\lambda}_1(\xi), \dots, \hat{\lambda}_{m_L}(\xi)$  be the sequence  $\lambda_1(\xi), \dots, \lambda_{m_L}(\xi)$  from (2.20) listed in decreasing order. Under Assumption 1.1, there is an  $A > 0$  such that the event*

$$\forall k \in \{1, \dots, m_L\}: \quad \hat{\lambda}_1(\xi) - \hat{\lambda}_k(\xi) < A \Rightarrow |\lambda_{D_L}^{(k)}(\xi) - \hat{\lambda}_k(\xi)| < 4d \left(1 + \frac{A}{2d}\right)^{1-2R_L} \quad (2.21)$$

occurs with probability tending to one as  $L \rightarrow \infty$ .

The proof is based on Theorem 2.1 and estimates of the probability that a component with an appreciable principal eigenvalue intersects the boundary of a partition box, or lies within  $N_L$  of the boundary of  $D_L$ . (These are the reasons for the restrictions in (2.18).)

*STEP 4: Identifying max-order class.* Theorem 2.3 brings the proof of the Poisson statistics in Theorem 1.2 and Corollary 1.3 to the realm of standard extreme-order limit theory (see again de Haan and Ferreira [23]). Naturally, one starts by defining the centering sequence  $a_L$  as

$$\mathbb{P}(\lambda_{B_{N_L}}^{(1)} \geq a_L) = \left(\frac{N_L}{L}\right)^d. \quad (2.22)$$

(Such an  $a_L$  exists because the  $\xi$ 's, and also the  $\lambda_{B_{N_L}}^{(1)}(\xi)$ 's, are continuously distributed.) In order to determine the correct scaling of the spacings between the eigenvalues, and generally place the limit distribution in the Gumbel max-order class, it remains to prove:

**Theorem 2.4 (Max-order class of local eigenvalues).** *Suppose Assumption 1.1 and let  $b_L$  obey*

$$b_L \log L \xrightarrow{L \rightarrow \infty} \frac{\rho}{d}. \quad (2.23)$$

Then for each  $s \in \mathbb{R}$ ,

$$\mathbb{P}(\lambda_{B_{N_L}}^{(1)} \geq a_L + sb_L) = e^{-s} \left(\frac{N_L}{L}\right)^d (1 + o(1)), \quad (2.24)$$

with  $o(1) \rightarrow 0$  as  $L \rightarrow \infty$  uniformly on compact sets of  $s$ .

The proof is based on regularity of the probability density of  $\xi$  supplied by Assumption 1.1. In simple terms, the change of the eigenvalue by  $sb_L$  can be achieved by shifting the whole  $\xi$  configuration by the same amount. A catch is that this would be too costly (i.e., inefficient) to perform in the entire  $N_L$ -box; rather one has to do this only in those parts of the box where the relevant contribution comes from.

Before we proceed to the next step, note that, on the basis of Theorems 2.3 and 2.4, we are already able to conclude the limit statement (1.15). (The condition on the left of (2.17) ensures that the error in (2.21) is much smaller than the spacing between eigenvalues.)

*STEP 5: Eigenfunction localization.* The final task before us is a control of the spatial localization of the eigenfunctions. As is well known, the main obstruction to localization is degeneracy of eigenvalues. Two techniques exist for dealing with this problem: averaging and multiscale analysis. In our context, we are able to address the problem directly by developing a deterministic link between the spatial decay of an eigenfunction and the distance of the associated eigenvalue to other eigenvalues. A key novel fact is that the (still needed) non-degeneracy of the eigenvalues will be supplied by the already-proved extreme-order limit theorem.

Let  $\mathfrak{C}_{R,A} := \mathfrak{C}_{R,A}(\xi)$  denote the set of connected components of  $D_{R,A}(\xi)$  and, for  $V \subset \mathbb{Z}^d$ , let  $\partial V$  mark the set of vertices outside  $V$  that have an edge into  $V$ . We will measure the decay of the eigenfunctions in terms of a distance-like object  $d(x, \mathcal{C})$ , indexed by vertices  $x \in \mathbb{Z}^d$  and components  $\mathcal{C} \in \mathfrak{C}_{R,A}$ , on which we impose the following requirements:

- (D0)  $d(z, \mathcal{C}) \geq 0$  for all  $z$  and  $\mathcal{C}$  with  $d(z, \mathcal{C}) = 0$  whenever  $z \in \mathcal{C}$ .
- (D1) For all  $z \in D \setminus D_{R,A}(\xi)$ , all  $y \in \partial B(z, R)$  and all  $\mathcal{C} \in \mathfrak{C}_{R,A}$ , we have  $d(z, \mathcal{C}) \leq d(y, \mathcal{C}) + R$ .
- (D2) For all  $\mathcal{C}' \neq \mathcal{C}$ , all  $z \in \mathcal{C}'$  and all  $y \in \partial \mathcal{C}'$  we have  $d(z, \mathcal{C}) \leq d(y, \mathcal{C}) + 1$ .

An example of such  $d(\cdot, \cdot)$  is constructed as follows: Define a graph by contracting all components in  $\mathfrak{C}_{R,A}$  to a single vertex while keeping the edges between the (new) vertex corresponding to component  $\mathcal{C}$  and all (old) vertices on  $\partial \mathcal{C}$  — which, by the fact that  $\mathcal{C}$  is in  $\mathfrak{C}_{R,A}$  do not lie in another component in  $\mathfrak{C}_{R,A}$ . Then set  $d(z, \mathcal{C})$  to the corresponding graph-theoretical distance from  $z$  to the (vertex corresponding to) component  $\mathcal{C}$ .

**Theorem 2.5 (Eigenfunction decay).** *Assume  $R \geq 1$  and  $A > 0$  are such that we have  $\varepsilon_R := 2d(1 + \frac{A}{2d})^{1-2R} < A/2$ . Let  $\lambda$ , resp.,  $\psi$  be a Dirichlet eigenvalue, resp., a corresponding eigenfunction of  $H_{D,\xi}$  such that*

$$\lambda \geq \lambda_D^{(1)}(\xi) - \frac{A}{2} + \varepsilon_R. \quad (2.25)$$

Assume the following:

- (1)  $\text{gap}(\lambda)$ , the distance of  $\lambda$  to the nearest eigenvalue of  $H_{D,\xi}$ , obeys  $\text{gap}(\lambda) > 10\epsilon_R$ .  
(2) there is  $h > 0$  such that, for any self-avoiding (nearest-neighbor) path  $x_1, \dots, x_R$  in  $D$ ,

$$\xi(x_j) < \lambda \text{ for all } j = 1, \dots, R \quad \Rightarrow \quad \prod_{j=1}^R \frac{2d}{2d + \lambda - \xi(x_j)} \leq e^{-hR}, \quad (2.26)$$

- (3) for some  $\delta \in (0, 1)$ ,

$$\frac{\text{gap}(\lambda) - 2\epsilon_R}{12d} \wedge 1 > 4e^{-(1-\delta)hR + \delta h} \sqrt{|\partial \mathcal{C}'|}, \quad \mathcal{C}' \in \mathfrak{C}_{R,A}. \quad (2.27)$$

Then there is  $\mathcal{C} \in \mathfrak{C}_{R,A}$  such that

$$|\psi(z)| \leq e^{-\delta h d(z, \mathcal{C})}, \quad z \in D. \quad (2.28)$$

In order to appreciate this general result, we again place ourselves in an  $L$ -dependent setting of domains  $D := D_L$ , with  $R := R_L$  satisfying (2.17). Under Assumption 1.1 on the upper tail of  $\xi(0)$ , we then have (2.11), which readily yields (2.26) for any given (fixed)  $h > 0$ . Since the component sizes are at most polylogarithmic in  $R_L$ , conditions (1) and (3) are satisfied as soon as  $\text{gap}(\lambda)$  is larger than exponentially-small in  $R_L$ . For the leading eigenvalues, the gap is typically order  $(\log L)^{-1}$ ; thanks to (2.17), the bound (2.28) thus applies to  $\psi := \psi_{D_L, \xi}^{(k)}$  for all  $k \geq 1$ . A straightforward comparison between  $d(z, \mathcal{C})$  and the Euclidean distance (see Lemma 8.2) then proves the decay estimate (1.21).

The remainder of this paper consists of proofs of the above claims in full technical detail. In particular, Section 4 deals with various deterministic spectral estimates underlying Steps 1 and 2 in the above scheme. Deterministic bounds underpinning eigenfunction localization (Step 5 above) appear in Section 5. Coupling to i.i.d. random variables (Step 3) is performed in Section 6. Step 4 is the subject of Section 7, where we also conclude the proofs of eigenvalue order statistics. In Section 8 this feeds into the proof of eigenfunction localization (the probabilistic part of Step 5) and concludes the proof of Theorems 1.2(2) and 1.4.

### 3. Connections and remarks

Before we move to actual proofs, let us pause to discuss some connections to the existing literature. References have so far been largely suppressed in order not to disturb the flow of explanations of results and ideas for proofs. We will not attempt to review all context of Anderson localization (some of it was done already in Section 2) but rather focus only on results dealing with spectral statistics.

Attempts to describe the statistics of the spectrum of random Schrödinger operators are as old as the subject itself. In the localization regime, the statistics was expected to be Poisson-like. In the Anderson model, this was proved by Molchanov [31] in dimension one and by Minami [30] in general dimension, based on techniques developed in Aizenman and Molchanov [1] and Wegner [37]. Killip and Nakano [27] later included also the localization centers of the eigenfunctions in the Poisson-process convergence. All these results are derived at parameter values where the fractional moment bound on the resolvent, one of the two main known methods for proving Anderson localization, applies.

A very appealing formulation has later been found in terms of convergence of *unfolded* eigenvalues to a homogeneous Poisson process. Here an unfolded eigenvalue is the quantity  $I(\lambda_D^{(k)})$  where  $I$  is the integrated density of states (see Carmona and Lacroix [11] or Veselić [36] for detailed definitions). The currently most complete result of this kind seems to have been derived by Germinet and Klopp with statements that apply both in the bulk of the spectrum [19] and, in  $d = 1$ , also close to the spectral edge [20] (extensions to arbitrary  $d$  require a modified kinetic term). We refer to the original papers for more details. Incidentally, our formulation of Corollary 1.3 can be thought of as a result in this vein although, to keep the paper reasonably short, we do not link this to the integrated density of states for our class of potentials.

Our work is focused on unbounded potentials and is thus closer in spirit to the studies of the potential tails that are heavier than doubly exponential; e.g., Grenkova, Molchanov and Sudarev [21,22], Austraškas [4,5], van der Hofstad, Mörters and Sidorova [25], König, Lacoïn, Mörters and Sidorova [29], Sidorova and Twarowski [33] and Fiodorov and Muirhead [12]. Austraškas' paper [5] includes also a discussion of doubly-exponential tails (1.5) — despite his vigorous insistence on the contrary throughout the abstract and introduction — but only for  $\rho$  very large. A common feature of all of these works is the corresponding eigenfunctions are localized more or less at a *single* lattice site; namely, a high excess value of the random potential. For the doubly exponential tails with general value of  $\rho$  this is no longer the case and this is exactly what makes these tails a challenge.

We note that in Gärtner, König and Molchanov [16], the asymptotic shape of the potential in the localization regions in a large box  $D$  was identified as the one of the maximizers in (1.17). Furthermore, an explicit form of exponential localization was proved for the principal eigenfunctions of  $H_{D,\xi}$ , after removing the top values of  $\xi$  in all the other localization regions; the shape of these eigenfunctions was identified as well. However, the method there was based on a tedious random walk enumeration technique, which we do not follow here. To keep the present paper self-contained, we also do not use any partial results from [16].

The class of doubly exponential tails was identified rather early in the studies of the parabolic Anderson problem (Gärtner and Molchanov [17,18], Gärtner and den Hollander [14], Gärtner and König [15] and Gärtner, König and Molchanov [16]) as that for which the leading eigenfunctions spread beyond a single point but no scaling with the size of the underlying set is necessary. Other “universal” classes of potential upper tails have been identified later (Biskup and König [8,9], van der Hofstad, König and Mörters [24]; see also the forthcoming book by König [28]). In contrast with doubly-exponential tails, in these cases the mass of the leading eigenfunctions in a set of side  $L$  spreads over regions whose size grows to infinity with  $L$ . Although the present paper deals only with doubly-exponential tails, we believe that the bulk of our method can be adapted to the other cases as well.

As already mentioned, our results are now being used to deal with the parabolic version of the problem as well (Biskup, König and dos Santos [10]).

#### 4. Domain truncation and spectral gap

The goal of this section is to establish Theorem 2.1 that underpins all subsequent derivations in this paper. In addition, we will prove bounds on the distance between the first and second leading eigenvalue, i.e., the spectral gap, as stated in Proposition 2.2.

*4.1. Martingale argument.* The proof of Theorem 2.1 is based on the fact that eigenfunctions decay rapidly away from  $D_{R,A}(\xi)$ . We will control the rate of this decay by a martingale argument. Let  $Y := (Y_k)_{k \in \mathbb{N}_0}$  denote a discrete-time simple symmetric random walk on  $\mathbb{Z}^d$ . We will write  $P^x$ , resp.,  $E^x$  for the law, resp., expectation for the walk started from  $x \in \mathbb{Z}^d$  and let  $\mathcal{F}_n := \sigma(Y_0, \dots, Y_n)$  denote the canonical filtration associated with  $Y$ .

**Lemma 4.1.** *Let  $\lambda := \lambda_D(\xi)$ , resp.,  $\psi := \psi_{D,\xi}$  be a Dirichlet eigenvalue, resp., a corresponding eigenfunction of  $H_{D,\xi}$ . Define*

$$\tau := \inf\{k \in \mathbb{N}_0 : \xi(Y_k) \geq \lambda \text{ OR } Y_k \notin D\} \quad (4.1)$$

and denote  $M_0 := \psi(Y_0)$  and, for  $1 \leq n \leq \tau$ ,

$$M_n := \psi(Y_n) \prod_{k=0}^{n-1} \frac{2d}{2d + \lambda - \xi(Y_k)}. \quad (4.2)$$

Then, under  $P^x$  for any  $x \in \mathbb{Z}^d$ , the process  $M^\tau = (M_{\tau \wedge n})_{n \in \mathbb{N}_0}$  is a martingale for the canonical filtration  $(\mathcal{F}_n)_{n \in \mathbb{N}_0}$ .

*Proof.* If  $x \notin D$ , then  $\tau = 0$  and  $M_{\tau \wedge n} = 0$  for any  $n$ ,  $P^x$ -a.s. For  $x \in D$  the following holds  $P^x$ -a.s.: On  $\{\tau \geq n\}$  we have  $\lambda - \xi(Y_k) \geq 0$  for  $k \leq n-1$  and hence  $|M_n| \leq |\psi(Y_n)| \leq \max_x |\psi(x)|$ ; i.e.,  $M_n$  is bounded. On  $\{\tau > n\}$ , the conditional expectation of  $\psi(Y_{n+1})$  given  $Y_0, \dots, Y_n$  equals  $\psi(Y_n) + \frac{1}{2d}(\Delta \psi)(Y_n)$ . Writing this using  $(\Delta + \xi)\psi = \lambda \psi$  shows that, on  $\{\tau > n\}$ ,

$$\begin{aligned} E^x(M_{n+1} | Y_0, \dots, Y_n) &= E^x(\psi(Y_{n+1}) | Y_0, \dots, Y_n) \prod_{k=0}^n \frac{2d}{2d + \lambda - \xi(Y_k)} \\ &= \left[ \psi(Y_n) + \frac{1}{2d}(\Delta \psi)(Y_n) \right] \prod_{k=0}^n \frac{2d}{2d + \lambda - \xi(Y_k)} \\ &= \psi(Y_n) \left[ 1 + \frac{1}{2d}(\lambda - \xi(Y_n)) \right] \prod_{k=0}^n \frac{2d}{2d + \lambda - \xi(Y_k)} = M_n. \end{aligned} \quad (4.3)$$

It follows that  $M^\tau$  is a martingale.  $\square$

The next lemma expresses the desired consequence of the martingale property. (The set  $D'$  will later be taken to be  $D \setminus U$  with  $U$  as in Theorem 2.1. Note that for  $\lambda$  being the principal eigenvalue  $\lambda_D^{(1)}(\xi)$ , (4.4) below reads alternatively as  $D' \cap D_{R,A/2} = \emptyset$ .)

**Lemma 4.2.** *Let  $\lambda$  and  $\psi$  be as in Lemma 4.1. Given  $A' \geq A > 0$  and  $R \in \mathbb{N}$ , let  $D' \subset D$  be such that  $\xi \leq \lambda - A'$  on  $D'$  and*

$$x \in D \quad \text{AND} \quad \xi(x) \geq \lambda - A \quad \implies \quad \text{dist}(x, D') \geq R, \quad (4.4)$$

where “dist” denotes the  $\ell^1$ -distance on  $\mathbb{Z}^d$ . Then

$$\sum_{x \in D'} |\psi(x)|^2 \leq \left(1 + \frac{A}{2d}\right)^{2-2R} \left(1 + \frac{A'}{2d}\right)^{-2} \|\psi\|_2^2. \quad (4.5)$$

*Proof.* As the square of a bounded martingale is a submartingale, we have

$$|\psi(x)|^2 = E^x |M_{\tau \wedge 0}|^2 \leq E^x |M_{\tau \wedge R}|^2, \quad x \in D'. \quad (4.6)$$

By our assumptions on  $D'$ , any path of the simple random walk started at  $x \in D'$  will either leave  $D$  or stay in the region where  $\xi < \lambda - A$  for at least  $R - 1$  steps. This implies that, on the event  $\{\tau < R\}$ , we necessarily have  $Y_{\tau \wedge R} = Y_\tau \notin D$  with  $P^x$ -probability one. Hence,  $M_{\tau \wedge R} = 0$  on  $\{\tau < R\}$  and so  $E^x(\mathbf{1}\{\tau < R\} |M_{\tau \wedge R}|^2) = 0$ .

On the other hand, on  $\{\tau \geq R\}$ , each term in the product in (4.2) is bounded by  $(1 + \frac{A}{2d})^{-1}$  with that for  $k = 0$  bounded even by  $(1 + \frac{A'}{2d})^{-1}$ . From (4.6), we thus get

$$|\psi(x)|^2 \leq \left(1 + \frac{A}{2d}\right)^{2-2R} \left(1 + \frac{A'}{2d}\right)^{-2} E^x(\mathbf{1}\{\tau \geq R\} |\psi(Y_R)|^2), \quad x \in D'. \quad (4.7)$$

The reversibility of the simple random walk implies  $\sum_{x \in \mathbb{Z}^d} P^x(Y_R = y) = 1$  and so

$$\sum_{x \in D'} E^x(\mathbf{1}\{\tau \geq R\} |\psi(Y_R)|^2) \leq \sum_{x \in D'} \sum_{y \in D} P^x(Y_R = y) |\psi(y)|^2 \leq \|\psi\|_2^2, \quad (4.8)$$

whereby the claim follows.  $\square$

*4.2. Rank-one perturbations.* To apply the above *a priori* bounds in the proof of Theorem 2.1, we also need to tie this with rank-one perturbation arguments. First we prove a continuity statement:

**Lemma 4.3.** *For  $U \subset D$ , let  $\xi_s := \xi - s \mathbf{1}_{D \setminus U}$ . For any  $k \in \{1, \dots, |U|\}$ , the map  $s \mapsto \lambda_D^{(k)}(\xi_s)$  is non-increasing, Lipschitz continuous (with Lipschitz constant one) and*

$$\lambda_D^{(k)}(\xi_s) \xrightarrow{s \rightarrow \infty} \lambda_U^{(k)}(\xi). \quad (4.9)$$

*Proof.* Since the eigenvalues are labeled in a decreasing order, the Minimax Theorem reads

$$\lambda_D^{(k)}(\xi) = \inf_{\substack{\mathcal{H}_k: \mathcal{H}_k \subset \mathbb{C}^{|D|} \\ \dim(\mathcal{H}_k) = k-1}} \sup_{\substack{\phi \in \mathcal{H}_k^\perp \\ \|\phi\|_2 = 1}} \langle \phi, H_{D, \xi} \phi \rangle, \quad (4.10)$$

where the supremum goes over all  $(k-1)$ -dimensional linear subspaces of  $\mathbb{C}^{|D|}$  and, here and henceforth,  $\langle \psi, \phi \rangle := \sum_{x \in D} \psi(x)^* \phi(x)$  denotes the inner product in  $\mathbb{C}^{|D|}$ . From

$$H_{D, \xi_{s'}} = H_{D, \xi_s} + (s - s') \mathbf{1}_{D \setminus U}, \quad (4.11)$$

we thus immediately get

$$0 \leq \lambda_D^{(k)}(\xi_{s'}) - \lambda_D^{(k)}(\xi_s) \leq s - s', \quad 0 \leq s' < s. \quad (4.12)$$

This proves the stated monotonicity and continuity.

Before we address the limit (4.9), we observe

$$\lambda_D^{(k)}(\xi_s) \geq \lambda_U^{(k)}(\xi), \quad k = 1, \dots, |U|. \quad (4.13)$$

Obviously, it suffices to prove this for  $s := 0$ . To see why this holds we recall that the infimum in (4.10) is achieved by  $\mathcal{H}_k$  being the span of the first  $k-1$  eigenvalues of  $H_{D,\xi}$ . This and the fact that  $\langle \phi, H_{D,\xi} \phi \rangle = \langle \phi, H_{U,\xi} \phi \rangle$  whenever  $\phi$  vanishes outside  $U$  yield

$$\begin{aligned} \lambda_D^{(k)}(\xi) &= \sup_{\substack{\phi \perp \psi_{D,\xi}^{(1)}, \dots, \psi_{D,\xi}^{(k-1)} \\ \|\phi\|_2=1}} \langle \phi, H_{D,\xi} \phi \rangle \geq \sup_{\substack{\phi \perp \psi_{D,\xi}^{(1)}, \dots, \psi_{D,\xi}^{(k-1)} \\ \|\phi\|_2=1, \phi=0 \text{ on } D \setminus U}} \langle \phi, H_{D,\xi} \phi \rangle \\ &= \sup_{\substack{\phi \perp \psi_{D,\xi}^{(1)}, \dots, \psi_{D,\xi}^{(k-1)} \\ \|\phi\|_2=1, \phi=0 \text{ on } D \setminus U}} \langle \phi, H_{U,\xi} \phi \rangle \geq \inf_{\substack{\mathcal{H}_k : \mathcal{H}_k \subset \mathbb{C}^{|U|} \\ \dim(\mathcal{H}_k) \leq k-1}} \sup_{\substack{\phi \in \mathcal{H}_k^\perp \\ \|\phi\|_2=1}} \langle \phi, H_{U,\xi} \phi \rangle, \end{aligned} \quad (4.14)$$

where the inequalities follow from the fact that the projection of the linear span of the vectors  $\psi_{D,\xi}^{(1)}, \dots, \psi_{D,\xi}^{(k-1)}$  onto  $\mathbb{C}^{|U|}$  has dimension at most  $k-1$  (in the first inequality this ensures that the supremum is over a non-empty set). By (4.10) again, the last infimum is actually achieved by some  $\mathcal{H}_k$  with  $\dim(\mathcal{H}_k) = k-1$  and so we get (4.13) as desired.

To prove the limit (4.9), consider a sequence  $s_n \rightarrow \infty$  such that  $\lambda_D^{(k)}(\xi_{s_n}) \rightarrow \lambda_k$  as well as  $\psi_{D,\xi_{s_n}}^{(k)}(x) \rightarrow \psi_k(x)$  exist. Since  $\lambda_k$ , with  $k = 1, \dots, |U|$ , are finite by the above reasoning, we can immediately conclude that  $(\Delta + \xi)\psi_k = \lambda_k \psi_k$  on  $U$  while  $\psi_k(x) = 0$  for  $x \in D \setminus U$ . As  $k \mapsto \lambda_k$  is non-increasing, we must have  $\lambda_U^{(k)}(\xi) = \lambda_k$ , regardless of the sequence  $s_n$ .  $\square$

The use of Lemma 4.2 will be aided by the following observation:

**Lemma 4.4.** *Fix  $A > 0$ ,  $R \in \mathbb{N}$  and, for  $U$  as in Theorem 2.1 and  $s \geq 0$ , let  $\xi_s := \xi - s \mathbf{1}_{D \setminus U}$ . Fix  $k \in \{1, \dots, |D|\}$  and let  $s > 0$  be such that  $\lambda_D^{(k)}(\xi_s) \geq \lambda_D^{(1)}(\xi) - A$ . Then*

$$0 \leq \lambda_D^{(k)}(\xi) - \lambda_D^{(k)}(\xi_s) \leq \int_0^s ds' \left(1 + \frac{A}{2d}\right)^{2-2R} \left(1 + \frac{A+s'}{2d}\right)^{-2}. \quad (4.15)$$

*Proof.* The left inequality follows from Lemma 4.3. The right inequality would be a simple consequence of the fact that  $\frac{d}{ds} \lambda_D^{(k)}(\xi_s) = -\|\psi_{D,\xi_s}^{(k)} \mathbf{1}_{D \setminus U}\|_2^2$  holds whenever the eigenvalue is non-degenerate. However, to get around the issue of degeneracy, we need to work a bit harder.

Abbreviate the integrand in (4.15) as

$$g(s) := \left(1 + \frac{A}{2d}\right)^{2-2R} \left(1 + \frac{A+s}{2d}\right)^{-2}. \quad (4.16)$$

We claim that it suffices to prove

$$\lambda_D^{(k)}(\xi_s) \geq \lambda_D^{(1)}(\xi) - A \quad \Rightarrow \quad \limsup_{s' \uparrow s} \frac{\lambda_D^{(k)}(\xi_{s'}) - \lambda_D^{(k)}(\xi_s)}{s - s'} \leq g(s). \quad (4.17)$$

This follows by the following general observation: If  $f: \mathbb{R} \rightarrow \mathbb{R}$  is non-increasing and Lipschitz continuous, and  $a < b$ , then the Fundamental Theorem of Calculus applies

to  $f$  on  $[a, b]$ . Explicitly, we have

$$\begin{aligned}
f(a) - f(b) &= \lim_{h \downarrow 0} \frac{1}{h} \left( \int_{a-h}^a f(x) dx - \int_{b-h}^b f(x) dx \right) \\
&= \lim_{h \downarrow 0} \frac{1}{h} \left( \int_{a-h}^{b-h} f(x) dx - \int_a^b f(x) dx \right) \\
&= \lim_{h \downarrow 0} \int_a^b \frac{f(x-h) - f(x)}{h} dx \leq \int_a^b \limsup_{h \downarrow 0} \frac{f(x-h) - f(x)}{h} dx,
\end{aligned} \tag{4.18}$$

where the last step follows by Fatou's Lemma and the boundedness of the integrand and the integration domain. So if (4.17) holds on an interval of  $s$ , the difference of the eigenvalues at the endpoints of this interval is bounded by the corresponding integral of  $g$ , as claimed in (4.15).

To establish (4.17), consider a sequence  $s'_n \uparrow s$  saturating the *limes superior* and such that the limits

$$\lambda_k := \lim_{n \rightarrow \infty} \lambda_D^{(k)}(\xi_{s'_n}) \quad \text{and} \quad \psi_k(x) := \lim_{n \rightarrow \infty} \psi_{D, \xi_{s'_n}}^{(k)}(x) \tag{4.19}$$

exist for all  $k \in \{1, \dots, |D|\}$  and all  $x \in D$ . Passing to the limit in the eigenfunction equation shows that  $(\lambda_k, \psi_k)$  is an eigenvalue/eigenfunction pair for the field  $\xi_s$  and, in light of the eigenvalues being ordered,  $\lambda_k = \lambda_D^{(k)}(\xi_s)$ . Both  $\psi_k$  and  $\psi_{D, \xi_{s'_n}}^{(k)}$  are eigenfunctions and so, by (4.11),

$$\begin{aligned}
\lambda_k \langle \psi_k, \psi_{D, \xi_{s'_n}}^{(k)} \rangle &= \langle \psi_k, H_{D, \xi_s} \psi_{D, \xi_{s'_n}}^{(k)} \rangle \\
&= \lambda_D^{(k)}(\xi_{s'_n}) \langle \psi_k, \psi_{D, \xi_{s'_n}}^{(k)} \rangle + (s' - s) \langle \psi_k, \mathbf{1}_{D \setminus U} \psi_{D, \xi_{s'_n}}^{(k)} \rangle.
\end{aligned} \tag{4.20}$$

It follows that

$$\limsup_{s' \uparrow s} \frac{\lambda_D^{(k)}(\xi_{s'}) - \lambda_D^{(k)}(\xi_s)}{s - s'} = \frac{\langle \psi_k, \mathbf{1}_{D \setminus U} \psi_k \rangle}{\langle \psi_k, \psi_k \rangle} \tag{4.21}$$

and so, to get (4.17), it suffices to verify (4.5) for  $\psi := \psi_k$ ,  $D' := D \setminus U$  and  $A' := A + s$ .

First note that  $\lambda_D^{(k)}(\xi_s) \geq \lambda_D^{(1)}(\xi) - A$  and  $U \supset D_{R,A}(\xi)$  force  $\xi < \lambda_D^{(1)}(\xi) - 2A \leq \lambda_D^{(k)}(\xi_s) - A$  on  $D \setminus U$ . Hence,  $\xi_s \leq \lambda_k - A'$  on  $D'$ . Similarly,  $\xi_s(x) \geq \lambda_k - A$  forces  $\xi(x) \geq \lambda_D^{(1)}(\xi) - 2A$  which by  $U \supset D_{R,A}(\xi)$  implies  $\text{dist}(x, D') \geq R$ . The conditions of Lemma 4.2 are thus met for the field  $\xi := \xi_s$  and eigenvalue  $\lambda := \lambda_k$ , and so (4.5) holds for  $\psi := \psi_k$  as desired.  $\square$

The above facts are now assembled into the control of truncation from  $D$  to  $U$ :

*Proof of Theorem 2.1.* Pick  $A > 0$  and  $R \in \mathbb{N}$  such that the right-hand side of (2.6) is  $\leq A/2$ . Fix a  $k \in \{1, \dots, |U|\}$  such that (2.5) holds, i.e.,  $\lambda_D^{(k)}(\xi) \geq \lambda_D^{(1)}(\xi) - \frac{A}{2}$ . For  $s \geq 0$ , introduce the shifted field  $\xi_s := \xi - s \mathbf{1}_{D \setminus U}$  and define

$$\tilde{s} := \inf \{ s \geq 0 : \lambda_D^{(k)}(\xi_s) < \lambda_D^{(1)}(\xi) - A \}. \tag{4.22}$$

By continuity of  $s \mapsto \lambda_D^{(k)}(\xi_s)$ , we have  $\tilde{s} \in (0, \infty]$ . Our aim is to show that  $\tilde{s} = \infty$ .



Suppose on the contrary that  $\tilde{s} < \infty$ . The bound (4.15) then shows that  $\lambda_D^{(k)}(\xi) - \lambda_D^{(k)}(\xi_{\tilde{s}})$  is strictly less than the right-hand side of (2.6) which by our assumption is  $\leq A/2$ . Therefore,

$$\tilde{s} < \infty \implies \lambda_D^{(k)}(\xi_{\tilde{s}}) > \lambda_D^{(k)}(\xi) - \frac{A}{2} \geq \lambda_D^{(1)}(\xi) - A. \quad (4.23)$$

This contradicts (4.22) because by continuity and monotonicity of  $s \mapsto \lambda_D^{(k)}(\xi_s)$  we would have  $\lambda_D^{(k)}(\xi_{s'}) \geq \lambda_D^{(1)}(\xi) - A$  also for an interval of  $s'$  slightly larger than  $\tilde{s}$ . Hence  $\tilde{s} = \infty$  as claimed.

To complete the proof we note that, in light of (4.9), the difference  $|\lambda_D^{(k)}(\xi) - \lambda_{D'}^{(k)}(\xi)|$  is bounded by the integral in (4.15) with  $s := \infty$ . A calculation then yields (2.6).  $\square$

*4.3. Spectral gap and potential profiles.* As an aside of the general arguments above, we will give a proof of Proposition 2.2 and establish versions of the inclusions (2.13) and (2.15) that link local eigenvalues with potential profiles.

Let us first address Proposition 2.2. A main tool of the proof is an inequality between the second eigenvalue and the principal eigenvalues in nodal domains of the second eigenfunction. Note that  $\psi_{C,\xi}^{(2)}$ , the second eigenfunction of  $H_{C,\xi}$ , has at least one negative and one positive value, since it is assumed orthogonal to the principal eigenfunction which is known to be of one sign.

**Lemma 4.5.** *Let  $C \subset \mathbb{Z}^d$  be finite and define*

$$B := \{x \in C : \psi_{C,\xi}^{(2)}(x) \geq 0\}. \quad (4.24)$$

Then

$$\lambda_C^{(2)}(\xi) \leq \min\{\lambda_B^{(1)}(\xi), \lambda_{C \setminus B}^{(1)}(\xi)\}. \quad (4.25)$$

*Proof.* Abbreviate  $\psi := \psi_{C,\xi}^{(2)}$ . As  $\psi$  is orthogonal to the first eigenfunction, both  $B$  and  $C \setminus B$  are non-empty. The eigenvalue equation  $(\Delta + \xi)\psi = \lambda_C^{(2)}(\xi)\psi$  implies

$$\begin{aligned} \lambda_C^{(2)}(\xi) \|\psi \mathbf{1}_B\|_2^2 &= \sum_{x \in B} \psi(x)(\Delta + \xi)\psi(x) \\ &= \sum_{x \in B} \left( \xi(x)\psi(x)^2 + \sum_{y: |y-x|=1} \psi(x)(\psi(y) - \psi(x)) \right). \end{aligned} \quad (4.26)$$

Let  $\hat{\psi}$  be equal to  $\psi$  on  $B$  and zero on  $C \setminus B$ . Then

$$\psi(x)(\psi(y) - \psi(x)) \leq \hat{\psi}(x)(\hat{\psi}(y) - \hat{\psi}(x)) \quad (4.27)$$

for all pairs  $x \in B$  and  $y$  with  $|y-x|=1$ . By the Minimax Theorem, the sum in (4.26) with  $\hat{\psi}$  instead of  $\psi$  is bounded by  $\lambda_B^{(1)}(\xi) \|\hat{\psi}\|_2^2$  and since  $\|\hat{\psi}\|_2^2 = \|\psi \mathbf{1}_B\|_2^2 > 0$ , we thus get

$$\lambda_C^{(2)}(\xi) \leq \lambda_B^{(1)}(\xi). \quad (4.28)$$

The bound in terms of  $\lambda_{C \setminus B}^{(1)}(\xi)$  is completely analogous.  $\square$

*Proof of Proposition 2.2.* Let  $B$  be as in (4.24). As  $\mathcal{L}_C(\xi) = \mathcal{L}_B(\xi) + \mathcal{L}_{C \setminus B}(\xi)$ , we may assume without loss of generality that  $\mathcal{L}_B(\xi) \leq \frac{1}{2}\mathcal{L}_C(\xi)$ . Then

$$\lambda_C^{(2)}(\xi) - \rho \log \mathcal{L}_C(\xi) \leq \lambda_B^{(1)}(\xi) - \rho \log \mathcal{L}_B(\xi) - \rho \log 2 \leq -\chi_B - \rho \log 2. \quad (4.29)$$

Invoking (2.9) and  $\chi_B \geq \chi_C$ , which follows from (2.8) by way of Lemma 4.3, we get (2.10).  $\square$

Next let us move to the inclusion (2.13). We will in fact derive a stronger version by relating local eigenvalues to the quantity

$$\mathcal{L}_{C,A}(\varphi) := \sum_{x \in C} e^{\varphi(x)/\rho} \mathbf{1}_{\{\varphi(x) \geq -2A\}}, \quad (4.30)$$

which is better suited for our later needs. Clearly, for  $A := \infty$  this degenerates to  $\mathcal{L}_C(\varphi)$ .

**Lemma 4.6.** *For all  $a \in \mathbb{R}$ , all finite  $C \subset \mathbb{Z}^d$  and all  $A \geq \chi_C$  satisfying  $A(1 + \frac{A}{4d}) \geq 4d$ ,*

$$\{\xi : \lambda_C^{(1)}(\xi) \geq a\} \subseteq \{\xi : \mathcal{L}_{C,A}(\xi - a - \chi_C) \geq e^{-\eta(A)/\rho}\}, \quad (4.31)$$

where  $\eta(A) := 2d(1 + \frac{A}{4d})^{-1}$ .

*Proof.* Suppose  $\lambda_C^{(1)}(\xi) \geq a$ , let  $\varphi := \xi - a - \chi_C$  and note that  $\lambda_C^{(1)}(\varphi) \geq -\chi_C$ . Let us first address the case  $A := \infty$ . We claim that  $\mathcal{L}_C(\varphi) \geq 1$ . Indeed, if we had  $\mathcal{L}_C(\varphi) < 1$  then we could find an  $\varepsilon > 0$  such that  $\tilde{\varphi} := \varphi + \varepsilon$  would obey  $\mathcal{L}_C(\tilde{\varphi}) \leq 1$  and yet  $\lambda_C^{(1)}(\tilde{\varphi}) > -\chi_C$ . This would contradict (2.8). Hence (4.31) holds for  $A := \infty$ .

Now take  $A \geq \chi_C$ , let  $\varphi := \xi - a - \chi_C$  and set  $C' := \{x \in C : \varphi(x) \geq -2A\}$ . As  $A(1 + \frac{A}{4d}) \geq 4d$  holds and  $\lambda_C^{(1)}(\varphi) \geq -\chi_C$  implies  $\varphi \leq \lambda_C^{(1)}(\varphi) - A$  on  $C \setminus C'$ , Theorem 2.1 can be applied for the choices  $D := C$ ,  $U := C'$ ,  $R := 1$  and  $A$  replaced by  $A/2$  to show  $\lambda_{C'}^{(1)}(\varphi) \geq \lambda_C^{(1)}(\varphi) - \eta(A)$ . Therefore,  $\lambda_C^{(1)}(\xi) \geq a$  gives  $\lambda_{C'}^{(1)}(\xi) \geq a - \eta(A)$  and so, by the  $A := \infty$  part of the claim, we have  $\mathcal{L}_{C'}(\xi - a - \chi_{C'} + \eta(A)) \geq 1$ . In conjunction with  $\chi_{C'} \geq \chi_C$ , this yields

$$\begin{aligned} \mathcal{L}_{C,A}(\xi - a - \chi_C) &= \mathcal{L}_{C'}(\xi - a - \chi_C) \\ &\geq \mathcal{L}_{C'}(\xi - a - \chi_{C'} + \eta(A)) e^{-\eta(A)/\rho} \geq e^{-\eta(A)/\rho}, \end{aligned} \quad (4.32)$$

as desired.  $\square$

As an aside we recall from, e.g., Gärtner and den Hollander [14], that (2.8) implies that, for any finite  $C \subset \mathbb{Z}^d$ , there is  $\varphi_C : C \rightarrow \mathbb{R}$  such that  $\mathcal{L}_C(\varphi_C) \leq 1$  and  $\lambda_C^{(1)}(\varphi_C) = -\chi_C$ . Therefore,

$$\{\xi : \lambda_C^{(1)}(\xi) \geq a\} \supseteq \{\xi : \xi \geq a + \chi_C + \varphi_C\}. \quad (4.33)$$

This provides a bound that is in a sense opposite to (4.31).

Moving to the inclusion (2.15), similarly to (4.31) we will restate this using the truncated functional  $\mathcal{L}_{C,A}$  as follows:

**Lemma 4.7.** For  $C \subset \mathbb{Z}^d$  finite, all  $a, a' \in \mathbb{R}$  and all  $A \geq 0$  sufficiently large,

$$\lambda_C^{(1)}(\xi) \geq a' \text{ AND } \lambda_C^{(1)}(\xi) - \lambda_C^{(2)}(\xi) \leq \frac{1}{2}\rho \log 2 \implies \mathcal{L}_{C,A}(\xi - a - \chi_C) \geq u, \quad (4.34)$$

where  $u$  is defined by

$$\log u := \frac{a' - a - \eta(A)}{\rho} + \frac{1}{2} \log 2 \quad (4.35)$$

with  $\eta(A) := 2d(1 + \frac{A}{4d})^{-1}$ .

*Proof.* Let  $\xi$  be such that the conditions on the left of (2.15) apply and, similarly as in the proof of Lemma 4.6, let  $\varphi := \xi - a - \chi_C$  and set  $C' := \{x \in C : \varphi(x) \geq -2A\}$ . For  $A$  large Theorem 2.1 can be used; which then implies  $|\lambda_C^{(2)}(\xi) - \lambda_{C'}^{(2)}(\xi)| \leq \eta(A)$ . With the help of (4.29), this yields

$$\begin{aligned} a' &\leq \lambda_C^{(1)}(\xi) \leq \lambda_{C'}^{(2)}(\xi) + \frac{1}{2}\rho \log 2 \\ &\leq \lambda_{C'}^{(2)}(\xi) + \eta(A) + \frac{1}{2}\rho \log 2 \\ &\leq \rho \log \mathcal{L}_{C'}(\xi) - \chi_{C'} + \eta(A) - \frac{1}{2}\rho \log 2. \end{aligned} \quad (4.36)$$

Using  $\chi_{C'} \geq \chi_C$  and  $\mathcal{L}_{C'}(\xi) = e^{(a+\chi_C)/\rho} \mathcal{L}_{C,A}(\varphi)$ , we get the claim.  $\square$

By a variant of estimates used in Lemma 4.6, we will now control the spatial concentration of the fields that are near optimizers of (2.14). This will be useful in the derivation of spatial decay of the corresponding eigenfunctions.

**Lemma 4.8.** Define  $A_0$  by  $A_0(1 + \frac{A_0}{4d}) = 4d$  and suppose the quantities  $A, \delta > 0$  and  $A' := -\frac{1}{2}\rho \log(2 \sinh \delta)$  satisfy  $A \geq A' \geq d + A_0$  and  $\eta(A)/\rho \leq \delta$ . There is  $c = c(A, \delta) \in (0, \infty)$  such that for any  $C \subset \mathbb{Z}^d$  finite, any  $a \in \mathbb{R}$ , any  $r \geq 1$  and any  $\xi : C \rightarrow \mathbb{R}$  with

$$\mathcal{L}_{C,A}(\xi - a - \chi_C) \leq e^\delta \quad \text{and} \quad \lambda_C^{(1)}(\xi) \geq a + 2d \left(1 + \frac{A' - d}{2d}\right)^{1-2r}, \quad (4.37)$$

there is  $x \in C$  with the property

$$z \in C \ \& \ |z - x| \geq cr \implies \xi(z) \leq \lambda_C^{(1)}(\varphi) - A' + \chi_C. \quad (4.38)$$

*Proof.* Abbreviate  $\varphi(z) := \xi(z) - a - \chi_C$  and set  $S := \{z \in C : \varphi(z) > -2A'\}$ . As  $A' \leq A$ , every point of  $S$  contributes to  $\mathcal{L}_{C,A}(\varphi)$ . Hence,  $e^\delta \geq \mathcal{L}_{C,A}(\varphi) \geq e^{-2A'/\rho} |S|$  and so  $|S| \leq e^{\delta + 2A'/\rho}$ . Our goal is to use this to show that  $S$  also has a bounded diameter.

Given  $r \geq 1$  as above, suppose, for the sake of contradiction, that  $S$  has diameter larger than  $2|S|r$ . Then  $S$  can be split into two parts,  $S = S_1 \cup S_2$ , such that  $\text{dist}(S_1, S_2) > 2r$ . Pick  $x \in S_1$  and set  $C' := \{z \in C : \text{dist}(z, S_2) \leq r\}$ . Since  $0 < \text{dist}(z, S_2) \leq 2r$  implies

$$\xi(z) \leq a + \chi_C - 2A' \leq \lambda_C^{(1)}(\xi) + \chi_C - 2A' \leq \lambda_C^{(1)}(\xi) - 2(A' - d), \quad (4.39)$$

and  $A' - d \geq A_0$ , we may use Theorem 2.1 to conclude

$$\lambda_{C'}^{(1)}(\xi) \geq \lambda_C^{(1)}(\xi) - 2d \left(1 + \frac{A' - d}{2d}\right)^{1-2r}. \quad (4.40)$$

By (4.37) the right-hand side is at least  $a$  and so  $\mathcal{L}_{C',A}(\varphi) \geq e^{-\eta(A)/\rho} \geq e^{-\delta}$  thanks to Lemma 4.6, where the second bound comes from  $\delta \geq \eta(A)/\rho$ . But (4.37) and  $x \in C \setminus C'$  also yield

$$\mathcal{L}_{C',A}(\varphi) \leq \mathcal{L}_{C,A}(\varphi) - e^{\varphi(x)/\rho} \leq e^\delta - e^{\varphi(x)/\rho}, \quad (4.41)$$

whereby we get  $\varphi(x) \leq \rho \log(2 \sinh \delta) = -2A'$ . This contradicts  $x \in S$  and so  $\text{diam } S \leq 2|S|r$  must hold after all. Setting  $c := 2e^{\delta+2A/\rho}$ , the claim follows.  $\square$

*Remark 4.9.* We note that considerable effort has been devoted to the study of the minimizers in the variational problem (1.17); cf Gärtner and den Hollander [14]. In spite of that, uniqueness of the minimizer remains open for small  $\rho$ . In the same study, decay estimates for the eigenfunctions associated with the minimizing potential are provided.

## 5. Eigenvector localization: deterministic estimates

Our discussion of deterministic estimates proceeds by giving the proof of Theorem 2.5. We will rely heavily on Theorem 2.1 so, given  $A > 0$  and  $R \in \mathbb{N}$ , let us write

$$\varepsilon_R := 2d \left(1 + \frac{A}{2d}\right)^{1-2R} \quad (5.1)$$

for the error bound in (2.6). There are two main inputs into our proof of Theorem 2.5. The first of these is an inequality between the distance to the nearest eigenvalue and the ratio of masses that the eigenfunction puts on the boundary of a set relative to what it puts inside.

**Proposition 5.1 (Boundary mass vs gap).** *For the setting of Theorem 2.5 and “dist” denoting the  $\ell^1$ -distance on  $\mathbb{Z}^d$ , suppose that  $U \subset D$  is such that*

$$z \in D \text{ AND } \text{dist}(z, \partial U) \leq R \implies \xi(z) \leq \lambda_D^{(1)}(\xi) - 2A, \quad (5.2)$$

Put  $U' := D \setminus (U \cup \partial U)$  (hence  $\partial U = \partial U'$ ). Then

$$\max \left\{ \frac{\|\psi 1_{\partial U}\|_2}{\|\psi 1_U\|_2}, \frac{\|\psi 1_{\partial U'}\|_2}{\|\psi 1_{U'}\|_2} \right\} \geq \frac{\text{gap}(\lambda) - 2\varepsilon_R}{12d} \wedge 1. \quad (5.3)$$

The proof will require a specific comparison of the eigenvalues of  $\Delta + \xi$  in  $U$  with different boundary conditions on  $\partial U$ .

**Lemma 5.2 (Removing boundary condition).** *Let  $U \subset \mathbb{Z}^d$  be finite and let  $\tilde{\psi}: \partial U \rightarrow \mathbb{R}$  obey for simplicity  $\|\tilde{\psi} 1_{\partial U}\|_2 \leq 1$ . Recall that  $\lambda_U^{(k)}(\xi)$  is the  $k$ -th largest eigenvalue for operator  $\Delta + \xi$  in  $U$  with zero boundary condition, and let  $\tilde{\lambda}_U^{(k)}(\xi)$  be the  $k$ -th largest eigenvalue for the same operator with boundary condition  $\tilde{\psi}$ . Then*

$$|\lambda_U^{(k)}(\xi) - \tilde{\lambda}_U^{(k)}(\xi)| \leq 6d \|\tilde{\psi} 1_{\partial U}\|_2, \quad k = 1, \dots, |U|. \quad (5.4)$$

*Proof.* Let  $\tilde{\Delta}$  denote the Laplace operator on  $U$  with boundary condition  $\tilde{\psi}$  on  $\partial U$ . Then, for any function  $\psi: U \rightarrow \mathbb{R}$ ,

$$\langle \psi, (\tilde{\Delta} + \xi)\psi \rangle - \langle \psi, (\Delta + \xi)\psi \rangle = \sum_{\substack{x \in U, y \in \partial U \\ |x-y|=1}} (\tilde{\psi}(y)^2 - 2\psi(x)\tilde{\psi}(y)). \quad (5.5)$$

Assuming  $\|\psi\|_2 = 1$ , the Cauchy-Schwarz inequality and  $\|\tilde{\psi}1_{\partial U}\|_2 \leq 1$  tell us

$$\begin{aligned} & \left| \langle \psi, (\tilde{\Delta} + \xi)\psi \rangle - \langle \psi, (\Delta + \xi)\psi \rangle \right| \\ & \leq 2d\|\tilde{\psi}1_{\partial U}\|_2^2 + 4d\|\psi\|_2\|\tilde{\psi}1_{\partial U}\|_2 \leq 6d\|\tilde{\psi}1_{\partial U}\|_2. \end{aligned} \quad (5.6)$$

By the Minimax Theorem (see (4.10)),  $6d\|\tilde{\psi}1_{\partial U}\|_2$  bounds the difference between the  $k$ -th largest eigenvalue of operators  $\tilde{\Delta} + \xi$  and  $\Delta + \xi$ . Hence, (5.4) follows.  $\square$

*Proof of Proposition 5.1.* Let  $k \in \mathbb{N}$  be such that  $\lambda = \lambda_D^{(k)}(\xi)$ . Then the following facts hold:

- (1) By (5.2), the set  $\partial U$  is at least  $R$  steps of the simple random walk from any point where  $\xi$  exceeds  $\lambda_D^{(k)}(\xi) - 2A$ . Since (2.25) holds, Theorem 2.1 implies

$$|\lambda_D^{(k)}(\xi) - \lambda_{D \setminus \partial U}^{(k)}(\xi)| \leq \varepsilon_R. \quad (5.7)$$

- (2) Restricting the eigenvalue relation on  $D$  to  $U$ , resp.,  $U'$ , there are  $\ell_1, \ell_2 \in \mathbb{N}$  such that

$$\lambda_D^{(k)}(\xi) = \tilde{\lambda}_U^{(\ell_1)}(\xi) = \tilde{\lambda}_{U'}^{(\ell_2)}(\xi), \quad (5.8)$$

where  $\tilde{\lambda}_U^{(\ell_1)}(\xi)$ , resp.,  $\tilde{\lambda}_{U'}^{(\ell_2)}(\xi)$  is the  $\ell_1$ -th, resp.,  $\ell_2$ -th eigenvalue of  $\Delta + \xi$  in  $U$ , resp.,  $U'$  with boundary condition  $\tilde{\psi} := \psi/\|\psi1_U\|_2$ , resp.,  $\tilde{\psi}' := \psi/\|\psi1_{U'}\|_2$  on  $\partial U = \partial U'$ .

- (3) If the left-hand side of (5.3) is  $\geq 1$ , then there is nothing to prove, so let us assume that it is  $< 1$ . Lemma 5.2 then tells us

$$|\tilde{\lambda}_U^{(\ell_1)}(\xi) - \lambda_U^{(\ell_1)}(\xi)| \leq 6d \frac{\|\psi1_{\partial U}\|_2}{\|\psi1_U\|_2} \quad (5.9)$$

and

$$|\tilde{\lambda}_{U'}^{(\ell_2)}(\xi) - \lambda_{U'}^{(\ell_2)}(\xi)| \leq 6d \frac{\|\psi1_{\partial U}\|_2}{\|\psi1_{U'}\|_2}. \quad (5.10)$$

- (4) The Dirichlet eigenvalues in  $D \setminus \partial U$  consist of the union of Dirichlet eigenvalues in  $U$  and  $U'$ . It follows that there are  $k_1, k_2 \in \mathbb{N}$  such that

$$\lambda_U^{(\ell_1)}(\xi) = \lambda_{D \setminus U}^{(k_1)}(\xi) \quad \text{and} \quad \lambda_{U'}^{(\ell_2)}(\xi) = \lambda_{D \setminus U}^{(k_2)}(\xi). \quad (5.11)$$

Our goal is to show that  $k \in \{k_1, k_2\}$  and  $k_1 \neq k_2$ .

First, Lemma 4.2 tells us  $\|\psi1_{\partial U}\|_2 \leq \varepsilon_R/2d$  and so  $\|\psi1_U\|_2 \vee \|\psi1_{U'}\|_2 \geq \frac{1}{2}(1 - \varepsilon_R/2d) \geq 1/4$ . A calculation now shows that at least one of the right-hand sides in (5.9–5.10) is  $\leq 8\varepsilon_R$  — say the first one. But then

$$|\lambda_D^{(k)}(\xi) - \lambda_{D \setminus U}^{(k_1)}(\xi)| = |\tilde{\lambda}_U^{(\ell_1)}(\xi) - \lambda_U^{(\ell_1)}(\xi)| \leq 8\varepsilon_R < \text{gap}(\lambda) - 2\varepsilon_R. \quad (5.12)$$

Since by (5.7)  $s \mapsto \lambda_D^{(k)}(\xi_s)$  stays at least  $\text{gap}(\lambda) - 2\varepsilon_R$  away from other eigenvalues as  $s$  slides off to infinity, we must have  $k_1 = k$ . But if also  $k_2 = k$ , then  $H_{D \setminus \partial U}$  would have two (degenerate) eigenvalues equal to  $\lambda_{D \setminus \partial U}^{(k)}(\xi)$  and so, by (5.7),  $H_{D, \xi}$  would have another eigenvalue within  $2\varepsilon_R$  of  $\lambda_D^{(k)}(\xi)$ . This is again impossible because  $\text{gap}(\lambda) > 2\varepsilon_R$  and so  $k_1 \neq k_2$ .

The rest of the proof now boils down to the estimates:

$$\begin{aligned}
\text{gap}(\lambda) - 2\varepsilon_R &\leq |\lambda_{D \setminus \partial U}^{(k)}(\xi) - \lambda_{D \setminus \partial U}^{(k_2)}(\xi)| = |\lambda_U^{(\ell_1)}(\xi) - \lambda_U^{(\ell_2)}(\xi)| \\
&\leq |\tilde{\lambda}_U^{(\ell_1)}(\xi) - \lambda_U^{(\ell_1)}(\xi)| + |\tilde{\lambda}_U^{(\ell_2)}(\xi) - \lambda_U^{(\ell_2)}(\xi)| \\
&\leq 6d \frac{\|\psi \mathbf{1}_{\partial U}\|_2}{\|\psi \mathbf{1}_U\|_2} + 6d \frac{\|\psi \mathbf{1}_{\partial U}\|_2}{\|\psi \mathbf{1}_{U'}\|_2} \\
&\leq 12d \times \text{l.h.s. of (5.3)},
\end{aligned} \tag{5.13}$$

whereby (5.3) is finally proved.  $\square$

The second main input into the proof of Theorem 2.5 is a continuity argument which we again state in general terms as follows:

**Proposition 5.3 (Continuity argument).** *Fix  $\delta \in (0, 1)$  and  $h > 0$  so that the conditions expressed in (2.26–2.27) hold. Fix  $d(x, \mathcal{C})$  satisfying (D0–D2) and suppose that a normalized eigenfunction  $\psi$  corresponding to eigenvalue  $\lambda$  of  $H_{D, \xi}$  is such that, for some  $\mathcal{C} \in \mathcal{C}_{R, A}$ ,*

$$|\psi(z)| \leq e^{-\delta h d(z, \mathcal{C})}, \quad z \in D \setminus \mathcal{C}, \tag{5.14}$$

and, in addition,  $\|\psi \mathbf{1}_{\mathcal{C}}\|_2 \geq 1/4$ . Then we have  $\|\psi \mathbf{1}_{\mathcal{C}}\|_2 > 1/2$  and

$$|\psi(z)| < e^{-\delta h d(z, \mathcal{C})}, \quad z \in D \setminus \mathcal{C}. \tag{5.15}$$

*Proof.* Let us first consider  $z \in D \setminus D_{R, A}(\xi)$  and set  $U := B(z, R) \cap D$ . For a path  $(Y_j)_{j \in \mathbb{N}_0}$  of the simple random walk started at  $z$ , let  $\tau_R$  denote the hitting time of  $\partial U$ . Recall the martingale  $M^\tau = (M_{\tau \wedge n})_{n \in \mathbb{N}_0}$  from Lemma 4.1; we will consider it at time  $n := \tau_R$ . Note that if  $Y_{\tau \wedge \tau_R} \in \partial D$ , then  $\psi(Y_{\tau_R}) = 0$ ; otherwise, the path  $Y := (Y_0, Y_1, \dots, Y_{\tau \wedge \tau_R - 1})$  stays inside the set where  $\xi < \lambda$ . Since the loop-erasure of  $Y$  is an  $R$ -step self-avoiding nearest-neighbor path from  $z$  to  $\partial B(z, R)$ , the product of the terms in (4.2) is at most  $e^{-hR}$ , by (2.26). (All the terms that we drop by loop-erasing are  $\leq 1$ .) The Optional Stopping Theorem implies

$$|\psi(z)| \leq E^z(|M_{\tau \wedge \tau_R}|) \leq e^{-hR} \max_{y \in \partial B(z, R)} |\psi(y)|. \tag{5.16}$$

The assumption (5.14) and (D1) allow us to bound the maximum by  $e^{-\delta h[d(z, \mathcal{C}) - R]}$ . (Note that we can seamlessly extend the maximum to all of  $\partial B(z, R)$  because  $\psi$  vanishes outside  $D$ .) Rearranging terms we thus get

$$|\psi(z)| \leq e^{-(1-\delta)hR} e^{-\delta h d(z, \mathcal{C})}, \quad z \in D \setminus D_{R, A}(\xi). \tag{5.17}$$

As  $1 - \delta > 0$  and  $h > 0$ , this is even stronger than (5.15).

Our next goal is to boost the lower bound on  $\|\psi \mathbf{1}_{\mathcal{C}}\|_2$  from (non-strict)  $1/4$  to (strict)  $1/2$ . To this end let  $\kappa$  abbreviate the right-hand side of (5.3) and note that, by (5.3) with  $U := \mathcal{C}$  and  $U' := D \setminus (\mathcal{C} \cup \partial \mathcal{C})$ ,

$$\min\{\|\psi \mathbf{1}_{\mathcal{C}}\|_2, \|\psi \mathbf{1}_{D \setminus (\mathcal{C} \cup \partial \mathcal{C})}\|_2\} \leq \frac{1}{\kappa} \|\psi \mathbf{1}_{\partial \mathcal{C}}\|_2. \tag{5.18}$$

Since (5.17) applies to all  $z \in \partial \mathcal{C}$ , we can estimate  $\|\psi \mathbf{1}_{\partial \mathcal{C}}\|_2$  as

$$\frac{1}{\kappa} \|\psi \mathbf{1}_{\partial \mathcal{C}}\|_2 \leq \frac{1}{\kappa} \sqrt{|\partial \mathcal{C}'|} \max_{y \in \partial \mathcal{C}'} |\psi(y)| \leq \frac{1}{\kappa} \sqrt{|\partial \mathcal{C}'|} e^{-(1-\delta)hR}. \tag{5.19}$$

Invoking (2.27), this is less than  $\frac{1}{4}$ . By our assumption  $\|\psi 1_{\mathcal{C}}\|_2 \geq \frac{1}{4}$ , we thus conclude from (5.18) that  $\|\psi 1_{D \setminus (\mathcal{C} \cup \partial \mathcal{C})}\|_2 \leq \frac{1}{4}$ . But the same bound and  $\kappa \leq 1$  yield

$$\|\psi 1_{D \setminus \mathcal{C}}\|_2^2 = \|\psi 1_{D \setminus (\mathcal{C} \cup \partial \mathcal{C})}\|_2^2 + \|\psi 1_{\partial \mathcal{C}}\|_2^2 \leq \frac{1}{8}, \quad (5.20)$$

which implies  $\|\psi 1_{\mathcal{C}}\|_2 \geq \sqrt{7/8} > 1/2$ , as desired.

We are now ready to prove (5.15) for  $z \in D_{R,A}(\xi) \setminus \mathcal{C}$ . Let  $\mathcal{C}'$  denote the component that  $z$  belongs to. We will apply Proposition 5.1 with  $U := \mathcal{C}'$  and  $U' := D \setminus (\mathcal{C}' \cup \partial \mathcal{C}')$ . Since we already know  $\|\psi 1_{\mathcal{C}}\|_2 > 1/2$ , we have  $\|\psi 1_U\|_2 < \|\psi 1_{U'}\|_2$  and so the maximum in (5.3) is achieved by the term corresponding to  $U$ . Therefore

$$|\psi(z)| \leq \|\psi 1_{\mathcal{C}'}\|_2 \leq \frac{1}{\kappa} \|\psi 1_{\partial \mathcal{C}'}\|_2 \leq \frac{1}{\kappa} \sqrt{|\partial \mathcal{C}'|} \max_{y \in \partial \mathcal{C}'} |\psi(y)|. \quad (5.21)$$

By (5.17) we have  $|\psi(y)| \leq e^{-(1-\delta)hR - \delta h d(y, \mathcal{C})}$  for each  $y \in \partial \mathcal{C}'$  and invoking condition (D2) we can estimate  $d(y, \mathcal{C}) \geq d(z, \mathcal{C}) - 1$ . Putting the terms together, we get

$$|\psi(z)| \leq \left( \frac{1}{\kappa} \sqrt{|\partial \mathcal{C}'|} e^{\delta h - (1-\delta)hR} \right) e^{-\delta h d(z, \mathcal{C})}. \quad (5.22)$$

By (2.27) the prefactor in the large parentheses is  $\leq \frac{1}{4}$  for all  $\mathcal{C}'$ . The claim follows.  $\square$

*Proof of Theorem 2.5.* We begin by identifying the component  $\mathcal{C}$ . Let  $k$  be such that  $\lambda = \lambda_D^{(k)}(\xi)$ , let  $U := D \setminus D_{R,A}(\xi)$  and consider the deformation  $\xi_s := \xi - s 1_U$  with  $s \in (-\infty, \infty]$ . By the assumption that  $\text{gap}(\lambda) > 2\varepsilon_R$ , the eigenvalue  $\lambda_D^{(k)}(\xi_s)$  stays non-degenerate for all  $s \in [0, \infty]$ , since this eigenvalue and its two neighbors  $\lambda_D^{(k+1)}(\xi_s)$  and  $\lambda_D^{(k-1)}(\xi_s)$  change by less than  $\varepsilon_R$  as  $s$  slides from  $s := 0$  to  $s := \infty$ . In other words, there is no eigenvalue crossing along the path.

It follows that also the corresponding eigenfunction  $\psi_s$  changes continuously with  $s$  and its limit as  $s \rightarrow \infty$  exists and defines an eigenfunction  $\tilde{\psi}$  for  $\Delta + \xi$  in  $D \setminus U$  corresponding to  $\lambda_{D \setminus U}^{(k)}(\xi)$ . Clearly, there is a unique component of  $D \setminus U$  where  $\tilde{\psi}$  puts all of its mass, because otherwise  $\lambda_{D \setminus U}^{(k)}(\xi)$  would be at least two-fold degenerate. We let  $\mathcal{C}$  denote this component.

The bounds (5.14) for  $\psi$  replaced by  $\psi_s$  and  $\|\psi_s 1_{\mathcal{C}}\|_2 \geq 1/4$  are satisfied at  $s := \infty$ ; thanks to continuity in  $s$  they also hold for all  $s$  sufficiently large. Let  $s_0$  be the supremum of  $s \in \mathbb{R}$  where any of these bounds fails. We claim that  $s_0 < 0$ . Indeed, (assuming  $s_0 > -\infty$ ) the bounds still hold (by continuity) at  $s = s_0$ , and if  $s_0 \geq 0$ , then Proposition 5.3 would imply even the stronger bound  $\|\psi_{s_0} 1_{\mathcal{C}}\|_2 \geq 1/2$ . By continuity again, one could find  $\varepsilon > 0$  such that at least the bound  $\|\psi_s 1_{\mathcal{C}}\|_2 \geq 1/4$  would hold for all  $s \in [s_0 - \varepsilon, s_0]$ , in contradiction with the definition of  $s_0$ . In particular, (5.14) and  $\|\psi 1_{\mathcal{C}}\|_2 \geq 1/4$  hold. As the eigenfunction is normalized and  $d(z, \mathcal{C}) = 0$  for  $z \in \mathcal{C}$ , we have (2.28), as desired.  $\square$

## 6. Coupling to i.i.d. random variables

Having dealt with the deterministic estimates that underly the proof of our main theorems, we now move on to the corresponding set of probabilistic arguments. Our specific task here is to establish a coupling to i.i.d. random variables as stated in Theorem 2.3. We assume throughout that  $\xi = (\xi(x) : x \in \mathbb{Z}^d)$  are i.i.d. random variables satisfying Assumption 1.1.

6.1. *Extreme values of the fields.* We begin by discussing the extreme order statistics of the potential field. Since  $\xi(0)$  is continuously distributed, there is a unique  $\widehat{a}_L$  such that

$$\mathbb{P}(\xi(0) \geq \widehat{a}_L) = \frac{1}{L^d}. \quad (6.1)$$

Assumption 1.1 forces  $F(s) := \log \log[\mathbb{P}(\xi(0) > s)^{-1}]$  to grow, in the leading order, as a linear function with slope  $1/\rho$ . Simple estimates then show

$$\widehat{a}_L = (\rho + o(1)) \log \log L, \quad L \rightarrow \infty. \quad (6.2)$$

The quantity  $\widehat{a}_L$  plays the role of a centering sequence for the extreme order statistics of the field  $\xi$  in boxes of volume  $L^d$ . Indeed, since  $e^{F(\widehat{a}_L)} = d \log L$ , setting

$$b_L := \frac{\rho}{d \log L} \quad (6.3)$$

and using that  $F(\widehat{a}_L + sb_L) = F(\widehat{a}_L) + \frac{1}{\rho} sb_L(1 + o(1))$  (by Assumption 1.1), we get (for  $D \in \mathfrak{D}$  and  $D_L$  its scaled-up lattice version from (1.4))

$$\begin{aligned} \mathbb{P}\left(\max_{x \in D_L} \xi(x) \leq \widehat{a}_L + sb_L\right) &= (1 - \exp\{-e^{F(\widehat{a}_L + sb_L)}\})^{|D_L|} \\ &= \exp\{-|D_L| \exp\{-e^{F(\widehat{a}_L + sb_L)}\}(1 + o(1))\} \\ &= \exp\left\{-|D_L| \exp\left\{-e^{F(\widehat{a}_L) + \frac{1}{\rho} sb_L(1 + o(1))}\right\}(1 + o(1))\right\} \\ &\xrightarrow{L \rightarrow \infty} \exp\{-|D|e^s\}. \end{aligned} \quad (6.4)$$

It follows that the extreme-order law of  $\xi$  lies in the Gumbel universality class.

Our ultimate goal is to arrive at a similar conclusion also for the statistics of the maximal eigenvalues in  $D_L$ . However, this requires more regularity than the above estimates, for two reasons: First, to make a (local) eigenvalue relevant requires arranging a whole “profile” of  $\xi$  values. Second, none of these  $\xi$  values will reach into the extreme-order range — i.e., within  $O(b_L)$  of  $\widehat{a}_L$  — rather, they will all lie a positive deterministic constant below  $\widehat{a}_L$ ; cf (1.16). It turns out that we will need:

**Lemma 6.1.** *Suppose Assumption 1.1. Then*

(1) *For any  $\varepsilon > 0$  there is  $r_0 < \infty$  such that for all  $r, r'$  with  $r' \geq r \geq r_0$ ,*

$$e^{(\frac{1}{\rho} - \varepsilon)(r' - r)} \leq \frac{\log \mathbb{P}(\xi(0) > r')}{\log \mathbb{P}(\xi(0) > r)} \leq e^{(\frac{1}{\rho} + \varepsilon)(r' - r)}. \quad (6.5)$$

(2) *The law of  $\xi(0)$  has a density  $f$  with respect to the Lebesgue measure which then satisfies*

$$\lim_{r \rightarrow \infty} \frac{f(r + se^{-F(r)})}{f(r)} = e^{-s/\rho} \quad (6.6)$$

*locally uniformly in  $s \in \mathbb{R}$ .*



*Proof.* For (6.5), the ratio of the logs equals the exponential of  $F(r') - F(r)$ . Using the Fundamental Theorem of Calculus and Assumption 1.1, once  $r', r$  are sufficiently large and  $r' > r$ ,

$$\left(\frac{1}{\rho} - \varepsilon\right)(r' - r) \leq F(r') - F(r) \leq \left(\frac{1}{\rho} + \varepsilon\right)(r' - r). \quad (6.7)$$

This implies part (1) of the claim. For part (2), a calculation shows that

$$f(r) := F'(r) \mathbb{P}(\xi(0) > r) \log \mathbb{P}(\xi(0) > r) \quad (6.8)$$

is the probability density of  $\xi(0)$ , and so

$$\begin{aligned} & \frac{f(r + se^{-F(r)})}{f(r)} \\ &= \frac{F'(r + se^{-F(r)})}{F'(r)} \frac{\mathbb{P}(\xi(0) > r + se^{-F(r)})}{\mathbb{P}(\xi(0) > r)} \frac{\log \mathbb{P}(\xi(0) > r + se^{-F(r)})}{\log \mathbb{P}(\xi(0) > r)}. \end{aligned} \quad (6.9)$$

As  $\lim_{r \rightarrow \infty} F(r) = \infty$  for  $r \rightarrow \infty$  by (1.7), the first ratio on the right tends to one by (1.7) and the same applies to the third ratio by (6.7). As for the middle ratio, we note

$$\frac{\mathbb{P}(\xi(0) > r + se^{-F(r)})}{\mathbb{P}(\xi(0) > r)} = \exp\left\{-e^{F(r)}(e^{F(r+se^{-F(r)})-F(r)} - 1)\right\}. \quad (6.10)$$

The claim follows by invoking  $F(r + se^{-F(r)}) - F(r) = \frac{1}{\rho} se^{-F(r)}(1 + o(1))$ .  $\square$

With the help of (6.5) we now get the asymptotic formula (2.11). This yields the identity in (2.12) which implies that the only “profiles” of the field that we can realistically expect to see in  $D_L$  are those for which  $\varphi := \xi - \widehat{a}_L$  obeys  $\mathcal{L}(\varphi) \lesssim 1$ .

**6.2. Local eigenvalue estimates.** Many of our arguments that are to follow will require the following (rather crude) bounds for the principal eigenvalues in rectangular boxes of sublogarithmic size in  $L$ .

**Proposition 6.2.** *Let  $R_L \rightarrow \infty$  with  $R_L = o(\log L)$ . Then for each  $\delta > 0$  sufficiently small there is  $\varepsilon > 0$  such that*

$$\mathbb{P}(\lambda_{B_{R_L}}^{(1)}(\xi) \geq \widehat{a}_L - \chi + \delta) \leq L^{-d-\varepsilon} \quad (6.11)$$

and

$$\mathbb{P}(\lambda_{B_{R_L}}^{(1)}(\xi) \geq \widehat{a}_L - \chi - \delta) \geq L^{-d+\varepsilon} \quad (6.12)$$

for all  $L$  sufficiently large.

For the regime in (6.12), we will also need a statement with the quantifiers interchanged:

**Proposition 6.3.** *Let  $R_L \rightarrow \infty$  with  $R_L = o(\log L)$ . Then for each  $\varepsilon > 0$  there is  $\delta > 0$  such that*

$$\mathbb{P}(\lambda_{B_{R_L}}^{(1)}(\xi) \geq \widehat{a}_L - \chi - \delta) \leq L^{-d+\varepsilon} \quad (6.13)$$

Let us start with the upper bounds (6.11) and (6.13). Their proof will be based on the set inclusion (4.31) established in Lemma 4.6. For that we will need to show that the event on the right of (4.31) is dominated by configurations with nearly minimal value of  $\mathcal{L}_{C,A}(\xi - a - \chi_C)$ .

**Lemma 6.4.** *Let  $\rho > 0$  and  $A > 0$  be given. There is a constant  $c_{A,\rho}$  such that for all  $\alpha \in (0, 1/2)$ , all  $u \geq 2/3$ , all finite  $C \subset \mathbb{Z}^d$ , all  $L$  large enough and  $d' := d2^{-2\alpha/3}e^{-A\alpha/(3\rho)}$ ,*

$$\mathbb{P}(\mathcal{L}_{C,A}(\xi - \widehat{a}_L) \geq u) \leq c_{A,\rho} \left( \frac{4}{\alpha} \frac{A + \rho \log(4/3)}{\rho \log(4/3)} \right)^{|C|} L^{-d'u^{1-\alpha}}. \quad (6.14)$$

*Proof.* Fix  $\rho > 0$ ,  $A > 0$  and  $\varepsilon > 0$ . For  $u > 0$ , define

$$F_u := \{\xi : \mathcal{L}_{C,A}(\xi - \widehat{a}_L) \in [u, 2u]\}. \quad (6.15)$$

The event in (6.14) is covered by  $\bigcup_{n \geq 0} F_{2^n u}$  so it suffices to derive a good estimate on  $\mathbb{P}(F_u)$ . Note that  $2u > 1$  because we assumed  $u \geq 2/3$ .

Our first task is to derive a version of (2.11) without the  $o(1)$ -term. For this, let  $r_0$  be such that (6.5) holds. We will also assume that  $L$  is so large that  $\widehat{a}_L - A \geq r_0$ . We claim that then

$$-A \leq s \leq \rho \log(2u) \quad \Rightarrow \quad \mathbb{P}(\xi(0) \geq \widehat{a}_L + s) \leq (L^{-d(2u)^{-2\rho\varepsilon}e^{-A\varepsilon}})^{e^{s/\rho}}. \quad (6.16)$$

Indeed, let  $\rho'$  be defined by  $1/\rho' := 1/\rho + \varepsilon$ . Then (6.5) bounds the probability by  $L^{-d\theta}$  with  $\theta := e^{s/\rho - \varepsilon|s|}$ . Since  $s \leq \rho \log(2u)$  implies  $|s| \leq 2\rho \log(2u) - s$ , this is further bounded by  $L^{-d\theta'}$  where  $\theta' := (2u)^{-2\rho\varepsilon}e^{s/\rho'}$ . As  $e^{s/\rho'} \geq e^{-A\varepsilon + s/\rho}$ , (6.16) follows.

Next we will discretize the set of possible potential values to cover  $F_u$  by a finite union of sets. Let  $\delta := \varepsilon\rho^2 \log(2u)$  and  $S_{A,u} := \{-A + m\delta : m \in \mathbb{N}_0\} \cap (-\infty, \rho \log(2u)]$ . Define

$$\varphi_\varepsilon(x) := -A + \delta \lfloor (\xi(x) + A)/\delta \rfloor. \quad (6.17)$$

Then  $\varphi$  takes values in  $S_{A,u}$  and

$$\mathcal{L}_{C,A}(\xi - \widehat{a}_L) \geq u \quad \Rightarrow \quad \mathcal{L}_{C,A}(\varphi_\varepsilon) \geq ue^{-\delta/\rho} \geq u^{1-\rho\varepsilon}. \quad (6.18)$$

Let  $\{C_k : k = 1, \dots, 2^{|C|} - 1\}$  be an enumeration of all non-empty subsets of  $C$  and let  $\{\varphi_{k,j}\}$  be an enumeration of all functions  $\varphi_{k,j} : C_k \rightarrow S_{A,u}$ . (6.18) then gives

$$F_u \subseteq \bigcup_k \bigcup_{j : \mathcal{L}_{C_k}(\varphi_{k,j}) \geq ue^{-\delta/\rho}} \{\xi : \xi \geq \widehat{a}_L + \varphi_{k,j} \text{ on } C_k\}. \quad (6.19)$$

Denoting  $d'' := d2^{-2\rho\varepsilon}e^{-A\varepsilon}$ , the condition  $\mathcal{L}_{C,A}(\varphi_{k,j}) \geq ue^{-\delta/\rho}$  implies

$$\mathbb{P}(\xi : \xi \geq \widehat{a}_L + \varphi_{k,j} \text{ on } C_k) \leq L^{-d''u^{1-3\rho\varepsilon}}. \quad (6.20)$$

The total number of pairs  $(k, j)$  contributing to the union is at most

$$\left(1 + \frac{A + \rho \log(2u)}{\delta}\right)^{|C|} = \left(1 + \frac{A + \rho \log(4/3)}{\rho^2 \varepsilon \log(4/3)}\right)^{|C|}. \quad (6.21)$$

So if  $\varepsilon$  (so far arbitrary) is linked to  $\alpha$  via  $\alpha := 3\rho\varepsilon$ , we get

$$\mathbb{P}(F_u) \leq \left( \frac{4}{\alpha} \frac{A + \rho \log(4/3)}{\rho \log(4/3)} \right)^{|C|} L^{-d'u^{1-\alpha}} \quad (6.22)$$

with  $d'$  related to  $\alpha$  as in the claim. Since the prefactor is independent of  $u$ , the desired bound follows by summing (6.22) over  $u$  taking values in  $\{2^n u : n \geq 0\}$ . (Note that the restriction on  $L$  was independent of  $u$  and  $\alpha$ .)  $\square$

We can now prove the upper bounds in Propositions 6.2 and 6.3:

*Proof of (6.11) and (6.13).* For  $A \geq \chi_C$  and  $\delta \in \mathbb{R}$ , the inclusion (4.31) and  $\chi_C \geq \chi$  show

$$\{\lambda_C^{(1)}(\xi) \geq \widehat{a}_L - \chi + \delta\} \subseteq \{\xi : \mathcal{L}_{C,A}(\xi - \widehat{a}_L) \geq u\} \quad \text{for } u := e^{(\delta - \eta(A))/\rho}. \quad (6.23)$$

Applying (6.14) with  $C := B_{R_L}$ , and noting that the term exponential in  $|C|$  is  $L^{o(1)}$ , the desired probability is at most  $L^{d'u^{1-\alpha+o(1)}}$ , for any  $\alpha \in (0, 1/2)$ , provided  $\delta$  is small and  $A$  large so that  $u \geq 2/3$ . Now for  $A$  large and  $\alpha > 0$  small,  $d - d'$  can be made arbitrary small (positive) while  $u$ , which satisfies  $u > 1$  for  $\delta > 0$  and  $u < 1$  for  $\delta < 0$ , can be made as close to 1 as desired by choosing  $\delta$  small. This proves the desired bounds.  $\square$

Concerning the lower bound in (6.12), we first state:

**Lemma 6.5.** *For each  $\rho' > \rho$  there is  $K_{\rho'} < \infty$  such that for any finite  $C \subset \mathbb{Z}^d$ , any  $\delta \in [0, 1)$  and any  $L$ , we have*

$$\mathbb{P}(\lambda_C^{(1)} \geq \widehat{a}_L - \chi_C - \delta) \geq L^{-d\theta} \quad \text{for } \theta := e^{-\delta/\rho'} (1 + K_{\rho'} |C|^{\rho/\rho'} e^{-\widehat{a}_L/\rho'}). \quad (6.24)$$

*Proof.* Recall (4.33) and the notation used therein. Fix  $\rho' > \rho$  and let  $r_0$  be such that the bound on the right (6.5) holds for all  $r, r' \geq r_0$  with  $1/\rho - \varepsilon$  equal to  $1/\rho'$ . Setting  $\varphi'_C := \varphi_C \vee (r_0 - \widehat{a}_L + \delta)$  and noting that  $\varphi'_C \geq \varphi_C$ , we get

$$\begin{aligned} \mathbb{P}(\lambda_C^{(1)} \geq \widehat{a}_L - \chi_C - \delta) &\geq \mathbb{P}(\xi \geq \widehat{a}_L - \delta + \varphi'_C) \\ &\geq \exp\left\{-\log(L^d) e^{-\delta/\rho'} \sum_{x \in C} e^{\varphi'_C(x)/\rho'}\right\} \end{aligned} \quad (6.25)$$

But  $\rho' > \rho$  implies  $\sum_{x \in C} e^{\varphi(x)/\rho'} \leq \mathcal{L}_C(\varphi)^{\rho/\rho'}$  and for  $\mathcal{L}_C(\varphi'_C)$  we get

$$\mathcal{L}_C(\varphi'_C) \leq \mathcal{L}_C(\varphi_C) + |C| e^{(r_0 + \delta - \widehat{a}_L)/\rho}. \quad (6.26)$$

In light of  $\mathcal{L}_C(\varphi_C) \leq 1$  and a simple convexity estimate, we get (6.24) with the choice  $K_{\rho'} := e^{(r_0 + 1)/\rho'}$ .  $\square$

*Proof of (6.12).* Since  $\chi_{B_{R_L}} \downarrow \chi$  and  $|B_{R_L}| = (\log L)^{o(1)}$  while  $\widehat{a}_L \asymp \log \log L$ , the bound (6.24) with  $C := B_{R_L}$  yields (6.12) for  $L$  large as soon as  $d - \varepsilon > d e^{-\delta/\rho}$ .  $\square$

6.3. *Approximation by i.i.d. process.* We are now ready to assemble the arguments needed for the proof of Theorem 2.3. Let  $R_L$  be a sequence of integers subject to (2.17). As before, let  $\mathfrak{C}_{R_L, A}$  denote the set of connected components of  $D_{R_L, A}(\xi)$  for  $D := D_L$ .

**Lemma 6.6.** *For any  $A > 0$  there is an integer  $n_A < \infty$  such that*

$$\text{diam } \mathcal{C} \leq n_A R_L, \quad \forall \mathcal{C} \in \mathfrak{C}_{R_L, A}, \quad (6.27)$$

*occurs with probability tending to one as  $L \rightarrow \infty$ .*

*Proof.* This is a consequence of (2.11) and a straightforward union bound. Indeed, let  $F_{L, n}(x)$  denote the event that the box  $B_{nR_L}(x)$  contains at least  $n$  vertices  $z$  for which  $\xi(z) \geq \widehat{a}_L - \chi - 2A$ . Then

$$\{\exists \mathcal{C} \in \mathfrak{C}_{R_L, A} : \text{diam } \mathcal{C} > nR_L\} \subseteq \bigcup_{x \in B_{nL}(0)} F_{L, n}(x). \quad (6.28)$$

By (2.11) and a union bound we obtain

$$\mathbb{P}(F_{L, n}(x)) \leq |B_{nR_L}|^n L^{-dn\theta + o(1)} \quad \text{where } \theta := e^{-(\chi + 2A)/\rho}. \quad (6.29)$$

Since  $R_L = L^{o(1)}$ , as soon as  $n$  is so large that  $n\theta > 1$ , the probability of the union on the right of (6.28) will tend to zero as  $L \rightarrow \infty$ .  $\square$

Next we will focus attention on components where the eigenvalue is close to the optimal threshold  $\widehat{a}_L - \chi$ . For these components we get:

**Lemma 6.7.** *Given  $A > 0$  sufficiently large, there is  $\delta > 0$  such that the following holds with probability tending to one as  $L \rightarrow \infty$ : For any  $\mathcal{C} \in \mathfrak{C}_{R_L, A}$  that obeys  $\lambda_{\mathcal{C}}^{(1)}(\xi) \geq \widehat{a}_L - \chi - \delta$ ,*

- (1)  $\lambda_{\mathcal{C}}^{(1)}(\xi) - \lambda_{\mathcal{C}}^{(2)}(\xi) \geq \frac{1}{2}\rho \log 2$ ,
- (2)  $\mathcal{C} \subset B_{N_L}(x)$  for some  $x \in ((N_L + 1)\mathbb{Z})^d$  and, in addition,  $\text{dist}(\mathcal{C}, D_L^c) > N_L$ .

*If, in addition,  $\lambda_{\mathcal{C}'}^{(1)}(\xi) \geq \widehat{a}_L - \chi - \delta$  for some  $\mathcal{C}' \in \mathfrak{C}_{R_L, A}$  then*

- (3) *either  $\mathcal{C}' = \mathcal{C}$  or  $\text{dist}(\mathcal{C}', \mathcal{C}) > N_L$ .*

*Proof.* Assume that  $A$  is large and pick  $\mathcal{C} \in \mathfrak{C}_{R_L, A}$  that obeys  $\lambda_{\mathcal{C}}^{(1)}(\xi) \geq \widehat{a}_L - \chi - \delta$ . By Lemma 6.6, we may also assume  $\text{diam } \mathcal{C} \leq n_A R_L$ . For the claim in (1), if we had  $\lambda_{\mathcal{C}}^{(1)}(\xi) - \lambda_{\mathcal{C}}^{(2)}(\xi) \leq \frac{1}{2}\rho \log 2$ , then Lemma 4.7 with the choices

$$a := \widehat{a}_L - \chi_{\mathcal{C}} \quad \text{and} \quad a' := \widehat{a}_L - \chi - \delta \quad (6.30)$$

would yield  $\mathcal{L}_{\mathcal{C}, A}(\xi - \widehat{a}_L) \geq \sqrt{2}e^{-\eta(A)/\rho} =: u$ . By Lemma 6.4 and the fact that  $|\mathcal{C}| = O(R_L^d) = o(\log L)$ , the probability that a given set  $\mathcal{C}$  is a component with these properties is at most  $L^{-d\sqrt{u} + o(1)}$ . But there are at most  $L^d e^{O(R^d)} = L^{d+o(1)}$  ways to choose such a connected component in  $D_L$  and so (1) follows by a union bound and  $u > 1$ .

For (2) and (3), let us abbreviate  $r_L := n_A R_L$ . Given  $\varepsilon > 0$  let  $\delta > 0$  be as in Proposition 6.2. Since  $\mathcal{C} \subset B_{r_L}$  implies  $\lambda_{B_{r_L}}^{(1)}(\xi) \geq \lambda_{\mathcal{C}}^{(1)}(\xi)$ , from (6.13) we immediately have

$$\mathbb{P}\left(\exists \mathcal{C} \in \mathfrak{C}_{R_L, A} : x \in \mathcal{C}, \text{diam}(\mathcal{C}) \leq r_L, \lambda_{\mathcal{C}}^{(1)}(\xi) \geq \widehat{a}_L - \chi - \delta\right) \leq |B_{r_L}| L^{-d+\varepsilon} \quad (6.31)$$

for any  $x \in B_{nL}(0)$ . Now if (2) fails for some component  $\mathcal{C} \in \mathfrak{C}_{R_L, A}$ , then  $\mathcal{C}$  contains a vertex either in  $D_L \setminus \bigcup_{x \in ((N_L+1)\mathbb{Z})^d} B_{N_L}(x)$  or in  $\{x \in D_L : \text{dist}(x, D_L^c) \leq N_L\}$ . The former set has cardinality of order  $L^d N_L^{-1}$  while the latter has cardinality of order  $N_L L^{d-1}$  (indeed, thanks to the smoothness of  $\partial D$  we have  $|\partial D_L| = O(L^{d-1})$ ). Hence, for some constant  $c_1$ ,

$$\mathbb{P}((2) \text{ fails}) \leq c_1 (L^d N_L^{-1} + N_L L^{d-1}) R_L^d L^{-d+\varepsilon}, \quad (6.32)$$

For (3) a similar argument yields

$$\mathbb{P}(((3) \text{ fails})) \leq c_2 L^d N_L^d (R_L^d L^{-d+\varepsilon})^2. \quad (6.33)$$

Using (2.17–2.18), both of these tend to zero as  $L \rightarrow \infty$  once  $\varepsilon$  is small enough.  $\square$

Finally, we also need a (slightly more explicit) version of (1.16):

**Lemma 6.8.** *We have*

$$\lambda_{D_L}^{(1)}(\xi) - \widehat{a}_L \xrightarrow{L \rightarrow \infty} -\chi, \quad \text{in } \mathbb{P}\text{-probability.} \quad (6.34)$$

*Proof.* Fix  $\varepsilon > 0$  and consider the event that  $\lambda_{D_L}^{(1)}(\xi) \geq \widehat{a}_L - \chi - \varepsilon$ . Cover  $D_L$  by order  $(L/R_L)^d$  disjoint translates of  $B_{R_L}$ . By (6.12) and  $R_L = L^{o(1)}$ , with probability tending to one as  $L \rightarrow \infty$ , at least in one of these boxes the principal Dirichlet eigenvalue exceeds  $\widehat{a}_L - \chi - \delta$ . Since  $\lambda_{D_L}^{(1)}$  dominates all these eigenvalues, we get a lower bound in (6.34).

Next let us examine the event  $F := \{\lambda_{D_L}^{(1)}(\xi) \geq \widehat{a}_L - \chi + \varepsilon\}$ . Let  $A > 0$  and fix  $R_L \rightarrow \infty$  with  $R_L = o(\log L)$ . Assume, with the help of Lemma 6.6, that all components of  $\mathfrak{C}_{R_L, A}$  have diameter less than  $n_A R_L$ . Theorem 2.1 thus implies

$$\lambda_{D_L}^{(1)}(\xi) \leq \max_{x \in D_L} \lambda_{B_{n_A R_L}(x)}^{(1)}(\xi) + \frac{\varepsilon}{2} \quad (6.35)$$

with probability tending to one as  $L \rightarrow \infty$ . So on  $F$ , at least one of the boxes  $B_{n_A R_L}(x)$ , with  $x \in D_L$ , has  $\lambda_{B_{n_A R_L}(x)}^{(1)}(\xi) \geq \widehat{a}_L + \varepsilon/2$ . By Proposition 6.2, this has probability  $o(L^{-d})$  and, since  $D_L = O(L^d)$ , also an upper bound in (6.34) holds.  $\square$

We are now finally ready to establish the coupling of the top part of the spectrum in  $D_L$  to a collection of i.i.d. random variables.

*Proof of Theorem 2.3.* Fix  $A > 0$  large and let  $\delta > 0$  be such that the conclusions of Lemmas 6.6, 6.7 and 6.8 hold. Let  $U$  denote the set  $D_{R_L, A}$  for  $D := D_L$ . Take  $\delta' < \min\{\delta/2, A/2\}$ . Theorem 2.1 then implies, for all  $k = 1, \dots, |D_L|$ ,

$$\lambda_{D_L}^{(k)}(\xi) \geq \widehat{a}_L - \chi - \delta' \quad \Rightarrow \quad |\lambda_{D_L}^{(k)}(\xi) - \lambda_U^{(k)}(\xi)| \leq 2d \left(1 + \frac{A}{2d}\right)^{1-2R_L}. \quad (6.36)$$

But the spectrum in  $U$  is the union of the spectra in the components in  $\mathfrak{C}_{R_L, A}$  and, once conclusions (2-3) in Lemma 6.7 are in place, we only need to pay attention to components that are entirely contained, and single of that kind, in one of the boxes  $B_{N_L}^{(i)}$ ,  $i = 1, \dots, m_L$ . Since Lemma 6.7(1) tells us that we can also disregard all but the principal eigenvalue, if  $\mathcal{C} \subset B_{N_L}^{(i)}$  is such a component, Theorem 2.1 yields

$$|\lambda_{\mathcal{C}}^{(1)}(\xi) - \lambda_k(\xi)| \leq 2d \left(1 + \frac{A}{2d}\right)^{1-2R_L}. \quad (6.37)$$

Combining (6.36–6.37), the claim follows.  $\square$

## 7. Eigenvalue order statistics

Our next item of business is a proof of extreme order statistics for eigenvalues in  $D_L$  as  $L \rightarrow \infty$ . Having coupled the eigenvalues at the top of the spectrum of  $H_{D_L, \xi}$  to a collection of i.i.d. random variables — namely the principal eigenvalues in disjoint subboxes of side  $N_L$  — the argument is reduced to identifying the max-order class that these variables fall into.

*7.1. Determining the max-order class.* Our strategy is to first identify the max-order class for eigenvalues in boxes of side  $R_L$  and only then relate this to the eigenvalues in boxes of side  $N_L$ . The bulk of proof of Theorem 2.4 will be supplied by:

**Proposition 7.1.** *Suppose Assumption 1.1 holds and, for  $R_L \rightarrow \infty$  with  $R_L = (\log L)^{o(1)}$ , let  $a_L$  be as defined in (2.22). Let  $b_L$  obey (2.23), i.e.,*

$$b_L \log L \xrightarrow[n \rightarrow \infty]{} \frac{\rho}{d}. \quad (7.1)$$

Then, with  $o(1) \rightarrow 0$  as  $L \rightarrow \infty$  uniformly on compact set of  $s, r \in \mathbb{R}$ ,

$$\mathbb{P}(\lambda_{B_{R_L}}^{(1)} \geq a_L + rb_L) = e^{-s+o(1)} \mathbb{P}(\lambda_{B_{R_L}}^{(1)} \geq a_L + (r-s)b_L). \quad (7.2)$$

*Remark 7.2.* It is the proof of this proposition that requires us to assume that the law of  $\xi(0)$  has a density with respect to the Lebesgue measure. Although this restriction can be overcome to some extent, we have not succeeded in formulating a more general condition that would yield a comparably easy proof of the asymptotic (7.2). A natural idea how to deal with discontinuous laws would be to first approximate the spectrum by that of a continuously-distributed field and then apply the present approach.

The main idea of the proof of Proposition 7.1 is to compensate for a shift in the eigenvalue by way of a rigid shift of the field configuration. In order to keep the action confined to the asymptotic regime, we will only shift the values of  $\xi$  that are close to  $\widehat{a}_L$ . Given  $A > 0$  and  $L \geq 1$ , consider the continuous function  $g_{L,A}: \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$  given by

$$g_{L,A}(\xi, s) := \begin{cases} \xi - s, & \text{if } \xi \geq \widehat{a}_L - A, \\ \xi, & \text{if } \xi \leq \widehat{a}_L - 2A, \\ \text{linear,} & \text{else.} \end{cases} \quad (7.3)$$

Clearly, for  $s < A$ , the map  $\xi \mapsto g_{L,A}(\xi, s)$  is strictly increasing. The deterministic part of the change-of-measure argument is provided by:

**Lemma 7.3.** *Given a finite  $C \subset \mathbb{Z}^d$ , a configuration  $(\xi(x))_{x \in C}$  and  $A > 0$ , abbreviate*

$$\tilde{\xi}_s(x) := g_{L,A}(\xi(x), s). \quad (7.4)$$

Then for all  $a \in \mathbb{R}$  and  $s \geq 0$ ,

$$\{\lambda_C^{(1)}(\xi) \geq a + s\} \subseteq \{\lambda_C^{(1)}(\tilde{\xi}_s) \geq a\} \quad (7.5)$$

and, for all  $a \geq \widehat{a}_L - A/2$  and all  $s \geq 0$ ,

$$\{\lambda_C^{(1)}(\xi) \geq a + s\} \supseteq \{\lambda_C^{(1)}(\tilde{\xi}_{s'}) \geq a\}, \quad (7.6)$$

where  $s' := s/[1 - 2d(1 + \frac{A}{4d})^{-2}]$ .

*Proof.* Abbreviate  $O := \{x \in C : \xi(x) \geq \widehat{a}_L - A\}$  and note that  $\tilde{\xi}_s = \xi(x) - s$  for  $x \in O$ . Since  $s \geq 0$ , the variational characterization of the principal eigenvalue tells us

$$\lambda_C^{(1)}(\tilde{\xi}_s) + s \geq \lambda_C^{(1)}(\xi) \geq \lambda_C^{(1)}(\tilde{\xi}_s) + s \sum_{x \in O} |\psi_{C, \tilde{\xi}_s}^{(1)}(x)|^2. \quad (7.7)$$

The inequality on the left then immediately yields (7.5).

For (7.6), let  $s'$  be as given and let us assume  $\lambda_C^{(1)}(\tilde{\xi}_{s'}) \geq a$ . Then  $a \geq \widehat{a}_L - A/2$  and  $\tilde{\xi}_{s'} \leq \widehat{a}_L - A$  on  $C \setminus O$  imply  $\tilde{\xi}_{s'} \leq a - A/2 \leq \lambda_C^{(1)}(\tilde{\xi}_{s'}) - A/2$  on  $C \setminus O$  and thus by Lemma 4.2 with  $D' := C \setminus O$ ,  $A' := A$  and  $R := 1$  and  $A$  replaced by  $A/2$ ,

$$\sum_{x \in O} |\psi_{C, \tilde{\xi}_{s'}}^{(1)}(x)|^2 \geq 1 - 2d \left(1 + \frac{A}{4d}\right)^{-2}. \quad (7.8)$$

The inequality on the right of (7.7) with  $s$  replaced by  $s'$  yields the sequence of bounds  $\lambda_C^{(1)}(\xi) \geq \lambda_C^{(1)}(\tilde{\xi}_{s'}) + s \geq a + s$ .  $\square$

The shift of the field will give rise to a term reflecting the change in the underlying measure. This term can be evaluated rather explicitly. As already pointed out, the function  $\xi \mapsto g_{L,A}(\xi, s)$  is strictly increasing for  $s < A$  so we can define its inverse,  $h_{L,A}(\xi, s)$ , by

$$g_{L,A}(h_{L,A}(\xi, s), s) = \xi. \quad (7.9)$$

Then we have:

**Lemma 7.4.** *Let  $f$  be the probability density of  $\xi(0)$ . For any event  $G$  depending only on  $\{\xi(x)\}_{x \in C}$ , any  $A > 0$ , any  $s \in [0, A)$  and all  $L$  sufficiently large,*

$$\mathbb{P}(\tilde{\xi}_s \in G) = \mathbb{E} \left( \mathbf{1}_G \left( \frac{A}{A-s} \right)^{K_{L,A}(\xi, s)} \prod_{x \in C} \frac{f(h_{L,A}(\xi(x), s))}{f(\xi(x))} \right), \quad (7.10)$$

where  $K_{L,A}(\xi, s) := \#\{x \in C : A < \widehat{a}_L - h_{L,A}(\xi(x), s) < 2A\}$ .

*Proof.* Let  $L$  be so large that the probability density  $f$  is well defined and positive for all arguments larger than  $\widehat{a}_L - 2A$ . Notice the change of variables  $\xi \mapsto \tilde{\xi}_s$ , with explicit form  $\xi = h_{L,A}(\tilde{\xi}_s, s)$ , incurs the Radon-Nikodym derivative

$$\frac{d\xi}{d\tilde{\xi}_s} = \left( \frac{A}{A-s} \right)^{1_{\widehat{a}_L - \xi \in (A, 2A)}} \quad (7.11)$$

for the corresponding Lebesgue measures. Multiplying this by the ratio of the probability densities,  $f(\xi)/f(\tilde{\xi}_s)$  gives us the Radon-Nikodym derivative of the law of  $\xi$  with respect to the law of  $\tilde{\xi}_s$ . The result thus follows by writing  $\mathbb{P}(\tilde{\xi}_s \in G)$  as an integral with respect to the Lebesgue measure  $\prod_{x \in C} d\xi(x)$  and changing variables using  $\xi(x) = h_{L,A}(\tilde{\xi}_s(x), s)$  for each  $x$ .  $\square$

We will now proceed to deal with the Radon-Nikodym terms in (7.10). The ratios of the probability densities will be controlled using (6.6) as follows:

**Lemma 7.5.** *Let  $f$  denote the probability density of  $\xi(0)$ . For any finite  $C \subset \mathbb{Z}^d$ , any  $\varphi = (\varphi(x))_{x \in C}$  and  $\alpha = (\alpha(x))_{x \in C}$  with  $\alpha \geq 0$ , there exists a quantity  $o(1)$  such that*

$$\prod_{x \in C} \frac{f(\widehat{a}_L + \varphi(x) - b_L \alpha(x))}{f(\widehat{a}_L + \varphi(x))} = \exp \left\{ (1 + o(1)) \sum_{x \in C} \alpha(x) e^{\varphi(x)/\rho} \right\}. \quad (7.12)$$

Moreover,  $o(1) \rightarrow 0$  as  $L \rightarrow \infty$  uniformly in  $C$  and  $\varphi, \alpha \in I^C$ ,  $\alpha \geq 0$ , for any compact set  $I \subset \mathbb{R}$ .

*Proof.* For  $F$  be as in Assumption 1.1 and  $t \in \mathbb{R}$  such that  $\widehat{a}_L + t$  lies in the domain of  $F$ , let  $h_L(t)$  be defined by

$$\rho(1 + h_L(t))e^{t/\rho} := b_L e^{F(\widehat{a}_L + t)}. \quad (7.13)$$

Thanks to (2.23), Assumption 1.1 and the Mean-Value Theorem for  $F$ , we have that  $h_L(t) \rightarrow 0$  as  $L \rightarrow \infty$  locally uniformly in  $t$ . Next, for  $u \geq 0$  such that  $f(\widehat{a}_L + t - b_L u)$  and  $f(\widehat{a}_L + t)$  are well-defined and positive, let  $q_L(t, u)$  be defined by

$$\frac{f(\widehat{a}_L + t - b_L u)}{f(\widehat{a}_L + t)} =: \exp \left\{ (1 + q_L(t, u)) u e^{t/\rho} \right\}. \quad (7.14)$$

Using (6.6) with  $r := \widehat{a}_L + t$  and  $s := b_L u e^{F(\widehat{a}_L + t)}$ , and applying (7.13), we get that  $q_L(t, u) \rightarrow 0$  locally uniformly in  $t$  and  $u \geq 0$ .

Thanks to  $u \geq 0$ , (7.14) can be written as upper/lower bounds valid for  $L$  large once  $t, u$  are confined to compact sets in  $\mathbb{R}$ . Setting  $t := \varphi(x)$  and  $u := \alpha(x)$  and applying this bound to the product in (7.12), the claim follows.  $\square$

We are now ready to prove the main claim of this subsection:

*Proof of Proposition 7.1.* Fix  $\varepsilon > 0$  and  $A > 0$  and set

$$\begin{aligned} G := & \left\{ \xi : \lambda_{B_{R_L}}^{(1)}(\xi) \geq a_L + r b_L \right\} \cap \left\{ \xi : \mathcal{L}_{B_{R_L}, A}(\xi - \widehat{a}_L) \geq 1 - \varepsilon \right\} \\ & \cap \left\{ \xi : \mathcal{L}_{B_{R_L}, 2A}(\xi - \widehat{a}_L) \leq 1 + \varepsilon \right\} \cap \left\{ \xi : \max_{x \in B_{R_L}} \xi(x) \leq \widehat{a}_L + A \right\}. \end{aligned} \quad (7.15)$$

Our ultimate goal is to show that the right hand side of (7.10) with this  $G$  and  $s$  replaced by  $s b_L$  is asymptotically equal to  $e^{-s+o(1)} \mathbb{P}(\xi \in G)$ .

The first term in the expectation in (7.10) is bounded directly: Since  $s \geq 0$  we have, for  $L$  sufficiently large and all  $\xi$ , that

$$1 \leq \left( \frac{A}{A - s} \right)^{K_{L,A}(\xi, s)} \leq e^{O(1) s b_L R_L^d}, \quad (7.16)$$

where we estimated  $K_{L,A}(\xi, s)$  by the total volume of  $B_{R_L}$ . Since  $R_L = o(\log L)$  while  $b_L = O(1/\log L)$ , the right-hand side tends to one uniformly on compact sets of  $s$ .

For the product of ratios of probability densities, we will apply Lemma 7.5. Given a configuration  $\xi$ , let us abbreviate  $\varphi(x) := \xi(x) - A$  and define  $\alpha(x)$  implicitly by the formula  $\xi(x) - s b_L \alpha(x) := h_{L,A}(\xi(x), s b_L)$ . As is easy to check,  $\alpha(x) \in [0, 1]$  while  $\varphi(x) \in [-2A, A]$  for all  $\xi \in G$  where  $h_L(\xi(x), s b_L) \neq \xi(x)$ . Lemma 7.5 thus implies

$$\prod_{x \in B_{R_L}} \frac{f(h_{L,A}(\xi(x), s b_L))}{f(\xi(x))} = \exp \left\{ s(1 + o(1)) \sum_{x \in B_{R_L}} \alpha(x) e^{\varphi(x)/\rho} \right\}, \quad (7.17)$$



where  $o(1) \rightarrow 0$  as  $L \rightarrow \infty$  uniformly on  $G$ . Concerning the sum in the exponential on the right, here we note that  $\alpha(x) = 1$  when  $\xi(x) \geq \widehat{a}_L - A$  while  $\alpha(x) = 0$  when  $\xi(x) < \widehat{a}_L - 2A$ . Hence,

$$\mathcal{L}_{B_{R_L}A}(\xi - \widehat{a}_L) \leq \sum_{x \in B_{R_L}} \alpha(x) e^{\varphi(x)/\rho} \leq \mathcal{L}_{B_{R_L}, 2A}(\xi - \widehat{a}_L). \quad (7.18)$$

On  $G$  the left-hand side is at least  $1 - \varepsilon$  while the right-hand side is at most  $1 + \varepsilon$ . We conclude

$$\mathbb{P}(\tilde{\xi}_{sb_L} \in G) = e^{-s+O(\varepsilon)} \mathbb{P}(\xi \in G), \quad (7.19)$$

where  $O(\varepsilon)$  is bounded by a constant times  $\varepsilon$  uniformly on compact sets of  $s$ , for all  $A > 0$  sufficiently large (and larger than  $s$ ).

We are ready to put all the above together and extract the desired claim. First, Lemmas 6.4, 6.8 and 4.6 and the bound  $\mathbb{P}(\max_{x \in B_{R_L}} \xi(x) > \widehat{a}_L + A) = o(L^{-d})$  yield

$$\mathbb{P}(\xi \in G) = \mathbb{P}(\lambda_{B_{R_L}}^{(1)}(\xi) \geq a_L + rb_L) + o(L^{-d}), \quad L \rightarrow \infty. \quad (7.20)$$

Since  $\xi \geq \tilde{\xi}_s \geq \xi - sb_L$ , we similarly get

$$\mathbb{P}(\tilde{\xi}_{sb_L} \in G) = \mathbb{P}(\lambda_{B_{R_L}}^{(1)}(\tilde{\xi}_{sb_L}) \geq a_L + rb_L) + o(L^{-d}), \quad L \rightarrow \infty. \quad (7.21)$$

Plugging these into (7.19), invoking the inclusions (7.5–7.6) and noting that  $s'$  in (7.6) can be made arbitrarily close to  $s$  by increasing  $A$ , we conclude the claim for  $s \geq 0$ . For  $s < 0$  the claim follows by symmetry.  $\square$

**7.2. Stability with respect to partition size.** In order to conclude the proof of Theorem 2.4, we need to relate the upper tails of the law of the principal eigenvalues in  $B_{R_L}$  and  $B_{N_L}$ . Related to this is the question on how much does  $a_L$ , defined in (2.22), depend on the (rather arbitrary) choice of the sequence  $N_L$ . As attested by the next lemma, one direction is quite easy:

**Lemma 7.6.** *There exists a constant  $c = c(d) \in (0, \infty)$  such that for any  $N \geq R$  and any  $a \in \mathbb{R}$ ,*

$$-\log(1 - \mathbb{P}(\lambda_{B_N}^{(1)} \geq a)) \geq (1 - cR/N) \left(\frac{N}{R}\right)^d \mathbb{P}(\lambda_{B_R}^{(1)} \geq a). \quad (7.22)$$

*Proof.* Let us cover  $\mathbb{Z}^d$  by disjoint translates of  $B_R$  and let  $B_R^{(i)}$ ,  $i = 1, \dots, n$ , denote those translates that are contained in the box  $B_N$ . Then  $\lambda_{B_N}^{(1)}(\xi) \geq \lambda_{B_R^{(i)}}^{(1)}(\xi)$  for every  $i$  and since  $\lambda_{B_R^{(i)}}^{(1)}(\xi)$  are independent and equidistributed to  $\lambda_{B_R}^{(1)}(\xi)$ , we thus have

$$\begin{aligned} \mathbb{P}(\lambda_{B_N}^{(1)} \geq a) &\geq 1 - \mathbb{P}(\lambda_{B_R}^{(1)} < a)^n \\ &\geq 1 - \exp\{-n\mathbb{P}(\lambda_{B_R}^{(1)} < a)\}. \end{aligned} \quad (7.23)$$

The claim follows by taking a log and using that  $n \geq (1 - cR/N)(N/R)^d$  for some constant  $c > 0$ .  $\square$

Notice that (7.22) implies that once  $\mathbb{P}(\lambda_{B_{N_L}}^{(1)} \geq t_L) \rightarrow 0$  and  $R_L/N_L \rightarrow 0$  as  $L \rightarrow \infty$ , for some sequences  $R_L$ ,  $N_L$  and  $t_L$ , then also

$$\mathbb{P}(\lambda_{B_{N_L}}^{(1)} \geq t_L) \geq (1 + o(1)) \left(\frac{N_L}{R_L}\right)^d \mathbb{P}(\lambda_{B_{R_L}}^{(1)} \geq t_L), \quad L \rightarrow \infty. \quad (7.24)$$

(Indeed, just expand the log into a power series and dominate it by the first-order term.)

The bound in the opposite direction will require introducing an auxiliary scale  $R'_L$  as follows: Suppose, for the sake of present section, that  $R_L$  and  $N_L$  are sequences of integers such that

$$\lim_{L \rightarrow \infty} \frac{R_L}{\log \log L} = \infty, \quad \lim_{L \rightarrow \infty} \frac{R_L}{N_L} = 0 \quad \text{and} \quad \lim_{L \rightarrow \infty} \frac{N_L}{L} = 0 \quad (7.25)$$

and let  $R'_L$  be a sequence of integers satisfying

$$\lim_{L \rightarrow \infty} \frac{R'_L}{R_L} = 0. \quad (7.26)$$

Then we have:

**Lemma 7.7.** *For any  $A > 0$  and any sequence  $t_L \geq -A$  there is  $c > 0$  such that*

$$\mathbb{P}(\lambda_{B_{N_L}}^{(1)} \geq \widehat{a}_L + t_L) \leq o(L^{-d}) + (1 + o(1)) \left(\frac{N_L}{R_L}\right)^d \mathbb{P}(\lambda_{B_{R_L}}^{(1)} \geq \widehat{a}_L + t_L - e^{-cR'_L}), \quad (7.27)$$

as  $L \rightarrow \infty$ .

*Proof.* Pick  $A > 0$  such that  $t_L \geq -A$  and consider the set  $\mathcal{C}$  of connected components of the union of balls  $B_{R'_L}(x)$  for  $x \in B_{N_L}$  such that  $\xi(x) \geq \widehat{a}_L - 3A$ . By Lemma 6.6 (with  $R_L$  replaced by  $R'_L$ ) and  $N_L \leq L$ , there is an integer  $n_A > 0$  such that

$$\mathbb{P}(\max_{\mathcal{C} \in \mathcal{C}} \text{diam } \mathcal{C} > n_A R'_L) = o(L^{-d}). \quad (7.28)$$

Now consider a partition of  $\mathbb{Z}^d$  into disjoint translates of  $B_{R_L}$  and let  $B_{R_L}^{(i)}$ ,  $i = 1, \dots, m_L$ , denote those boxes in the covering that have at least one vertex in common with  $B_{N_L}$ . Considering the set  $S$  of all vertices on the inner boundary of these boxes, let  $B_{2n_A R'_L}^{(j)}$ ,  $j = 1, \dots, k_L$ , denote a covering thereof by translates of  $B_{2n_A R'_L}$  centered at these vertices such that no vertex in  $S$  lies in more than two boxes from these. The key point is that, on the event

$$G := \{\xi : \max_{\mathcal{C} \in \mathcal{C}} \text{diam } \mathcal{C} \leq n_A R'_L\}, \quad (7.29)$$

each component  $\mathcal{C} \in \mathcal{C}$  is entirely contained in one of the above boxes  $B_{R_L}^{(i)}$  or  $B_{2n_A R'_L}^{(j)}$ .

Since  $\xi(x) \geq \lambda_{B_{N_L}}^{(1)}(\xi) - 2A$  and  $\lambda_{B_{N_L}}^{(1)}(\xi) \geq \widehat{a}_L + t_L$  imply  $\xi(x) \geq \widehat{a}_L - 3A$ , Theorem 2.1 can be used for the set  $U := \bigcup_{\mathcal{C} \in \mathcal{C}} \mathcal{C}$ . Thereby we get

$$\lambda_{B_{N_L}}^{(1)}(\xi) \leq \max_{\mathcal{C} \in \mathcal{C}} \lambda_{\mathcal{C}}^{(1)}(\xi) + \varepsilon_{R'_L}, \quad (7.30)$$

where  $\varepsilon_{R'_L} := 2d(1 + \frac{A}{2d})^{1-2R'_L}$ . The monotonicity of  $C \mapsto \lambda_C^{(1)}(\xi)$  then shows that

$$\lambda_{B_{N_L}}^{(1)}(\xi) \leq \varepsilon_{R'_L} + \max \left\{ \max_{i=1, \dots, m_L} \lambda_{B_{R'_L}^{(i)}}^{(1)}(\xi), \max_{j=1, \dots, k_L} \lambda_{B_{2n_A R'_L}^{(j)}}^{(1)}(\xi) \right\} \quad (7.31)$$

holds on  $G \cap \{\lambda_{B_{N_L}}^{(1)}(\xi) \geq \widehat{a}_L + t_L\}$ .

Applying the union bound (7.31) and (7.28) yield

$$\begin{aligned} \mathbb{P}(\lambda_{B_{N_L}}^{(1)} \geq \widehat{a}_L + t_L) &\leq m_L \mathbb{P}(\lambda_{B_{R'_L}}^{(1)} \geq \widehat{a}_L + t_L - \varepsilon_{R'_L}) \\ &\quad + k_L \mathbb{P}(\lambda_{B_{2n_A R'_L}}^{(1)} \geq \widehat{a}_L + t_L - \varepsilon_{R'_L}) + o(L^{-d}) \\ &\leq (m_L + k_L) \mathbb{P}(\lambda_{B_{R'_L}}^{(1)} \geq \widehat{a}_L + t_L - \varepsilon_{R'_L}) + o(L^{-d}), \end{aligned} \quad (7.32)$$

where the last inequality holds because  $2n_A R'_L \leq R_L$ . Since  $R'_L \ll R_L$ ,

$$m_L = (1 + o(1)) \left(\frac{N_L}{R_L}\right)^d \quad \text{and} \quad k_L \leq O(1) \frac{N_L}{R_L} \left(\frac{N_L}{R_L}\right)^{d-1} = o(m_L) \quad (7.33)$$

the claim follows by noting that  $\varepsilon_{R'_L} \leq e^{-cR'_L}$  for some  $c > 0$ .  $\square$

**7.3. Proof of eigenvalue order statistics.** First we establish the Gumbel max-order tail for the principal eigenvalues:

*Proof of Theorem 2.4.* Choose  $r_L$  so that  $r_L / \log \log L \rightarrow \infty$  and (7.26) hold. Then  $e^{-cr_L} = o(b_L)$  as  $L \rightarrow \infty$ , for any  $c > 0$ . Combining Lemma 7.7 and Proposition 7.1, the desired claim follows.  $\square$

This then implies the extreme order law for eigenvalues:

*Proof of Corollary 1.3.* This is a direct consequence of Theorems 2.3 and 2.4, the standard results about max-order statistics and the facts that  $R_L / \log \log L \rightarrow \infty$  and  $\log |D_L| = (d + o(1)) \log L$  as  $L \rightarrow \infty$  for any  $D \in \mathfrak{D}$ .  $\square$

## 8. Eigenfunction decay

In this short section we will provide the arguments needed in the proof of eigenfunction localization. Recall our notation  $\varepsilon_R := 2d(1 + \frac{A}{2d})^{1-2R}$ . A key observation is:

**Lemma 8.1.** *Let  $R_L / \log \log L \rightarrow \infty$ . Then for each  $A > 0$  and each  $k \geq 2$ ,*

$$\frac{1}{\varepsilon_{R_L}} \min_{\ell=1, \dots, k-1} (\lambda_{D_L}^{(\ell)}(\xi) - \lambda_{D_L}^{(\ell+1)}(\xi)) \xrightarrow{L \rightarrow \infty} \infty \quad (8.1)$$

*in probability.*

*Proof.* By the convergence to the Poisson point process established in Corollary 1.3,

$$[\lambda_{D_L}^{(\ell)}(\xi) - \lambda_{D_L}^{(\ell+1)}(\xi)] \log L \xrightarrow{L \rightarrow \infty} \frac{\rho}{d} \log \frac{Z_1 + \dots + Z_{\ell+1}}{Z_1 + \dots + Z_\ell}, \quad (8.2)$$

where  $Z_1, Z_2, \dots$  are i.i.d. exponentials with parameter one. Since the right-hand side is positive with probability one, and  $\varepsilon_{R_L} \log L \rightarrow 0$ , the result follows.  $\square$

Next let us consider the distance  $d(x, \mathcal{C})$  defined, as an example, right before Theorem 2.5 and let  $\text{dist}(x, y)$  stand for the  $\ell^1$ -distance between  $x$  and  $y$  on  $\mathbb{Z}^d$ . Clearly,  $d(x, \mathcal{C}) \leq \text{dist}(x, \mathcal{C})$ . Our next item of business is to show:

**Lemma 8.2 (Comparison of distances).** *For each  $A > 0$  there are  $c_0, c_1, c_2 \in (0, \infty)$  such that for any  $R_L \rightarrow \infty$  that satisfies  $R_L \leq c_0 \log L$ , we have*

$$d(x, \mathcal{C}) \geq c_1 \text{dist}(x, \mathcal{C}) - c_2 R_L, \quad x \in D_L, \mathcal{C} \in \mathfrak{C}_{R_L, A}, \quad (8.3)$$

holds with probability tending to one as  $L \rightarrow \infty$ .

*Proof.* By its definition,  $d(x, y)$  is the (graph-theoretical) distance on the graph obtained by contracting each connected components  $\mathcal{C}$  to a single vertex. So it suffices to prove

$$d(x, y) \geq c_1 \text{dist}(x, y) - c_2 R_L, \quad x, y \in D_L. \quad (8.4)$$

Let  $G := \{\xi : \text{diam } \mathcal{C} \leq n_A R_L \forall \mathcal{C}\}$ , where the diameter is in the  $\ell^1$ -distance on  $\mathbb{Z}^d$ . Given  $x$  and  $y$ , consider a path  $\pi$  on this contracted graph achieving  $d(x, y)$ . This can be extended into a path on  $D_L$  by concatenating with paths inside the components, which yields

$$\text{dist}(x, y) \leq d(x, y) + Y n_A R_L, \quad \text{on } G, \quad (8.5)$$

where  $Y$  denotes the number of connected components encountered by  $\pi$ .

To estimate  $Y$ , consider any vertex self-avoiding path from  $x$  to  $y$  and let  $\mathcal{K}$  denote the union of  $B_{R_L}(z)$  for all  $z$  on this path. Clearly,  $|\mathcal{K}| \leq c \text{dist}(x, y) R_L^d$  for some constant  $c > 0$ . By a union bound and (2.11),

$$\mathbb{P}(Y \geq n) \leq |\mathcal{K}| L^{-n\theta} \leq c \text{dist}(x, y) R_L^d L^{-nd\theta} \leq L^{-n\theta'} \quad (8.6)$$

for some  $\theta, \theta' > 0$  and all  $n \geq 1$ , where we used that  $\text{dist}(x, y) R_L^d \leq L^{1+o(1)}$ . Hence, for any  $\eta > 0$ ,

$$\mathbb{P}\left(Y \geq \eta \frac{\text{dist}(x, y)}{R_L}\right) \leq \exp\left\{-\theta' \eta \frac{\log L}{R_L} \text{dist}(x, y)\right\}. \quad (8.7)$$

Summing this over  $x, y \in D_L$  with  $\text{dist}(x, y) \geq R_L/\eta^2$ , the result will tend to zero with  $L \rightarrow \infty$  provided  $\eta$  is sufficiently small. As also  $G$  has probability tending to one, we get

$$\text{dist}(x, y) \leq d(x, y) + \eta \frac{\text{dist}(x, y)}{R_L} n_A R_L, \quad (8.8)$$

implying  $d(x, y) \geq (1 - \eta n_A) \text{dist}(x, y)$  as soon as  $\text{dist}(x, y) > R_L/\eta^2$ , with probability tending to one as  $L \rightarrow \infty$ . As  $\eta$  can be chosen so that  $\eta n_A < 1$ , we are done.  $\square$

We are now ready to establish the eigenfunction decay starting first with decay over long distances:

*Proof of Theorem 1.4, large distances.* We will prove (1.21) at distances in excess of  $\log L / \log \log L$ . Our aim is to apply Theorem 2.5 for  $\lambda := \lambda_{D_L}^{(k)}(\xi)$ , any  $A > 0$  and  $R := R_L$ , where

$$R_L := \lfloor \log L / \log \log L \rfloor. \quad (8.9)$$

We will now check the conditions (1-3) of Theorem 2.5.

First, since  $\varepsilon_{R_L} = o(1/\log L)$ , condition (1) holds thanks to Lemma 8.1. Concerning condition (2), we note that as soon as

$$R_L \frac{\log(\mathbb{P}(\xi(0) > t_L)^{-1})}{\log L} \xrightarrow{L \rightarrow \infty} \infty \quad (8.10)$$

for some sequence  $t_L$ , then the following holds with probability tending to one as  $L \rightarrow \infty$ : For any self-avoiding path  $(x_1, \dots, x_n)$  in  $D_L$  of length  $n \geq R_L$ ,

$$\#\{i = 1, \dots, n: \xi(x_i) \geq t_L\} \leq \frac{n}{2}. \quad (8.11)$$

Assumption 1.1 tells us that  $\log(\mathbb{P}(\xi(0) > a)^{-1}) = e^{F(a)}$  with  $F(a) = (1/\rho + o(1))a$  as  $a \rightarrow \infty$ , we easily check that (8.10) holds for, say,  $t_L := \widehat{a}_L/2$ . Since  $\lambda = \widehat{a}_L - o(1)$ , the condition in (2.26) is valid with  $h := c \log \log L$  for some  $c > 0$ .

Concerning (3), by Lemma 6.6 (which holds as soon as  $R_L = o(L)$ ) all components of  $\mathcal{C}_{R_L, A}$  have diameter at most  $n_A R_L$ , with probability tending to one as  $L \rightarrow \infty$ . Condition (3) is then readily checked for any  $\delta \in (0, 1)$ .

Since the premises of Theorem 2.5 hold, we know that there is a connected component  $\mathcal{C} \in \mathcal{C}_{R_L, A}$  such that  $\psi := \psi_{D_L, \xi}^{(k)}$  obeys

$$|\psi(z)| \leq e^{-c\delta(\log \log L)d(z, \mathcal{C})}, \quad z \in D_L. \quad (8.12)$$

In particular,  $X_k$ , defined by (1.9), must satisfy  $\text{dist}(X_k, \mathcal{C}) \leq R_L$ . As Lemma 8.2 is at our disposal, we further conclude

$$\begin{aligned} d(z, \mathcal{C}) &\geq c_1 \text{dist}(z, \mathcal{C}) - c_2 R_L \\ &\geq c_1 |z - X_k| - (c_1 n_A + c_2 + 1) R_L, \end{aligned} \quad (8.13)$$

where we used that, by Lemma 6.6,  $\text{diam } \mathcal{C} \leq n_A R_L$  with probability tending to one as  $L \rightarrow \infty$ . Hereby we get

$$|z - X_k| \geq \frac{c_1 n_A + c_2 + 1}{2c_1} R_L \quad \Rightarrow \quad |\psi(z)| \leq e^{-c'(\log \log L) \text{dist}(z, \mathcal{C})}, \quad (8.14)$$

for  $c'$  given by  $c' := c\delta c_1/2$ . In particular, (1.21) is true.  $\square$

Before we move on to the short distances, let us abbreviate  $R'_L := \frac{c_1 n_A + c_2 + 1}{2c_1} R_L$  for  $R_L$  as in the previous proof and notice that (8.14) yields

$$\sum_{|z - X_k| \geq R'_L} \xi(z) |\psi(z)|^2 = o(1), \quad (8.15)$$

with probability tending to one as  $L \rightarrow \infty$  because  $\xi(z)$  is at most a constant times  $\log \log L$  in this limit. In particular, we have

$$\lambda_{B(X_k, R'_L)}^{(1)}(\xi) = \lambda - o(1) \quad (8.16)$$

by the variational characterization of the principal eigenvalue.

*Proof of Theorem 1.4, short distances.* It remains to prove exponential decay for distances less than order  $R_L$ . Set  $R'_L$  as right before this proof. Our aim is to use Lemma 4.8 to show that all but a finite number of values in  $B(X_k, R'_L)$  are more than a positive constant below  $\lambda$ . We will now proceed to verify the premises of Lemma 4.8.

Pick  $\varepsilon > 0$ . Thanks to Lemmas 6.4 and 6.8, and  $\chi_{B(X_k, R'_L)} \downarrow \chi$ , the following holds with probability one as  $L \rightarrow \infty$ : Once  $A$  is sufficiently large,

$$\mathcal{L}_{C,A}(\xi - (a_L - \varepsilon) - \chi_C) \leq e^{\varepsilon/(2\rho)} \quad (8.17)$$

holds for  $C$  ranging over all translates of  $B(X_k, R'_L)$  that intersect  $D_L$ . Set  $\delta := \varepsilon/(2\rho)$  and abbreviate  $C := B(X_k, R'_L)$ . Next observe that for  $A' := -\frac{1}{2}\rho \log(2 \sinh \delta)$  and  $\varepsilon$  small, there is  $r \geq 1$  so that

$$2d \left(1 + \frac{A' - d}{2d}\right)^{1-2r} \leq \frac{\varepsilon}{2}. \quad (8.18)$$

(Indeed,  $A' = \frac{1}{2} \log(2\delta) + O(1)$  so setting  $r$  to be proportional to a constant times  $\log(1/\varepsilon)$  will do once  $\varepsilon$  is small enough.) As  $\lambda = a_L + o(1)$ , the bound (8.16) tells us

$$\lambda_C^{(1)}(\xi) \geq a_L - \frac{\varepsilon}{2} \geq a_L - \varepsilon + 2d \left(1 + \frac{A' - d}{2d}\right)^{1-2r}. \quad (8.19)$$

By taking  $\delta$  small and  $A$  large, and setting  $A'' := A - d$  (which is less than  $A - \chi_C$ ) the premises of Lemma 4.8 are thus satisfied and so we conclude that  $\xi(z) \leq \lambda_C^{(1)}(\varphi) - A''$  for all  $z \in C$  that are at least  $cr$  away from  $X_k$ . As  $A'' > 0$ , Lemma 4.2 then shows

$$\sum_{|z - X_k| \geq a} |\psi(z)|^2 \leq \left(1 + \frac{A''}{2d}\right)^{-2a} \|\psi\|_2^2 \quad (8.20)$$

for all  $a > cr$ . In particular,  $\psi(z)$  decays exponentially with distance from  $X_k$ .  $\square$

Finally, we will supply a formal proof of our principal result:

*Proof of Theorem 1.2.* Part (1) is a direct consequence of the bounds in Theorem 1.4. For part (2) notice that argument producing the coupling to i.i.d. random variables in Theorem 2.3 is such that  $\lambda_{D_L}^{(k)}$  is coupled exactly to  $\lambda_{B_{N_L}^{(i)}}(\xi)$  for  $i$  such that  $X_k \in B_{N_L}^{(i)}$ . In the notation of this theorem, since  $N_L = o(L)$ , the collection (1.14) is well approximated by

$$\left\{ \left( \frac{z_i}{L}, \frac{1}{\rho} (\lambda_i(\xi) - a_L) \log |D_L| \right) : i = 1, \dots, m_L \right\}, \quad (8.21)$$

where  $z_i$  denotes the point at the center of  $B_{N_L}^{(i)}$ . As  $m_L \rightarrow \infty$  while the number of points with the second coordinate above a given value is with high probability bounded, this is in turn well approximated by sampling the first coordinate uniformly from  $\{z_i : i = 1, \dots, m_L\}$ , independently of the second coordinate. The last approximating process converges to a Poisson point process with intensity measure  $1_D dx \otimes e^{-\lambda} d\lambda$ .  $\square$

## Acknowledgments

This research has been partially supported by DFG Forschergruppe 718 “Analysis and Stochastics in Complex Physical Systems,” NSF grants DMS-0949250, DMS-1106850, NSA grant H98230-11-1-0171 and GAČR project P201-11-1558. The final parts of this paper were written when M.B. was visiting RIMS at Kyoto University, whose hospitality is gratefully acknowledged.

## References

1. M. Aizenman and S.A. Molchanov, Localization at large disorder and at extreme energies: an elementary derivation, *Commun. Math. Phys.* **157** (1993), no. 2, 245–278.
2. M. Aizenman, J.H. Schenker, R.M. Friedrich and D. Hundertmark, Finite-volume fractional-moment criteria for Anderson localization, *Commun. Math. Phys.* **224** (2001) 219–253.
3. P.W. Anderson, Absence of diffusion in certain random lattices, *Phys. Rev.* **109** (1958) 1492–1505.
4. A. Austraškas, Poisson-type limit theorems for eigenvalues of finite-volume Anderson hamiltonians, *Acta Appl. Math.* **96** (2007) 3–15.
5. A. Austraškas, Extremal theory for spectrum of random discrete Schrödinger operator. I. Asymptotic expansion formulas, *J. Statist. Phys.* **131** (2008) 867–916.
6. A.A. Balkema and L. de Haan, Limit distribution for order statistics, I., *Theory Prob. Appl.* **23** (1978), no. 1, 77–92.
7. A.A. Balkema and L. de Haan, Limit distribution for order statistics, II., *Theory Prob. Appl.* **23** (1978), no. 2, 341–358.
8. M. Biskup and W. König, Long-time tails in the parabolic Anderson model with bounded potential, *Ann. Probab.* **29** (2001), no. 2, 636–682.
9. M. Biskup and W. König, Screening effect due to heavy lower tails in one-dimensional parabolic Anderson model, *J. Statist. Phys.* **102** (2001), no. 5/6, 1253–1270.
10. M. Biskup, W. König and R. dos Santos, Mass concentration in one island for the parabolic Anderson model with doubly exponential tails, in preparation.
11. R. Carmona and J. Lacroix, *Spectral Theory of Random Schrödinger Operators*, Probability and Its Applications, Birkhäuser, Boston, 1990.
12. A. Fiodorov and S. Muirhead, Complete localisation and exponential shape of the parabolic Anderson model with Weibull potential field, preprint (arXiv:1311.7634).
13. J. Fröhlich and T.C. Spencer, Absence of diffusion in the Anderson tight binding model for large disorder or low energy, *Commun. Math. Phys.* **88** (1983) 151–184.
14. J. Gärtner and F. den Hollander, Correlation structure of intermittency in the parabolic Anderson model, *Probab. Theory Rel. Fields* **114** (1999) 1–54.
15. J. Gärtner and W. König, The parabolic Anderson model. In: J.-D. Deuschel and A. Greven (Eds.), *Interacting Stochastic Systems*, pp. 153–179, Springer, 2005.
16. J. Gärtner, W. König and S. Molchanov, Geometric characterization of intermittency in the parabolic Anderson model, *Ann. Probab.* **35** (2007), no. 2, 439–499.
17. J. Gärtner and S. Molchanov, Parabolic problems for the Anderson model I. Intermittency and related topics, *Commun. Math. Phys.* **132** (1990) 613–655.
18. J. Gärtner and S. Molchanov, Parabolic problems for the Anderson model II. Second-order asymptotics and structure of high peaks, *Probab. Theory Rel. Fields* **111** (1998) 17–55.
19. F. Germinet and F. Klopp, Spectral statistics for random Schrödinger operators in the localized regime, *J. Europ. Math. Soc.* **16:9** (2014) 1967–2031.
20. F. Germinet and F. Klopp, Enhanced Wegner and Minami estimates and eigenvalue statistics of random Anderson models at spectral edges, *Ann. Henri Poincaré* **14** (2013), no. 5, 1263–1285.
21. L.N. Grenkova, S. A. Molchanov, Yu. N. Sudarev, On the basic states of one-dimensional disordered structures, *Commun. Math. Phys.* **90** (1983) 101–124.
22. L.N. Grenkova, S. A. Molchanov, Yu. N. Sudarev, Structure of spectrum edge for multidimensional Anderson model, *Teoret. Mat. Fiz.* **85** (1990), no. 1, 32–40.
23. L. de Haan and A. Ferreira, *Extreme Value Theory. An Introduction*. Springer Series in Operations Research and Financial Engineering. Springer, New York, 2006.
24. R. van der Hofstad, W. König and P. Mörters, The universality classes in the parabolic Anderson model, *Commun. Math. Phys.* **267** (2006), no. 2, 307–353.
25. R. van der Hofstad, P. Mörters and N. Sidorova, Weak and almost sure limits for the parabolic Anderson model with heavy-tailed potential, *Ann. Appl. Probab.* **18** (2008) 2450–2494.

26. D. Hundertmark, A short introduction to Anderson localization. In: *Analysis and stochastics of growth processes and interface models*, 194-218, Oxford Univ. Press, Oxford, 2008.
27. R. Kilip and F. Nakano, Eigenfunction statistics in the localized Anderson model, *Ann. Henri Poincaré* **8** (2007), no. 1, 27–36.
28. W. König, *The Parabolic Anderson Model*, Springer Briefs Series (to appear).
29. W. König, H. Lacoïn, P. Mörters and N. Sidorova, A two cities theorem for the parabolic Anderson model, *Ann. Probab.* **37** (2009), no. 1, 347–392.
30. N. Minami, Local fluctuation of the spectrum of a multidimensional Anderson tight binding model, *Commun. Math. Phys.* **177** (1996), no. 3, 709–725.
31. S. Molchanov. The local structure of the spectrum of the one-dimensional Schrödinger operator, *Commun. Math. Phys.* **78**, 429–446 (1981).
32. L. Pastur and A. Figotin, *Spectra of Random and Almost-Periodic Operators*, Springer-Verlag, Berlin, 1992.
33. N. Sidorova and A. Twarowski, Localisation and ageing in the parabolic Anderson model with Weibull potential, *Ann. Probab.* **42** (2014), no. 4, 1666–1698.
34. B. Simon and T. Wolff, Singular continuous spectrum under rank one perturbations and localization for random Hamiltonians, *Commun. Pure Appl. Math.* **39** (1986), no. 1, 75–90.
35. P. Stollmann, *Caught by Disorder. Bound States in Random Media*, Progress in Mathematical Physics vol. 20, Birkhäuser Boston, Inc., Boston, MA, 2001.
36. I. Veselić, *Existence and Regularity Properties of the Integrated Density of States of Random Schrödinger Operators*, Lecture Notes in Mathematics, vol. 1917, Springer-Verlag, Berlin, x+142 pp, 2008.
37. F. Wegner, Bounds on the density of states in disordered systems. *Z. Phys. B* **44** (1981) 9–15.

Communicated by L. Erdős