TOP OF THE SPECTRUM OF DISCRETE ANDERSON HAMILTONIANS WITH CORRELATED GAUSSIAN POTENTIALS

GIUSEPPE CANNIZZARO, CYRIL LABBÉ, AND WILLEM VAN ZUIJLEN

ABSTRACT. We investigate the top of the spectrum of discrete Anderson Hamiltonians with correlated Gaussian noise in the large volume limit. The class of Gaussian noises under consideration allows for long-range correlations. We show that the largest eigenvalues converge to a Poisson point process and we obtain a very precise description of the associated eigenfunctions near their localisation centres. We also relate these localisation centres with the locations of the maxima of the noise. Actually, our analysis reveals that this relationship depends in a subtle way on the behaviour near 0 of the covariance function of the noise: in some situations, the largest eigenfunctions are *not* associated with the largest values of the noise.

AMS 2010 subject classifications: Primary 60H25; Secondary 82B44, 60G70. Keywords: Schrödinger operator; Anderson Hamiltonian; Gaussian field; order statistics; eigenvalue; eigenfunction; correlation

This version has some additional explanations and calculations in this colour.

CONTENTS

1.	Introducti	ion and main results	1
2.	Local behaviour of correlated Gaussian fields		11
3.	. Statistics of the maxima of correlated Gaussian fields		22
4.	The mesoscopic eigenproblem		30
5.	5. The macroscopic eigenproblem		42
Appendix A. Gaussian estimates		49	
Ap	pendix B.	Basic properties and estimates on the quadratic forms	51
References		53	

1. INTRODUCTION AND MAIN RESULTS

The present article is concerned with the behaviour of the top eigenvalues / eigenfunctions of random operators of the form $\Delta + \xi$ on $Q_L \stackrel{\text{def}}{=} [-L/2, L/2]^d \cap \mathbb{Z}^d$ in the limit $L \to \infty$. Here ξ is a random potential on \mathbb{Z}^d and Δ is the discrete Laplacian:

$$\Delta f(x) = \sum_{y \in \mathbb{Z}^d: y \sim x} (f(y) - f(x)) , \quad x \in \mathbb{Z}^d .$$

Such operators are often called random Schrödinger operators, or Anderson Hamiltonians. They are considered in physics to model the Hamiltonian of a quantum particle evolving in a crystal subject to defects or impurities. They are named after P.W. Anderson due to his seminal paper [And58] which discusses the

localisation of the quantum particle for large enough disorder of the potential and had a profound and lasting impact on the field. We refer to [CL90, Kir08, AW15] for some references which address a part of this literature, in particular that regarding Anderson localisation, i.e., the property of having a pure point spectrum and exponentially decaying eigenfunctions.

These operators also naturally arise in the mathematical study of the so-called parabolic Anderson model:

$$\partial_t u = \Delta u + \xi u$$
, $u(0, \cdot) = \delta_0(\cdot)$.

Indeed, the behaviour of the solution u at a large time t can be well-approximated by the solution of the same stochastic partial differential equation but restricted to a finite ball of growing size L = L(t), so that the top of the spectrum of the operator typically provides an accurate description of the growth and spreading of this solution, see for instance [GKM07, KLMS09, ST14, FM14] and the book of König [Kön16]. We also refer to [GKM00, KPvZ22, GY23] for articles on this topic in the continuous setting.

Most of the literature on these questions concern potentials ξ made of i.i.d. random variables with common distribution μ . It is now well-understood that the right tail of μ plays a prominent role in the behaviour of the top of the spectrum, in particular: the heavier the right tail of μ is, the more localised the top eigenfunctions are. To illustrate this, let us present informally two important classes of distributions:

- (Single peak): $\mu([x,\infty))$ decays "slowly" as $x \to \infty$ (for instance Gaussian, exponential or Pareto distributions). In the limit $L \to \infty$, the top eigenfunctions are asymptotically given by Dirac masses localised at i.i.d. uniform r.v.'s drawn from $[-L/2, L/2]^d \cap \mathbb{Z}^d$.
- (Doubly-exponential): $\mu([x, \infty))$ behaves like $\exp(-Ce^{x/\varrho})$ for some $C, \varrho > 0$ as $x \to \infty$. In the limit $L \to \infty$, the top eigenfunctions vary at scale 1 and are "supported" on balls of unbounded radius centred at i.i.d. uniform r.v., see [BK16].

As it will be useful for later comparisons, let us mention a special class of laws, the Weibull distributions, which are such that $\mu([x,\infty)) = \exp(-Cx^q), x \ge 0$, for some q > 1 and C > 0. They fall into the Single peak case, and precise results on the top of the spectrum of the Anderson Hamiltonian were established in [Ast08, Ast16].

The relationship between the localisation centres of the top eigenfunctions and the successive maxima of the potential ξ was investigated by Astrauskas [Ast13]. For Weibull tails (and more generally, in the Single peak class), a natural guess would be that, in the limit $L \to \infty$, the localisation center $x_{k,L} \in Q_L$ of the k-th eigenfunction is such that $\xi(x_{k,L})$ is the k-th largest value reached by ξ on Q_L . The situation is actually subtler: if we denote by $\ell_L(k)$ the integer such that $\xi(x_{k,L})$ is the $\ell_L(k)$ -th largest value of ξ over Q_L , then with large probability as $L \to \infty$:

- if q < 3, $\ell_L(k) = k$,

- if q = 3, $\ell_L(k)$ is a non-trivial r.v. of order 1,
- if q > 3, the r.v. $\ell_L(k)$ goes to ∞ .

Heuristically, when the right tail of μ is not so heavy $(q \ge 3)$, one has to take into account the behaviour of ξ at the nearest neighbours of the successive maxima:

the negligible mass that the eigenfunction puts on these neighbouring sites may produce a shift in the eigenvalue that compensates for the difference between successive maxima and, thereby, makes the correspondence between successive maxima and successive eigenvalues / eigenfunctions non trivial. Let us mention that in the article [Ast13], there are no precise statements that explain how this shift is produced.

Very little is known on the top of the spectrum of the Anderson Hamiltonian when ξ is a correlated field: in [Ast03], a few results were collected on the asymptotic behaviour of the potential and of the Anderson Hamiltonian for a Gaussian correlated field, while in [GM00] the asymptotic of the moments of the parabolic Anderson model with a correlated field were investigated. Let us also cite [GKM00] for the almost sure asymptotic of the parabolic Anderson model with a correlated Gaussian field in the continuum.

In the present article, we initiate a comprehensive study of the Anderson Hamiltonian with a correlated Gaussian field and we aim at answering the following questions:

- (1) What features of the covariance function of the field are relevant to determine the statistics of the top of the spectrum?
- (2) How do the top eigenvalues / eigenfunctions behave?
- (3) What is the relationship between the top eigenvalues / eigenfunctions, and the successive maxima of the field?

Actually, we consider a more general framework where the potential is allowed to depend on the size L at which we consider the Anderson Hamiltonian: more precisely, we give ourselves a sequence $(\xi_L)_{L\geq 1}$ of Gaussian potentials, and we investigate the above questions on the operator $\mathcal{H}_L \stackrel{\text{def}}{=} \Delta + \xi_L$ on Q_L . We work under two main assumptions on our field. The first condition concerns the longrange decay: roughly speaking, the covariance function is required to decay at infinity faster than $1/\log |x|$. This condition ensures that the statistics of our field behave in a way similar to that of the i.i.d. case. However, to encompass such long-range correlations in the study of the Anderson Hamiltonian requires substantial technical work. The second condition concerns the short-range decay: the covariance function is assumed to decay fast enough near 0. Actually, our study reveals that the behaviour near the origin of the covariance function of ξ_L has a subtle impact on Question (3), and we identify non-trivial relationships between the top of the spectrum and the maxima of the field.

1.1. The potential. Let us begin by rigorously introducing the (family of) Gaussian field(s) the present paper is concerned with.

Definition 1.1. For any integer $L \geq 1$, let $(\xi_L(x))_{x \in \mathbb{Z}^d}$ be a centred Gaussian field, stationary in law under spatial shifts, with unit variance at every point, and non-negative covariance function v_L on \mathbb{Z}^d . Further, we assume that v_L is such that

(I) (Long-Range decay) its tails \mathcal{T}_{v_L} satisfy

$$\mathcal{T}_{v_L} \stackrel{\text{def}}{=} \sup_{|x| \ge \exp(\sqrt{\ln L})} v_L(x) \ln |x| \longrightarrow 0 , \quad \text{as } L \to \infty .$$
(1.1)

(II) (Short-Range decay) there exist $\mathfrak{c}, \mathfrak{c}' > 0$ such that for all $L \ge 1$

$$1 - \frac{e^{\mathfrak{c}'|x|}}{d_L} \le v_L(x) \le 1 - \frac{\mathfrak{c}}{d_L} , \quad \forall x \in \mathbb{Z}^d \setminus \{0\} , \qquad (1.2)$$

where $d_L > 0$ is defined through

$$\sup_{|x|=1} v_L(x) = 1 - \frac{1}{d_L} .$$
(1.3)

While the structural assumptions on the field ξ_L are somewhat standard, conditions (I) and (II) require some justification. The former controls the long-range behaviour of the potential and determines the minimal speed of decay of its correlations. As written, (1.1) is very mild. Indeed, for ξ_L independent of L, it is equivalent to $v_L(x) = v(x) = o(1/\ln |x|)$ as $|x| \to \infty$, which is a well-known condition in the study of extreme values of correlated Gaussian fields (see [LLR83, Chapter 4]).

For L-independent potentials, the assumption (II) only imposes that v is strictly below 1 outside the origin. On the other hand, in the L-dependent case, the assumption (II) is non-trivial as (1.2) ensures that the parameter d_L in (1.3) is a faithful control of the decay of v_L near the origin.

Let us present a few examples of potentials ξ_L that satisfy Definition 1.1. We start with *L*-independent potentials:

(1) *L*-independent correlated Gaussian field. Take $\xi_L = \xi$ to be a centred, stationary Gaussian field with a covariance function $v_L = v$ independent of *L*, satisfying v(0) = 1, $v(x) \ln |x| \to 0$ as $|x| \to \infty$ and $\sup_{x \neq 0} v(x) < 1$. Then d_L is independent of *L* and finite. This covers the i.i.d. case (where $v(x) = \mathbb{1}_{x=0}$) and, for instance, the discrete Gaussian Free Field in dimension $d \geq 3$ (where *v* is the Green function associated to the discrete Laplacian).

We now present examples of potentials whose laws depend on L, and that arise by discretising a continuum Gaussian potential on a grid. More precisely, we start from a Gaussian potential $\zeta \stackrel{\text{def}}{=} u * \eta$ obtained by convolving a white noise η on \mathbb{R}^d with some function $u \colon \mathbb{R}^d \to \mathbb{R}$. We give ourselves a sequence of $m_L \ge 1$ that converges to ∞ and we set $\xi_L(x) \stackrel{\text{def}}{=} \zeta(x/m_L)$ for all $x \in \mathbb{Z}^d$. As we will see, the regularity of u has a subtle impact on the top of the spectrum of the Anderson Hamiltonian, and therefore we distinguish two cases:

- (2) **Smooth.** Let $u : \mathbb{R}^d \to \mathbb{R}$ be a smooth, compactly supported, radial function with unit L^2 -norm. Then, ξ_L satisfies the requirements listed above and d_L is of order m_L^2 .
- (3) **Indicator.** Let u be the indicator of the centred ball of \mathbb{R}^d with radius 1/2, normalised so that it has unit L^2 -norm. Then, ξ_L satisfies the requirements listed above and d_L is of order m_L .

For our purposes, we need to collect some properties on ξ_L restricted to the domain $Q_L \stackrel{\text{def}}{=} [-L/2, L/2]^d \cap \mathbb{Z}^d$. First of all, let us introduce the order of magnitude of the maximum of ξ_L over Q_L , which is given by the parameter a_L

4

implicitly defined through¹

$$\mathbb{P}(\xi_L(0) > a_L) = \frac{1}{L^d} \,. \tag{1.4}$$

5

It is elementary to check that $a_L = \sqrt{2d \ln L} (1 + o(1))$ as $L \to \infty$.

Let us now introduce an approximation of ξ_L near one of its large peaks as it will be instrumental in this article. Let x_0 be a site such that $\xi_L(x_0) \ge a_L - \theta$, where $\theta > 0$ is some fixed, arbitrary value². We will show in Section 2 that ξ_L can be well-approximated, in a neighborhood of x_0 , as follows

$$\xi_L(x) \approx \xi_L(x_0) - \mathcal{S}_L(x - x_0) + \zeta_{L,x_0}(x) , \qquad (1.5)$$

where \mathcal{S}_L is the so-called *shape* defined by

$$\mathcal{S}_L(x) \stackrel{\text{der}}{=} a_L(1 - v_L(x)) \ge 0 , \quad x \in \mathbb{Z}^d , \qquad (1.6)$$

and ζ_{L,x_0} is a Gaussian field independent of $\xi_L(x_0)$, that we will dub *fluctuation* field. The shape should be seen as the first order description of the variation of the potential near a large peak, while the fluctuation field provides a random correction to this deterministic shape.

At this point, we observe that (1.2) implies that S_L is of order a_L/d_L in the vicinity of the origin. We are thus naturally led to introduce the following assumption.

Assumption 1.2. Let d_L and a_L be respectively defined by (1.3) and (1.4). We assume that

$$d_L \ll a_L . \tag{1.7}$$

Under this assumption, the shape is very steep so that we should be close to the Single peak case of the i.i.d. setting. On the other hand, when d_L is of order a_L , the shape is of order 1 and this should correspond to the Doubly Exponential case of the i.i.d. setting.

Our first result determines the statistics of the largest peaks of the potential and of the locations where these are achieved. To state it, for any $1 \le k \le \#Q_L$, let $y_{k,L}$ be the site in Q_L where ξ_L reaches its k-th largest value.

Theorem 1.3. Under Assumption 1.2,

$$\left(\frac{y_{k,L}}{L}, a_L(\xi_L(y_{k,L}) - a_L)\right)_{1 \le k \le \#Q_L}$$

converges in law as $L \to \infty$ to a Poisson point process on $[-1, 1]^d \times \mathbb{R}$ of intensity $dx \otimes e^{-u} du$.

Remark 1.4. Observe that the value a_L is the same for any potential ξ_L as in Definition 1.1, because $\xi_L(0) \sim \mathcal{N}(0, 1)$. Theorem 1.3 shows that the statistics of the largest peaks are asymptotically the same both in the i.i.d. case and in any of the correlated cases considered here (which a posteriori justifies the comparison with the Single Peak class).

 $^{{}^{1}}L^{d}$ should be interpreted as the cardinality of Q_{L} , although the exact value of the cardinality is slightly different but asymptotically equivalent.

²We will take $\theta = 2d + 1$ later on for definiteness.

Actually, in Sections 2 and 3 we will gather much more information on the Gaussian field as we will not only study its largest values but also other functionals (including the fluctuation field) which are instrumental in the study of the top of the spectrum of the Anderson Hamiltonian.

Let us finally mention that the study of extrema of Gaussian fields has been the topic of a large literature. Let us in particular cite the recent works [BL16, BL18] of Biskup and Louidor on the discrete Gaussian Free Field in dimension 2. It should be observed that for a discrete Gaussian Free Field, the variance at a given point blows up in dimension 2 while it remains bounded in dimension $d \ge 3$. As mentioned above, the discrete Gaussian Free Field in $d \ge 3$ falls into our framework.

1.2. Main result. We consider the random operator $\mathcal{H}_L \stackrel{\text{def}}{=} \Delta + \xi_L$ on $Q_L = [-L/2, L/2]^d \cap \mathbb{Z}^d$ endowed with Dirichlet boundary conditions³. This operator is finite dimensional and self-adjoint: we let $(\lambda_{k,L})_{1 \leq k \leq \#Q_L}$ be its successive eigenvalues in non-increasing order, and $(\varphi_{k,L})_{1 \leq k \leq \#Q_L}$ be the associated eigenfunctions normalised in ℓ^2 . We also denote by $x_{k,L}$ the⁴ point in Q_L that maximizes $|\varphi_{k,L}|$ for any $1 \leq k \leq \#Q_L$, and w.l.o.g. we can take $\varphi_{k,L}$ positive at $x_{k,L}$.

Our analysis of the top of the spectrum of \mathcal{H}_L relies on a splitting scheme where we analyse the operator restricted to *mesoscopic boxes* of side-length $1 \ll R_L \ll L$ (see below for further details). A key step consists in obtaining a fine description of the top eigenvalue, $\lambda_1(Q_{R_L}, \xi_L)$, of the operator $\Delta + \xi_L$ restricted to Q_{R_L} , when there is a point $x_0 \in Q_{R_L}$ at which $\xi_L(x_0) \ge a_L - \theta$. Recall that the approximation of the field near x_0 given in (1.5) displays three terms: the value of the field at x_0 , the shape and the fluctuation field. While the impact on $\lambda_1(Q_{R_L}, \xi_L)$ of the value of the field at x_0 is rather straightforward as it amounts to a (random) shift by $\xi_L(x_0)$, those of the shape and the fluctuation fields are more subtle. Let us consider the deterministic operator

$$\bar{\mathcal{H}}_L \stackrel{\text{def}}{=} \Delta - \mathcal{S}_L, \quad \text{on } Q_{r_L} \stackrel{\text{def}}{=} [-r_L/2, r_L/2]^d \cap \mathbb{Z}^d$$
(1.8)

endowed with Dirichlet b.c., where $1 \ll r_L \ll R_L$ will be introduced later on. We let $\bar{\lambda}_L$ be the largest eigenvalue of this operator, and $\bar{\varphi}_L$ be its associated normalised eigenfunction (taken non-negative w.l.o.g.). We will show that

$$\bar{\lambda}_L = -2d + \mathcal{O}(\frac{d_L}{a_L}), \quad L \to \infty$$

while

$$\bar{\varphi}_L(0) = 1 - \mathcal{O}(\frac{d_L}{a_L})$$
, and $\bar{\varphi}_L(x) \asymp \frac{d_L}{a_L}$ for $x \in \mathbb{Z}^d$ with $|x| = 1$. (1.9)

Our ansatz, which is detailed in Section 4, is that the main eigenfunction of $\Delta + \xi_L$ on Q_{R_L} should be well approximated by $\bar{\varphi}_L(\cdot - x_0)$, and consequently the variational characterisation of the principal eigenvalue together with (1.5) suggests that $\lambda_1(Q_{R_L}, \xi_L)$ should satisfy

$$\lambda_1(Q_{R_L},\xi_L) \approx \langle \bar{\varphi}_L(\cdot - x_0), \mathcal{H}_L \bar{\varphi}_L(\cdot - x_0) \rangle = \langle \bar{\varphi}_L(\cdot - x_0), (\Delta + \xi_L) \bar{\varphi}_L(\cdot - x_0) \rangle$$

³The domain of \mathcal{H}_L is the set of all functions $f: Q_L \to \mathbb{R}$, extended outside Q_L by setting them to 0, and the value of $\mathcal{H}_L f(x)$ is simply given by $\Delta f(x) + \xi_L(x) f(x)$ for all $x \in Q_L$.

⁴If this point is not unique, take the smallest one for some arbitrary total order on \mathbb{Z}^d .

$$\approx \langle \bar{\varphi}_L(\cdot - x_0), (\Delta + \xi_L(x_0) - \mathcal{S}_L(x - x_0) + \zeta_{L,x_0}(\cdot))\bar{\varphi}_L(\cdot - x_0) \rangle$$

= $\xi_L(x_0) + \langle \bar{\varphi}_L, \bar{\mathcal{H}}_L \bar{\varphi}_L \rangle + \langle \bar{\varphi}_L(\cdot - x_0), \zeta_{L,x_0}(\cdot)\bar{\varphi}_L(\cdot - x_0) \rangle$
= $\xi_L(x_0) + \bar{\lambda}_L + \sum_{x \in Q_{r_L}} \bar{\varphi}_L^2(x - x_0)\zeta_{L,x_0}(x)$. (1.10)

In (1.10), we observe a competition between the randomness coming from the first and third terms. More precisely, the first term fluctuates at scale $1/a_L$ around a leading order a_L (as shown by Theorem 1.3), while the third term is of order τ_L where

$$\tau_L^2 \stackrel{\text{def}}{=} \operatorname{Var}\left(\sum_{x \in Q_{r_L}} \bar{\varphi}_L^2(x - x_0)\zeta_{L, x_0}(x)\right).$$
(1.11)

Therefore, we should expect that the relative values of $1/a_L$ and τ_L play an important role. We will work under the following assumption on the strength of τ_L .

Assumption 1.5. The L-dependent constant τ_L in (1.11) satisfies

$$\tau_L \ll \frac{1}{a_L} \left(\frac{a_L}{d_L}\right)^{\frac{1}{2}} . \tag{1.12}$$

The restriction (1.12) is intimately related to the replacement of $\lambda_1(Q_{R_L}, \xi_L)$ by the quantity in (1.10). More precisely, it guarantees that the error made in the approximation is negligible. If this assumption is not satisfied, one certainly needs to take into account many more terms induced by the fluctuation field than the sole projection $\sum_{x \in Q_{r_L}} \bar{\varphi}_L^2(x - x_0)\zeta_{L,x_0}(x)$. We leave this task for future investigations.

That said, Assumption 1.5 is satisfied in most cases: when the covariance function v_L does not depend on L (Example (1)) or when the covariance function depends on L but is "regular" enough (Example (2)), then it holds. On the other hand, it fails when the covariance function is not regular enough (Example (3)) and d_L is "close enough" to a_L .

Our first main result concerns the eigenvalue order statistics and the localisation centers.

Theorem 1.6 (Eigenvalue order statistics). Under Assumptions 1.2 and 1.5, the point process

$$\left(\frac{x_{k,L}}{L}, a_L\left(\lambda_{k,L} - a_L\sqrt{1 + \tau_L^2} - \bar{\lambda}_L\right)\right)_{1 \le k \le \#Q_L}$$

converges in law as $L \to \infty$ towards a Poisson point process on $[-1,1]^d \times \mathbb{R}$ of intensity $dx \otimes e^{-u} du$.

If $\tau_L \ll 1/a_L$, one can replace $\sqrt{1+\tau_L^2}$ by 1 without altering the result. However, when τ_L is order $1/a_L$ or larger, then the correction is crucial and hints at the fact that the relationship of the top eigenvalues with the largest values of ξ_L is no longer trivial, see Theorem 1.8.

We now address the localisation properties of the main eigenfunctions. Recall that $\bar{\varphi}_L$ is a deterministic function which is almost a Dirac mass at 0 (see (1.9)).

Theorem 1.7 (Localisation). Under Assumptions 1.2 and 1.5, for any $k \ge 1$, the r.v.

$$\frac{d_L}{d_L} \left\| \varphi_{k,L}(\cdot) - \bar{\varphi}_L(\cdot - x_{k,L}) \right\|_{\ell^2(Q_L)},$$

converges to 0 in probability.

We now relate the top eigenvalues/eigenfunctions with the maxima of the field. For any $k \geq 1$, we define the random variable $\ell_L(k)$ through

$$x_{k,L} = y_{\ell_L(k),L} \,, \tag{1.13}$$

where, as a reminder, $y_{k,L}$ is the site in Q_L where ξ_L reaches its k-th largest value. The random variable $\ell_L(k)$ provides the rank of the maximum of ξ_L at which the k-th eigenfunction is localised.

To state our result, we need to introduce, for any given parameter b > 0, a random permutation $(\ell_{\infty,b}(k))_{k>1}$ of $(1, 2, \ldots)$.

- Let $u_1 > u_2 > \dots$ be distributed according to a Poisson point process of intensity $e^{-u} du$.
- Draw an independent sequence $(v_i)_{i\geq 1}$ of i.i.d. $\mathcal{N}(0,b)$ r.v.
- Let $(p_i)_{i\geq 1}$ be the (non-increasing) order statistics of the decorated Poisson point process $(u_i + v_i)_{i>1}$.

Then, for any $k \ge 1$, define $\ell_{\infty,b}(k)$ according to

$$p_k = u_{\ell_{\infty,b}(k)} + v_{\ell_{\infty,b}(k)} \,. \tag{1.14}$$

Theorem 1.8 (Relationship with the maxima of ξ_L). Under Assumptions 1.2 and 1.5, it holds:

- (a) if $\tau_L \ll \frac{1}{a_L}$ then for any given $k \ge 1$, $\mathbb{P}(\ell_L(k) = k) \to 1$ as $L \to \infty$, (b) if $\tau_L \sim \sqrt{b} \frac{1}{a_L}$ for some constant b > 0 then $(\ell_L(k))_{k\ge 1}$ converges in law to $(\ell_{\infty,b}(k))_{k\geq 1}$,
- (c) if $\tau_L \gg \frac{1}{a_L}$ then for any given $k \ge 1$, $\ell_L(k) \to \infty$ in probability.

Let us point out the analogy with the i.i.d. Weibull case presented in the introduction: the regime $\tau_L \ll 1/a_L$ corresponds to q < 3, the regime $\tau_L \sim b/a_L$ corresponds to q = 3 and $\tau_L \gg \frac{1}{a_L}$ to q > 3. Let us mention that the law of the permutation was not identified in the literature in the Weibull case with q = 3.

To conclude the introduction, we note that each of the three scenarios detailed in the above statement do indeed realise. For the specific examples of Section 1.1, we have:

- (1) L-independent correlated Gaussian field: τ_L is of order $1/a_L^2$ so that Assumption 1.5 is satisfied and the relationship with the maxima is given by (a).
- (2) Smooth: τ_L is of order d_L/a_L^2 so that Assumption 1.5 is satisfied and the relationship with the maxima is given by (a).
- (3) Indicator: τ_L is of order $d_L^{3/2}/a_L^2$ so that Assumption 1.5 is satisfied provided $d_L \ll a_L^{\frac{3}{4}}$. According to whether $d_L \ll a_L^{2/3}$, $d_L \simeq a_L^{2/3}$ or $d_L \gg a_L^{2/3}$, the relationship with the maxima is respectively given by (a), (b) or (\tilde{c}) .

From now on and throughout the article, we will always work under Assumptions 1.2 and 1.5 unless otherwise stated.

1.3. Strategy of proof and structure of the article. The study of the orderstatistics of (the sequence of) Gaussian field(s) ξ_L in Definition 1.1 and the spectral properties of the Hamiltonian \mathcal{H}_L follow distinct but interdependent routes that we now outline. Both rely on a suitable localisation procedure (or splitting scheme) whose aim is to reduce the analysis from the *macroscopic* box of side-length L to a collection of *mesoscopic* boxes of side-length R_L , which grows with L but is much smaller than L, and ultimately, for the Hamiltonian, to an even smaller box of side-length r_L . Let us fix the sequences $(R_L)_{L\geq 1}$ and $(r_L)_{L\geq 1}$ of positive constants in such a way that they satisfy⁵

$$a_L \ll \ln R_L \ll a_L \frac{a_L}{d_L} , \qquad (1.15)$$

$$\ln a_L \le \ln r_L \ll \sqrt{a_L} , \qquad (1.16)$$

where a_L and d_L are defined in (1.4) and (1.3), respectively, and where, for two sequences $(u_L)_{L\geq 1}, (v_L)_{L\geq 1}$ in $(0,\infty)$, the notation $u_L \ll v_L$ means $u_L/v_L \to 0$ as $L \to \infty$.

Now, the aforementioned localisation procedure can be roughly visualised in the following diagram

$$Q_L \xrightarrow{(a)} U_L \xrightarrow{(b)} Q_{R_L} \xrightarrow{(c)} Q_{r_L}$$
 (1.17)

where, for a > 0, we denoted by $Q_a \stackrel{\text{def}}{=} [-a/2, a/2]^d \cap \mathbb{Z}^d$, and U_L is obtained by peeling off suitable strips from Q_L and is thus given by the union of disjoint and well-separated boxes of side-length R_L . More precisely, we consider a covering of Q_L into boxes of side-length $R_L + \sqrt{R_L}$ whose interiors are disjoint

$$Q_L = \bigcup_{j=1}^{n_L} Q_{R_L + \sqrt{R_L}, z_{j,L}}, \quad \text{for} \quad n_L \stackrel{\text{def}}{=} \frac{\# Q_L}{\# Q_{R_L + \sqrt{R_L}}}, \quad (1.18)$$

where⁶ $(z_{j,L})_{j=1,...,n_L}$ forms implicitly a lattice of points at distance at least $R_L + \sqrt{R_L}$ from each other, and then we peel off a boundary layer of size $\sqrt{R_L}$ by setting

$$U_{L} \stackrel{\text{def}}{=} \bigcup_{j=1}^{n_{L}} Q_{R_{L}, z_{j,L}} .$$
 (1.19)

Remark 1.9. In classical references on the Anderson Hamiltonian with i.i.d. potential (see e.g. [BK16]), the localisation procedure (or splitting scheme) is "random" in that the set U_L is chosen to be the union of mesoscopic boxes centred around the large peaks of the potential. While possibly we could have proceeded similarly, it would have added an additional layer of difficulty as the presence of correlations makes such procedure much more complicated and the way in which (the already challenging) step (b) in (1.17) should be approached much less transparent.

⁵Concerning R_L , the lower bound is needed in the proof of Lemma 4.8 while the upper bound in that of Proposition 2.5. For r_L instead, the lower bound is needed in the proof of Lemma 4.8, while the upper bound in the of Lemma 5.3.

⁶For notational convenience, we assume that n_L is an integer. To treat the general case, it suffices to adapt the splitting scheme.

10

Getting back to the cartoon in (1.17), step (a) is relatively simple: at the level of the Gaussian field, since $\sqrt{R_L} \ll R_L$, the cardinality of $Q_L \setminus U_L$ is negligible compared to that of Q_L so that, with large probability, ξ_L does not display large peaks in this set (see Section 3.2) and thus its order statistics are unaffected by its value therein. As we will see in Section 5.2, this also implies that the top eigenfunctions put an exponentially small amount of mass on $Q_L \setminus U_L$ and thus the top eigenpairs of \mathcal{H}_L on Q_L and on U_L (asymptotically) coincide.

Step (b) for ξ_L is one of the main novelties of the present work. The advantage of U_L over Q_L is that any two distinct boxes $Q_{R_L,z_{j,L}}$ and $Q_{R_L,z'_{j,L}}$ lie at a distance at least $\sqrt{R_L} \ge \exp(\sqrt{\ln L})$ so that the restrictions of ξ_L to these boxes display negligible correlations thanks to (1.1). This suggests that it should be possible to regard these as independent, but making this rigorous is technically challenging as it amounts to determine non-trivial *decorrelation estimates* (whose nature, in case U_L were chosen as in Remark 1.9, is unclear) to which Section 3.3 is dedicated. In view of these decorrelation estimates, the analysis of ξ_L and \mathcal{H}_L on U_L is rather standard and presented in Sections 3.1 and 5.1 respectively.

Thanks to the steps (a) and (b), all that remains to do is to study the behaviour of the Gaussian field and the Anderson Hamiltonian on the mesoscopic box Q_{R_I} , and this is the second main novelty of the present paper. For the former, it consists of, first, formalising the description of the noise close to a large peak in terms of the shape S_L and the fluctuation field $\zeta_{L,.}$ as in (1.5), and this is carried out in Section 2.2, and then deducing its implications as done in Section 2.3. Such a description is then employed in the study of the principal eigenpair of \mathcal{H}_L on Q_{R_L} . In Section 4, we show that, thanks to an apriori estimate on the exponential decay of the eigenfunctions, we can further localise the eigenproblem to $Q_{r_{L}}$ (step (c) in (1.17)) and then, more importantly, that the main eigenvalue and eigenfunction are respectively well-approximated by (1.10) and by the main eigenfunction of the deterministic operator $\overline{\mathcal{H}}_L$ in (1.8). It is thanks to step (c) that \mathcal{H}_L can be taken to be independent of the point at which ξ_L achieves its maximum. The approximation on Q_{r_L} relies on a simple and effective convex analysis argument applied to the local quadratic form (see Lemma 4.5) that crucially allows to identify the correction due to the fluctuation field and that we believe could be of independent interest.

1.4. Notation and basic Gaussian estimates. Here we introduce (and recall) some notation and conventions we will be using throughout the paper. For $x \in \mathbb{Z}^d$, we denote by $|x| \stackrel{\text{def}}{=} (\sum_{i=1}^d |x_i|^2)^{1/2}$ the ℓ^2 -norm of $x \in \mathbb{R}^d$. As above, we write Q_a for the $(\ell^{\infty}$ -)box of side-length a, i.e. $Q_a \stackrel{\text{def}}{=} [-a/2, a/2]^d \cap \mathbb{Z}^d$, and, for $x, y \in \mathbb{Z}^d$, $Q_{a,x} \stackrel{\text{def}}{=} x + Q_a$ and $Q_{a,x}^{\neq y} \stackrel{\text{def}}{=} Q_{a,x} \setminus \{y\}$. For any subset $C \subset \mathbb{Z}^d$ and any function $V: C \to \mathbb{R}$, we let $\mathcal{H}_{C,V}$ be the operator $\Delta + V$ on C endowed with Dirichlet boundary conditions.

For two sequences $(u_L)_{L\geq 1}$, $(v_L)_{L\geq 1}$ in $(0,\infty)$, we write $u_L \ll v_L$ to express that $u_L/v_L \to 0$ as $L \to \infty$, we write $u_L \gg v_L$ to express that $v_L \ll u_L$, we write $u_L \asymp v_L$ if $0 < \liminf_{L\to\infty} u_L/v_L \leq \limsup_{L\to\infty} u_L/v_L < \infty$, and we write $u_L \sim v_L$ if u_L/v_L converges to 1 as $L \to \infty$. At last, recall that if X is a standard Gaussian r.v., i.e. $X \stackrel{\text{law}}{=} \mathcal{N}(0, 1)$, then for all x > 0

$$\mathbb{P}(X \ge x) \le \frac{1}{\sqrt{2\pi x}} e^{-\frac{x^2}{2}},$$
 (1.20)

and

$$\mathbb{P}(X \ge x) = \frac{1}{\sqrt{2\pi}x} e^{-\frac{x^2}{2}} (1 + \mathcal{O}(1/x^2)), \quad x \to \infty.$$
(1.21)

These immediately provide a more explicit characterisation of a_L in (1.4), i.e.

$$\frac{1}{\sqrt{2\pi}a_L}e^{-\frac{a_L^2}{2}} \sim \frac{1}{L^d} , \qquad \text{as } L \to \infty, \tag{1.22}$$

and a perturbative result for the tail of X around a_L , namely, for any sequence $(b_L)_{L\geq 1}$ satisfying $|b_L| \ll a_L$, it holds

$$\mathbb{P}(X \ge a_L + b_L) \sim \frac{1}{L^d} e^{-a_L b_L - \frac{b_L^2}{2}}, \quad L \to \infty.$$
 (1.23)

By Gaussian scaling, it is immediate to adapt the above to the case in which X has variance $\sigma > 0$.

Acknowledgements. The work of C. L. was partially supported by the ANR project Smooth ANR-22-CE40-0017, and by the Institut Universitaire de France. G. C. gratefully acknowledges financial support via the UKRI FL fellowship "Large-scale universal behaviour of Random Interfaces and Stochastic Operators" MR/W008246/1.

2. Local behaviour of correlated Gaussian fields

This section is devoted to the study of the potential ξ_L introduced in Section 1.1 and the Gaussian field introduced in (1.10). We will begin by spelling out a useful orthogonal decomposition of ξ_L and stating a few useful properties thereof (Section 2.1). Then, we will study the behaviour of the field on a mesoscopic box of diameter $R_L \ll L$: we will first establish an approximation of ξ_L near a large peak (Section 2.2) and use it to identify the behaviour of the maxima of ξ_L and of associated random fields (Section 2.3).

2.1. Orthogonal decomposition and basic properties. Let $(\xi_L(x))_{x\in\mathbb{Z}^d}$ be a Gaussian field that satisfies Definition 1.1. At first, we devise the orthogonal decomposition to ξ_L and rigorously introduce the fluctuation field alluded to in (1.5). For any $x_0 \in \mathbb{Z}^d$, the latter is the (Gaussian) field ζ_{L,x_0} defined through

$$\xi_L(x) = \xi_L(x_0)v_L(x - x_0) + \zeta_{L,x_0}(x), \quad x \in \mathbb{Z}^d .$$
(2.1)

Its main properties are summarised in the next lemma.

Lemma 2.1. For any given $x_0 \in \mathbb{Z}^d$, ζ_{L,x_0} is a centred Gaussian field independent of $\xi_L(x_0)$, satisfying $\zeta_{L,x_0}(x_0) = 0$. Its covariance is

$$\mathbb{E}[\zeta_{L,x_0}(x)\zeta_{L,x_0}(y)] = v_L(x-y) - v_L(x-x_0)v_L(y-x_0) , \qquad (2.2)$$

while its variance is bounded above by 1 and there exists c'' > 0 such that for all $x \in \mathbb{Z}^d$

$$\operatorname{Var}[\zeta_{L,x_0}(x)] = 1 - v_L(x - x_0)^2 \le \frac{e^{c''|x - x_0|}}{d_L}.$$
(2.3)

Proof. The fact that ζ_{L,x_0} is a centred Gaussian field, independent of $\xi_L(x_0)$ and such that $\zeta_{L,x_0}(x_0) = 0$ is obvious by (2.1) and the definition of v_L . The expression of the covariance follows from a straightforward computation. Since v_L is bounded by 1 by (1.2), so is $\operatorname{Var}[\zeta_{L,x_0}(x)]$. For (2.3), we immediately have

$$v_L(x-x_0)^2 = (1+v_L(x-x_0)-1)^2 \ge 1-2(1-v_L(x-x_0))$$

and thus, by (1.2), we deduce that there exists c'' > 0 such that for all $x \neq x_0$

$$1 - v_L(x - x_0)^2 \le 2(1 - v_L(x - x_0)) \le 2\frac{e^{c'|x - x_0|}}{d_L} \le \frac{e^{c''|x - x_0|}}{d_L} \ .$$

As discussed in Section 1.2, our analysis of the Hamiltonian aims at showing that its main eigenvalues can be described in terms of the sum of $\xi_L(\cdot)$ and of an additional term whose definition only involves the fluctuation field $\zeta_{L,\cdot}$ in (2.1). For $y \in \mathbb{Z}^d$, the latter is given by

$$\Phi_L(y) \stackrel{\text{def}}{=} \sum_{x \in Q_{r_L}^{\neq 0}} \bar{\varphi}_L(x)^2 \zeta_{L,y}(x+y) , \qquad (2.4)$$

where r_L is as in (1.16) and $\bar{\varphi}_L$ is the principal normalised eigenfunction of the operator $\bar{\mathcal{H}}_L$ in (1.8).

Lemma 2.2. Let $(\Phi_L(x))_{x\in\mathbb{Z}^d}$ be defined according to (2.4). Then, the field

$$\Xi_L(x) \stackrel{def}{=} \xi_L(x) + \Phi_L(x), \qquad x \in \mathbb{Z}^d$$
(2.5)

is Gaussian with variance $1 + \tau_L^2$ at every point, for τ_L as in (1.11), translation invariant and, under Assumptions 1.2 and 1.5, its covariance function $v_L^{\Xi} : \mathbb{Z}^d \to \mathbb{R}_+$ satisfies

$$\mathcal{T}_{v_L^{\Xi}} \stackrel{def}{=} \sup_{|x| \ge \exp(\sqrt{\ln L})} v_L^{\Xi}(x) \ln |x| \longrightarrow 0, \qquad as \ L \to \infty.$$
(2.6)

Furthermore, for any $z \in \mathbb{Z}^d$ such that $|z| > \sqrt{d} r_L$, we have

$$\operatorname{Cov}(\xi_L(z), \Phi_L(0)) \vee \operatorname{Cov}(\Phi_L(z), \Phi(0)) \lesssim \sup_{|x| \ge |z| - \sqrt{d}r_L} v_L(x) \,. \tag{2.7}$$

Proof. The fact that Ξ_L is Gaussian, translation invariant and at every point has variance $1 + \tau_L^2$ is an immediate consequence of the definition of ξ_L and Lemma 2.1. A direct computation (using translation invariance) shows that the covariance function v_L^{Ξ} equals

$$\begin{split} v_L^{\Xi}(z) &= \operatorname{Cov}(\xi_L(z), \xi_L(0)) + 2\operatorname{Cov}(\xi_L(z), \Phi_L(0)) + \operatorname{Cov}(\Phi_L(z), \Phi_L(0)) \\ &= \mathbb{E}[\xi_L(z)\xi_L(0)] + 2\mathbb{E}[\xi_L(z)\Phi_L(0)] + \mathbb{E}[\Phi_L(z)\Phi_L(0)] \\ &= v_L(z) + 2\sum_{x \in Q_{r_L}^{\neq 0}} \bar{\varphi}_L(x)^2 \Big(v_L(z-x) - v_L(z)v_L(x) \Big) \\ &+ \sum_{x,y \in Q_{r_L}^{\neq 0}} \bar{\varphi}_L(x)^2 \bar{\varphi}_L(y)^2 \Big(v_L(z+x-y) - 2v_L(x+z)v_L(y) + v_L(z)v_L(x)v_L(y) \Big) \,. \end{split}$$

13

To control its decay and establish both (2.6) and (2.7), let us point out that, for $|z| \ge \sqrt{d} r_L$ and any $|w| \le \sqrt{d} r_L$, we clearly have

$$v_L(z-w) \le \sup_{|x|\ge |z|-\sqrt{d}r_L} v_L(x) \,.$$

Therefore, we immediately deduce

$$\operatorname{Cov}(\xi_L(z), \Phi_L(0)) \lesssim \sup_{|x| \ge |z| - \sqrt{d} r_L} v_L(x) \sum_{y \in Q_{r_L}^{\neq 0}} \bar{\varphi}_L(y)^2 \le \sup_{|x| \ge |z| - \sqrt{d} r_L} v_L(x)$$

where we used that $\overline{\varphi}_L$ is normalised. Arguing similarly (and recalling that $v_L \leq 1$) for $\text{Cov}(\Phi_L(z), \Phi_L(0))$, (2.7) follows at once.

The very same procedure also implies that

$$\mathcal{T}_{v_L^{\Xi}} \lesssim \mathcal{T}_{v_L} + \sup_{|z| \ge \exp(\sqrt{\ln L})} \ln |z| \sup_{|x| \ge |z| - \sqrt{d} r_L} v_L(x) \le 2\mathcal{T}_{v_L}$$

and, by (1.1), the r.h.s. converges to 0, which completes the proof.

In what follows, we will derive the order statistics of both fields ξ_L and Ξ_L , and study how they relate to each other, which in particular requires to identify the order of magnitude of their maxima and the size of the fluctuations around these. To do so, it is crucial to understand the mechanism that produces high peaks and what behaviour ξ_L must display for Ξ_L to be large. Notice that, in view of Lemma 2.1, for every $x \in \mathbb{Z}^d$, $\Xi_L(x)$ is the sum of two independent Gaussian random variables, $\xi_L(x)$ and $\Phi_L(x)$, of variance 1 and τ_L^2 , respectively.

In the next lemma, we address the afore-mentioned questions for generic Gaussian random variables satisfying these features and after its statement we will translate its content in our context. The proof of the lemma is postponed to Appendix A.

Lemma 2.3. Let X, Y_L be two independent centred Gaussian random variables of variance 1 and τ_L^2 , respectively. Then, for any $s \in \mathbb{R}$, as $L \to \infty$

$$\mathbb{P}\left(X+Y_L \ge a_L \sqrt{1+\tau_L^2} + \frac{s}{a_L}\right) \sim \frac{1}{L^d} e^{-s} , \qquad (2.8)$$

and then

$$\limsup_{C \to \infty} \limsup_{L \to \infty} L^d \mathbb{P} \left(X + Y_L \ge a_L \sqrt{1 + \tau_L^2} + \frac{s}{a_L}; X \notin I_L(C) \right) = 0 , \quad (2.9)$$

where, for L, C > 0, $I_L(C)$ is the interval

$$I_L(C) \stackrel{def}{=} \left[\frac{a_L}{\sqrt{1 + \tau_L^2}} - C \max\{\frac{1}{a_L}, \tau_L\}, \frac{a_L}{\sqrt{1 + \tau_L^2}} + C \max\{\frac{1}{a_L}, \tau_L\} \right].$$
(2.10)

As a consequence, for any sequence $(\theta_L)_{L\geq 1}$ such that $\max\{a_L\tau_L^2, a_L^{-1}\} \ll \theta_L$, we have

$$\lim_{L \to \infty} L^d \mathbb{P} \left(X + Y_L \ge a_L \sqrt{1 + \tau_L^2} + \frac{s}{a_L}; |X - a_L| > \theta_L \right) = 0.$$
 (2.11)

In terms of Ξ_L and ξ_L , we can infer from the above a number of useful insights. First, in view of (2.8) and (1.4) we set

$$a_L^{\Xi} = a_L \sqrt{1 + \tau_L^2} \tag{2.12}$$

as it is the counterpart of a_L for Ξ_L , and the fluctuations around it are of order a_L^{-1} (i.e. the same as those of ξ_L , see (1.23)). Second, the estimate (2.9) shows that, for any $x \in \mathbb{Z}^d$, for $\Xi_L(x)$ to be of order a_L^{Ξ} , $\xi_L(x)$ must be of order $a_L/\sqrt{1+\tau_L^2}$ up to an event of probability negligible compared to $1/L^d$.

2.2. Deterministic shape around local maxima. In this section, we study the behaviour of the potential ξ_L around its maxima in a mesoscopic box of sidelength R_L . To do so, we introduce an event around which our analysis revolves. Recall the definition of the shape S_L in (1.6), i.e. $S_L(\cdot) \stackrel{\text{def}}{=} a_L(1 - v_L(\cdot))$, and of the fluctuation field $\zeta_{L,\cdot}$ in (2.1).

Definition 2.4. Set $\theta \stackrel{\text{def}}{=} 2d + 1$. For $x_0 \in \mathbb{Z}^d$ and a constant $\kappa \in (0, 1/3)$, we define $E_{L,x_0} = E_{L,x_0}(\kappa)$ as the intersection of the three events E_{L,x_0}^i , i = 1, 2, 3, respectively given by

$$E_{L,x_0}^1 \stackrel{\text{def}}{=} \{ |\xi_L(x_0) - a_L| < \theta \}, \qquad (2.13)$$

$$E_{L,x_0}^2 \stackrel{\text{def}}{=} \{ |\zeta_{L,x_0}(x)| \le \frac{1}{10} \mathcal{S}_L(x-x_0) \quad \forall \ x \in Q_{2R_L,x_0} \}, \qquad (2.14)$$

and

$$E_{L,x_0}^{3} \stackrel{\text{def}}{=} \bigcap_{x \in Q_{R_L,x_0}^{\neq x_0}} \left\{ \frac{|\zeta_{L,x_0}(x)|}{\sqrt{\operatorname{Var}[\zeta_{L,x_0}(x)]}} \le \left(\frac{a_L}{d_L}\right)^{\kappa|x-x_0|} \sqrt{1 \lor (|\xi_L(x_0) - a_L|a_L)} \right\}.$$
(2.15)

Let us make some comments on this definition. The event E_{L,x_0}^1 forces ξ_L to display a large peak at x_0 . The requirement $\xi_L(x_0) < a_L + \theta$ was added for convenience only and, in any case, its complement is unlikely (*i.e.*, its probability is negligible compared to $1/L^d$) and can thus be excluded. The event E_{L,x_0}^2 ensures that the fluctuation field remains "small" around x_0 (the value 1/10at the r.h.s. is arbitrary and anything sufficiently small would do) so that, as we will see in Proposition 2.5, $\xi_L(x_0)$ is a local maximum and $\xi_L(x)$ remains below $\xi_L(x_0)$ around x_0 . Finally, E_{L,x_0}^3 prescribes the order of ζ_{L,x_0} on the box Q_{R_L,x_0} . It morally requires that $\zeta_{L,x_0} \approx \sqrt{\operatorname{Var}(\zeta_{L,x_0})}$ up to an error which suitably depends on the size of the fluctuations of $\xi_L(x_0)$ around a_L (which in turn are expected to be $\mathcal{O}(a_L^{-1})$, see Theorem 2.9 below).

The specific choice of θ and the control over the ζ_{L,x_0} 's will become clearer in the proof of the next two results as well as those of Lemmas 4.8, 4.9.

Proposition 2.5. For $x_0 \in \mathbb{Z}^d$, let E_{L,x_0} be the event in Definition 2.4. Then, there exists an $L_0 \geq 1$ such that for all $L \geq L_0$,

(1) on E_{L,x_0} , ξ_L admits a unique maximum over Q_{2R_L,x_0} which is attained at x_0 and the following bound holds

$$\xi_L(x) - \xi_L(x_0) \le -\frac{\mathfrak{c}}{2} \frac{a_L}{d_L} , \quad \forall x \in Q_{2R_L, x_0}^{\neq x_0} , \qquad (2.16)$$

where $\mathfrak{c} > 0$ is as in (1.2).

(2) for any $y_0 \neq x_0$ such that $|x_0 - y_0| \leq \sqrt{d} r_L$, $E_{L,x_0} \cap E_{L,y_0} = \emptyset$.

Furthermore, there is C > 0 such that for all $L \ge L_0$ and $x_0 \in \mathbb{Z}^d$, we have

$$\mathbb{P}(\xi_L(x_0) \ge a_L - \theta; E_{L,x_0}^{\complement}) \le \frac{1}{L^d} e^{-C\left(\frac{a_L}{d_L}\right)^{2\kappa}}, \qquad (2.17)$$

and

$$\lim_{L \to \infty} \left(\frac{L}{R_L}\right)^d \mathbb{P}\left(\max_{x \in Q_{R_L} + r_L} \xi_L(x) \ge a_L - \theta; \left(\bigcup_{x_0 \in Q_{R_L}} E_{L,x_0}\right)^{\complement}\right) = 0.$$
(2.18)

Remark 2.6. The rule of thumb underlying our probabilistic analysis is that we can neglect any event "based at x_0 " whose probability is negligible compared to $1/L^d$. Indeed, the union over $x_0 \in Q_L$ of such events has a probability which is then negligible compared to 1. Also, (2.17) combined with (1.23), ensures that

$$\mathbb{P}(E_{L,x_0}) = \mathbb{P}(\xi_L(x_0) \ge a_L - \theta; E_{L,x_0}) \sim \mathbb{P}(\xi_L(x_0) \ge a_L - \theta) \sim \frac{e^{\theta a_L - \frac{\theta^2}{2}}}{L^d} .$$

Proof. We start with (2.16). Equations (2.1), (2.14) and (1.6) together with the fact that $\xi_L(x_0) > a_L - \theta$ imply that on the event E_{L,x_0} for any $x \in Q_{2R_L,x_0}^{\neq x_0}$

$$\begin{split} \xi_L(x) - \xi_L(x_0) &\leq \xi_L(x_0) \left(v_L(x - x_0) - 1 \right) + \frac{1}{10} \mathcal{S}_L(x - x_0) \\ &\leq -\frac{\xi_L(x_0)}{a_L} \mathcal{S}_L(x - x_0) + \frac{1}{10} \mathcal{S}_L(x - x_0) \\ &\leq - \left(1 - \frac{\theta}{a_L} - \frac{1}{10} \right) \mathcal{S}_L(x - x_0) \leq -\frac{\mathfrak{c}}{2} \frac{a_L}{d_L} \,. \end{split}$$

where we used (1.2) and, in the last step, that, for L large enough, the quantity in parenthesis is larger than 1/2. This establishes (2.16), which in turn implies both properties ((1)) and ((2)).

Now, assume (2.17) and observe that

$$\mathbb{P}\left(\max_{x\in Q_{R_L+r_L}}\xi_L(x) \ge a_L - \theta; \left(\bigcup_{x_0\in Q_{R_L}}E_{L,x_0}\right)^{\mathsf{L}}\right) \\
\le \sum_{x\in Q_{R_L+r_L}\setminus Q_{R_L}}\mathbb{P}(\xi_L(x) \ge a_L - \theta) + \sum_{x_0\in Q_{R_L}}\mathbb{P}(\xi_L(x_0) \ge a_L - \theta; E_{L,x_0}^{\mathsf{C}}) \\
\lesssim r_L R_L^{d-1}\frac{e^{a_L\theta}}{L^d} + \frac{R_L^d}{L^d}e^{-C\left(\frac{a_L}{d_L}\right)^{2\kappa}},$$

where we have also used (1.23). Since $\ln r_L \ll a_L \ll \ln R_L$, we deduce that the r.h.s. is negligible compared to $(R_L/L)^d$. Hence, it remains to argue (2.17).

For this, we show that each of the summands at the r.h.s. of

$$\mathbb{P}\big(\xi_L(x_0) \ge a_L - \theta; (E_{L,x_0})^{\complement}\big) \le \sum_{i=1}^3 \mathbb{P}\big(\xi_L(x_0) \ge a_L - \theta; (E_{L,x_0}^i)^{\complement}\big)$$

can be bounded above by a term of the desired order. Now, the first can be controlled via (1.23), which gives

$$\mathbb{P}\big(\xi_L(x_0) \ge a_L - \theta; (E_{L,x_0}^1)^{\complement}\big) = \mathbb{P}\big(\xi_L(x_0) \ge a_L + \theta\big) \lesssim \frac{e^{-\theta a_L}}{L^d} (1 + o(1))$$

and since $a_L \gg (a_L/d_L)^{2\kappa}$, the r.h.s. is bounded above by the r.h.s. of (2.17).

For the second, we recall that, by Lemma 2.1, $\xi_L(x_0)$ and ζ_{L,x_0} are independent, thus so are the events $\{\xi_L(x_0) \ge a_L - \theta\}$ and E^2_{L,x_0} . Hence,

$$\mathbb{P}\big(\xi_L(x_0) \ge a_L - \theta; (E_{L,x_0}^2)^{\complement}\big) = \mathbb{P}\big(\xi_L(x_0) \ge a_L - \theta\big)\mathbb{P}\big((E_{L,x_0}^2)^{\complement}\big)$$

and we only need to focus on the latter factor. A union bound yields

$$\mathbb{P}((E_{L,x_0}^2)^c) \le \sum_{x \in Q_{2R_L,x_0}} \mathbb{P}(|\zeta_{L,x_0}(x)| \ge \frac{1}{10} \mathcal{S}_L(x-x_0)).$$

Since $v_L \leq 1$, we find

16

$$\frac{\mathcal{S}_L(x-x_0)^2}{\operatorname{Var}[\zeta_{L,x_0}(x)]} = \frac{a_L^2 \left[1 - v_L(x-x_0)\right]^2}{(1 - v_L(x-x_0)^2)} = \frac{a_L^2 \left[1 - v_L(x-x_0)\right]}{(1 + v_L(x-x_0))} \ge \frac{\mathfrak{c}}{2} \frac{a_L^2}{d_L}$$

the last step being a consequence of (1.2). Note that the r.h.s. is larger than 1 for all L large enough, so that (1.20) and the definition of R_L yield

$$\sum_{x \in Q_{2R_L,x_0}} \mathbb{P}\Big(|\zeta_{L,x_0}(x)| \ge \frac{1}{10} \mathcal{S}_L(x-x_0)\Big) \le \sum_{x \in Q_{2R_L,x_0}} \exp\Big(-\frac{\mathcal{S}_L(x-x_0)^2}{200 \text{Var}[\zeta_{L,x_0}(x)]}\Big) \le \#Q_{2R_L} e^{-\mathfrak{c}\frac{a_L^2}{400d_L}} \le e^{-C'\frac{a_L}{d_L}a_L}$$

for some constant C' > 0, where we used (1.15).

We turn to the event E_{L,x_0}^3 . Let $\sigma_{L,x_0}(x) \stackrel{\text{def}}{=} \sqrt{\operatorname{Var}[\zeta_{L,x_0}(x)]}$. Then

$$\mathbb{P}\left(\xi_{L}(x_{0}) \geq a_{L} - \theta; (E_{L,x_{0}}^{3})^{\complement}\right)$$

$$\leq \sum_{x \in Q_{R_{L},x_{0}}^{\neq x_{0}}} \mathbb{P}\left(\xi_{L}(x_{0}) \geq a_{L} - \theta; \frac{|\zeta_{L,x_{0}}(x)|}{\sigma_{L,x_{0}}(x)} > \left(\frac{a_{L}}{d_{L}}\right)^{\kappa|x-x_{0}|} \sqrt{1 \vee \left(|\xi_{L}(x_{0}) - a_{L}|a_{L}\right)}\right)$$
(2.19)

By translation invariance, the probability at the r.h.s. does not depend on x_0 . The independence of $\zeta_{L,x_0}(x)$ and $\xi_L(x_0)$ then implies

$$\mathbb{P}\Big(\xi_L(x_0) \ge a_L - \theta; \frac{|\zeta_{L,x_0}(x)|}{\sigma_{L,x_0}(x)} > \Big(\frac{a_L}{d_L}\Big)^{\kappa|x-x_0|} \sqrt{1 \lor (|\xi_L(x_0) - a_L|a_L)}\Big) \\
= \int_{a_L-\theta}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \mathbb{P}\Big(\frac{|\zeta_{L,x_0}(x)|}{\sigma_{L,x_0}(x)} > \Big(\frac{a_L}{d_L}\Big)^{\kappa|x-x_0|} \sqrt{1 \lor (|z-a_L|a_L)}\Big) dz \\
\le \int_{a_L-\theta}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\Big(-\frac{z^2}{2} - \frac{1}{2}\Big(\frac{a_L}{d_L}\Big)^{2\kappa|x-x_0|} 1\lor (|z-a_L|a_L)\Big) dz.$$

where we have used (1.20) to go from the second to the third line. We now split the domain of integration into $I_1 \stackrel{\text{def}}{=} [a_L - \theta, a_L - \frac{1}{a_L}]$ and $I_2 \stackrel{\text{def}}{=} (a_L - \frac{1}{a_L}, \infty)$. Using (1.23), the integral over I_2 is bounded by

$$e^{-\frac{1}{2}(\frac{a_L}{d_L})^{2\kappa|x-x_0|}} \mathbb{P}(\mathcal{N}(0,1) \ge a_L - \frac{1}{a_L}) \lesssim \frac{1}{L^d} e^{1 - \frac{1}{2}(\frac{a_L}{d_L})^{2\kappa|x-x_0|}}$$

while the integral over I_1 can be rewritten as $(take \ y = (a_L - z)a_L)$

$$\int_{1}^{\theta a_{L}} \frac{1}{\sqrt{2\pi}a_{L}} \exp\left(-\frac{a_{L}^{2}}{2} - \frac{y^{2}}{2a_{L}^{2}} - y\left(\frac{1}{2}\left(\frac{a_{L}}{d_{L}}\right)^{2\kappa|x-x_{0}|} - 1\right)\right) \mathrm{d}y$$
$$\lesssim \frac{1}{L^{d}} \int_{1}^{\infty} \exp\left(-\frac{y}{4}\left(\frac{a_{L}}{d_{L}}\right)^{2\kappa|x-x_{0}|}\right) \mathrm{d}y \lesssim \frac{1}{L^{d}} \exp\left(-\frac{1}{4}\left(\frac{a_{L}}{d_{L}}\right)^{2\kappa|x-x_{0}|}\right).$$

thanks to $d_L \ll a_L$ and (1.22). Plugging these estimates into (2.19) we are left to control

$$\frac{1}{L^d} \sum_{x \in Q_{R_L, x_0}^{\neq x_0}} \exp\left(-\frac{1}{4} \left(\frac{a_L}{d_L}\right)^{2\kappa |x-x_0|}\right) = \frac{1}{L^d} \sum_{x \in Q_{R_L}^{\neq 0}} \exp\left(-\frac{1}{4} \left(\frac{a_L}{d_L}\right)^{2\kappa |x|}\right),$$

from which (2.17) follows at once.

For the last estimate: There exists a C > 0 such that for all $b \in (0, 1)$ and R > 0

$$\sum_{x \in Q_R^{\neq 0}} b^{|x|} \le Cb.$$

We can estimate the above sum by a multiple of following:

$$\int_{1}^{\infty} r^{d-1} e^{r \ln b} dr = \frac{1}{\ln b} [r^{d-1} e^{r \ln b}]|_{1}^{\infty} - (d-1) \int_{1}^{\infty} r^{d-2} e^{r \ln b} dr$$
$$= \frac{b}{\ln b} - \frac{d-1}{(\ln b)^{2}} \int_{1}^{\infty} r^{d-2} \frac{d}{dr} e^{r \ln b} dr,$$

so we get something like $bP((\ln b)^{-1})$, where P is a polynomial. As $(\ln b)^{-1} < 1$, $P((\ln b)^{-1}) \leq 1$.

We now derive two properties of the maxima of the fields ξ_L and Ξ_L (whose definition is in (2.5)) on Q_{R_L} , which are implied by the previous lemma. Before that, since thanks to (1.23) and (2.8), we know that the maxima of ξ_L and Ξ_L are of respective sizes a_L in (1.4) and a_L^{Ξ} in (2.12), and that the order of the fluctuations around them is a_L^{-1} , let us introduce the rescaled fields

$$\Theta_L^{\xi}(x) \stackrel{\text{def}}{=} a_L(\xi_L(x) - a_L), \quad \Theta_L^{\Xi}(x) \stackrel{\text{def}}{=} a_L(\Xi_L(x) - a_L^{\Xi}), \qquad x \in \mathbb{Z}^d.$$
(2.20)

At first, we want to show that on a box of size R_L both ξ_L and Ξ_L can have at most one point at which Θ_L^{ξ} and Θ_L^{Ξ} are order 1 (and this is the point at which they achieve their maxima). For this, for $z \in \mathbb{Z}^d$ and $s \in \mathbb{R}$, we will show that the events

$$\mathcal{A}_{L,z}^{\chi}(s) \stackrel{\text{def}}{=} \bigcup_{x \neq y \in Q_{R_L,z}} \left\{ \Theta_L^{\chi}(x) \ge -s \; ; \; \Theta_L^{\chi}(y) \ge -s \right\}, \tag{2.21}$$

for χ either ξ or Ξ , are asymptotically negligible.

To phrase the second property, for $z \in \mathbb{Z}^d$, let

$$w_{L,z} \stackrel{\text{def}}{=} \underset{Q_{R_{L,z}}}{\operatorname{arg\,max}} \xi_L \quad \text{and} \quad w_{L,z}^{\Xi} \stackrel{\text{def}}{=} \underset{Q_{R_{L,z}}}{\operatorname{arg\,max}} \Xi_L \,.$$
 (2.22)

Then, we want to verify that, provided the maximum of Ξ_L is "large", $w_L = w_L^{\Xi}$ with high probability.

For both features, an important step is to show that if the maximum of Ξ_L is of order a_L^{Ξ} , then the maximum of ξ_L must be of order a_L . This is the first point of the next lemma. The argument we will exploit uses the exponential decay of the principal normalised eigenfunction $\bar{\varphi}_L$ of the operator $\bar{\mathcal{H}}_L$ in (1.8), which is a standard fact (independent of the specific setting of the present paper) and will

17

anyway be detailed in Lemma 4.4. For the reader's convenience, let us anticipate that this amounts to say that there exists a constant C > 0 such that

$$\bar{\varphi}_L(x)^2 \le C \left(\frac{a_L}{d_L}\right)^{-2|x|} \qquad \forall x \in Q_{r_L}^{\neq 0}.$$
(2.23)

Lemma 2.7. The following statements hold:

(i) For $\theta = 2d+1$ as in Definition 2.4 and for any $s \in \mathbb{R}$, there exists $L_0 \ge 1$ such that for every $L \ge L_0$ we have

$$\left\{\max_{x\in Q_{R_L,z}}\Xi_L(x)\geq a_L^{\Xi}-\frac{s}{a_L}\right\}\cap\left\{\max_{x\in Q_{R_L+r_L,z}}\xi_L(x)\leq a_L-\theta\right\}=\emptyset,\qquad(2.24)$$

(ii) For every $z \in \mathbb{Z}^d$ and every $s \in \mathbb{R}$, it holds

$$\lim_{L \to \infty} \left(\frac{L}{R_L}\right)^d \mathbb{P}\left(\mathcal{A}_{L,z}^{\chi}(s)\right) = 0, \qquad (2.25)$$

where the events $\mathcal{A}_{L,z}^{\chi}(s)$ are defined according to (2.21) and χ is either ξ or Ξ ,

(iii) For every $z \in \mathbb{Z}^d$, we have with $w_{L,z}$ and $w_{L,z}^{\Xi_L}$ as in (2.22), we have

$$\lim_{L \to \infty} \left(\frac{L}{R_L}\right)^d \mathbb{P}\left(\max_{x \in Q_{R_L,z}} \Xi_L(x) \ge a_L^{\Xi} - \frac{s}{a_L}; w_{L,z} \ne w_{L,z}^{\Xi}\right) = 0.$$
(2.26)

Remark 2.8. The reason why in (2.24) the maximum of ξ_L is taken on $Q_{R_L+r_L}$, while that of Ξ_L on Q_{R_L} , is that, by definition, see (2.5) and (2.4), Ξ_L depends on the values of ξ_L on a box of side-length $R_L + r_L$.

Proof. W.l.o.g. we can (and will) take z = 0 throughout the proof and omit the corresponding subscript, i.e. $\mathcal{A}_L^{\chi}(s) = \mathcal{A}_{L,0}^{\chi}(s), w_L = w_{0,L}$, etc. We begin by proving (i). On $\{\max_{Q_{R_L}+r_L} \xi_L \leq a_L - \theta\}$, using that by definition, $\Xi_L = \xi_L + \Phi_L$ (2.5), where Φ_L is as in (2.4), and recalling $\zeta_{L,}$ from (2.1), we can control Ξ_L at any $y \in Q_{R_L}$ as

$$\Xi_L(y) = \xi_L(y) \Big(1 - \sum_{x \in Q_{r_L}^{\neq 0}} \bar{\varphi}_L(x)^2 v_L(x) \Big) + \sum_{x \in Q_{r_L}^{\neq 0}} \bar{\varphi}_L(x)^2 \xi_L(y+x) \\ \leq (a_L - \theta) \Big(1 + \sum_{x \in Q_{r_L}^{\neq 0}} \bar{\varphi}_L(x)^2 (1 - v_L(x)) \Big) \,.$$

Now, the lower bound in (1.2) and the bound in (2.23) ensure that the remaining sum can be bounded by

$$\sum_{x \neq 0} \bar{\varphi}_L(x)^2 (1 - v_L(x)) \le \frac{1}{d_L} \sum_{x \in Q_{r_L}^{\neq 0}} \left(\frac{d_L}{a_L}\right)^{2|x|} e^{\mathfrak{c}'|x|} \lesssim \frac{d_L}{a_L^2} \,.$$

Since further $d_L \ll a_L$ by Assumption 1.2, we deduce for large L

$$\Xi_L(y) \le (a_L - \theta) \left(1 + C \left(\frac{d_L}{a_L^2} \right) \right) \le a_L - \theta + C \left(\frac{d_L}{a_L} \right)$$
$$< a_L - \frac{\theta}{2} < a_L \sqrt{1 + \tau_L^2} - \frac{s}{a_L} = a_L^{\Xi} - \frac{s}{a_L} .$$

Henceforth, if $\max \xi_L \leq a_L - \theta$, also $\max \Xi_L < a_L^{\Xi} - s/a_L$ and (i) is proved.

Let us begin by proving (2.25).

Notice that for L large enough, \mathcal{A}_L^{χ} is contained in $\{\max_{Q_{R_L}+r_L} \xi_L \geq a_L - \theta\}$: for $\chi = \xi$ this holds by definition, while for $\chi = \Xi$ it follows from $\mathcal{A}_L^{\Xi} \subset \{\max_{Q_{R_L}} \Xi_L \geq a_L^{\Xi} - s/a_L\}$ and (2.24). As a consequence,

$$\mathbb{P}\Big(\mathcal{A}_{L}^{\chi}\cap\Big(\bigcup_{x_{0}\in Q_{R_{L}}}E_{L,x_{0}}\Big)^{\complement}\Big)\leq\mathbb{P}\Big(\max_{x\in Q_{R_{L}+r_{L}}}\xi_{L}(x)\geq a_{L}-\theta\,;\,\Big(\bigcup_{x_{0}\in Q_{R_{L}}}E_{L,x_{0}}\Big)^{\complement}\Big)\,,$$

and, by (2.18) since $r_L \ll R_L$, the r.h.s. is negligible compared to $(R_L/L)^d$. Thus,

$$\mathbb{P}(\mathcal{A}_{L}^{\chi}) = \mathbb{P}\left(\mathcal{A}_{L}^{\chi} \cap \bigcup_{x_{0} \in Q_{R_{L}}} E_{L,x_{0}}\right) + o\left(\frac{R_{L}^{a}}{L^{d}}\right).$$
(2.27)

Now, for $\chi = \xi$, we exploit property (2) from Proposition 2.5 which implies that

$$\mathbb{P}\Big(\mathcal{A}_{L}^{\xi} \cap \bigcup_{x_{0} \in Q_{R_{L}}} E_{L,x_{0}}\Big) = \sum_{x_{0} \in Q_{R_{L}}} \mathbb{P}(\mathcal{A}_{L}^{\xi} \cap E_{L,x_{0}}).$$

We have $\mathcal{A}_{L}^{\xi} \cap E_{L,x_{0}} = \emptyset$. Indeed, on $E_{L,x_{0}}$, ξ_{L} has a unique maximum at x_{0} and ξ_{L} is "small" nearby due to (2.16), and we conclude that (2.25) holds. On $E_{L,x_{0}}$ we have for x with $1 \leq |x - x_{0}| \leq 2R_{L}$, and for large L,

$$\xi_L(x) \le \xi_L(x_0) - \frac{\mathfrak{c}}{2} \frac{a_L}{d_L} \le a_L + \theta - \frac{\mathfrak{c}}{2} \frac{a_L}{d_L} < a_L - \frac{s}{a_L}$$

Hence, no other x in Q_{R_L} than x_0 itself can be such that $\xi_L(x) \ge a_L - \frac{s}{a_L}$, i.e. $\Theta_L^{\xi}(x) \ge -s$.

For $\chi = \Xi$ instead, let us introduce the event \mathcal{F}_L which is defined by

$$\mathcal{F}_L \stackrel{\text{def}}{=} \bigcup_{y \in Q_{R_L}} \left\{ \Xi_L(y) \ge a_L^{\Xi} - \frac{s}{a_L}; \, \xi_L(y) < a_L - \theta \right\}.$$

Thanks to (2.11), its probability is negligible compared to $(R_L/L)^d$, so that

$$\mathbb{P}\left(\mathcal{A}_{L}^{\Xi} \cap \bigcup_{x_{0} \in Q_{R_{L}}} E_{L,x_{0}}\right) \leq \mathbb{P}(\mathcal{F}_{L}) + \mathbb{P}\left(\mathcal{A}_{L}^{\Xi} \cap \bigcup_{x_{0} \in Q_{R_{L}}} E_{L,x_{0}} \cap \mathcal{F}_{L}^{\complement}\right)$$
$$\leq o\left(\frac{R_{L}^{d}}{L^{d}}\right) + \sum_{x_{0} \in Q_{R_{L}}} \mathbb{P}\left(\mathcal{A}_{L}^{\Xi} \cap E_{L,x_{0}} \cap \mathcal{F}_{L}^{\complement}\right).$$

By (2.16), on E_{L,x_0} , $\xi_L(y) < a_L + \theta - \frac{c}{2}(a_L/d_L) < a_L - \theta$ for L large enough and any $y \neq x_0$, so that on $E_{L,x_0} \cap \mathcal{F}_L^{\complement}$, we must necessarily have that for every $y \neq x_0, \ \Xi_L(y) < a_L^{\Xi} - s/a_L$. But this means that $\mathcal{A}_L^{\Xi} \cap E_{L,x_0} \cap \mathcal{F}_L^{\complement} = \emptyset$ which, together with (2.27), implies that (2.25) holds also for $\chi = \Xi$.

At last, we show (iii). We already know from (2.24) that $\{\max \Xi_L \geq a_L^{\Xi} - s/a_L; w_L \neq w_L^{\Xi_L}\}$ is contained in $\{\max \xi_L \geq a_L - \theta\}$, so that, by (2.18), the probability of the former is the same as that of its intersection with the union of the $E_{L,\cdot}$ up to an error negligible compared to R_L^d/L^d . Moreover, by (2.25) for $\chi = \Xi$ and (2.11), we have that $\mathbb{P}(\mathcal{A}_L^{\Xi})$ and $\mathbb{P}(\mathcal{F}_L)$ are negligible compared to $(R_L/L)^d$. Putting all these together, we deduce

 $\mathbb{P}\Big(\max_{x \in Q_{R_L}} \Xi_L \ge a_L^{\Xi} - \frac{s}{a_L} \, ; \, w_L \neq w_L^{\Xi_L}\Big)$

19

$$= \mathbb{P}\Big(\max_{x \in Q_{R_L}} \Xi_L \ge a_L^{\Xi} - \frac{s}{a_L}; w_L \neq w_L^{\Xi_L}; \bigcup_{x_0 \in Q_{R_L}} E_{L,x_0}; (\mathcal{A}_L^{\Xi})^{\complement}; \mathcal{F}_L^{\complement}\Big) + o\Big(\frac{R_L^d}{L^d}\Big)$$
$$= \sum_{x_0 \in Q_{R_L}} \mathbb{P}\Big(\max_{x \in Q_{R_L}} \Xi_L \ge a_L^{\Xi} - \frac{s}{a_L}; x_0 \neq w_L^{\Xi_L}; E_{L,x_0}; (\mathcal{A}_L^{\Xi})^{\complement}; \mathcal{F}_L^{\complement}\Big) + o\Big(\frac{R_L^d}{L^d}\Big).$$

But now, each of the summands above is 0. Indeed, on $(\mathcal{A}_L^{\Xi})^{\complement}$ there is at most one point on Q_{R_L} at which Ξ_L is above $a_L^{\Xi} - s/a_L$, and, for every $x_0 \in Q_{R_L}$, on $E_{L,x_0} \cap \mathcal{F}_L^{\complement}$, $\max_{y \neq x_0} \Xi_L(y) < a_L^{\Xi} - s/a_L$, which implies that the only point at which Ξ_L can be above $a_L^{\Xi} - s/a_L$ is x_0 . Thus, $w_L^{\Xi_L} = x_0$ and the intersection of the events is empty. \Box

2.3. Tail distribution of the maximum. A first major consequence of the analysis carried out in the previous section is that it provides a rather simple way to determine the tail distributions of $\xi_L(w_L)$, $\Xi_L(w_L)$ and of the couple $(\xi_L(w_L), \Phi_L(w_L))$ where ξ_L is our potential, Ξ_L and Φ_L are the fields respectively defined in (2.5) and (2.4), and w_L is the point in Q_{R_L} at which ξ_L achieves its maximum, i.e. $w_L = w_{L,0}$ and the latter is given in (2.22). The next theorem enucleates the rigorous statement we are after.

Theorem 2.9. As $L \to \infty$, the Radon measures

$$\left(\frac{L}{R_L}\right)^d \mathbb{P}\left(a_L\left(\xi_L(w_L) - a_L\right) \in \mathrm{d}u\right), \text{ and } \left(\frac{L}{R_L}\right)^d \mathbb{P}\left(a_L\left(\Xi_L(w_L) - a_L^{\Xi}\right) \in \mathrm{d}u\right)$$
(2.28)

on $(-\infty, \infty]$ converge vaguely to e^{-u} du. Furthermore, if $a_L \tau_L \sim \sqrt{b}$ for $b \ge 0$, then the Radon measure on $(-\infty, \infty] \times [-\infty, \infty]$ given by

$$\left(\frac{L}{R_L}\right)^d \mathbb{P}\left(\left(a_L\left(\xi_L(w_L) - a_L\right), a_L\Phi_L(w_L)\right) \in \mathrm{d}u \otimes \mathrm{d}v\right), \qquad (2.29)$$

converges vaguely to $e^{-u} du \otimes \frac{1}{\sqrt{2\pi b}} e^{-\frac{v^2}{2b}} dv$.

Remark 2.10. The restriction $a_L \tau_L \sim \sqrt{b}$ for $b \ge 0$ in the second part of the statement can be easily lifted (and the proof would be unaffected) to cover also the case $\tau_L \gg a_L^{-1}$. This would require to scale the second component in (2.29) to be $\Phi_L(w_L)/\tau_L$ instead of $a_L \Phi_L(w_L)$ and take b = 1 in the limiting measure.

The reason why we stated the result as such is that the joint convergence will only be needed when $\tau_L = \mathcal{O}(a_L^{-1})$, in which case the scaling proposed is more meaningful (see the proof of point (3) in Theorem 5.1).

Remark 2.11. The convergence in the second part of the statement means that for any bounded continuous function $f: (-\infty, \infty] \times [-\infty, \infty] \to \mathbb{R}$ that vanishes outside some set $(-c, \infty] \times [-\infty, \infty]$ for some c > 0, we have

$$\left(\frac{L}{R_L}\right)^d \mathbb{E}\left[f\left(a_L\left(\xi_L(w_L) - a_L\right), a_L\Phi_L(w_L)\right)\right] \to \int f(u, v)e^{-u} \mathrm{d}u \otimes \frac{1}{\sqrt{2\pi b}}e^{-\frac{v^2}{2b}} \mathrm{d}v ,$$

and similarly for the first part of the statement.

Proof. We present a joint proof of the convergence of the measures in (2.28) and in (2.29). Let Θ_L^{ξ} and Θ_L^{Ξ} be as in (2.20) and similarly set

$$\Theta_L^{(\xi,\Phi)}(x) \stackrel{\text{def}}{=} \left(a_L(\xi_L(x) - a_L), a_L \Phi_L(x) \right), \qquad x \in \mathbb{Z}^d.$$

20

Below χ will denote either ξ, Ξ or (ξ, Φ) .

Now, in either case the limit measures have no atoms, thus it suffices to determine the behaviour of the measures when evaluated at sets of the form $I^{\chi} = (u, \infty]$, for $\chi = \xi, \Xi$, and $I^{\chi} = (u, \infty] \times (v, \infty]$ for $\chi = (\xi, \Phi)$, with $u, v \in \mathbb{R}$. We claim that

$$\mathbb{P}\Big(\Theta_L^{\chi}(w_L) \in I^{\chi}\Big) = \sum_{x_0 \in Q_{R_L}} \mathbb{P}\Big(\Theta_L^{\chi}(x_0) \in I^{\chi}\Big) + o\Big(\frac{R_L^d}{L^d}\Big).$$
(2.30)

Before proving (2.30), let us see how it implies the result. For $\chi = \xi$ or Ξ this is an immediate consequence of the fact that the law of $\xi_L(x_0)$ and $\Xi_L(x_0)$ is independent of x_0 , and of (1.23) and (2.8) respectively. For the other, the independence of $\xi_L(x_0)$ and ζ_{L,x_0} stated in Lemma 2.1 implies

$$\mathbb{P}\Big(\Theta^{\xi_L}(x_0) > u; a_L \Phi_L(x_0) > v\Big) = \mathbb{P}\Big(a_L(\xi_L(x_0) - a_L) > u\Big)\mathbb{P}\Big(a_L \Phi_L(x_0) > v\Big) \\ = \frac{1}{L^d} e^{-u}(1 + o(1)) \int_v^\infty \frac{1}{\sqrt{2\pi b}} e^{-\frac{w^2}{2b}} dw,$$

where in the last equality, we used (1.23) on the first factor, and the fact that $a_L \Phi_L(x_0)$ is a centred Gaussian random variable with variance $(a_L \tau_L)^2 \sim b$ on the second.

Thus, it remains to show the claim. Note that $\{\Theta_L^{\chi}(w_L) \in I^{\chi}\} \subset \{\max_{Q_{R_L}+r_L} \xi_L \ge a_L - \theta\}$, for $\chi = \Xi$, by Lemma 2.7 while for $\chi = \xi$, (ξ, Φ) , by definition. Thus, thanks to (2.18) we have

$$\mathbb{P}(\Theta_{L}^{\chi}(w_{L}) \in I^{\chi}) = \mathbb{P}\left(\{\Theta_{L}^{\chi}(w_{L}) \in I^{\chi}\} \cap \bigcup_{x_{0} \in Q_{R_{L}}} E_{L,x_{0}}\right)$$
$$+ \mathbb{P}\left(\{\Theta_{L}^{\chi}(w_{L}) \in I^{\chi}\} \cap \left(\bigcup_{x_{0} \in Q_{R_{L}}} E_{L,x_{0}}\right)^{\complement}\right) \qquad (2.31)$$
$$= \sum_{x_{0} \in Q_{R_{L}}} \mathbb{P}\left(\Theta_{L}^{\chi}(x_{0}) \in I^{\chi}; E_{L,x_{0}}\right) + o\left(\frac{R_{L}^{d}}{L^{d}}\right).$$

Now we argue separately for the three measures. For $\chi = \xi$ or (ξ, Φ) , we write

$$\mathbb{P}\Big(\Theta_L^{\chi}(x_0) \in I^{\chi}; E_{L,x_0}\Big) = \mathbb{P}\Big(\Theta_L^{\chi}(x_0) \in I^{\chi}\Big) - \mathbb{P}\Big(\Theta_L^{\chi}(x_0) \in I^{\chi}; E_{L,x_0}^{\complement}\Big)$$

and the second term is bounded by $\mathbb{P}(\xi_L(x_0) \ge a_L - \theta; E_{L,x_0}^{\complement})$ which is negligible compared to $1/L^d$ by (2.17), so that the claim follows. For $\chi = \Xi$, we write

$$\mathbb{P}\left(\Theta_{L}^{\Xi}(x_{0}) \in I^{\Xi}; E_{L,x_{0}}\right) = \mathbb{P}\left(\Theta_{L}^{\Xi}(x_{0}) \in I^{\Xi}; |\xi_{L}(x_{0}) - a_{L}| \leq \theta\right)$$

$$-\mathbb{P}\left(\Theta_{L}^{\Xi}(x_{0}) \in I^{\Xi}; |\xi_{L}(x_{0}) - a_{L}| \leq \theta; (E_{L,x_{0}})^{\complement}\right)$$

$$(2.32)$$

and the second term can be bounded by $\mathbb{P}(\xi_L(x_0) \ge a_L - \theta; (E_{L,x_0})^{\complement})$ which, once again, is negligible compared to $1/L^d$ by (2.17). Regarding the first term

$$\mathbb{P}(\Theta_L^{\Xi}(x_0) \in I^{\Xi}; |\xi_L(x_0) - a_L| \le \theta) = \mathbb{P}(\Theta_L^{\Xi}(x_0) \in I^{\Xi}) - \mathbb{P}(\Theta_L^{\Xi}(x_0) \in I^{\Xi}; |\xi_L(x_0) - a_L| > \theta)$$

where the second summand is again negligible compared to $1/L^d$ in view of (2.11) since, by Assumption 1.5, $\theta = 2d + 1 \gg \max\{a_L \tau_L^2, a_L^{-1}\}$, and thus the proof is complete.

3. STATISTICS OF THE MAXIMA OF CORRELATED GAUSSIAN FIELDS

The primary goal of this section is to identify the statistics of the maxima of the potential ξ_L on Q_L as $L \to \infty$ and thus establish Theorem 1.3. Actually, we will not only consider the maxima of ξ_L , but also of the fields Ξ_L and (ξ_L, Φ_L) , as these quantities are instrumental in the determination of the top of the spectrum of the Anderson Hamiltonian on Q_L , and its relation to the (location of the) maxima of ξ_L .

Our analysis will rely on the splitting scheme introduced in Section 1.3. We will restrict ourselves to the study of the maxima of the fields on U_L , since, on $Q_L \setminus U_L$, they remain "small" with large probability. As already mentioned in Section 1.3, U_L is a union of mesoscopic boxes which lie at a distance at least $\sqrt{R_L}$ from one another. A crucial step of our analysis will be to establish suitable decorrelation estimates, which, roughly speaking, allow to regard the restrictions of ξ_L to the mesoscopic boxes $Q_{R_L,z_{j,L}}$, $j \in \{1,\ldots,n_L\}$, as independent. As these are technically challenging, we will postpone their statement and proof to the end of this section, in Section 3.3.

3.1. Convergence of the maxima and fluctuations on U_L . We will deal with the rescaled fields Θ_L^{ξ} , Θ_L^{Ξ} (see (2.20)) and $\Theta_L^{(\xi,\Phi)}$, where

$$\Theta_L^{\xi}(x) = a_L \big(\xi_L(x) - a_L\big), \quad \Theta_L^{\Xi}(x) = a_L \big(\Xi_L(x) - a_L^{\Xi}\big), \qquad x \in \mathbb{Z}^d,$$

for Ξ_L as in (2.5) and a_L^{Ξ} in (2.12), and

$$\Theta_L^{(\xi,\Phi)}(x) \stackrel{\text{def}}{=} \left(a_L \big(\xi_L(x) - a_L \big), a_L \Phi_L(x) \big), \qquad x \in \mathbb{Z}^d.$$
(3.1)

for Φ_L as in (2.4) and τ_L in (1.11). For any $j \in \{1, \ldots, n_L\}$, denote by $w_{j,L}$ the point in $Q_{R_L, z_{j,L}}$ where ξ_L achieves its maximum (recall the definition of $z_{j,L}$ and n_L in (1.18)).

For $\chi = \xi$, Ξ or (ξ, Φ) , the random (point) measures of interest are

$$\mathcal{P}_{L}^{\chi} \stackrel{\text{def}}{=} \sum_{j=1}^{n_{L}} \delta_{\left(\frac{z_{j,L}}{L},\Theta_{L}^{\chi}(w_{j,L})\right)}, \qquad (3.2)$$

and our first goal is to establish their vague convergence.

Proposition 3.1. As $L \to \infty$, each random measure \mathcal{P}_L^{χ} in (3.2) converges in law to a Poisson random measure $\mathcal{P}_{\infty}^{\chi}$ where

- for $\chi = \xi$ or Ξ , the convergence holds for the topology of vague convergence on $[-1,1]^d \times (-\infty,\infty]$, and, in both cases, the limiting Poisson measure has intensity $dx \otimes e^{-u} du$,
- for $\chi = (\xi, \Phi)$, we further assume that $a_L \tau_L \sim \sqrt{b}$ for some $b \ge 0$. Then the convergence holds for the topology of vague convergence on $[-1, 1]^d \times$ $(-\infty, \infty] \times [-\infty, \infty]$, and the limiting Poisson measure has intensity $dx \otimes$ $e^{-u} du \otimes \frac{1}{\sqrt{2\pi b}} e^{-\frac{v^2}{2b}} dv$.

Remark 3.2. The restriction $a_L \tau_L \sim \sqrt{b}$ in case $\chi = (\xi, \Phi)$ can be lifted following the same changes discussed in Remark 2.10.

To define the notion of convergence stated in the above theorem, set $I^{\chi} \stackrel{\text{def}}{=} [-1,1]^d \times (-\infty,\infty]$ for $\chi = \xi$ or Ξ , and $I^{\chi} \stackrel{\text{def}}{=} [-1,1]^d \times (-\infty,\infty] \times [-\infty,\infty]$ for $\chi = (\xi, \Phi)$. Then, \mathcal{P}_L^{χ} is said to converge in law in the topology of vague convergence to $\mathcal{P}_{\infty}^{\chi}$ provided that for any continuous function $g: I^{\chi} \to \mathbb{R}$ with compact support, the real-valued random variable $\mathcal{P}_L^{\chi}(g)$ converges in law to $\mathcal{P}_{\infty}^{\chi}(g)$. We refer to [Kal02, Theorem 16.16 and Theorem A2.3] for further details on this topology.

Proof. Let $\chi = \xi, \Xi$ or (ξ, Φ) . Let $g: I^{\chi} \to \mathbb{R}$ with compact support, non-negative and of class \mathcal{C}^2 . Thanks to [Kal02, Theorem 16.16]⁷, if we show that for all $\lambda \geq 0$

$$\mathbb{E}[\exp(-\lambda \mathcal{P}_L^{\chi}(g))] \to \mathbb{E}[\exp(-\lambda \mathcal{P}_{\infty}^{\chi}(g))], \quad \text{as } L \to \infty,$$
(3.3)

then we conclude that $\mathcal{P}_L^{\chi}(g)$ converges in law to $\mathcal{P}_{\infty}^{\chi}(g)$. Since g was arbitrary, the very definition of the vague convergence of point measures ensures that this is enough to establish the statement.

To prove (3.3), let us fix g and λ as above. First of all, observe that

$$\mathbb{E}[\exp(-\lambda \mathcal{P}_L^{\chi}(g))] = \mathbb{E}\left[\prod_{j=1}^{n_L} \exp(-\lambda g(\frac{z_{j,L}}{L}, \Theta_L^{\chi}(w_{j,L})))\right]$$

We claim that

$$\mathbb{E}\Big[\prod_{j=1}^{n_L} \exp(-\lambda g(\frac{z_{j,L}}{L}, \Theta_L^{\chi}(w_{j,L}))\Big] = \prod_{j=1}^{n_L} \mathbb{E}\Big[\exp(-\lambda g(\frac{z_{j,L}}{L}, \Theta_L^{\chi}(w_{j,L}))\Big] + o(1) ,$$
(3.4)

where o(1) is a quantity that vanishes as $L \to \infty$. Given (3.4), we have

$$\mathbb{E}[\exp(-\lambda \mathcal{P}_L^{\chi}(g))] = \exp\left(\sum_{j=1}^{n_L} \ln\left(1 - \mathbb{E}[1 - \exp(-\lambda g(\frac{z_{j,L}}{L}, \Theta_L^{\chi}(w_{j,L}))]\right)\right) + o(1) .$$

By Theorem 2.9, uniformly over all j the expectation on the r.h.s. is of order $(R_L/L)^d$. Since n_L is of order $(L/R_L)^d$, we deduce that the last term equals

$$\exp\left(-\sum_{j=1}^{n_L} \mathbb{E}[1-\exp(-\lambda g(\frac{z_{j,L}}{L},\Theta_L^{\chi}(w_{j,L}))])\right)(1+o(1))$$
$$=\exp\left(-\sum_{j=1}^{n_L} \int \left(1-\exp(-\lambda g(\frac{z_{j,L}}{L},q))\right) \mathbb{P}(\Theta_L^{\chi}(w_{j,L}) \in \mathrm{d}q)\right)(1+o(1)) .$$

Invoking Theorem 2.9 once again, the latter converges to

$$\exp\left(-\int\int\left(1-\exp(-\lambda g(x,u))\right)\mathrm{d}x\otimes e^{-u}\mathrm{d}u\right)\,,$$

if $\chi = \xi$ or Ξ , and to

$$\exp\left(-\int\int\left(1-\exp(-\lambda g(x,u,v))\right)\mathrm{d}x\otimes e^{-u}\mathrm{d}u\otimes\frac{1}{\sqrt{2\pi b}}e^{-\frac{v^2}{2b}}\mathrm{d}v\right),\,$$

⁷Actually, in the above-mentioned reference g is not assumed to be of class C^2 but merely continuous. However, a straightforward approximation procedure guarantees that one can restrict to C^2 functions.

if $\chi = (\xi, \Phi)$. In all cases, this equals $\mathbb{E}[\exp(-\lambda \mathcal{P}^{\chi}_{\infty}(g))]$, and thus (3.3) follows.

We are left with proving (3.4). What we will show is that (3.4) holds provided the following decorrelation estimate does

$$\mathbb{E}\left[\prod_{j=1}^{n_L}\prod_{x_0\in Q_{R_L,z_{j,L}}}\exp(-\lambda g(\frac{z_{j,L}}{L},\Theta_L^{\chi}(x_0)))\right]$$

$$=\prod_{j=1}^{n_L}\mathbb{E}\left[\prod_{x_0\in Q_{R_L,z_{j,L}}}\exp(-\lambda g(\frac{z_{j,L}}{L},\Theta_L^{\chi}(x_0)))\right] + o(1),$$
(3.5)

which in turn will be proved in Proposition 3.3.

To see the relation between (3.4) and (3.5), let us begin by considering $\chi = (\xi, \Phi)$. Let c > 0 be such that g vanishes on $[-1, 1]^d \times (-\infty, -c] \times [-\infty, \infty]$. On the complement of the event $\mathcal{A}_{z_{j,L},L}^{\xi}(c)$ in (2.21), there is at most one point $x_0 \in Q_{R_L, z_{j,L}}$ where $\xi_L(x_0) \ge a_L - \frac{c}{a_L}$ (thus $\Theta^{(\xi, \Phi)}(x_0) \in [-c, \infty] \times [-\infty, \infty]$), and necessarily, if such a x_0 exists then $x_0 = w_{j,L}$. Since g is a non-negative function supported on $[-1, 1]^d \times [-c, \infty] \times [-\infty, \infty]$, we deduce that

$$\left|\exp\left(-\lambda g\left(\frac{z_{j,L}}{L},\Theta_{L}^{(\xi,\Phi)}(w_{j,L})\right)-\prod_{x_{0}\in Q_{R_{L},z_{j,L}}}\exp\left(-\lambda g\left(\frac{z_{j,L}}{L},\Theta_{L}^{(\xi,\Phi)}(x_{0})\right)\right)\right|\leq \mathbb{1}_{\mathcal{A}_{L,z_{j,L}}^{\xi}(c)},$$

and

24

$$\mathbb{E}\Big[\exp(-\lambda g(\frac{z_{j,L}}{L},\Theta_L^{(\xi,\Phi)}(w_{j,L}))\Big] \\ -\mathbb{E}\Big[\prod_{x_0\in Q_{R_L,z_{j,L}}}\exp(-\lambda g(\frac{z_{j,L}}{L},\Theta_L^{(\xi,\Phi)}(x_0)))\Big]\Big| \le \mathbb{P}\big(\mathcal{A}_{L,z_{j,L}}^{\xi}(c)\big),$$

Using the identity $\prod_{j=1}^{n_L} a_j - \prod_{j=1}^{n_L} b_j = \sum_{k=1}^{n_L} a_1 \cdots a_{k-1} (a_k - b_k) b_{k+1} \cdots b_{n_L}$, and the fact that each factor is bounded by 1, we get

$$\left|\prod_{j=1}^{n_L} \mathbb{E}\left[\exp\left(-\lambda g\left(\frac{z_{j,L}}{L},\Theta_L^{(\xi,\Phi)}(w_{j,L})\right)\right)\right] - \prod_{j=1}^{n_L} \mathbb{E}\left[\prod_{x_0 \in Q_{R_L,z_{j,L}}} \exp\left(-\lambda g\left(\frac{z_{j,L}}{L},\Theta_L^{(\xi,\Phi)}(x_0)\right)\right)\right]\right| \le \sum_{j=1}^{n_L} \mathbb{P}(\mathcal{A}_{L,z_{j,L}}^{\xi}(c)) ,$$

and

$$\left| \mathbb{E} \left[\prod_{j=1}^{n_L} \exp\left(-\lambda g\left(\frac{z_{j,L}}{L}, \Theta_L^{(\xi,\Phi)}(w_{j,L})\right)\right) \right] - \mathbb{E} \left[\prod_{j=1}^{n_L} \prod_{x_0 \in Q_{R_L, z_{j,L}}} \exp\left(-\lambda g\left(\frac{z_{j,L}}{L}, \Theta_L^{(\xi,\Phi)}(x_0)\right)\right) \right] \right| \le \sum_{j=1}^{n_L} \mathbb{P}(\mathcal{A}_{L, z_{j,L}}^{\xi}(c)) .$$

By (2.25), the sums at the r.h.s. of the last two inequalities go to 0 as $L \to \infty$, and thus, the triangle inequality immediately implies that the proof of (3.4) can be reduced to that of (3.5).

For $\chi = \xi$ or Ξ , the argument is virtually identical, the only difference in the case $\chi = \Xi$ is that $\mathcal{A}_{L,z_{j,L}}^{\xi}(c)$ has to be replaced by

$$\mathcal{A}_{L,z_{j,L}}^{\Xi}(c) \cup \{ \max_{x \in Q_{R_L,z_{j,L}}} \Xi_L \ge a_L^{\Xi} - \frac{c}{a_L} \, ; \, w_{L,z_{j,L}} \neq w_{L,z_{j,L}}^{\Xi} \} \, .$$

To bound the sum over j of the probability of their union, it suffices to use (2.25) and (2.26), which once again implies that (3.4) holds provided (3.5) does.

3.2. **Proof of Theorem 1.3.** Thanks to the results in the previous section and in particular Proposition 3.1 for $\chi = \xi$, we can complete the proof of the first theorem stated in the introduction.

Proof of Theorem 1.3. Define the random measure

$$\mathcal{M}_L \stackrel{\text{def}}{=} \sum_{k=1}^{\#Q_L} \delta_{\left(\frac{y_{k,L}}{L}, \Theta_L^{\xi}(y_{k,L})\right)} ,$$

for $\Theta_L^{\xi}(\cdot) = a_L(\xi_L(\cdot) - a_L)$ as in (2.20) and $y_{k,L}$ the point on Q_L at which ξ_L reaches its k-th largest maximum. Theorem 1.3 states that \mathcal{M}_L converges in law towards a Poisson random measure \mathcal{M}_{∞} of intensity $dx \otimes e^{-u}du$, for the topology of vague convergence of Radon measures on $[-1,1]^d \times (-\infty,\infty]$. By [Kal02, Theorem 16.16], it suffices to prove that for any continuous function $g: [-1,1]^d \times (-\infty,\infty] \to \mathbb{R}_+$ with compact support, $\mathcal{M}_L(g) \to \mathcal{M}_{\infty}(g)$ in law as $L \to \infty$. Fix such a function g and let c > 0 be such that g vanishes outside $[-1,1]^d \times (-c,\infty]$. Set

$$\mathcal{B}_L(c) \stackrel{\text{def}}{=} \left\{ \exists x \in Q_L \backslash U_L \colon \Theta_L^{\xi}(x) \ge -c \right\} \,,$$

that is, $\mathcal{B}_L(c)$ is the event on which ξ_L is "large" in $Q_L \setminus U_L$. By definition of U_L , $|Q_L \setminus U_L| \leq n_L \sqrt{R_L} R_L^{d-1}$ and thus by (1.23) and (1.18), we get

$$\mathbb{P}(\mathcal{B}_L(c)) \le |Q_L \setminus U_L| \,\mathbb{P}(\Theta_L^{\xi}(0) \ge -c) \lesssim n_L \sqrt{R_L} R_L^{d-1} \frac{e^c}{L^d} \lesssim \frac{1}{\sqrt{R_L}} \,, \qquad (3.6)$$

and the r.h.s. vanishes as $L \to \infty$.

Recall the definition of $z_{j,L}$ and $\mathcal{A}_{L,\cdot}^{\xi}(c)$ in (1.18) and (2.21), respectively, and set $\mathcal{A}_{L}^{\xi}(c) = \bigcup_{j=1}^{n_{L}} \mathcal{A}_{L,z_{j,L}}^{\xi}(c)$. On the event $\mathcal{A}_{L}^{\xi}(c)^{\complement}$, for every $j \in \{1, \ldots, n_{L}\}$, the box $Q_{R_{L},z_{j,L}}$ contains at most one point where ξ_{L} lies above $a_{L} - c/a_{L}$, and if such a point exists it must be $w_{j,L}$.

This implies that on the event $\mathcal{A}_{L}^{\xi}(c)^{\complement} \cap \mathcal{B}_{L}(c)^{\complement}$ the set of points $\{y_{k,L}: 1 \leq k \leq \#Q_{L}, \xi_{L}(y_{k,L}) \geq a_{L} - c/a_{L}\}$ coincides with the set of points $\{w_{j,L}: 1 \leq j \leq n_{L}, \xi_{L}(w_{j,L}) \geq a_{L} - c/a_{L}\}$. Therefore, using the notation of (3.2), on $\mathcal{A}_{L}^{\xi}(c)^{\complement} \cap \mathcal{B}_{L}(c)^{\complement}$, we get

$$\mathcal{P}_{L}^{\xi}(g) - \mathcal{M}_{L}(g) = \sum_{j=1}^{n_{L}} \left(g\left(\frac{z_{j,L}}{L}, \Theta_{L}^{\xi}(w_{j,L})\right) - g\left(\frac{w_{j,L}}{L}, \Theta_{L}^{\xi}(w_{j,L})\right) \right)$$

Now, g is uniformly continuous in its first coordinate, so that, since $|z_{j,L} - w_{j,L}| \le \sqrt{d} R_L$, we deduce

$$|\mathcal{P}_L^{\xi}(g) - \mathcal{M}_L(g)| \lesssim \omega(\frac{\sqrt{d}R_L}{L})\mathcal{P}_L^{\xi}([-1,1]^d \times (-c,\infty]) .$$

25

where $\omega(\cdot)$ is the modulus of continuity of g in its first coordinate. The prefactor $\omega(\sqrt{d} R_L/L)$ vanishes as $L \to \infty$, while $\mathcal{P}_L^{\xi}([-1,1]^d \times (-c,\infty])$ converges in law to a finite limit thanks to Proposition 3.1. As the probability of $\mathcal{A}_L^{\xi}(c)^{\complement} \cap \mathcal{B}_L(c)^{\complement}$ converges to 1 by (2.25) and (3.6), we deduce that $|\mathcal{P}_L^{\xi}(g) - \mathcal{M}_L(g)|$ goes to 0 in (probability and thus in) law, and therefore, invoking once again Proposition 3.1, the statement follows at once.

3.3. **Decorrelation estimates.** In order to deal with the long-range correlations of our field, we now prove the decorrelation estimates which were exploited in the proof of Proposition 3.1.

Proposition 3.3. For $\chi = \xi, \Xi$, or (ξ, Φ) , let $g: I^{\chi} \to \mathbb{R}$ (where I^{χ} is defined after Proposition 3.1) be a compactly supported non-negative function of class \mathcal{C}^2 . If $\chi = (\xi, \Phi)$, further assume that $a_L \tau_L \sim \sqrt{b}$ for some $b \ge 0$. Then, as $L \to \infty$

$$\mathbb{E}\left[\prod_{j=1}^{n_L}\prod_{x_0\in Q_{R_L,z_{j,L}}}\exp\left(-\lambda g\left(\frac{z_{j,L}}{L},\Theta_L^{\chi}(x_0)\right)\right)\right] - \prod_{j=1}^{n_L}\mathbb{E}\left[\prod_{x_0\in Q_{R_L,z_{j,L}}}\exp\left(-\lambda g\left(\frac{z_{j,L}}{L},\Theta_L^{\chi}(x_0)\right)\right)\right] = o(1).$$
(3.7)

The proof is inspired by [LLR83, Theorem 4.2.1]. Let us highlight a few differences. First, in that reference the estimate is established for a function g which only depends on the second coordinate and which is the indicator of a semi-infinite interval. It turns out that dealing with regular functions g makes the proof somewhat easier. Second, and more importantly, we establish here a "long-range" decorrelation estimate: indeed, in the second term on the l.h.s. of (3.7) the product over all $x_0 \in Q_{R_L,z_{j,L}}$ remains inside the expectation (and so the decorrelation is proved for disjoint boxes) while in that reference, there is no such partial decorrelation. This is because in our setting, at small distances the r.v.'s at stake may have a complicated correlation structure that we do not try to disentangle.

Proof. We will present the proof in detail for $\chi = (\xi, \Phi)$, and, since that for $\chi = \xi, \Xi$ is similar (and actually simpler), we will limit ourselves to outline the main (minor) differences at the very end.

Let us introduce some notation. We rename the Gaussian vector of interest as

$$(\eta_L^{1,0}(x_0))_{x_0 \in U_L} \stackrel{\text{def}}{=} ((\eta_{1,L}^{1,0}(x_0), \eta_{2,L}^{1,0}(x_0))_{x_0 \in U_L} \stackrel{\text{def}}{=} (\xi_L(x_0), a_L \Phi_L(x_0))_{x_0 \in U_L})$$

(the reason for the double superscript will be clarified soon). Let $\Sigma^{1,0}$ be its covariance matrix and index its entries by (x_0, i) with $x_0 \in U_L$ and $i \in \{1, 2\}$. For instance, $\Sigma^{1,0}_{(x_0,2),(y_0,1)}$ is the covariance of $a_L \Phi_L(x_0)$ and $\xi_L(y_0)$.

As mentioned at the beginning of the section, the statement boils down to show that the error made by replacing $\eta_L^{1,0}$ with a Gaussian vector $\eta_L^{0,0}$ such that, for every $j = 1, \ldots, n_L$, $(\eta_L^{1,0}(x))_{x \in Q_{R_L, z_{j,L}}} \stackrel{\text{law}}{=} (\eta_L^{0,0}(x))_{x \in Q_{R_L, z_{j,L}}}$, and $(\eta_L^{0,0}(x))_{x \in Q_{R_L, z_{j_1,L}}}$ is independent of $(\eta_L^{0,0}(x))_{x \in Q_{R_L, z_{j_2,L}}}$ for every $j_1 \neq j_2$, is negligible as $L \to \infty$. Let $\Sigma^{0,0}$ be the covariance matrix of $\eta_L^{0,0}$ and notice that it is given by

$$\Sigma_{(x_0,i),(y_0,i')}^{0,0} \stackrel{\text{def}}{=} \begin{cases} \Sigma_{(x_0,i),(y_0,i')}^{1,0} & \text{if there exists } j \text{ s.t. } x_0, y_0 \in Q_{R_L,z_{j,L}}, \\ 0 & \text{else.} \end{cases}$$
(3.8)

For the reader's convenience, let us split the (quite involved) proof into four steps: *deceneracy*, *interpolation*, *density estimates* and *decay*, whose names will be justified along the way.

Step 1: Degeneracy. The problem with the Gaussian vectors $\eta_L^{1,0}$ and $\eta_L^{0,0}$ is that they may be degenerate and thus might not admit a density with respect to the Lebesgue measure. To overcome the issue, we will slightly perturb them: let $\varepsilon > 0$ and $(\gamma_1(x_0), \gamma_2(x_0))_{x_0 \in U_L}$ be an independent centred Gaussian vector of i.i.d. $\mathcal{N}(0, \varepsilon)$ r.v.'s and, for $\alpha = 0, 1$, define $\eta_L^{\alpha, \varepsilon}$ according to

$$(\eta_{1,L}^{\alpha,\varepsilon}(x_0),\eta_{2,L}^{\alpha,\varepsilon}(x_0)) \stackrel{\text{def}}{=} (\eta_{1,L}^{\alpha,0}(x_0) + \gamma_1(x_0),\eta_{2,L}^{\alpha,0}(x_0) + \gamma_2(x_0)), \qquad x_0 \in U_L.$$

Notice that $\eta_L^{\alpha,\varepsilon}$ is again a Gaussian vector but, this time, has a full rank covariance matrix, which we denote by $\Sigma^{\alpha,\varepsilon}$.

We now claim that provided we show

$$\limsup_{L \to \infty} \limsup_{\varepsilon \to 0} |I_L^{\varepsilon}| = 0 \tag{3.9}$$

27

where I_L^{ε} is defined according to

$$I_{L}^{\varepsilon} \stackrel{\text{def}}{=} \mathbb{E} \left[\prod_{j=1}^{n_{L}} \prod_{x_{0} \in Q_{R_{L}, z_{j,L}}} \exp\left(-\lambda g(\frac{z_{j,L}}{L}, a_{L}(\eta_{1,L}^{1,\varepsilon}(x_{0}) - a_{L}), \eta_{2,L}^{1,\varepsilon}(x_{0}))\right) \right]$$

$$- \prod_{j=1}^{n_{L}} \mathbb{E} \left[\prod_{x_{0} \in Q_{R_{L}, z_{j,L}}} \exp\left(-\lambda g(\frac{z_{j,L}}{L}, a_{L}(\eta_{1,L}^{0,\varepsilon}(x_{0}) - a_{L}), \eta_{2,L}^{0,\varepsilon}(x_{0}))\right) \right],$$
(3.10)

then (3.7) follows. Indeed, since I_L^0 coincides with the l.h.s. of (3.7) and the vectors $\eta_L^{1,\varepsilon}$ and $\eta_L^{0,\varepsilon}$ converge in law as $\varepsilon \downarrow 0$ respectively to $\eta_L^{1,0}$ and $\eta_L^{0,0}$, for any fixed L we have $\lim_{\varepsilon \downarrow 0} I_L^{\varepsilon} = I_L^0$, so that (3.9) implies the statement. We are left to prove (3.9) to which the next steps are devoted.

Step 2: Interpolation. We now introduce an interpolation between the covariance matrices $\Sigma^{1,\varepsilon}$ and $\Sigma^{0,\varepsilon}$, i.e. for $h \in [0,1]$ we define

$$\Sigma^{h,\varepsilon} \stackrel{\text{def}}{=} h \Sigma^{1,\varepsilon} + (1-h) \Sigma^{0,\varepsilon} .$$
(3.11)

This is still a positive definite matrix, and therefore it is the covariance matrix of a non-degenerate Gaussian vector $\eta_L^{h,\varepsilon}$. Let $f_h(s)$ for $s = (s_{(x_0,i)})_{(x_0,i)\in U_L\times\{1,2\}} \in \mathbb{R}^{2|U_L|}$ be the associated Gaussian density at s and set

$$A(s) \stackrel{\text{def}}{=} \prod_{j=1}^{n_L} \prod_{x_0 \in Q_{R_L, z_{j,L}}} \exp(-\lambda g(\frac{z_{j,L}}{L}, a_L(s_{(x_0,1)} - a_L), s_{(x_0,2)})) , \qquad (3.12)$$

and, for any $h \in [0, 1]$,

$$F(h) \stackrel{\text{def}}{=} \mathbb{E} \left[\prod_{j=1}^{n_L} \prod_{x_0 \in Q_{R_L, z_{j,L}}} \exp\left(-\lambda g\left(\frac{z_{j,L}}{L}, a_L(\eta_{1,L}^{h,\varepsilon}(x_0) - a_L), \eta_{2,L}^{h,\varepsilon}(x_0)\right)\right) \right]$$

GIUSEPPE CANNIZZARO, CYRIL LABBÉ, AND WILLEM VAN ZUIJLEN

$$=\int A(s)f_h(s)\mathrm{d}s$$

Notice in particular that this means

$$I_L^{\varepsilon} = F(1) - F(0) = \int_0^1 F'(h) dh = \int_0^1 \int A(s) \partial_h f_h(s) ds dh.$$
 (3.13)

Hence, to obtain (3.9) we need to bound the r.h.s. by a quantity independent of ε and that vanishes as $L \to \infty$.

Step 3: Density Estimates. Let \leq be an arbitrary total order on $\mathbb{R}^{|U_L|}$, and with a slight abuse of notation let us extend it into a total order on $\mathbb{R}^{|U_L|} \times \{1, 2\}$ by setting

$$(x_0, i) \le (y_0, i') \Leftrightarrow (x_0 < y_0) \text{ or } (x_0 = y_0, i \le i')$$
.

The dependence of f_h on h only goes through $\Sigma^{h,\varepsilon}$. Since this matrix is symmetric, we will only consider its entries "above the diagonal", that is $(\Sigma^{h,\varepsilon}_{(x_0,i),(y_0,i')}: (x_0,i) \leq (y_0,i'))$. By (3.11) and the definition of $\Sigma^{0,\varepsilon}$ in (3.8), the derivative of $\Sigma^{h,\varepsilon}$ in h reads

$$\begin{aligned} \partial_h \Sigma_{(x_0,i),(y_0,i')}^{h,\varepsilon} &= \Sigma_{(x_0,i),(y_0,i')}^{1,\varepsilon} - \Sigma_{(x_0,i),(y_0,i')}^{0,\varepsilon} \\ &= \begin{cases} 0 & \text{if there exists } j \text{ s.t. } x_0, y_0 \in Q_{R_L,z_{j,L}}, \\ \Sigma_{(x_0,i),(y_0,i')}^{1,\varepsilon} & \text{else.} \end{cases} \end{aligned}$$

On the other hand, using the identities

$$\frac{\partial \det \Sigma}{\partial \Sigma_{(x_0,i),(y_0,i')}} = 2(\det \Sigma) \Sigma_{(x_0,i),(y_0,i')}^{-1} , \quad \frac{\partial \Sigma^{-1}}{\partial \Sigma_{(x_0,i),(y_0,i')}} = -\Sigma^{-1} \frac{\partial \Sigma}{\partial \Sigma_{(x_0,i),(y_0,i')}} \Sigma^{-1} ,$$

where Σ^{-1} denotes the inverse of Σ and $\Sigma_{(x_0,i),(y_0,i')}^{-1}$ its $((x_0,i),(y_0,i'))$ -entry, a straightforward computation yields

$$\frac{\partial f_h}{\partial \Sigma^{h,\varepsilon}_{(x_0,i),(y_0,i')}} = \frac{\partial^2 f_h}{\partial s_{(x_0,i)} \partial s_{(y_0,i')}}$$

Therefore,

$$\begin{split} F'(h) &= \int A(s)\partial_h f_h(s) \mathrm{d}s = \sum_{(x_0,i) \le (y_0,i')} \int A(s)\partial_h \Sigma^{h,\varepsilon}_{(x_0,i),(y_0,i')} \frac{\partial f_h(s)}{\partial \Sigma^{h,\varepsilon}_{(x_0,i),(y_0,i')}} \mathrm{d}s \\ &= \sum_{(x_0,i) \le (y_0,i')} \Sigma^{1,\varepsilon}_{(x_0,i),(y_0,i')} \int A(s) \frac{\partial^2 f_h(s)}{\partial s_{(x_0,i)} \partial s_{(y_0,i')}} \mathrm{d}s \\ &= \sum_{(x_0,i) \le (y_0,i')} \Sigma^{1,\varepsilon}_{(x_0,i),(y_0,i')} \int \frac{\partial^2 A(s)}{\partial s_{(x_0,i)} \partial s_{(y_0,i')}} f_h(s) \mathrm{d}s \;, \end{split}$$

where the sum is only over x_0 and y_0 that do not fall within the same box.

We now need to estimate both $\Sigma^{1,\varepsilon}$ and A. For the former, since γ_1 , γ_2 and (ξ_L, Φ_L) are independent of each others, it is immediate to see that for any $(x_0, i), (y_0, i') \in U_L \times \{1, 2\}, \Sigma^{1,\varepsilon}_{(x_0,i),(y_0,i')} = \operatorname{Cov}(\xi_L(x_0), \xi_L(y_0))$ if i = i' = 1, $a_L \operatorname{Cov}(\xi_L(x_0), \Phi_L(y_0))$ if i = 1, i' = 2 and $a_L^2 \operatorname{Cov}(\Phi_L(x_0), \Phi_L(y_0))$ if i = i' = 2.

28

Therefore, provided x_0, y_0 do not belong to the same mesoscopic box (so that in particular $|x_0 - y_0| > \sqrt{R_L}$), (2.7) implies

$$|\Sigma_{(x_0,i),(y_0,i')}^{1,\varepsilon}| \lesssim a_L^{i+i'-2} \sup_{|x| \ge |x_0-y_0| - \sqrt{d} r_L} v_L(x) \,.$$

For A, by assumption $g \in C^2$ is compactly supported, so let c > 0 be such that the support of g is contained in $[-1, 1]^d \times [-c, \infty] \times [-\infty, \infty]$. Then, evaluating the second derivative of A in (3.12) gives

$$\left|\frac{\partial^2 A(s)}{\partial s_{(x_0,i)} \partial s_{(y_0,i')}}\right| \le C a_L^{4-i-i'} \mathbb{1}_{\{s_{(x_0,1)} \wedge s_{(y_0,1)} \ge a_L - \frac{c}{a_L}\}},$$

for some constant C > 0 independent of L (the variables in the indicator are both with i = 1!). Putting everything together, we obtain

$$\begin{aligned} |F'(h)| &\lesssim \sum_{(x_0,i) \le (y_0,i')} \sup_{|x| \ge |x_0 - y_0| - \sqrt{d} \, r_L} v_L(x) \, a_L^2 \int \mathbbm{1}_{\{s_{(x_0,1)} \land s_{(y_0,1)} \ge a_L - \frac{c}{a_L}\}} f_h(s) \mathrm{d}s \\ &\lesssim \sum_{x_0 \le y_0} \sup_{|x| \ge |x_0 - y_0| - \sqrt{d} \, r_L} v_L(x) \, a_L^2 \int \mathbbm{1}_{\{s_{x_0} \land s_{y_0} \ge a_L - \frac{c}{a_L}\}} \tilde{f}_h(s_{x_0}, s_{y_0}) \mathrm{d}s_{x_0} \mathrm{d}s_{y_0} \end{aligned}$$

where in the last step we used that the summand only depends on x_0, y_0 and not on i, i' so that, with a slight abuse, we suppressed the latter from the notation, and we denoted by \tilde{f}_h the marginal of f_h restricted to the two coordinates $s_{x_0} = s_{x_0,1}, s_{y_0} = s_{y_0,1}$. In other words, \tilde{f}_h is the density of the Gaussian pair $(\xi_L(x_0) + \gamma^1(x_0), \xi_L(y_0) + \gamma^2(y_0))$ and thus is given by

$$\tilde{f}_h(s_{x_0}, s_{y_0}) = \frac{\exp\left(-\frac{(1+\varepsilon)s_{x_0}^2 - 2v_L(x_0 - y_0)s_{x_0}s_{y_0} + (1+\varepsilon)s_{y_0}^2}{2((1+\varepsilon)^2 - v_L(x_0 - y_0)^2)}\right)}{2\pi\left((1+\varepsilon)^2 - v_L(x_0 - y_0)^2\right)^{1/2}} \le \frac{\exp\left(-\frac{s_{x_0}^2 + s_{y_0}^2}{2(1+\varepsilon+v_L(x_0 - y_0))}\right)}{2\pi\left((1+\varepsilon)^2 - v_L(x_0 - y_0)^2\right)^{1/2}}$$

as follows by applying $a^2 + b^2 \ge 2ab$. Therefore, using the above and the basic Gaussian estimate (1.20), we deduce

$$\begin{aligned} a_L^2 \int \mathbb{1}_{\{s_{x_0} \wedge s_{y_0} \ge a_L - \frac{c}{a_L}\}} \tilde{f}_h(s_{x_0}, s_{y_0}) \mathrm{d}s_{x_0} \mathrm{d}s_{y_0} \\ & \leq \frac{a_L^2}{\left((1+\varepsilon)^2 - v_L(x_0 - y_0)^2\right)^{1/2}} \left(\int_{a_L - \frac{c}{a_L}}^{\infty} \frac{\exp\left(-\frac{t^2}{2(1+\varepsilon+v_L(x_0 - y_0))}\right)}{\sqrt{2\pi}} \mathrm{d}t\right)^2 \\ & \lesssim \frac{1}{\left((1+\varepsilon)^2 - v_L(x_0 - y_0)^2\right)^{1/2}} \exp\left(-\frac{a_L^2}{1+\varepsilon+v_L(x_0 - y_0)}\right). \end{aligned}$$

Plugging all the previous estimates into (3.13), since the r.h.s. of the bounds obtained so far are independent of h, we finally obtain

$$\begin{split} \limsup_{\varepsilon \to 0} |I_L^{\varepsilon}| &\lesssim \sum_{x_0 \le y_0} \frac{\sup_{|x| \ge |x_0 - y_0| - \sqrt{d} r_L} v_L(x)}{\left(1 - v_L(x_0 - y_0)^2\right)^{1/2}} \exp\left(-\frac{a_L^2}{1 + v_L(x_0 - y_0)}\right) \\ &\lesssim L^d \sum_{z \in Q_L \setminus Q_{2\exp(\sqrt{\ln L})}} \frac{\sup_{|x| \ge |z| - \sqrt{d} r_L} v_L(x)}{\left(1 - v_L(z)^2\right)^{1/2}} \exp\left(-\frac{a_L^2}{1 + v_L(z)}\right) \end{split}$$
(3.14)

where the last step holds as x_0 and y_0 belong to distinct boxes, so that $|x_0 - y_0| \ge \sqrt{R_L} \ge 2 \exp(\sqrt{\ln L})$. The last step consists of proving that the sum vanishes, which in turn is a consequence of the decay of v_L .

Step 4: Decay of Correlations. We split the sum at the r.h.s. of (3.14) into two parts. First, we consider the sum over $z \in Q_{L^{1/4}} \setminus Q_{2\exp(\sqrt{\ln L})}$, on which for all L large enough, (1.1) ensures that $v_L(z) \leq 1/2$ and $\sup_{|x| \geq |z| - \sqrt{d}r_L} v_L(x) \leq 1/2$. Thus,

$$\begin{split} L^d & \sum_{z \in Q_{L^{1/4}} \setminus Q_{2\exp(\sqrt{\ln L})}} \frac{\sup_{|x| \ge |z| - \sqrt{d} r_L} v_L(x)}{\left(1 - v_L(z)^2\right)^{1/2}} \exp\left(-\frac{a_L^2}{1 + v_L(z)}\right) \\ & \lesssim L^d \left|Q_{L^{1/4}} \setminus Q_{2\exp(\sqrt{\ln L})}\right| \, \exp\left(-\frac{a_L^2}{1 + 1/2}\right) \lesssim L^{d(1 + \frac{1}{4})} \exp\left(-\frac{2}{3}a_L^2\right) \\ & = L^{\frac{5}{4}d} \left(e^{-\frac{a_L^2}{2}}\right)^{\frac{4}{3}} \lesssim L^{\frac{5}{4}d} \left(\frac{a_L}{L^d}\right)^{\frac{4}{3}} = L^{-\frac{d}{12}} a_L^{\frac{4}{3}} \,, \end{split}$$

where we further used (1.22), and the r.h.s. vanishes as $L \to \infty$.

On the other hand, to control the sum over $x \in Q_L \setminus Q_{L^{1/4}}$, set $\varepsilon_L \stackrel{\text{def}}{=} \sup_{|x| \ge L^{1/4} - \sqrt{d}r_L} v_L(x)$ and write

$$L^{d} \sum_{z \in Q_L \setminus Q_{L^{1/4}}} \frac{\sup_{|x| \ge |z| - \sqrt{d} r_L} v_L(x)}{\left(1 - v_L(z)^2\right)^{1/2}} \exp\left(-\frac{a_L^2}{1 + v_L(z)}\right)$$
$$\lesssim L^{2d} \varepsilon_L e^{-\frac{a_L^2}{1 + \varepsilon_L}} \lesssim L^{2d} \varepsilon_L e^{-a_L^2} e^{a_L^2 \frac{\varepsilon_L}{1 + \varepsilon_L}} \lesssim a_L^2 \varepsilon_L e^{a_L^2 \varepsilon_L} ,$$

which goes to 0 since $a_L^2 \varepsilon_L$ goes to 0 by (1.1).

The proof of the statement is thus complete for $\chi = (\xi, \Phi)$. For $\chi = \xi$ or Ξ , one can follow the exact same steps. Note that in the latter case, one has to replace the properties of v_L with those of v_L^{Ξ} in Lemma 2.2, in particular, that $v_L^{\Xi}(0) = 1 + \tau_L^2$ and (2.6).

4. The mesoscopic eigenproblem

While in the previous section we completely characterised the asymptotics of the maxima of the potential, we now turn to the analysis of the Anderson Hamiltonian associated to it and, as in Section 2, we begin by studying the eigenproblem *locally* on a mesoscopic box of side-length R_L , with R_L as in (1.15). More specifically, we aim at understanding the behaviour of the principal eigenvalue

31

 $\lambda_1(Q_{R_L},\xi_L)$ and eigenfunction φ_{R_L} of $\mathcal{H}_{Q_{R_L},\xi_L}$. As mentioned in the introduction, their behaviour is intimately related to that of the deterministic eigenproblem associated to the Hamiltonian

$$\bar{\mathcal{H}}_L \stackrel{\text{def}}{=} \mathcal{H}_{Q_{r_L}, -\mathcal{S}_L} = \Delta - \mathcal{S}_L, \quad \text{on } Q_{r_L},$$

for r_L as in (1.16) and S_L the shape defined in (1.6), whose principal eigenfunction and eigenvalue are denoted by $\bar{\varphi}_L$ and $\bar{\lambda}_L$.

To state the main theorem of this section we need to introduce a few quantities. Recall that we denote by w_L the point in Q_{R_L} where ξ_L attains its maximum, by Ξ_L and Φ_L the fields in (2.5) and (2.4) respectively, by τ_L^2 the variance of $\Phi_L(y)$ for any given $y \in \mathbb{Z}^d$ and $a_L^{\Xi} = a_L \sqrt{1 + \tau_L^2}$. At last, the event(s) whose probability we want to determine is

$$\Lambda_L(s) \stackrel{\text{def}}{=} \left\{ \lambda_1(Q_{R_L}, \xi_L) \ge a_L^{\Xi} + \bar{\lambda}_L + \frac{s}{a_L} \right\}, \qquad s \in \mathbb{R}.$$
(4.1)

Theorem 4.1. There exists a sequence of positive constants $(\eta_L)_{L\geq 1}$ which vanishes in the limit $L \to \infty$ such that the following statements hold for any given $s \in \mathbb{R}$

(1) (Tail distribution of the main eigenvalue)

$$\lim_{L \to \infty} \left(\frac{L}{R_L}\right)^d \mathbb{P}(\Lambda_L(s)) = e^{-s} .$$
(4.2)

(2) (Approximation of the main eigenvalue)

$$\lim_{L \to \infty} \left(\frac{L}{R_L}\right)^d \mathbb{P}\left(\Lambda_L(s); \left|\lambda_1(Q_{R_L}, \xi_L) - \left(\Xi_L(w_L) + \bar{\lambda}_L\right)\right| > \frac{\eta_L}{a_L}\right) = 0 , \quad (4.3)$$

and

$$\lim_{L \to \infty} \left(\frac{L}{R_L}\right)^d \mathbb{P}\left(\Xi_L(w_L) \ge a_L^{\Xi} + \frac{s}{a_L}; \left|\lambda_1(Q_{R_L}, \xi_L) - \left(\Xi_L(w_L) + \bar{\lambda}_L\right)\right| > \frac{\eta_L}{a_L}\right) = 0.$$

$$(4.4)$$

(3) (Magnitude of the maximum)

$$\limsup_{C \to \infty} \limsup_{L \to \infty} \left(\frac{L}{R_L}\right)^d \mathbb{P}\left(\Lambda_L(s); \xi_L(w_L) \notin I_L(C)\right) = 0 , \qquad (4.5)$$

where, for L, C > 0, $I_L(C)$ is the interval defined in (2.10).

(4) (Large spectral gap) there exists a C' > 0 independent of s such that

$$\lim_{L \to \infty} \left(\frac{L}{R_L}\right)^d \mathbb{P}\left(\Lambda_L(s); \lambda_2(Q_{R_L}, \xi_L) > a_L^{\Xi} + \bar{\lambda}_L - C' \frac{a_L}{d_L}\right) = 0.$$
(4.6)

(5) (Behaviour of the eigenfunction)

$$\lim_{L \to \infty} \left(\frac{L}{R_L}\right)^d \mathbb{P}\left(\Lambda_L(s); \|\varphi_{R_L} - \bar{\varphi}_L(\cdot - w_L)\|_{\ell^2(Q_{R_L})} > \frac{d_L}{a_L}\eta_L\right) = 0.$$
(4.7)

The rest of the section is devoted to the proof of this theorem. The crucial step in our analysis is the identification of the expansion for the eigenvalue in point (2) above. As it is one of the major technical novelties of our work, we dedicate to it the next section.

4.1. Approximating the eigenproblem. In Section 2.2, we have seen that whenever the potential is larger than $a_L - \theta$ at some point x_0 , it induces a local (deterministic) shape in a neighbourhood of x_0 and the fluctuations around such shape are encoded via ζ_{L,x_0} . In this section, we want to understand how this influences the behaviour of the main eigenvalue and eigenfunction of the Anderson Hamiltonian on Q_{R_L} . To this purpose, notice first that for any $x_0 \in Q_{R_L}$, it holds that

$$\lambda_1(Q_{R_L},\xi_L) = \xi_L(x_0) + \lambda_1(Q_{R_L},V_{L,x_0})$$
(4.8)

where we set $V_{L,x_0} \stackrel{\text{def}}{=} \xi_L - \xi_L(x_0)$. For $x \in Q_{R_L}$, by (2.1), V_{L,x_0} satisfies

$$V_{L,x_0}(x) = \xi_L(x_0)(v_L(x-x_0)-1) + \zeta_{L,x_0}(x) .$$
(4.9)

Our goal now is twofold. On the one hand we want to prove that, since on the event E_{L,x_0} , $\xi_L(x_0)$ is the unique maximum and is of order a_L , in V_{L,x_0} , we can replace the first summand by $-S_L$. On the other hand, we will show that the first non-trivial contribution of the fluctuation field ζ_{L,x_0} to the main eigenvalue on Q_{R_L} is given by the r.v. $\Phi_L(x_0)$. Let us state the theorem which rigorously details what we just explained.

Theorem 4.2. There exists a constant C > 0 and an integer $L_0 \ge 1$ such that for all $L \ge L_0$ and $x_0 \in Q_{R_L-r_L}$, on the event $E_{L,x_0} = E_{L,x_0}(\kappa)$ (for $\kappa \in (0, 1/3)$) from Definition 2.4, we have

$$\lambda_1(Q_{R_L},\xi_L) - \bar{\lambda}_L - \Xi_L(x_0) \Big| \le C \frac{d_L}{a_L} \Big(\Big(\frac{d_L}{a_L}\Big)^{1-2\kappa} |\xi_L(x_0) - a_L| + \frac{1}{a_L} \Big), \quad (4.10)$$

where we recall that $\Xi_L(x_0) = \xi_L(x_0) + \Phi_L(x_0)$ and the latter is as in (2.4), and

$$\|\varphi_{R_L} - \bar{\varphi}_L(\cdot - x_0)\|_{\ell^2(Q_{R_L})} \le C \frac{d_L}{a_L} \sqrt{\left(\frac{d_L}{a_L}\right)^{1-2\kappa} |\xi_L(x_0) - a_L| + \frac{1}{a_L}} , \quad (4.11)$$

where we extended $\bar{\varphi}_L$ by setting it to be zero outside Q_{r_L} .

In order to prove the above statement, we need to show that (1) we can localise the eigenproblem to a ball of size r_L centred around the maximum of ξ on Q_{R_L} , which, since we will be working on E_{L,x_0} , is at $x_0 \in Q_{R_L-r_L}$, and (2) the local eigenproblem on such ball is close to the deterministic one associated to the operator $\overline{\mathcal{H}}_L$. For these, two main ingredients are required, namely, suitable a-priori estimates on the decay of the main eigenfunctions, and a basic (but very useful) technical lemma on convex functionals, which in particular applies to the Dirichlet forms associated to $\mathcal{H}_{Q_{R_L},\xi_L}$ and $\overline{\mathcal{H}}_L$. For the former, we will use [BK16, Lemma 4.2] whose statement is recalled below in a slightly different formulation which better suits our purposes.

Lemma 4.3. [BK16, Lemma 4.2] Let $V: \mathbb{Z}^d \to \mathbb{R}$ and $D \subset \mathbb{Z}^d$. Let λ, φ be an eigenvalue and eigenfunction (normalised in $\ell^2(D)$) of $\mathcal{H}_{D,V}$ with Dirichlet boundary conditions. Assume $D' \subset D$, $A' \geq A > 0$ and $R \geq 1$ is an integer, such that

- (1) for all $x \in D'$, $V(x) \leq \lambda A'$,
- (2) for all $x \in D$ such that $\min_{y \in D'} |x y|_1 < R$, $V(x) < \lambda A$,

where $|x|_1 \stackrel{def}{=} \sum_{i=1}^d |x_i|$ denotes the ℓ^1 -norm. Then,

$$\sum_{x \in D'} |\varphi(x)|^2 \le \left(1 + \frac{A}{2d}\right)^{2-2R} \left(1 + \frac{A'}{2d}\right)^{-2}.$$
(4.12)

33

Let us see what type of information the previous lemma provides in our context.

Lemma 4.4. In the setting of Theorem 4.2, there exists a constant $c_d > 0$ such that for L large enough, on the event $E_{L,x_0}^1 \cap E_{L,x_0}^2$ (see (2.13) and (2.14)), we have

$$\varphi_{R_L}(x)^2 \le \left(1 + c_d \frac{a_L}{d_L}\right)^{-2|x-x_0|} \qquad \forall x \in Q_{R_L},$$
(4.13)

and

$$\varphi_{R_L}(x_0)^2 \ge 1 - \left(1 + c_d \frac{a_L}{d_L}\right)^{-2}.$$
 (4.14)

Furthermore, both (4.13) and (4.14) hold with φ_{R_L} and x_0 replaced by $\bar{\varphi}_L$ and 0 (the restriction to the event $E^1_{L,x_0} \cap E^2_{L,x_0}$ clearly being unnecessary in this case).

Proof. Without loss of generality, we take $x_0 = 0$ throughout this proof, and omit the corresponding subscript from the notation (so that e.g. $V_L = V_{L,0}$ and so on). Moreover, note that φ_{R_L} is also the main eigenfunction of $\mathcal{H}_{Q_{R_L},V_L}$, for V_L as in (4.9), associated to the eigenvalue $\lambda_1(Q_{R_L}, V_L) = \lambda_1(Q_{R_L}, \xi_L) - \xi_L(0)$.

If we establish the bound

$$\sum_{\substack{y \in Q_{R_L}: \\ y|_1 \ge |x|_1}} \varphi_{R_L}(y)^2 \le \left(1 + c_d \frac{a_L}{d_L}\right)^{-2|x|} \quad \forall x \in Q_{R_L}, \quad (4.15)$$

then both (4.13) and (4.14) follow (for the latter recall that φ_{R_L} and $\bar{\varphi}_L$ are normalised in $\ell^2(Q_{R_L})$). We thus prove (4.15). On the event $E_L^1 \cap E_L^2$, by Proposition 2.5 provided L is large enough

$$V_L(y) \le -\frac{\mathfrak{c}}{2} \frac{a_L}{d_L}, \qquad \text{for all } y \in Q_{R_L}^{\neq 0}.$$

$$(4.16)$$

By (1.2), the same holds for $-S_L$ with $\mathfrak{c}/2$ replaced by \mathfrak{c} . Since $V_L(0) = 0 =$ $\mathcal{S}_L(0)$, we deduce from the min-max formula that $\lambda_1(Q_{R_L}, V_L) \wedge \overline{\lambda}_L \geq -2d$. Hence, upon setting $R = |x|_1$, $D = Q_{R_L}$, $D' = \{y \in Q_{R_L} : |y|_1 \ge |x|_1\}$ and $A = A' = \mathfrak{c} d_L/(4a_L)$ (or $\mathfrak{c} d_L/(2a_L)$ if we deal with $\bar{\varphi}_L$), both hypothesis (1) and (2) in Lemma 4.3 hold provided L is large enough. Thus, (4.12) yields the bound in (4.15) but with the exponent at the r.h.s. given by $-2|x|_1$. However since $|y|_1 \geq |y|$ for any $y \in \mathbb{Z}^d$, the desired bound immediately follows and the proof is complete. \square

Before stating the next lemma detailing the second tool we need, let us briefly motivate it. Let \mathcal{H} be either of the operators $\mathcal{H}_{Q_{R_L},\xi_L} = -\Delta + \xi_L$ on Q_{R_L} or $\overline{\mathcal{H}}_L = -\Delta - \mathcal{S}_L$ on Q_{r_L} , λ and φ be its respective principal eigenvalue and eigenfunction (which we take normalised and non-negative to ensure uniqueness), r be either R_L or r_L . By the min-max theorem, we know that $\lambda = \max \mathcal{D}(\psi) =$ $\mathcal{D}(\varphi)$ where \mathcal{D} is either \mathcal{D}_{R_L} or $\overline{\mathcal{D}}_L$ and the latter are given by

$$\mathcal{D}_{R_L}(\psi) = \langle \psi, \mathcal{H}_{Q_{R_L}, \xi_L} \psi \rangle_{\ell^2(Q_{R_L})} , \quad \mathcal{D}_L(\psi) = \langle \psi, \mathcal{H}_L \psi \rangle_{\ell^2(Q_{r_L})} , \qquad (4.17)$$

34

the maximum carrying over all functions $\psi: Q_r \to \mathbb{R}$, normalised in $\ell^2(Q_r)$. Actually, we do not need to consider all such functions ψ , but only those that share the decay properties of φ as detailed in Lemma 4.4. Thus, we will view \mathcal{D} as a functional of $(r+1)^d - 1$ variables (as the value at 0 of the normalised, non-negative functions can be recovered from those elsewhere) defined on $\mathcal{Z}_r \subset \ell^2(Q_r^{\neq 0})$ whose elements ψ satisfy

$$|\psi(x)|^2 \le \left(c_d \frac{a_L}{d_L}\right)^{-2|x|}, \qquad \forall x \in Q_r^{\neq 0}.$$

$$(4.18)$$

 \mathcal{Z}_r is closed and convex. Note that, compared to (4.13), we imposed a slightly larger upper bound: this is to ensure that φ_{R_L} and $\bar{\varphi}_L$ lie in the interiors of \mathcal{Z}_{R_L} and \mathcal{Z}_{r_L} respectively.

The next lemma provides a general statement that suitably exploits convexity to derive estimates on the increments of functionals as above near their maximisers.

Lemma 4.5. Let $S \subset \mathbb{Z}^d$ be finite and $\mathcal{C} \subset \ell^2(S)$ be closed and convex. Assume that $G: \mathcal{C} \to \mathbb{R}$ is a strictly concave, twice continuously differentiable (on $\mathring{\mathcal{C}}$) functional for which there exists a constant H > 0, such that for all $z \in \mathring{\mathcal{C}}$, its Hessian Hess G at z satisfies

$$\langle y, \text{Hess}\,G(z)y\rangle_{\ell^2(S)} \le -H\|y\|_{\ell^2(S)}^2, \qquad \forall y \in \ell^2(S).$$
 (4.19)

Let x be the maximiser of G in C (which exists and is unique by (4.19)), and assume it lies in \mathring{C} . Then, for any $\bar{x} \in \mathring{C}$, we have

$$|G(x) - G(\bar{x})| \le \frac{1}{H} \|\nabla G(\bar{x})\|_{\ell^2(S)}^2, \qquad (4.20)$$

$$\|x - \bar{x}\|_{\ell^2(S)} \le \frac{1}{H} \|\nabla G(\bar{x})\|_{\ell^2(S)}.$$
(4.21)

Proof. Throughout the proof, the scalar product and the norm used are those on $\ell^2(S)$ thus, to lighten the notation, we omit the corresponding subscript, i.e. we write $\langle \cdot, \cdot \rangle$ and $\|\cdot\|$ in place of $\langle \cdot, \cdot \rangle_{\ell^2(S)}$ and $\|\cdot\|_{\ell^2(S)}$.

We first establish (4.21) and then use it to show (4.20). Let x be the maximiser of G in \mathcal{C} and $\bar{x} \neq x$ be another element of \mathcal{C} . Since x is a maximiser, $\nabla G(x) \equiv 0$ and therefore

$$\langle x - \bar{x}, \nabla G(\bar{x}) \rangle = -\langle x - \bar{x}, \nabla G(x) - \nabla G(\bar{x}) \rangle$$

= $-\int_0^1 \langle x - \bar{x}, \text{Hess } G(\bar{x} + t(x - \bar{x}))(x - \bar{x}) \rangle dt .$ (4.22)

By assumption \mathcal{C} is convex, so that $\bar{x} + t(x - \bar{x}) \in \mathcal{C}$ for any $t \in [0, 1]$, and we can use (4.19) to bound the r.h.s. of (4.22) from below by $H||x - \bar{x}||^2$. As a consequence, we deduce

$$||x - \bar{x}||^2 \le \frac{1}{H} \langle x - \bar{x}, \nabla G(\bar{x}) \rangle \le \frac{1}{H} ||x - \bar{x}|| ||\nabla G(\bar{x})||$$

from which (4.21) follows at once.

For (4.20), consider the map $f : [0,1] \ni t \mapsto G(\bar{x} + t(x - \bar{x}))$ which is concave and achieves its maximum at t = 1. Necessarily the maximum of its derivative is attained at t = 0 and therefore

$$|G(x) - G(\bar{x})| = |f(1) - f(0))| \le f'(0) = \langle x - \bar{x}, \nabla G(\bar{x}) \rangle \le ||x - \bar{x}|| ||\nabla G(\bar{x})||$$

and thus (4.20) follows by plugging (4.21) at the r.h.s.

With Lemmas 4.4 and 4.5 at our disposal, and anticipating some properties of \mathcal{D}_{R_L} and $\bar{\mathcal{D}}_L$ stated and shown in Appendix B, we are ready to prove Theorem 4.2.

Proof of Theorem 4.2. As soon as $x_0 \in Q_{R_L-r_L}$, the inclusion $Q_{r_L,x_0} \subset Q_{R_L}$ holds. As the arguments presented in this proof only rely on such inclusion, w.l.o.g., we can take $x_0 = 0$ and omit the corresponding index from the notation.

Let us first consider the l.h.s. of (4.10). Let $\mathcal{H}_{Q_{R_L},V_L}$ be the operator on Q_{R_L} given by $\Delta + V_L$ for $V_L = \xi_L - \xi_L(0)$. By (4.8) and (2.4), we have

$$\begin{split} \lambda_1(Q_{R_L},\xi_L) - \bar{\lambda}_L - \Xi_L(0) &= \lambda_1(Q_{R_L},\xi_L) - \xi_L(0) - \bar{\lambda}_L - \Phi_L(0) \\ &= \lambda_1(Q_{R_L},V_L) - \bar{\lambda}_L - \langle \bar{\varphi}_L, \zeta_L \bar{\varphi}_L \rangle_{\ell^2(Q_{r_L})} \\ &= \mathcal{D}_{R_L}(\varphi_{R_L}) - \bar{\mathcal{D}}_L(\bar{\varphi}_L) - \langle \bar{\varphi}_L, \zeta_L \bar{\varphi}_L \rangle_{\ell^2(Q_{r_L})}, \end{split}$$

where \mathcal{D}_{R_L} and $\overline{\mathcal{D}}_L$ are defined according to (4.17). Then, the r.h.s. coincides with the sum of two terms

(A)
$$\stackrel{\text{def}}{=} \mathcal{D}_{R_L}(\varphi_{R_L}) - \mathcal{D}_{R_L}(\bar{\varphi}_L),$$
 (4.23)

(B)
$$\stackrel{\text{def}}{=} \mathcal{D}_{R_L}(\bar{\varphi}_L) - \bar{\mathcal{D}}_L(\bar{\varphi}_L) - \langle \bar{\varphi}_L, \zeta_L \bar{\varphi}_L \rangle_{\ell^2(Q_{r_L})},$$
 (4.24)

which we will separately control.

Let us begin with (B). Since $\operatorname{supp}(\bar{\varphi}_L) \subset Q_{r_L}$, in the first summand the scalar product in the definition of \mathcal{D}_{R_L} in (4.17) can be taken in $\ell^2(Q_{r_L})$ instead of $\ell^2(Q_{R_L})$. Since all the scalar products appearing in this term are in $\ell^2(Q_{r_L})$, we lighten the presentation by omitting the corresponding subscript from the notation. Then, (4.9) and the definition of \mathcal{S}_L in (1.6) give

$$(B) = \langle \bar{\varphi}_L, \mathcal{H}_{Q_{R_L}, V_L} \bar{\varphi}_L \rangle - \langle \bar{\varphi}_L, \bar{\mathcal{H}}_L \bar{\varphi}_L \rangle - \langle \bar{\varphi}_L, \zeta_L \bar{\varphi}_L \rangle = (\xi_L(0) - a_L) \langle \bar{\varphi}_L, [v_L(\cdot) - 1] \bar{\varphi}_L \rangle .$$

On E_{L,x_0}^1 , $|\xi_L(0) - a_L| \le \theta$ which, together with (1.2), implies

$$|(\mathbf{B})| = |\xi_L(0) - a_L| \left| \sum_{x \in Q_{r_L}^{\neq 0}} [v_L(x) - 1] \bar{\varphi}_L(x)^2 \right| \le \frac{\theta}{d_L} \sum_{x \in Q_{r_L}^{\neq 0}} e^{\mathfrak{c}' |x|} \bar{\varphi}_L(x)^2 ,$$

the exclusion of 0 in the first sum is a consequence of $v_L(0) = 1$. Using the bound (4.13) on the decay of $\bar{\varphi}_L$, we easily deduce that

$$|(\mathbf{B})| \le \frac{\theta}{d_L} \sum_{x \in Q_{r_L}^{\neq 0}} e^{\mathfrak{c}'|x|} \left(1 + c_d \frac{a_L}{d_L}\right)^{-2|x|} \le \frac{C}{a_L} \left(\frac{d_L}{a_L}\right) \tag{4.25}$$

for some constant C > 0 independent of L.

We now turn to (A), for which we apply Lemma 4.5 with $S = Q_{R_L}^{\neq 0}$. More specifically, by Lemma B.1, the Hessian of \mathcal{D}_{R_L} satisfies (4.19) with $H = c_0 a_L/d_L$ and the $\ell^2(Q_{R_L}^{\neq 0})$ -norm of its gradient can be bounded by (B.3). As a consequence, (4.20) gives

$$|(\mathbf{A})| \lesssim \frac{d_L}{a_L} \left(\frac{1}{a_L} + \|\bar{\varphi}_L \zeta_L\|_{\ell^2(Q_{r_L})}^2\right)$$

which, together with (4.25), implies for some constant C > 0 and for all L large enough

$$\left|\lambda_1(Q_{R_L},\xi_L) - \bar{\lambda}_L - \Xi_L(0)\right| \le C \frac{d_L}{a_L} \Big(\frac{1}{a_L} + \|\bar{\varphi}_L\zeta_L\|_{\ell^2(Q_{r_L})}^2\Big).$$
(4.26)

Before completing the proof, let us consider the l.h.s. of (4.11), for which we argue as for (A) above invoking (4.21) instead of (4.20). Thus, we deduce

$$\|\varphi_{R_L} - \bar{\varphi}_L\|_{\ell^2(Q_{R_L}^{\neq 0})} \le C \frac{d_L}{a_L} \sqrt{\frac{1}{a_L}} + \|\bar{\varphi}_L \zeta_L\|_{\ell^2(Q_{r_L})}^2.$$

To control the difference of φ_{R_L} and $\bar{\varphi}_L$ at 0, we use the fact that $x \mapsto \sqrt{1-x^2}$ is Lipschitz on (-1/2, 1/2) to get

$$\begin{aligned} |\varphi_{R_{L}}(0) - \bar{\varphi}_{L}(0)| &= \left| \sqrt{1 - \|\varphi_{R_{L}}\|_{\ell^{2}(Q_{R_{L}}^{\neq 0})}^{2}} - \sqrt{1 - \|\bar{\varphi}_{L}\|_{\ell^{2}(Q_{R_{L}}^{\neq 0})}^{2}} \right| \\ &\lesssim \left| \|\varphi_{R_{L}}\|_{\ell^{2}(Q_{R_{L}}^{\neq 0})}^{2} - \|\bar{\varphi}_{L}\|_{\ell^{2}(Q_{R_{L}}^{\neq 0})}^{2} \right| \lesssim \|\varphi_{R_{L}} - \bar{\varphi}_{L}\|_{\ell^{2}(Q_{R_{L}}^{\neq 0})}^{2}, \end{aligned}$$
(4.27)

which ultimately gives (possibly for a different constant C > 0)

$$\|\varphi_{R_L} - \bar{\varphi}_L\|_{\ell^2(Q_{R_L})} \le C \frac{d_L}{a_L} \sqrt{\frac{1}{a_L} + \|\bar{\varphi}_L \zeta_L\|_{\ell^2(Q_{r_L})}^2} .$$
(4.28)

Thanks to (4.26) and (4.28), (4.10) and (4.11) follow provided we suitably estimate the $\ell^2(Q_{r_L})$ -norm of $\bar{\varphi}_L \zeta_L$. Notice that so far, we never used the bound provided by E_L^3 and this is the point at which it becomes essential. Indeed, on E_L^3 for $x \in Q_{r_L}$ the fluctuation field ζ_L is bounded above by

$$\begin{aligned} |\zeta_L(x)| &\leq \sqrt{\operatorname{Var}[\zeta_L(x)]} \left(\frac{a_L}{d_L}\right)^{\kappa|x|} \sqrt{1 \vee (|\xi(0) - a_L|a_L)} \\ &\lesssim \frac{e^{\frac{c''|x|}{2}}}{\sqrt{d_L}} \left(\frac{a_L}{d_L}\right)^{\kappa|x|} \sqrt{1 + |\xi(0) - a_L|a_L} \end{aligned}$$

where in the last step we also used (2.3). Note that $\zeta_L(0) = 0$. Using the exponential decay of $\bar{\varphi}_L$ stated in (4.13), we thus deduce

$$\begin{split} \|\bar{\varphi}_L\zeta_L\|_{\ell^2(Q_{r_L})}^2 &\lesssim \sum_{x \in Q_{r_L}^{\neq 0}} \bar{\varphi}_L(x)^2 \frac{e^{c''|x|}}{d_L} \left(\frac{a_L}{d_L}\right)^{2\kappa|x|} (1+|\xi(0)-a_L|a_L) \\ &\lesssim \frac{1+|\xi(0)-a_L|a_L}{d_L} \sum_{x \in Q_{r_L}^{\neq 0}} e^{c''|x|} \left(c_d \frac{d_L}{a_L}\right)^{(2-2\kappa)|x|} \\ &\lesssim \frac{1+|\xi(0)-a_L|a_L}{d_L} \left(\frac{d_L}{a_L}\right)^{2-2\kappa} = \left(\frac{d_L}{a_L}\right)^{1-2\kappa} \left[\frac{1}{a_L}+|\xi_L(0)-a_L|\right]. \end{split}$$

Plugging this estimate into (4.26) and (4.28), the statement follows at once.

Before concluding this section we state and prove the next proposition which, together with Theorem 4.2, will be shown to imply a (diverging) spectral gap for the operator $\mathcal{H}_{Q_{R_L},\xi_L}$ on the event E_{L,x_0} in Definition 2.4.

Proposition 4.6. There exist a constant $C_{gap} > 0$ and an integer $L_0 > 1$ such that for all $L \ge L_0$ and all $x_0 \in Q_{R_L-r_L}$, on the event $E_{L,x_0}^1 \cap E_{L,x_0}^2$ as in Definition 2.4, we have

$$\lambda_2(Q_{R_L},\xi_L) \le \xi_L(x_0) - C_{\text{gap}} \frac{a_L}{d_L}.$$
 (4.29)

Proof. By the min-max formula, the second eigenvalue of $\mathcal{H}_{Q_{R_I},\xi_L}$ satisfies

$$\lambda_2(Q_{R_L},\xi_L) = \sup\{\mathcal{D}_{R_L}(\psi) \colon \|\psi\|_{\ell^2(Q_{R_L})} = 1, \, \langle\psi,\varphi_{R_L}\rangle = 0\}.$$
(4.30)

Notice first that, for any $\psi \in \ell^2(Q_{R_L})$ normalised to 1 such that $0 = \langle \psi, \varphi_{R_L} \rangle = \sum_{x \in Q_{R_L}} \psi(x) \varphi_{R_L}(x)$, we have

$$\psi(x_0) = -\frac{1}{\varphi_{R_L}(x_0)} \sum_{x \in Q_{R_L}^{\neq x_0}} \psi(x)\varphi_{R_L}(x)$$

the expression above being meaningful as (4.14) implies that, for L large enough on the event E_{L,x_0} , $\varphi_{R_L}(x_0)^2 \ge 1 - C(d_L/a_L)^2 \ge 1/4$. As a consequence, Cauchy-Schwarz gives

$$|\psi(x_0)| \le 2\|\varphi_{R_L}\|_{\ell^2(Q_{R_L}^{\neq x_0})} \|\psi\|_{\ell^2(Q_{R_L}^{\neq x_0})} \le 2\sqrt{1 - \varphi_{R_L}(x_0)^2} \le 2\sqrt{C} \frac{d_L}{a_L} \le \frac{1}{2}.$$
(4.31)

Now, let ψ be as above and consider the quadratic form at ψ , which is

$$\mathcal{D}_{R_L}(\psi) \le \langle \psi, \xi_L \psi \rangle \le \xi_L(x_0) \psi(x_0)^2 + \Big(\max_{x \in Q_{R_L}^{\neq x_0}} \xi_L(x)\Big) \|\psi\|_{\ell^2(Q_{R_L}^{\neq x_0})}^2$$

= $\xi_L(x_0) \psi(x_0)^2 + \Big(\max_{x \in Q_{R_L}^{\neq x_0}} \xi_L(x)\Big) (1 - \psi(x_0)^2).$ (4.32)

By Proposition 2.5, on the event E_{L,x_0} we have

$$\max_{x \in Q_{R_L}^{\neq x_0}} \xi_L(x) \le \xi_L(x_0) - \frac{\mathfrak{c}}{2} \frac{a_L}{d_L} \,.$$

Using (4.31) we find

$$\mathcal{D}_{R_L}(\psi) \le \xi_L(x_0) - \frac{\mathfrak{c}}{2} \frac{a_L}{d_L} (1 - \psi(x_0)^2) \le \xi_L(x_0) - \frac{3\mathfrak{c}}{8} \frac{a_L}{d_L} ,$$

thus concluding the proof.

4.2. **Proof of Theorem 4.1.** For Theorem 4.2 to provide a useful description of the principal eigenvalue of $\mathcal{H}_{Q_{R_L},\xi_L}$, we need to ensure that the random variable at the r.h.s. of (4.10), i.e.

$$\frac{d_L}{a_L} \left(\left(\frac{d_L}{a_L} \right)^{1-2\kappa} |\xi_L(x_0) - a_L| + \frac{1}{a_L} \right) , \qquad (4.33)$$

is negligible compared to the putative fluctuation scale of $\lambda_1(Q_{R_L}, \xi_L)$ itself, that is, a_L^{-1} . For this, in the next proposition, we show that we can restrict ourselves to the event $\{|\xi_L(x_0) - a_L| < \theta_L\}$ for some sequence $(\theta_L)_{L\geq 1}$ satisfying

$$\left(\frac{d_L}{a_L}\right)^{2-2\kappa} \theta_L \ll \frac{1}{a_L} \,. \tag{4.34}$$

 \square

38

As we will apply Lemma 2.3, in particular (2.11), we are not allowed to take θ_L arbitrarily small, but we need it to satisfy $\theta_L \gg \max\{a_L^{-1}, a_L \tau_L^2\}$. For concreteness, let us make a specific choice and from here on set

$$\theta_L \stackrel{\text{def}}{=} \left(\frac{a_L}{d_L}\right)^{\kappa} \max\{a_L^{-1}, a_L \tau_L^2\}, \qquad (4.35)$$

where κ is the small parameter appearing in the definition of the event E_{L,x_0} . As $\kappa < 1/3$ and recalling Assumption 1.5, (4.34) can be immediately checked to be satisfied.

Let us now state the above-mentioned proposition, whose proof is postponed to the end of the section.

Proposition 4.7. Let θ_L be defined according to (4.35). For any event G

$$\mathbb{P}\Big(G \cap \Lambda_L(s)\Big) \tag{4.36}$$

$$= \sum_{x_0 \in Q_{R_L - r_L}} \mathbb{P}\Big(G \cap \Lambda_L(s) \cap E_{L,x_0} \cap \{|\xi_L(x_0) - a_L| < \theta_L\}\Big) + o\Big(\frac{R_L^d}{L^d}\Big).$$

Before turning to the proof of Theorem 4.1, let us appreciate the advantage of the previous statement. What it guarantees is that, when studying the asymptotic behaviour of the probability of $\Lambda_L(s) \cap G$, for $s \in \mathbb{R}$ and some event G, it suffices to analyse that of $\Lambda_L(s) \cap G \cap E_{L,x_0} \cap \{|\xi_L(x_0) - a_L| < \theta_L\}$ for $x_0 \in Q_{R_L-r_L}$. On $E_{L,x_0} \cap \{|\xi_L(x_0) - a_L| < \theta_L\}$, Theorem 4.2 says that $\lambda_1(Q_{R_L}, \xi_L)$ satisfies

$$\lambda_1(Q_{R_L},\xi_L) \approx \bar{\lambda}_L + \Xi_L(x_0) = \bar{\lambda}_L + \xi_L(x_0) + \Phi_L(x_0)$$

up to an error strictly smaller than the size of the fluctuations a_L^{-1} (see (4.33) and (4.34)). Among the terms at the r.h.s., $\bar{\lambda}_L$ is deterministic while, for fixed $x_0, \xi_L(x_0)$ and $\Phi_L(x_0)$ are independent Gaussian random variables of variances 1 and τ_L^2 respectively (see Lemma 2.1 and (2.4)).

In other words, we managed to reduce the analysis of the fluctuations of the complicated object $\lambda_1(Q_{R_L}, \xi_L)$ to that of the sum of two independent Gaussian random variables which in turn was studied in detail in Lemma 2.3.

Proof of Theorem 4.1. We take θ_L as in (4.35). Let us begin by identifying a suitable sequence $(\eta_L)_{L\geq 1}$. For any $x_0 \in Q_{R_L-r_L}$, on the event E_{L,x_0} , Theorem 4.2 gives

$$\left|\lambda_1(Q_{R_L},\xi_L) - \bar{\lambda}_L - \Xi_L(x_0)\right| \le \frac{C}{a_L} \left(\left(\frac{d_L}{a_L}\right)^{2-2\kappa} \theta_L a_L + \frac{d_L}{a_L} \right) =: \frac{\tilde{\eta}_L}{a_L}, \quad (4.37)$$

$$\|\varphi_{R_L} - \bar{\varphi}_L(\cdot - x_0)\|_{\ell^2(Q_{R_L})} \le C \frac{d_L}{a_L} \sqrt{\left(\frac{d_L}{a_L}\right)^{1-2\kappa}} \theta_L + \frac{1}{a_L} =: \frac{d_L}{a_L} \tilde{\eta}'_L, (4.38)$$

and set $\eta_L \stackrel{\text{def}}{=} \tilde{\eta}_L \vee \tilde{\eta}'_L$. By (4.34), we see that η_L goes to 0 as $L \to \infty$. We now turn to the proof of each of the five points in which the statement is divided.

Point (1). Proposition 4.7 with G being the whole probability space, implies that the statement follows if we show that uniformly over all $x_0 \in Q_{R_L-r_L}$, we have

$$P_{L,x_0} \stackrel{\text{def}}{=} \mathbb{P}\Big(\Lambda_L(s) \cap E_{L,x_0} \cap \{|\xi_L(x_0) - a_L| < \theta_L\}\Big) \sim \frac{e^{-s}}{L^d} \,. \tag{4.39}$$

Thanks to (4.37), we immediately get

$$P_{L,x_{0}} \leq \mathbb{P}\left(\Xi_{L}(x_{0}) \geq a_{L}^{\Xi} + \frac{s - \eta_{L}}{a_{L}}\right),$$

$$P_{L,x_{0}} \geq \mathbb{P}\left(\left\{\Xi_{L}(x_{0}) \geq a_{L}^{\Xi} + \frac{s + \eta_{L}}{a_{L}}\right\} \cap E_{L,x_{0}} \cap \{|\xi_{L}(x_{0}) - a_{L}| \leq \theta_{L}\}\right),$$
(4.40)

and recall that $\Xi_L(x_0) = \xi_L(x_0) + \Phi_L(x_0)$, that is $\Xi_L(x_0)$ is the sum of independent mean-zero Gaussian random variables of variance 1 and τ_L^2 , respectively. Now, for the upper bound, we apply Lemma 2.3 and in particular (2.8). For the lower bound, we first remove the event E_{L,x_0} at a price negligible with respect to L^{-d} , which is allowed since $\theta_L \ll \theta$ and

$$\mathbb{P}(|\xi_L(x_0) - a_L| \le \theta_L; E_{L,x_0}^{\complement}) \le \mathbb{P}(\xi_L(x_0) \ge a_L - \theta; E_{L,x_0}^{\complement}) = o\left(\frac{1}{L^d}\right),$$

as implied by (2.17). Thus, additionally using (2.11) we deduce

$$P_{L,x_0} \ge \mathbb{P}\Big(\Xi_L(x_0) \ge a_L^{\Xi} + \frac{s + \eta_L}{a_L}; |\xi_L(x_0) - a_L| \le \theta_L\Big) + o\Big(\frac{1}{L^d}\Big)$$
$$= \mathbb{P}\Big(\Xi_L(x_0) \ge a_L^{\Xi} + \frac{s + \eta_L}{a_L}\Big)$$
$$- \mathbb{P}\Big(\Xi_L(x_0) \ge a_L^{\Xi} + \frac{s + \eta_L}{a_L}; |\xi_L(x_0) - a_L| > \theta_L\Big) + o\Big(\frac{1}{L^d}\Big)$$
$$= \mathbb{P}\Big(\Xi_L(x_0) \ge a_L^{\Xi} + \frac{s + \eta_L}{a_L}\Big) + o\Big(\frac{1}{L^d}\Big)$$

and, to the latter, we apply once again (2.8). Putting upper and lower bounds together, (4.39) follows.

Point (2). Take G in Proposition 4.7 to be the event $\{|\lambda_1(Q_{R_L},\xi_L) - \bar{\lambda}_L - \Xi_L(w_L)| > \eta_L/a_L\}$. Note that, for every $x_0 \in Q_{R_L-r_L}$, on the event E_{L,x_0} the maximum of ξ_L is achieved at x_0 so that $w_L = x_0$ by Proposition 2.5. Thus, (4.37) implies that each of the summands at the r.h.s. of (4.36) is 0 and (4.3) follows at once. Regarding (4.4), since

$$\left\{\Xi_L(w_L) > a_L^{\Xi} + \frac{s}{a_L}\right\} = \left\{\Theta_L^{\Xi} \ge s\right\},\,$$

the same argument as in (2.31) ensures that for any event H

$$\mathbb{P}\Big(\Xi_L(w_L) > a_L^{\Xi} + \frac{s}{a_L}; H\Big) = \sum_{x_0 \in Q_{R_L}} \mathbb{P}\Big(\Xi_L(x_0) > a_L^{\Xi} + \frac{s}{a_L}; H; E_{L,x_0}\Big) + o\Big(\frac{R_L^a}{L^d}\Big) \,.$$

Now taking $H \stackrel{\text{def}}{=} \{ |\lambda_1(Q_{R_L}, \xi_L) - (\Xi_L(w_L) + \bar{\lambda}_L)| > \eta_L/a_L \}$, and recalling that on E_{L,x_0} we have $w_L = x_0$, Theorem 4.2 ensures that each term in the sum over x_0 vanishes, thus completing the proof of (4.4).

Point (3). We choose the event G in (4.36) to be $G = \{\xi_L(w_L) \notin I_L(C)\}$. As before, it suffices to control the probability of the event $G \cap \Lambda_L(s) \cap E_{L,x_0} \cap \{|\xi_L(x_0) - a_L| < \theta_L\}$ uniformly in x_0 . Arguing as in (4.40), we get

$$\mathbb{P}\Big(G \cap \Lambda_L(s) \cap E_{L,x_0} \cap \{|\xi_L(x_0) - a_L| < \theta_L\}\Big)$$
$$\leq \mathbb{P}\Big(\Xi_L(x_0) \ge a_L^{\Xi} + \frac{s - \eta_L}{a_L}; \xi_L(x_0) \notin I_L(C)\Big)$$

39

_ _ _

and the quantity at the r.h.s. is independent of x_0 . The limit in (2.9) implies that for any given $\varepsilon > 0$, provided C is large enough, its $\lim_{L\to\infty}$ passes below ϵ/L^d and one can conclude.

Point (4). We choose G in (4.36) to be

40

$$G \stackrel{\text{def}}{=} \left\{ \lambda_2(Q_{R_L}, \xi_L) > a_L^{\Xi} + \bar{\lambda}_L - C' \frac{a_L}{d_L} \right\},\,$$

and, as argued before we only need to show that uniformly over all $x_0 \in Q_{R_L-r_L}$

$$\mathbb{P}\Big(G \cap \Lambda_L(s) \cap E_{L,x_0}\Big) = o\Big(\frac{1}{L^d}\Big).$$
(4.41)

Actually, an even stronger statement is true, namely, there exists a constant C' > 0 such that for L sufficiently large, the probability at the l.h.s. of (4.41) is simply equal to 0. Indeed, using Proposition 4.6, the definition of E_{L,x_0}^1 , the fact that $\bar{\lambda}_L \geq -2d$ (by the minmax formula) and that $\theta = 2d+1$ as in Definition 2.4, we know that on $E_{L,x_0} \cap \Lambda_L(s)$,

$$\begin{split} \lambda_2(Q_{R_L},\xi_L) &\leq \xi_L(x_0) - C_{\text{gap}} \frac{a_L}{d_L} \leq a_L + \theta - C_{\text{gap}} \frac{a_L}{d_L} \\ &\leq a_L \sqrt{1 + \tau_L^2} + \bar{\lambda}_L + 2d + \theta - C_{\text{gap}} \frac{a_L}{d_L} \leq a_L^{\Xi} + \bar{\lambda}_L + 2\theta - C_{\text{gap}} \frac{a_L}{d_L} \\ &< a_L^{\Xi} + \bar{\lambda}_L - C' \frac{a_L}{d_L} \,, \end{split}$$

provided $C' < C_{\text{gap}}$ and L is large enough. Hence G cannot hold on $E_{L,x_0} \cap \Lambda_L(s)$, and thus the l.h.s. of (4.41) is 0.

Point (5). Let *G* be the event $\{\|\varphi_{R_L} - \bar{\varphi}_L(\cdot - x_0)\|_{\ell^2(Q_{R_L})} > \eta_L d_L/a_L\}$, where η_L was defined right below (4.38). But then, by (4.38), for any x_0 , $\Lambda_L(s) \cap G \cap E_{L,x_0} = \emptyset$, so that each summand in the sum at the r.h.s. of (4.36) is 0 and (4.7) follows at once. Therefore, the proof of point (5) and of the statement are complete.

We now turn to the proof of Proposition 4.7. It is performed in two steps, summarised by the following two lemmas.

Lemma 4.8. For any event G, as $L \to \infty$, we have

$$\mathbb{P}\Big(G \cap \Lambda_L(s)\Big) = \sum_{x_0 \in Q_{R_L - r_L}} \mathbb{P}\Big(G \cap \Lambda_L(s) \cap E_{L,x_0}\Big) + o\Big(\frac{R_L^d}{L^d}\Big).$$
(4.42)

Proof. As an initial step, we want to localise $\Lambda_L(s)$ to an event in which the maximum of ξ_L over Q_{R_L} is of order a_L . To do so, we begin with two remarks. First, by the minmax formula, we have that

$$\lambda_1(Q_{R_L},\xi_L) \le \max_{x \in Q_{R_L}} \xi_L(x) \; .$$

Second, since $\lambda_L \geq -2d$, we immediately deduce that for L large enough

$$a_{L} - \theta = a_{L} - 2d - 1 < a_{L}\sqrt{1 + \tau_{L}^{2}} + \bar{\lambda}_{L} - 1 \le a_{L}^{\Xi} + \frac{s}{a_{L}} + \bar{\lambda}_{L}$$

with $\theta = 2d + 1$ as in Definition 2.4. Consequently, if $\max \xi_L < a_L - \theta$ then $\lambda_1(Q_{R_L}, \xi_L) < a_L - \theta \le a_L^{\Xi} + \frac{s}{a_L} + \overline{\lambda}_L$, which means that $\Lambda_L(s) \subset \{\max \xi_L \ge a_L - \theta\}$. Hence, by (2.18), we deduce that

$$\mathbb{P}\Big(G \cap \Lambda_L(s) \cap \Big(\bigcup_{x_0 \in Q_{R_L}} E_{L,x_0}\Big)^{\mathsf{L}}\Big) \le \mathbb{P}\Big(\{\max_{Q_{R_L}} \xi_L \ge a_L - \theta\} \cap \Big(\bigcup_{x_0 \in Q_{R_L}} E_{L,x_0}\Big)^{\mathsf{L}}\Big)$$

is negligible compared to R_L^d/L^d , which implies

$$\mathbb{P}(G \cap \Lambda_L(s)) = \sum_{x_0 \in Q_{R_L}} \mathbb{P}(G \cap \Lambda_L(s) \cap E_{L,x_0}) + o\left(\frac{R_L^d}{L^d}\right).$$

As a consequence, we are left to neglect the sum over $Q_{R_L} \setminus Q_{R_L-r_L}$, which in turn can be controlled by (1.23) as

$$\sum_{x_0 \in Q_{R_L} \setminus Q_{R_L - r_L}} \mathbb{P}(G \cap \Lambda_L(s) \cap E_{L, x_0}) \le \sum_{x_0 \in Q_{R_L} \setminus Q_{R_L - r_L}} \mathbb{P}(\xi_L(x_0) \ge a_L - \theta)$$
$$\lesssim r_L R_L^{d-1} \frac{1}{L^d} e^{\theta a_L}$$

which is also negligible compared to R_L^d/L^d since $\ln r_L \ll a_L \ll \ln R_L$ by (1.15) and (1.16).

Lemma 4.9. Let θ_L be defined according to (4.35). Then, for any event G and any $x_0 \in Q_{R_L-r_L}$, as $L \to \infty$, we have

$$\mathbb{P}\Big(G \cap \Lambda_L(s) \cap E_{L,x_0}\Big) = \mathbb{P}\Big(G \cap \Lambda_L(s) \cap E_{L,x_0} \cap \{|\xi_L(x_0) - a_L| < \theta_L\}\Big) + o\Big(\frac{1}{L^d}\Big)$$

which, together with Lemma 4.8, ultimately gives (4.36).

Proof. Our goal is to show that uniformly over all $x_0 \in Q_{R_L-r_L}$

$$\mathbb{P}\Big(G \cap \Lambda_L(s) \cap E_{L,x_0} \cap \{|\xi_L(x_0) - a_L| \ge \theta_L\}\Big) = o\Big(\frac{1}{L^d}\Big).$$
(4.43)

By a union bound, we can separately estimate the probability of the events $G \cap \Lambda_L(s) \cap E_{L,x_0} \cap \{\xi_L(x_0) \ge a_L + \theta_L\}$ and $G \cap \Lambda_L(s) \cap E_{L,x_0} \cap \{\xi_L(x_0) \le a_L - \theta_L\}$. For the former, (4.43) holds as can be seen by applying (1.23) to $\mathbb{P}(\xi_L(x_0) \ge a_L + \theta_L)$ and using that, by definition of θ_L in (4.35), we have $a_L \theta_L \gg 1$.

For the latter, notice that by (4.10), on $E_{L,x_0} \cap \{\xi_L(x_0) \leq a_L - \theta_L\}$ we have

$$\lambda_1(Q_{R_L},\xi_L) - \bar{\lambda}_L - \Xi_L(x_0) \le C\left(\left(\frac{d_L}{a_L}\right)^{2-2\kappa}(a_L - \xi_L(x_0)) + \frac{d_L}{a_L^2}\right),$$

so that on $\Lambda_L(s) \cap E_{L,x_0} \cap \{\xi_L(x_0) \le a_L - \theta_L\}$ we have

$$\xi_L(x_0) + \Phi_L(x_0) \ge a_L \sqrt{1 + \tau_L^2} - \frac{s + Cd_L/a_L}{a_L} - C\left(\frac{d_L}{a_L}\right)^{2-2\kappa} (a_L - \xi_L(x_0)) ,$$

which implies

$$\Phi_L(x_0) \ge a_L(\sqrt{1+\tau_L^2}-1) - \frac{s + Cd_L/a_L}{a_L} + (a_L - \xi_L(x_0))(1 - C(d_L/a_L)^{2-2\kappa})$$

$$\ge \frac{1}{2}(a_L - \xi_L(x_0)),$$

where we used that, by definition of θ_L , on $\{\xi_L(x_0) \leq a_L - \theta_L\}, |s + Cd_L/a_L|/a_L \ll \theta_L \leq a_L - \xi_L(x_0)$. We can now exploit the independence of $\xi_L(x_0)$ and $\Phi_L(x_0)$ (and the fact that the variance of the latter is τ_L^2) to deduce

$$\begin{split} \mathbb{P}\Big(G \cap \Lambda_L(s) \cap E_{L,x_0} \cap \{\xi_L(x_0) \le a_L - \theta_L\}\Big) \\ & \leq \mathbb{P}\Big(\Phi_L(x_0) \ge \frac{1}{2}(a_L - \xi_L(x_0)) \,;\, \xi_L(x_0) \le a_L - \theta_L\Big) \\ & = \int_{\theta_L}^{\infty} \frac{e^{-\frac{(y-a_L)^2}{2}}}{\sqrt{2\pi}} \mathbb{P}\Big(\Phi_L(x_0) \ge \frac{y}{2}\Big) \mathrm{d}y \lesssim \frac{a_L}{L^d} \int_{\theta_L}^{\infty} e^{ya_L} e^{-\frac{y^2}{8\tau_L^2}} \mathrm{d}y \\ & \leq \frac{a_L}{L^d} \int_{\theta_L}^{\infty} e^{-\frac{y^2}{16\tau_L^2}} \mathrm{d}y = 4\sqrt{\pi} \frac{a_L\tau_L}{L^d} \mathbb{P}\Big(\mathcal{N}(0,1) \ge \frac{\theta_L}{\sqrt{8}\tau_L}\Big) \lesssim \frac{1}{L^d} \frac{a_L\tau_L^2}{\theta_L} \,. \end{split}$$

In the third step, we neglected the term $e^{-y^2/2}$, and we used (1.22) and (1.20), since $\theta_L \gg \tau_L$. The fourth step relies on (4.35), as $y > \theta_L \gg a_L \tau_L^2$ implies $a_L y \leq y^2/(16\tau_L^2)$ for L large enough. The last step uses (1.21) (as $\theta_L \gg \tau_L$) and a basic exponential bound. By (4.35), the last quantity is negligible compared to L^{-d} , and thus the proof of the statement is complete.

5. The macroscopic eigenproblem

The goal of this section is to prove Theorems 1.6, 1.7 and 1.8. Recall the splitting scheme introduced in (1.19) and the definition of U_L in (1.19). Similarly to what was done in Section 3 for the potential, we will first (Section 5.1) establish the above mentioned theorems for the Hamiltonian restricted to U_L , and then, in Section 5.2, show that the difference in behaviour on U_L and Q_L is negligible. To carry out the first task, we will patch together the spectral information on the operator $\Delta + \xi_L$ on each mesoscopic box $Q_{R_L,z_{j,L}}$ in order to deduce the spectral behaviour of the same operator but on U_L .

5.1. Convergence of the top of the spectrum on U_L . Consider the operator \mathcal{H}_{U_L,ξ_L} and let $(\hat{\lambda}_{k,L}, \hat{\varphi}_{k,L})_{k\geq 1}$ be the sequence of its eigenvalues and normalised eigenfunctions in the non-increasing order of their first coordinates: this is nothing but the collection of all the eigenvalues and eigenfunctions of $\Delta + \xi_L$ on every mesoscopic box $Q_{R_L,z_{j,L}}$. We will argue below that only the principal eigenvalue / eigenfunction on each mesoscopic box may contribute to the top of the spectrum on U_L with large probability.

Let $\hat{y}_{k,L} \in U_L$ be the point where ξ_L reaches its k-th largest value on U_L . Denote by $\hat{x}_{k,L}$ the point where $|\hat{\varphi}_{k,L}|$ reaches its maximum, and assume w.l.o.g. that $\hat{\varphi}_{k,L}$ is positive at this point. Finally, let $\hat{\ell}_L(k)$ be defined through $\hat{x}_{k,L} = \hat{y}_{\hat{\ell}_L(k),L}$.

The main result of this section is the translation of Theorems 1.6, 1.7 and 1.8 for the operator \mathcal{H}_{U_L,ξ_L} .

Theorem 5.1. The following statements are satisfied.

(1) The point process

$$\left(\frac{\hat{x}_{k,L}}{L}, a_L \left(\hat{\lambda}_{k,L} - a_L \sqrt{1 + \tau_L^2} - \bar{\lambda}_L\right)\right)_{1 \le k \le \# U_L}$$

converges in law as $L \to \infty$ towards a Poisson point process on $[-1, 1]^d \times \mathbb{R}$ of intensity $dx \otimes e^{-u} du$,

43

(2) For any $k \ge 1$, the r.v.

$$\frac{a_L}{d_L} \left\| \hat{\varphi}_{k,L}(\cdot) - \bar{\varphi}_L(\cdot - \hat{x}_{k,L}) \right\|_{\ell^2(Q_L)}$$

converges to 0 in probability.

- (3) It holds:
 - (a) if $\tau_L \ll \frac{1}{a_L}$, then for any given $k \ge 1$, $\mathbb{P}(\hat{\ell}_L(k) = k) \to 1$ as $L \to \infty$,
 - (b) if $\tau_L \sim \sqrt{\tilde{b}} \frac{1}{a_L}$ for some constant b > 0, then $(\hat{\ell}_L(k))_{k \ge 1}$ converges in law to $(\ell_{\infty,b}(k))_{k \ge 1}$, the latter being defined according to (1.14),
 - (c) if $\tau_L \gg \frac{1}{a_L}$, then for any given $k \ge 1$, $\ell_L(k) \to \infty$ in probability.

For every $j \in \{1, \ldots, n_L\}$, let $w_{j,L}$ be the location of the maximum of ξ_L on $Q_{R_L, z_{j,L}}$, $\Phi_L(\cdot)$ be given as in (2.4), $a_L^{\Xi} = a_L \sqrt{1 + \tau_L^2}$ as in (2.12), $\varphi_{R_L, j}$ be the eigenfunction of the operator $\mathcal{H}_{Q_{R_L, z_{j,L}}, \xi_L}$ associated to $\lambda_1(Q_{R_L, z_{j,L}}, \xi_L)$ and $(\eta_L)_{L\geq 1}$ be the vanishing sequence as in Theorem 4.1.

Let $\Xi_{j,L} = \xi_L(w_{j,L}) + \Phi_L(w_{j,L})$ and $(j_k)_{1 \le k \le n_L}$ be the permutation of $(1, \ldots, n_L)$ corresponding to the order statistics of $(\Xi_{j,L})_j$, that is, $\Xi_{j_1,L} \ge \Xi_{j_2,L} \ge \ldots$. The next lemma shows, among other things, that these order statistics provide the ordering of the eigenvalues with large probability.

Lemma 5.2. Let $(C_L)_{L\geq 1}$ be an arbitrary sequence of non-negative numbers going to ∞ as $L \to \infty$. For any integer $k \geq 1$, let $\mathcal{V}_L(k)$ be the event on which

$$\hat{\lambda}_{k,L} = \lambda_1(Q_{R_L, z_{j_k,L}}, \xi_L) , \quad \hat{\varphi}_{k,L} = \varphi_{R_L, j_k} , \quad \hat{x}_{k,L} = w_{j_k,L} , \quad (5.1)$$

and

$$\begin{aligned} \left|\lambda_1(Q_{R_L,z_{j_k,L}},\xi_L) - (\Xi_{j_k,L} + \bar{\lambda}_L)\right| &\leq \frac{\eta_L}{a_L} ,\\ \frac{a_L}{d_L} \left\|\varphi_{R_L,j_k}(\cdot) - \bar{\varphi}_L(\cdot - w_{j_k,L})\right\|_{\ell^2(Q_L)} &\leq \eta_L ,\\ \xi_L(w_{j_k,L}) &\leq \frac{a_L}{\sqrt{1 + \tau_L^2}} + C_L \max\{\frac{1}{a_L}, \tau_L\} . \end{aligned}$$

$$(5.2)$$

Then, the probability of $\mathcal{V}_L(k)$ goes to 1 as $L \to \infty$.

We postpone the proof of this crucial lemma, and proceed with that of Theorem 5.1.

Proof of Theorem 5.1. Fix an integer $k_0 \ge 1$. Throughout the proof, we will use that, by Lemma 5.2, we have

$$\mathbb{P}\Big(\bigcap_{k=1}^{k_0} \mathcal{V}_L(k)\Big) \longrightarrow 1, \qquad \text{as } L \to \infty.$$
(5.3)

In particular (see (5.1), (5.2)), the k_0 pairs $(\hat{x}_{k,L}, \lambda_{k,L})_{1 \leq k \leq k_0}$ match with the k_0 pairs $(z_{j_k,L}, \Xi_{L,j_k} + \bar{\lambda}_L)_{1 \leq k \leq k_0}$, up to an error of at most R_L for the first coordinate and η_L/a_L for the second. Thus, the convergence of \mathcal{P}_L^{Ξ} stated in Proposition 3.1 ensures that $(\hat{x}_{k,L}/L, a_L(\hat{\lambda}_{k,L} - a_L^{\Xi} - \bar{\lambda}_L))_{1 \leq k \leq k_0}$ converges in law to the k_0 largest points (in the non-increasing order of their second coordinate) of $\mathcal{P}_{\Sigma}^{\Xi}$ and this completes the proof of (1).

Furthermore, for any integer $k \leq k_0$

$$\frac{a_L}{d_L} \left\| \hat{\varphi}_{k,L}(\cdot) - \bar{\varphi}_L(\cdot - \hat{x}_{k,L}) \right\|_{\ell^2(Q_L)} = \frac{a_L}{d_L} \left\| \varphi_{R_L,j_k}(\cdot) - \bar{\varphi}_L(\cdot - w_{j_k,L}) \right\|_{\ell^2(Q_L)} \le \eta_L ,$$

so that also the conclusion of (2) follows.

We turn to (3). For any $k \geq 1$, on $\mathcal{V}_L(k)$ the r.v. $\hat{\ell}_L(k)$ is the rank of the r.v. $\xi_L(w_{j_k,L})$ among the values taken by ξ_L on U_L in non-increasing order, that is

$$w_{j_k,L} = \hat{y}_{\hat{\ell}_L(k),L} \,. \tag{5.4}$$

Moreover, on $\mathcal{V}_L(k)$ we have

44

$$\xi_L(w_{j_k,L}) \le \frac{a_L}{\sqrt{1 + \tau_L^2}} + C_L \max\{\frac{1}{a_L}, \tau_L\} =: \varrho_L^+ .$$
(5.5)

If $\tau_L \gg a_L^{-1}$, upon taking $C_L = \sqrt{a_L \tau_L}$ and applying a Taylor expansion of the r.h.s. of the last inequality, we see that $a_L(\varrho_L^+ - a_L) \to -\infty$. The convergence of \mathcal{P}_L^{ξ} stated in Proposition 3.1 gives that $\mathcal{P}_L^{\xi}([-1, 1]^d \times [a_L(\varrho_L^+ - a_L), \infty])$ goes to ∞ in probability as $L \to \infty$, which means that the number of points in U_L where ξ_L lies above ϱ_L^+ diverges in probability. As a consequence the rank $\hat{\ell}_L(k)$ goes to ∞ in probability and this yields (3)-(c) of the theorem.

We now assume that $\tau_L = \mathcal{O}(a_L^{-1})$. To cover jointly (3)-(a) and (3)-(b), when $\tau_L \ll a_L^{-1}$, set b = 0 and $\ell_{\infty,0}(k) = k$ for any $k \ge 1$, while for $a_L \tau_L \sim \sqrt{b}$ and b > 0 recall that $\ell_{\infty,b}$ was defined in (1.14).

For any $k \ge 1$ let j'_k be the index of the mesoscopic box where the k-th largest value of ξ_L on U_L lies, that is, $\hat{y}_{k,L} \in Q_{R_L,z_{j'_L,L}}$. Then, define

$$u_{k,L} \stackrel{\text{def}}{=} a_L(\xi_L(\hat{y}_{k,L}) - a_L) , \quad v_{k,L} \stackrel{\text{def}}{=} a_L \Phi_L(w_{j'_k,L}) .$$

$$(5.6)$$

The convergence of $\mathcal{P}_L^{(\xi,\Phi)}$ stated in Proposition 3.1 implies that $(u_{k,L}, v_{k,L})_{k\geq 1}$ converges in law to $(u_k, v_k)_{k\geq 1}$ where $u_1 > u_2 > \ldots$ follow a Poisson point process of intensity $e^{-u} du$, and $(v_k)_{k\geq 1}$ are independent $\mathcal{N}(0, b)$ r.v.'s.

Recall that $(j_k)_{k\geq 1}$ is the permutation corresponding to the order statistics of $(\Xi_{j,L})_{j\geq 1}$ where $\Xi_{j,L} = \xi_L(w_{j,L}) + \Phi_L(w_{j,L})$. Now set

$$p_{k,L} \stackrel{\text{def}}{=} a_L \big(\xi_L(w_{j_k,L}) - a_L \big) + a_L \Phi_L(w_{j_k,L}) \,.$$

Combining (5.4) and (5.6), the integer $\hat{\ell}_L(k)$ is such that $j_k = j'_{\hat{\ell}(k)}$ and thus

$$p_{k,L} = u_{\hat{\ell}_L(k),L} + v_{\hat{\ell}_L(k),L} \,. \tag{5.7}$$

Hence, the r.h.s. of (5.7) converges in law to the r.h.s. of (1.14) and therefore the r.v.'s $(\hat{\ell}_L(k))_{k\geq 1}$ converge in law to $(\ell_{\infty,b}(k))_{k\geq 1}$, which gives both (3)-(a) and (3)-(b). The proof of the theorem is complete.

It remains to show Lemma 5.2 whose proof relies extensively on Theorem 4.1. Let us first introduce some additional notation. For any C > 0, let $F_{j,L}(C) = F_{j,L} \stackrel{\text{def}}{=} \bigcup_{i=1}^{4} F_{j,L}^{i}$, for $j \in \{1, \ldots, n_L\}$, be the union of the events whose probability is estimated in (4.3)-(4.7) but on boxes of side-length R_L centred at $z_{j,L}$, i.e.

$$F_{j,L}^{1} \stackrel{\text{def}}{=} \left\{ \left| \lambda_{1}(Q_{R_{L},z_{j,L}},\xi_{L}) - \left(\Xi_{L}(w_{j,L}) + \bar{\lambda}_{L}\right) \right| > \frac{\eta_{L}}{a_{L}} \right\},\$$

$$\begin{split} F_{j,L}^{2} &\stackrel{\text{def}}{=} \left\{ \xi_{L}(w_{j,L}) \notin \left[\frac{a_{L}}{\sqrt{1+\tau_{L}^{2}}} - C \max(\frac{1}{a_{L}},\tau_{L}), \frac{a_{L}}{\sqrt{1+\tau_{L}^{2}}} + C \max(\frac{1}{a_{L}},\tau_{L}) \right] \right\} \\ F_{j,L}^{3} &\stackrel{\text{def}}{=} \left\{ \lambda_{2}(Q_{R_{L},z_{j,L}},\xi_{L}) > a_{L}^{\Xi} + \bar{\lambda}_{L} - C' \frac{a_{L}}{d_{L}} \right\}, \\ F_{j,L}^{4} &\stackrel{\text{def}}{=} \left\{ \|\varphi_{R_{L},j}(\cdot) - \bar{\varphi}_{L}(\cdot - w_{j,L})\|_{\ell^{2}(Q_{R_{L},z_{j,L}})} > \frac{d_{L}}{a_{L}} \eta_{L} \right\}, \end{split}$$

where C' > 0 is fixed and chosen so that (4.6) holds. Set also

$$G_{j,L}(s) \stackrel{\text{def}}{=} \left\{ \Xi_L(w_{j,L}) \ge a_L^{\Xi} + \frac{s}{a_L}; \left| \lambda_1(Q_{R_L, z_{j,L}}, \xi_L) - \left(\Xi_L(w_{j,L}) + \bar{\lambda}_L \right) \right| > \frac{\eta_L}{a_L} \right\}.$$

Given C > 0, the event of interest is (the complement of)

$$B_L(s,C) \stackrel{\text{def}}{=} \bigcup_{j=1}^{n_L} B_{j,L}(s,C) , \qquad (5.8)$$

where

$$B_{j,L}(s,C) \stackrel{\text{def}}{=} \left(\left\{ \lambda_1(Q_{R_L,z_{j,L}},\xi_L) \ge a_L^{\Xi} + \bar{\lambda}_L + \frac{s}{a_L} \right\} \cap F_{j,L}(C) \right) \cup G_{j,L}(s) .$$

The probability of $B_{j,L}(s,C)$ is independent of j, thus

$$\mathbb{P}(B_L(s,C)) \le n_L \,\mathbb{P}(B_{1,L}(s,C)) \lesssim \left(\frac{L}{R_L}\right)^d \mathbb{P}(B_{1,L}(s,C))$$

and, thanks to (4.3)-(4.7) we obtain

$$\limsup_{C \to \infty} \limsup_{L \to \infty} \mathbb{P}(B_L(s, C)) = 0.$$
(5.9)

Proof of Lemma 5.2. Fix $k_0 \ge 1$ and $\epsilon > 0$. We will show that

$$\lim_{L\to\infty} \mathbb{P}\big(\mathcal{V}_L(1)\cap\ldots\cap\mathcal{V}_L(k_0)\big)\geq 1-\epsilon.$$

Consider the event

$$D_L(c) \stackrel{\text{def}}{=} \left\{ \Xi_{L,j_{k_0}} < a_L^{\Xi} - \frac{c}{a_L} \right\} \cup \bigcup_{k=1}^{k_0} \left\{ |\Xi_{L,j_{k+1}} - \Xi_{L,j_k}| \le 10 \frac{\eta_L}{a_L} \right\} \,.$$

From the convergence of \mathcal{P}_L^{Ξ} stated in Proposition 3.1, we deduce that, provided c > 0 is sufficiently big, for all L large enough $\mathbb{P}(D_L(c)) < \epsilon/2$. Furthermore, choosing also C > 0 sufficiently big, we deduce from (5.9) that for all L large enough $\mathbb{P}(B_L(-c-1,C)) < \epsilon/2$.

We now work on the event $B_L(-c-1,C)^{\complement} \cap D_L(c)^{\complement}$ whose probability is at least $1 - \epsilon$ for all L large enough, and will show that this event is contained in $\mathcal{V}_L(1) \cap \cdots \cap \mathcal{V}_L(k_0)$. Using the fact that we are on the complements of the events $F_{j_k,L}^1, F_{j_k,L}^3$ and $G_{j_k,L}(-c-1), 1 \leq k \leq k_0$, as well as the complement of $D_L(c)$, we deduce that there is a one-to-one correspondence between the k_0 largest eigenvalues / eigenfunctions of \mathcal{H}_{U_L,ξ_L} , and the k_0 largest principal eigenvalues / eigenfunctions over the mesoscopic boxes, namely for every $1 \leq k \leq k_0$

$$\lambda_{k,L} = \lambda_1(Q_{R_L, z_{j_k,L}}, \xi_L) , \quad \hat{\varphi}_{k,L} = \varphi_{R_L, j_k} .$$

The three bounds of (5.2) follow from the complements of the events $F_{j_k,L}^1$, $F_{j_k,L}^4$, and $F_{j_k,L}^2$ (note that C_L lies above C for all L large enough). Since $\bar{\varphi}_L$ is almost a Dirac mass at the origin, the second bound in (5.2) also implies that $\hat{\varphi}_{k,L}$ (which is equal to φ_{R_L,j_k}) admits its maximum at $w_{j_k,L}$ and therefore $\hat{x}_{k,L} = w_{j_k,L}$. \Box

45

5.2. Proof of the main results. This last section is devoted to the proof of Theorems 1.6, 1.7 and 1.8. Thanks to Theorem 5.1, what remains to show is that the eigenvalues and eigenfunctions of \mathcal{H}_L are sufficiently close to those of \mathcal{H}_{U_L,ξ_L} and that the localisation centres are the same. More precisely, we need to check that for any $k \in \mathbb{N}$, the random variables

$$a_L(\hat{\lambda}_{k,L} - \lambda_{k,L}), \quad \frac{a_L}{d_L} \|\hat{\varphi}_{k,L} - \varphi_{k,L}\|_{\ell^2(Q_L)},$$
 (5.10)

converge to 0 in probability as $L \to \infty$, and that the probability of the event $\{\hat{x}_{k,L} = x_{k,L}\}$ goes to 1 as $L \to \infty$.

Recall that R_L and r_L satisfy (1.15) and (1.16), respectively. As a preliminary step, fix $k_0 \geq 1$ and define the event $G_L \stackrel{\text{def}}{=} G_L^{(1)} \cup G_L^{(2)}$ as

(1) on $G_L^{(1)}$ the first k_0+1 eigenvalues of \mathcal{H}_{U_L,ξ_L} are larger than $a_L^{\Xi}+\bar{\lambda}_L-\frac{1}{\sqrt{a_L}}$, and all their spacings are at least $a_L^{-3/2}$, i.e.

$$\hat{\lambda}_{k_0+1,L} \ge a_L^{\Xi} + \bar{\lambda}_L - \frac{1}{\sqrt{a_L}} \quad \text{and} \quad \hat{\lambda}_{i,L} - \hat{\lambda}_{i+1,L} > a_L^{-3/2} ,$$

for all $i \in \{1, ..., k_0\}$,

(2) on $G_L^{(2)}$, for all $x \in Q_L$ we have (recall that $\theta = 2d + 1$)

$$x \notin \bigcup_{j=1}^{n_L} Q_{R_L - r_L, z_{j,L}} \implies \xi_L(x) < a_L - \theta$$
.

The lower bounds $a_L^{\Xi} + \bar{\lambda}_L - \frac{1}{\sqrt{a_L}}$ and $a_L^{-3/2}$ in the first bullet point are relatively arbitrary and chosen so that the next lemma holds.

Lemma 5.3. We have $\mathbb{P}(G_L) \to 1$ as $L \to \infty$.

Proof. By the first item of Theorem 5.1, we know that $(a_L(\hat{\lambda}_{k,L} - a_L^{\Xi} - \bar{\lambda}_L))_{k \geq 1}$ converges to a Poisson point process of intensity $e^{-u}du$. Therefore, the probability of $G_L^{(1)}$ goes to 1. Regarding $G_L^{(2)}$, note that the cardinality of $Q_L \setminus \bigcup_{j=1}^{n_L} Q_{R_L-r_L,z_{j,L}} = \bigcup_{j=1}^{n_L} Q_{R_L+\sqrt{R_L},z_{j,L}} \setminus Q_{R_L-r_L,z_{j,L}}$ (recall (1.18)) is of order

$$n_L (R_L - r_L)^{d-1} (\sqrt{R_L} + r_L) \lesssim L^d \frac{1}{\sqrt{R_L}}.$$

Hence, a union bound and (1.23) imply that for L large enough

$$\mathbb{P}\left((G_L^{(2)})^{\complement}\right) \leq \sum_{x \in Q_L \setminus \bigcup_{j=1}^{n_L} Q_{R_L - r_L, z_{j,L}}} \mathbb{P}(\xi_L(x) \geq a_L - \theta) \lesssim \frac{1}{\sqrt{R_L}} e^{\theta a_L}$$

and, since $\ln R_L \gg a_L$ by (1.15) and (1.16), the right-hand side vanishes as $L \to \infty$. Thus, the statement follows.

We are now ready to complete the proof of the main statements.

Proof of Theorems 1.6, 1.7 and 1.8. In view of Lemma 5.3, we can and will work on G_L throughout the proof. Our goal is to show that for every $k \in \{1, \ldots, k_0\}$, the r.v.'s in (5.10) vanish as $L \to \infty$. Once this is established, we can easily deduce by Theorem 5.1 that the localisation centres are the same with probability converging to 1. Indeed, recall that $x_{k,L}$, resp. $\hat{x}_{k,L}$, is the point at which $|\varphi_{k,L}|$, resp. $|\hat{\varphi}_{k,L}|$, achieves its maximum. By item (2) of Theorem 5.1, combined with the fact that $\bar{\varphi}_L(0) = 1 - \mathcal{O}(d_L/a_L)$ and $\bar{\varphi}_L(x) = \mathcal{O}(d_L/a_L)$ for $x \neq 0$ as shown in (4.14), we deduce that $\hat{\varphi}_{k,L}(\hat{x}_{k,L})$ converges to 1 in probability as $L \to \infty$, while $\hat{\varphi}_{k,L}(y)$ for $y \neq \hat{x}_{k,L}$ vanishes at rate d_L/a_L . Now, if $\frac{a_L}{d_L} \|\hat{\varphi}_{k,L} - \varphi_{k,L}\|_{L^2(Q_L)}$ goes to 0 in probability, then $\varphi_{k,L}$ behaves as $\hat{\varphi}_{k,L}$, which means that it converges to 1 at $\hat{x}_{k,L}$ and vanishes elsewhere, so that $\hat{x}_{k,L}$ is the unique maximum of $\varphi_{k,L}$ and thus the probability of $x_{k,L} = \hat{x}_{k,L}$ goes to 1.

Let us now turn to the convergence of the r.v.'s in (5.10). What we will prove is that these quantities are bounded above by a deterministic constant that goes to 0 as $L \to \infty$.

Denote by ∂U_L the inner boundary of U_L , that is, the set of points of U_L that admit at least one neighbour outside U_L . Similarly denote by $\partial(Q_L \setminus U_L)$ the set of points of $Q_L \setminus U_L$ that admit at least one neighbour in U_L .

Let (λ, φ) be an eigenvalue/eigenfunction of \mathcal{H}_L on Q_L which we assume to be such that $\lambda \geq \hat{\lambda}_{k_0,L}$. Recall that $\theta = 2d + 1$ and $\bar{\lambda}_L \geq -2d$. Note that, as we are on G_L , for all $x \in Q_L \setminus \bigcup_{j=1}^{n_L} Q_{R_L - r_L, z_{j,L}}$, we have for large L

$$\xi_L(x) < a_L - \theta \le a_L^{\Xi} + \bar{\lambda}_L - 1 \le \hat{\lambda}_{k_0,L} - 1 + \frac{1}{\sqrt{a_L}} \le \lambda - \frac{1}{2}$$

Therefore, Lemma 4.3 applied with $D' = (Q_L \setminus U_L) \cup \partial U_L$, A' = A = 1/2 and $R = r_L - 1$ yields

$$\sum_{x \in (Q_L \setminus U_L) \cup \partial U_L} |\varphi(x)|^2 \le \delta_L \stackrel{\text{def}}{=} \left(1 + \frac{1}{4d}\right)^{-2(r_L - 1)}.$$
 (5.11)

and, since $r_L \ge a_L$ by (1.16), the r.h.s. is negligible compared to a_L^{-n} for any given $n \ge 1$. In particular, φ puts negligible mass on the complement of U_L . What we want to do now is (a) use the above to show that the ℓ^2 -distance between φ and its normalised restriction to U_L is small, and (b) prove that there exists a unique k such that the latter is close to $\hat{\varphi}_{k,L}$.

Set

$$\psi \stackrel{\text{def}}{=} \frac{\varphi \mathbb{1}_{U_L}}{\|\varphi \mathbb{1}_{U_L}\|_2} ,$$

where, here and below, we write $\|\cdot\|_2$ for the $\ell^2(Q_L)$ -norm and $\|\cdot\|_{\ell^2(U_L)}$ for the $\ell^2(U_L)$ -norm. For (a), it suffices to note that, for all L large enough

$$\|\psi - \varphi\|_{2} = \|\psi - \varphi \mathbb{1}_{U_{L}} + \varphi \mathbb{1}_{U_{L}} - \varphi\|_{2} \le (1 - \|\varphi \mathbb{1}_{U_{L}}\|_{2}) + \|\varphi \mathbb{1}_{U_{L}^{c}}\|_{2} \le 2\sqrt{\delta_{L}} .$$
(5.12)

For (b) instead, the argument exploits the equation satisfied by φ and $\hat{\varphi}_{k,L}$, and the fact that $(\hat{\varphi}_{k,L})_{k\geq 1}$ forms an orthonormal basis of $\ell^2(U_L)$. By the former, we get

$$(\mathcal{H}_L - \lambda)\psi = \frac{1}{\|\varphi \mathbb{1}_{U_L}\|_2} \Big(\Delta(\varphi \mathbb{1}_{U_L}) + (\xi_L - \lambda)\varphi \mathbb{1}_{U_L} \Big) \\ = \frac{1}{\|\varphi \mathbb{1}_{U_L}\|_2} \Big(\Delta(\varphi \mathbb{1}_{U_L}) - \mathbb{1}_{U_L}\Delta\varphi + (\Delta\varphi + (\xi_L - \lambda)\varphi)\mathbb{1}_{U_L} \Big) \\ = \frac{1}{\|\varphi \mathbb{1}_{U_L}\|_2} \Big(\Delta(\varphi \mathbb{1}_{U_L}) - \mathbb{1}_{U_L}\Delta\varphi \Big) .$$

48

and the r.h.s. is 0 outside $\partial U_L \cup \partial (Q_L \setminus U_L)$. It is then easy to check that there exists a constant C > 0, independent of L, such that

$$\|(\mathcal{H}_L - \lambda)\psi\|_{\ell^2(U_L)}^2 \le \frac{C}{\|\varphi \mathbb{1}_{U_L}\|_2^2} \sum_{x \in (Q_L \setminus U_L) \cup \partial U_L} |\varphi(x)|^2 \le C \frac{\delta_L}{1 - \delta_L} \,. \tag{5.13}$$

On the other hand, we can expand ψ on the $\ell^2(U_L)$ basis provided by the eigenfunctions of \mathcal{H}_{U_L,ξ_L} thus yielding

$$\psi = \sum_{k \ge 1} \hat{\varphi}_{k,L} \langle \hat{\varphi}_{k,L}, \psi \rangle .$$

Since further $\mathcal{H}_L \hat{\varphi}_{k,L} = \mathcal{H}_{U_L,\xi_L} \hat{\varphi}_{k,L} = \hat{\lambda}_{k,L} \hat{\varphi}_{k,L}$ on U_L , we can write

$$\|(\mathcal{H}_L - \lambda)\psi\|_{\ell^2(U_L)}^2 = \sum_{k \ge 1} \langle \hat{\varphi}_{k,L}, \psi \rangle^2 |\hat{\lambda}_{k,L} - \lambda|^2 .$$

Now, by construction, $1 = \|\psi\|_{\ell^2(U_L)} = \sum_{k \ge 1} \langle \hat{\varphi}_{k,L}, \psi \rangle^2$, so that the sum at the r.h.s. is a convex combination of the $(|\hat{\lambda}_{k,L} - \lambda|^2)_{k \ge 1}$. Then, (5.13) implies that necessarily there exists a $k \ge 1$ such that

$$|\hat{\lambda}_{k,L} - \lambda|^2 \le C \frac{\delta_L}{1 - \delta_L} \,. \tag{5.14}$$

As, by (5.11), $\delta_L \ll a_L^{-3}$, $\lambda \geq \hat{\lambda}_{k_0,L}$ by assumption and, on G_L , the spacings between the $k_0 + 1$ first eigenvalues $\hat{\lambda}_{k,L}$ are all larger than $a_L^{-3/2}$, the integer k satisfying (5.14) belongs to $\{1, \ldots, k_0\}$ and is unique. Moreover, for any $\ell \neq k$, we must have

$$|\hat{\lambda}_{\ell,L} - \lambda|^2 \ge \frac{1}{2a_L^3}$$

because $(a+b)^2 \le 2(a^2+b^2)$ and $a^2 \ge \frac{1}{2}((a+b)^2-2b^2)$, we have

$$|\hat{\lambda}_{\ell,L} - \lambda|^2 \ge \frac{1}{2} (|\lambda_{k,L} - \lambda_{\ell,L}|^2 - 2|\hat{\lambda}_{k,L} - \lambda|^2) \ge \frac{1}{2} \Big(\frac{1}{a_L^3} - 2C \frac{\delta_L}{1 - \delta_L} \Big),$$

or equivalently $2a_L^3|\hat{\lambda}_{\ell,L} - \lambda|^2 \ge 1$. As a consequence,

$$q_{k,L}^2 \stackrel{\text{def}}{=} \sum_{\ell \neq k} \langle \hat{\varphi}_{\ell,L}, \psi \rangle^2 \le 2a_L^3 \sum_{\ell \neq k} \langle \hat{\varphi}_{\ell,L}, \psi \rangle^2 |\hat{\lambda}_{\ell,L} - \lambda|^2 \le C \frac{\delta_L}{1 - \delta_L} 2a_L^3 ,$$

that vanishes as $L \to \infty$ and thus gives

$$\begin{aligned} \|\psi - \hat{\varphi}_{k,L}\|_{\ell^{2}(U_{L})}^{2} &= (\langle \hat{\varphi}_{k,L}, \psi \rangle - 1)^{2} + \sum_{\ell \neq k} \langle \hat{\varphi}_{\ell,L}, \psi \rangle^{2} \\ &= (\sqrt{1 - q_{k,L}^{2}} - 1)^{2} + q_{k,L}^{2} \lesssim \delta_{L} a_{L}^{3} \,. \end{aligned}$$

Combining the previous with (5.12), we finally obtain

$$\|\varphi - \hat{\varphi}_{k,L}\|_2 \le \|\psi - \varphi\|_2 + \|\psi - \hat{\varphi}_{k,L}\|_{\ell^2(U_L)} \lesssim \sqrt{\delta_L a_L^3}.$$

Summarising, we have constructed, on the event G_L , a map that associates to any eigenvalue/eigenfunction (λ, φ) of \mathcal{H}_L such that $\lambda \geq \hat{\lambda}_{k_0,L}$, some $(\hat{\lambda}_{k,L}, \hat{\varphi}_{k,L})$ with $k \in \{1, \ldots, k_0\}$ such that we simultaneously have

$$|a_L(\hat{\lambda}_{k,L} - \lambda)|^2 \lesssim \delta_L a_L^2 , \quad \frac{a_L}{d_L} \|\varphi - \hat{\varphi}_{k,L}\|_2 \lesssim \sqrt{\delta_L a_L^3} \frac{a_L}{d_L} . \tag{5.15}$$

Note that this map is necessarily injective. Indeed, otherwise there would exist two orthonormal functions φ and $\tilde{\varphi}$ in $\ell^2(Q_L)$ such that for some k

$$\|\varphi - \hat{\varphi}_{k,L}\|_2 \lesssim \sqrt{\delta_L a_L^3} , \quad \|\tilde{\varphi} - \hat{\varphi}_{k,L}\|_2 \lesssim \sqrt{\delta_L a_L^3} ,$$

thus raising a contradiction.

By the variational formula, we know that there are at least k_0 eigenvalues of \mathcal{H}_L that lie above $\hat{\lambda}_{k_0,L}$, which means that the above map is also surjective, and thus bijective. From the ordering of the eigenvalues, this map necessarily sends $\lambda_{k,L}$ to $\hat{\lambda}_{k,L}$ for every $k \in \{1, \ldots, k_0\}$. Since δ_L is negligible compared to any negative power of a_L , the r.h.s.'s of (5.15) go to 0 as $L \to \infty$, and this ensures the convergence in probability to 0 of (5.10) and completes the proof.

APPENDIX A. GAUSSIAN ESTIMATES

Proof of Lemma 2.3. We set $u_L \stackrel{\text{def}}{=} a_L \sqrt{1 + \tau_L^2} + \frac{s}{a_L}$. At several places in the proof, we will use the inequality $1 - (2\tau_L^2/3) \leq (1 + \tau_L^2)^{-1/2} \leq 1 - (\tau_L^2/4)$ which holds true provided L is large enough. We start by proving (2.8). Since $X + Y_L$ is a standard Gaussian random variable with variance $1 + \tau_L^2$, a simple scaling argument applied to (1.23), combined with the fact that τ_L converges to 0 as $L \to \infty$, implies

$$\mathbb{P}\left(X+Y_L \ge u_L\right) = \mathbb{P}\left(X \ge \frac{u_L}{\sqrt{1+\tau_L^2}}\right) \sim \frac{1}{L^d} e^{-s}$$

Concerning (2.11), it follows from (2.9) and the fact that, for θ_L as in the statement, $I_L(C) \subset [a_L - \theta_L, a_L + \theta_L]$. Indeed, for the upper bound (the lower bound being analogous) we have

$$\frac{a_L}{\sqrt{1+\tau_L^2}} + C \max\{a_L^{-1}, \tau_L\} \le a_L - \frac{1}{8}a_L\tau_L^2 + C \max\{a_L^{-1}, \tau_L\} \ll a_L + \theta_L,$$

from which the result follows.

We are thus left with proving (2.9). As X and Y_L are independent, we have

$$\mathbb{P}\Big(X+Y_L \ge u_L \ ; \ X \notin I_L(C)\Big) = \int_{I_L(C)^c} \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}} \mathbb{P}(Y_L > u_L - x) \mathrm{d}x$$
$$= \int_{I_L(C)^c} \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}} \mathbb{P}\Big(\mathcal{N}(0,1) > \frac{u_L - x}{\tau_L}\Big) \mathrm{d}x =: J_L^1 + J_L^2$$

where the former is the integral over $x < a_L(1 + \tau_L^2)^{-1/2} - C \max(\frac{1}{a_L}, \tau_L) =: \varrho_L^$ while the latter that over $x > a_L(1 + \tau_L^2)^{-1/2} + C \max(\frac{1}{a_L}, \tau_L) =: \varrho_L^+$. We are going to show that $\limsup_{C \to \infty} \limsup_{L \to \infty} L^d J_L^i = 0, i = 1, 2$.

Let us begin with J_L^1 . Note that, provided C > |s| and L is large enough, for all $x < \varrho_L^-$

$$\frac{u_L - x}{\tau_L} \ge \frac{1}{2} a_L \tau_L \; .$$

We can apply (1.20) to deduce

$$J_L^1 \le \int_{-\infty}^{\varrho_L^-} \frac{e^{-\frac{x^2}{2}}}{2\pi} \frac{\tau_L}{u_L - x} e^{-\frac{(u_L - x)^2}{2\tau_L^2}} \mathrm{d}x \lesssim \frac{1}{a_L \tau_L} \int_{-\infty}^{\varrho_L^-} \frac{1}{2\pi} e^{-\frac{x^2}{2} - \frac{(u_L - x)^2}{2\tau_L^2}} \mathrm{d}x$$

Now,

$$\frac{x^2}{2} + \frac{(u_L - x)^2}{2\tau_L^2} = \frac{u_L^2}{2(1 + \tau_L^2)} + \frac{1 + \tau_L^2}{2\tau_L^2} \left(x - \frac{u_L}{1 + \tau_L^2}\right)^2$$

The first summand is independent of x and a simple computation combined with (1.23) yields

$$\frac{1}{\sqrt{2\pi}} e^{-\frac{u_L^2}{2(1+\tau_L^2)}} \lesssim \frac{a_L}{L^d} \,. \tag{A.1}$$

On the other hand, the change of variable $y = -\sqrt{(1 + \tau_L^2)/\tau_L^2}(x - \frac{u_L}{1 + \tau_L^2})$ and the fact that, provided $C > 2 \max\{|s|, 1\}, x < \varrho_L^-$ implies that y > C/2 yield

$$\int_{-\infty}^{\varrho_L^-} \frac{1}{\sqrt{2\pi}} e^{-\frac{1+\tau_L^2}{2\tau_L^2} (x - \frac{u_L}{1+\tau_L^2})^2} \mathrm{d}x \le \frac{\tau_L}{\sqrt{1+\tau_L^2}} \mathbb{P}\Big(\mathcal{N}(0,1) > \frac{C}{2}\Big) \le \frac{\tau_L}{\sqrt{1+\tau_L^2}} e^{-\frac{C^2}{8}} ,$$

where we used (1.20) at the last line. Putting everything together, we have shown that

$$J_L^1 \lesssim \frac{1}{L^d} \frac{1}{\sqrt{1 + \tau_L^2}} e^{-\frac{C^2}{8}} ,$$

so that $\limsup_{C\to\infty} \limsup_{L\to\infty} L^d J_L^1 = 0$. We turn to J_L^2 . First of all, we note that

$$V_{L} \stackrel{\text{def}}{=} \int_{a_{L} + \frac{C}{3a_{L}}}^{\infty} \frac{e^{-\frac{x^{2}}{2}}}{\sqrt{2\pi}} \mathbb{P}\Big(\mathcal{N}(0,1) > \frac{u_{L} - x}{\tau_{L}}\Big) \mathrm{d}x \le \int_{a_{L} + \frac{C}{3a_{L}}}^{\infty} \frac{e^{-\frac{x^{2}}{2}}}{\sqrt{2\pi}} \mathrm{d}x \lesssim \frac{1}{L^{d}} e^{-\frac{C}{3}} ,$$

where we used (1.23) at the last line. We thus deduce that the limsup first in $L \to \infty$ and then in $C \to \infty$ of $L^d V_L$ vanishes. Coming back to J_L^2 , we distinguish two cases. If $\tau_L \leq \sqrt{C}/a_L$ then

$$\varrho_L^+ \ge a_L - \frac{2}{3}a_L \tau_L^2 + \frac{C}{a_L} \ge a_L + \frac{C}{3a_L},$$

and therefore $J_L^2 \leq V_L$ and we can conclude. If $\tau_L > \sqrt{C}/a_L$ then, provided C is large enough compared to |s|,

$$u_L - \frac{1}{8}a_L\tau_L^2 \ge a_L + \frac{1}{4}a_L\tau_L^2 + \frac{s}{a_L} - \frac{1}{8}a_L\tau_L^2 \ge a_L + \frac{s + (C/8)}{a_L} \ge a_L + \frac{C}{10a_L}.$$

Consequently $J_L^2 = W_L + V_L$ where

$$W_L \stackrel{\text{def}}{=} \int_{\varrho_L^+}^{u_L - \frac{1}{8}a_L\tau_L^2} \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}} \mathbb{P}\Big(\mathcal{N}(0, 1) > \frac{u_L - x}{\tau_L}\Big) \mathrm{d}x$$

We now follow the same steps as for the bound on J_L^1 : provided C is large enough compared to |s| it holds

$$W_{L} \leq \int_{\varrho_{L}^{+}}^{u_{L} - \frac{1}{8}a_{L}\tau_{L}^{2}} \frac{e^{-\frac{x^{2}}{2}}}{2\pi} \frac{\tau_{L}}{u_{L} - x} e^{-\frac{(u_{L} - x)^{2}}{2\tau_{L}^{2}}} dx$$
$$\lesssim \frac{1}{a_{L}\tau_{L}} \int_{\varrho_{L}^{+}}^{u_{L} - \frac{1}{8}a_{L}\tau_{L}^{2}} \frac{1}{2\pi} e^{-\frac{x^{2}}{2} - \frac{(u_{L} - x)^{2}}{2\tau_{L}^{2}}} dx$$

50

$$\lesssim \frac{1}{L^d \tau_L} \int_{\varrho_L^+}^{\infty} \frac{1}{2\pi} e^{-\frac{1+\tau_L^2}{2\tau_L^2} (x - \frac{u_L}{1+\tau_L^2})^2} \mathrm{d}x$$
$$\lesssim \frac{1}{L^d} \mathbb{P} \left(\mathcal{N}(0, 1) > \frac{C}{2} \right) \,.$$

We can apply (1.20) and get $W_L \leq \frac{1}{L^d} e^{-\frac{C}{8}}$. Hence, $\lim_{C\to\infty} \lim_{L\to\infty} \lim_{L\to\infty} L^d J_L^2 = 0$, which, together with the same limit for J_L^1 , completes the proof. \Box

APPENDIX B. BASIC PROPERTIES AND ESTIMATES ON THE QUADRATIC FORMS

In this appendix, we state and prove some basic results concerning the quadratic forms \mathcal{D}_{R_L} and $\overline{\mathcal{D}}_L$ defined in (4.17), with \mathcal{H} and r given by $\mathcal{H}_{Q_{R_L},V_L}$ and R_L , and, $\overline{\mathcal{H}}_L$ and r_L , respectively. Recall that these operators are defined on $\mathcal{Z}_{R_L} \subset \ell^2(Q_{R_L}^{\neq 0})$ and $\mathcal{Z}_{r_L} \subset \ell^2(Q_{r_L}^{\neq 0})$, which are closed and convex (see (4.18)). With a slight abuse of notation, for ψ in either of the two sets, we write

$$\psi(0) \stackrel{\text{def}}{=} \sqrt{1 - \sum_{x \in Q_r^{\neq 0}} |\psi(x)|^2} . \tag{B.1}$$

In the following lemma, we collect the properties of \mathcal{D}_{R_L} and \mathcal{D}_L we will need.

Lemma B.1. The maps \mathcal{D}_{R_L} and $\overline{\mathcal{D}}_L$ are twice continuously differentiable and there exists $c_0 > 0$ such that for every $\psi \in \mathcal{Z}_{R_L}$ (resp. $\psi \in \mathcal{Z}_{r_L}$), the Hessian Hess $\mathcal{D}_{R_L}(\psi)$ (resp. Hess $\overline{\mathcal{D}}_L(\psi)$) satisfies on $E_L^1 \cap E_L^2$ (as in Definition 2.4)

$$\langle \varphi, \operatorname{Hess} \mathcal{D}_{R_L}(\psi) \varphi \rangle_{\ell^2(Q_{R_L}^{\neq 0})} \leq -c_0 \frac{a_L}{d_L} \|\varphi\|_{\ell^2(Q_{R_L}^{\neq 0})}^2, \quad \forall \varphi \in \ell^2(Q_{R_L})$$
(B.2)

(resp. $\langle \varphi, \operatorname{Hess} \bar{\mathcal{D}}_L(\psi) \varphi \rangle_{\ell^2(Q_{r_L}^{\neq 0})}$). In particular, \mathcal{D}_{R_L} (resp. $\bar{\mathcal{D}}_L$) has a unique maximiser and such maximiser is φ_{R_L} (resp. $\bar{\varphi}_L$). Furthermore, on $E_L^1 \cap E_L^2$, there exists a constant C > 0 such that for any L large enough

$$\|\nabla \mathcal{D}_{R_L}(\bar{\varphi}_L)\|_{\ell^2(Q_{R_L}^{\neq 0})} \le C\sqrt{\frac{1}{a_L} + \|\bar{\varphi}_L\zeta_L\|_{\ell^2(Q_{r_L})}^2}.$$
 (B.3)

At last, for any $x \in Q_1^{\neq 0}$ we have

$$\bar{\varphi}_L(x) = \frac{1}{\mathcal{S}_L(x)} (1 + o(1)) = \mathcal{O}\left(\frac{d_L}{a_L}\right), \tag{B.4}$$

where S_L is the shape in (1.6), and consequently

$$\bar{\lambda}_L = -2d + \sum_{x \in Q_1^{\neq 0}} \frac{1}{\mathcal{S}_L(x)} + o\left(\frac{d_L}{a_L}\right).$$
(B.5)

Proof. We start with the differentiability and convexity of \mathcal{D}_{R_L} . Observe that for $\psi \in \mathcal{Z}_{R_L}$, since $V_L(0) = 0$, the map \mathcal{D}_{R_L} is given by

$$\mathcal{D}_{R_L}(\psi) = -2d + 2\psi(0) \sum_{\substack{x \in Q_{R_L}^{\neq 0} \\ x \sim 0}} \psi(x) + \sum_{\substack{x, y \in Q_{R_L}^{\neq 0} \\ x \sim y}} \psi(x)\psi(y) + \sum_{\substack{x \in Q_{R_L}^{\neq 0} \\ x \sim y}} \psi(x)^2 V_L(x) \,.$$

A direct computation shows that the first derivative of \mathcal{D}_{R_L} in the direction $\psi(x)$ for $x \in Q_{R_L}^{\neq 0}$ is given by

$$\frac{\partial \mathcal{D}_{R_L}}{\partial \psi(x)}(\psi) = 2\psi(0)\mathbb{1}_{\{x \sim 0\}} + 2\frac{\partial \psi(0)}{\partial \psi(x)} \sum_{\substack{y \in Q_{R_L}^{\neq 0} \\ y \sim 0}} \psi(y) + 2\sum_{\substack{y \in Q_{R_L}^{\neq 0} \\ y \sim x}} \psi(y) + 2\psi(x)V_L(x) ,$$
(B.6)

and the second derivative reads

52

$$\frac{\partial^{2} \mathcal{D}_{R_{L}}}{\partial \psi(x)^{2}}(\psi) = 4 \frac{\partial \psi(0)}{\partial \psi(x)} \mathbb{1}_{\{x \sim 0\}} + 2 \frac{\partial^{2} \psi(0)}{\partial \psi(x)^{2}} \sum_{\substack{y \in Q_{R_{L}}^{\neq 0} \\ y \sim 0}} \psi(y) + 2V_{L}(x),$$

$$\frac{\partial^{2} \mathcal{D}_{R_{L}}}{\partial \psi(x') \partial \psi(x)}(\psi) = 2 \left(\frac{\partial \psi(0)}{\partial \psi(x')} \mathbb{1}_{\{x \sim 0\}} + \frac{\partial \psi(0)}{\partial \psi(x)} \mathbb{1}_{\{x' \sim 0\}} + \frac{\partial^{2} \psi(0)}{\partial \psi(x) \partial \psi(x')} \sum_{\substack{y \in Q_{R_{L}}^{\neq 0} \\ y \sim 0}} \psi(y) + \mathbb{1}_{\{x \sim x'\}} \right),$$
(B.7)

where the latter holds for $x \neq x'$. In the above expressions, the derivatives involving $\psi(0)$ equal

$$\frac{\partial\psi(0)}{\partial\psi(x)} = -\frac{\psi(x)}{\psi(0)} , \qquad \frac{\partial^2\psi(0)}{\partial\psi(x)\partial\psi(x')} = -\frac{\psi(x)\psi(x')}{\psi(0)^3} - \frac{\mathbb{1}_{\{x=x'\}}}{\psi(0)} .$$
(B.8)

Since any $\psi \in \mathcal{Z}_{R_L}$ satisfies the upper bound in (4.13) and $\psi(0)$ is given according to (B.1), it follows that $\psi(0) > 0$ so that $1/\psi(0)$ is well-defined. Hence, \mathcal{D}_{R_L} is twice continuously differentiable. It can now be checked that all terms appearing in the second order derivatives of \mathcal{D}_{R_L} are of order at most 1 except $V_L(x)$ which, on the event $E_L^1 \cap E_L^2$, is smaller than $-(\mathfrak{c}/2)a_L/d_L$ thanks to (2.16). Consequently, it is straightforward to check the existence of $c_0 > 0$ such that, uniformly in $\psi \in \mathcal{Z}_{R_L}$, Hess $\mathcal{D}_{R_L}(\psi)$ satisfies (B.2) for any $\varphi \in \ell^2(Q_{R_L})$. This in particular means that \mathcal{D}_{R_L} is strictly concave on \mathcal{Z}_{R_L} and its unique maximiser coincides with φ_{R_L} .

The arguments apply almost verbatim to $\bar{\mathcal{D}}_L$, the only specific input comes from the bound $-\mathcal{S}_L(x) \leq -\mathfrak{c}a_L/d_L$ which is a consequence of (1.2).

For (B.3), note that, since $\bar{\varphi}_L$ is a maximiser for $\bar{\mathcal{D}}_L$, $\nabla \bar{\mathcal{D}}_L(\bar{\varphi}_L) \equiv 0$. Viewing $\bar{\mathcal{D}}_L$ as a function from $Q_{R_L}^{\neq 0}$ into \mathbb{R} (which does not depend on $\psi(x)$ whenever $x \notin Q_{r_L}$), we further have $\frac{\partial \bar{\mathcal{D}}_L}{\partial \psi(x)}(\bar{\varphi}_L) = 0$ for all x. We now write

$$\begin{split} \frac{\partial \mathcal{D}_{R_L}}{\partial \psi(x)}(\bar{\varphi}_L) = & \frac{\partial \bar{\mathcal{D}}_L}{\partial \psi(x)}(\bar{\varphi}_L) + 2\mathbbm{1}_{\tilde{Q}_{r_L}}(x) \sum_{\substack{y \in Q_{r_L} \\ y \sim x}} \bar{\varphi}_L(y) + 2\bar{\varphi}_L(x)(V_L(x) + \mathcal{S}_L(x)) \\ = & 2\mathbbm{1}_{\tilde{Q}_{r_L}}(x) \sum_{\substack{y \in Q_{r_L} \\ y \sim x}} \bar{\varphi}_L(y) \\ & + 2\bar{\varphi}_L(x)(\xi_L(0) - a_L)[v_L(x) - 1] + 2\bar{\varphi}_L(x)\zeta_L(x) \,, \end{split}$$

53

where $\tilde{Q}_{r_L} \stackrel{\text{def}}{=} \{ x \notin Q_{r_L} \colon \exists y \in Q_{r_L} \text{ s.t. } |x - y| = 1 \}.$ Thus,

$$\begin{aligned} \|\nabla \mathcal{D}_{R_L}(\bar{\varphi}_L)\|_{\ell^2(Q_{R_L}^{\neq 0})}^2 &\lesssim \sum_{x \in \tilde{Q}_{r_L}} \left| \sum_{\substack{y \in Q_{r_L} \\ y \sim x}} \bar{\varphi}_L(y) \right|^2 \\ &+ |\xi_L(0) - a_L|^2 \sum_{x \in Q_{r_L}^{\neq 0}} \bar{\varphi}_L(x)^2 |v_L(x) - 1|^2 + \sum_{x \in Q_{r_L}^{\neq 0}} \bar{\varphi}_L(x)^2 |\zeta_L(x)|^2 \,. \end{aligned}$$

Let us bound the first two terms on the r.h.s. From the exponential decay (4.13) of $\bar{\varphi}_L$ we get

$$\sum_{x\in \tilde{Q}_{r_L}} \Big| \sum_{\substack{y\in Q_{r_L}\\y\sim x}} \bar{\varphi}_L(y) \Big|^2 \lesssim r_L^{d-1} \Big(c_d \frac{d_L}{a_L} \Big)^{2r_L} \le r_L^{d-1} \Big(\frac{1}{2} \Big)^{2r_L} \lesssim \frac{1}{r_L} \le \frac{1}{a_L}$$

for L large enough, where in the last step we used that, by (1.16), $r_L \ge a_L$.

We turn to the second term. Using the exponential decay (4.13) of $\bar{\varphi}_L$, (1.2), and the content of event E_L^1 , we deduce that

$$|\xi(0) - a_L|^2 \sum_{x \in Q_{r_L}^{\neq 0}} \bar{\varphi}_L(x)^2 |v_L(x) - 1|^2 \lesssim \sum_{x \in Q_{r_L}^{\neq 0}} \frac{e^{2c'|x|}}{d_L^2} \Big(\frac{d_L}{a_L}\Big)^{2|x|} \lesssim \frac{1}{d_L^2} \Big(\frac{d_L}{a_L}\Big)^2$$

which is negligible compared to a_L^{-1} . Putting everything together (B.3) follows at once.

In order to prove (B.4), note that $\nabla \bar{\mathcal{D}}_L(\bar{\varphi}_L) = 0$. Hence, similar to (B.6), for any $x \in Q_1^{\neq 0}$ we have

$$0 = \frac{\partial \mathcal{D}_L}{\partial \psi(x)}(\bar{\varphi}_L) = 2\bar{\varphi}_L(0) - 2\frac{\bar{\varphi}_L(x)}{\bar{\varphi}_L(0)} \sum_{\substack{y \in Q_{R_L}^{\neq 0}\\ y \sim 0}} \bar{\varphi}_L(y) + 2\sum_{\substack{y \in Q_{R_L}^{\neq 0}\\ y \sim x}} \bar{\varphi}_L(y) - 2\bar{\varphi}_L(x)\mathcal{S}_L(x),$$

which implies

$$\bar{\varphi}_L(x) = \frac{1}{\mathcal{S}_L(x)} \Big[1 + (\bar{\varphi}_L(0) - 1) - \frac{\bar{\varphi}_L(x)}{\bar{\varphi}_L(0)} \sum_{\substack{y \in Q_{R_L}^{\neq 0} \\ y \sim 0}} \bar{\varphi}_L(y) + \sum_{\substack{y \in Q_{R_L}^{\neq 0} \\ y \sim x}} \bar{\varphi}_L(y) \Big] \,.$$

By (4.14) and (4.13), the last three summands in the parenthesis are $\mathcal{O}(d_L/a_L)$. In addition, (1.2) and (1.3) imply that $1 - v_L(x) \approx 1/d_L$ for all $x \in Q_1^{\neq 0}$. Using (1.6), we thus deduce that $\mathcal{S}_L(x) \approx a_L/d_L$ for any $x \in Q_1^{\neq 0}$, from which (B.4) follows. To prove (B.5), it suffices to compute $\bar{\mathcal{D}}_L(\bar{\varphi}_L)$ using the estimates that we collected for $\bar{\varphi}_L$.

References

- [And58] P. W. Anderson. Absence of diffusion in certain random lattices. <u>Phys. Rev.</u>, 109:1492–1505, Mar 1958. 1
- [Ast03] Arvydas Astrauskas. On high-level exceedances of gaussian fields and the spectrum of random Hamiltonians. Acta Appl. Math., 78:35–42, 2003. 3
- [Ast08] A. Astrauskas. Extremal theory for spectrum of random discrete Schrödinger operator. I. Asymptotic expansion formulas. J. Stat. Phys., 131(5):867–916, 2008. 2
- [Ast13] A. Astrauskas. Extremal theory for spectrum of random discrete Schrödinger operator. III. Localization properties. J. Stat. Phys., 150(5):889–907, 2013. 2, 3

54 GIUSEPPE CANNIZZARO, CYRIL LABBÉ, AND WILLEM VAN ZUIJLEN

- [Ast16] Arvydas Astrauskas. From extreme values of i.i.d. random fields to extreme eigenvalues of finite-volume Anderson Hamiltonian. <u>Probab. Surv.</u>, 13:156–244, 2016. 2
- [AW15] Michael Aizenman and Simone Warzel. <u>Random operators</u>, volume 168 of <u>Graduate</u> <u>Studies in Mathematics</u>. American Mathematical Society, Providence, RI, 2015. Disorder effects on quantum spectra and dynamics. 1
- [BK16] M. Biskup and W. König. Eigenvalue order statistics for random Schrödinger operators with doubly-exponential tails. <u>Comm. Math. Phys.</u>, 341(1):179–218, 2016. 2, 9, 32
- [BL16] Marek Biskup and Oren Louidor. Extreme local extrema of two-dimensional discrete Gaussian free field. Comm. Math. Phys., 345(1):271–304, 2016. 6
- [BL18] Marek Biskup and Oren Louidor. Full extremal process, cluster law and freezing for the two-dimensional discrete Gaussian free field. <u>Adv. Math.</u>, 330:589–687, 2018. 6
- [CL90] René Carmona and Jean Lacroix. Spectral theory of random Schrödinger operators. Probability and its Applications. Birkhäuser Boston, Inc., Boston, MA, 1990. 1
- [FM14] Artiom Fiodorov and Stephen Muirhead. Complete localisation and exponential shape of the parabolic Anderson model with Weibull potential field. <u>Electron. J.</u> Probab., 19:no. 58, 27, 2014. 2
- [GKM00] J. Gärtner, W. König, and S. A. Molchanov. Almost sure asymptotics for the continuous parabolic Anderson model. <u>Probab. Theory Related Fields</u>, 118(4):547–573, 2000. 2, 3
- [GKM07] Jürgen Gärtner, Wolfgang König, and Stanislav Molchanov. Geometric characterization of intermittency in the parabolic Anderson model. <u>Ann. Probab.</u>, 35(2):439–499, 2007. 2
- [GM00] Jürgen Gärtner and Stanislav A. Molchanov. Moment asymptotics and Lifshitz tails for the parabolic Anderson model. In <u>Stochastic models (Ottawa, ON, 1998)</u>, volume 26 of <u>CMS Conf. Proc.</u>, pages 141–157. Amer. Math. Soc., Providence, RI, 2000. 3
- [GY23] Promit Ghosal and Jaeyun Yi. Fractal geometry of the PAM in 2D and 3D with white noise potential. arXiv e-prints, page arXiv:2303.16063, March 2023. 2
- [Kal02] Olav Kallenberg. Foundations of modern probability. Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2002. 23, 25
- [Kir08] Werner Kirsch. An invitation to random Schrödinger operators. In <u>Random</u> <u>Schrödinger operators</u>, volume 25 of <u>Panor. Synthèses</u>, pages 1–119. Soc. Math. France, Paris, 2008. With an appendix by Frédéric Klopp. 1
- [KLMS09] Wolfgang König, Hubert Lacoin, Peter Mörters, and Nadia Sidorova. A two cities theorem for the parabolic Anderson model. Ann. Probab., 37(1):347–392, 2009. 2
- [Kön16] Wolfgang König. <u>The parabolic Anderson model</u>. Pathways in Mathematics. Birkhäuser/Springer, [Cham], 2016. Random walk in random potential. 2
- [KPvZ22] Wolfgang König, Nicolas Perkowski, and Willem van Zuijlen. Longtime asymptotics of the two-dimensional parabolic Anderson model with white-noise potential. <u>Ann.</u> Inst. Henri Poincaré Probab. Stat., 58(3):1351–1384, 2022. 2
- [LLR83] M. R. Leadbetter, Georg Lindgren, and Holger Rootzén. <u>Extremes and related</u> properties of random sequences and processes. Springer Series in Statistics. Springer-Verlag, New York-Berlin, 1983. 4, 26
- [ST14] Nadia Sidorova and Aleksander Twarowski. Localisation and ageing in the parabolic Anderson model with Weibull potential. <u>Ann. Probab.</u>, 42(4):1666–1698, 2014. 2

UNIVERSITY OF WARWICK, MATHEMATICAL SCIENCES BUILDING, COVENTRY CV4 7AL, UK

 $Email \ address: \verb"giuseppe.cannizzaro@warwick.ac.uk"$

UNIVERSITÉ PARIS CITÉ, LABORATOIRE DE PROBABILITÉS, STATISTIQUE ET MODÉLISATION, UMR 8001, F-75205 PARIS, FRANCE AND INSTITUT UNIVERSITAIRE DE FRANCE (IUF). *Email address*: clabbe@lpsm.paris

WIAS BERLIN, MOHRENSTRASSE 39, 10117 BERLIN, GERMANY *Email address:* vanzuijlen@wias-berlin.de