

# THE EFFECTIVE RADIUS OF SELF REPELLING ELASTIC MANIFOLDS

CARL MUELLER AND EYAL NEUMAN

ABSTRACT. We study elastic manifolds with self-repelling terms and estimate their effective radius. This class of manifolds is modelled by a self-repelling vector-valued Gaussian free field with Neumann boundary conditions over the domain  $[-N, N]^d \cap \mathbb{Z}^d$ , that takes values in  $\mathbb{R}^d$ . Our main result states that in two dimensions ( $d = 2$ ), the effective radius  $R_N$  of the manifold is approximately  $N$ . This verifies the conjecture of Kantor, Kardar and Nelson [8] up to a logarithmic correction. Our results in  $d \geq 3$  give a similar lower bound on  $R_N$  and an upper of order  $N^{d/2}$ . This result implies that self-repelling elastic manifolds undergo a substantial stretching at any dimension.

## 1. INTRODUCTION

**1.1. Motivation.** The Gaussian free field (GFF) has become a central object in probability, studied for its own sake and with connections to several areas of physics. In pure probability, GFF plays the role of Brownian motion with multidimensional time. In Euclidean field theory, GFF represents a field of noninteracting particles, with interacting models such as the  $\phi^4$  field theory arising through a change of measure, often requiring renormalization. In fact GFF over  $\mathbb{R}^d$  with  $d \geq 2$  is valued in the space of Schwartz distributions and cannot be realized as a function, so its fourth power  $\phi^4$  is undefined in the usual sense (see Biskup [4]).

In this paper we take inspiration from a different set of physical models, elastic manifolds (see Mezard and Parisi [10] and Balents and Fisher [1]). Here GFF is vector-valued, and represents the position of a random surface. To avoid using Schwartz distributions as mentioned above, we discretize the domain of GFF, and use the notation DGFF (discrete GFF) for the resulting model.

---

2020 *Mathematics Subject Classification.* Primary, 60G60; Secondary, 60G15.

*Key words and phrases.* Gaussian free field, self-avoiding, elastic manifold.

CM was partially supported by Simons Foundation grant 513424.

If the domain of DGFF is one dimensional, then we have a well-known model of a random polymer. In this context, it is common to include a self-repelling term which reflects the fact that different parts of the polymer cannot occupy the same position. A typical object of study is the end-to-end distance of such a polymer, or the closely related concept of effective radius. There is a vast literature on such problems, see Bauerschmidt, Duminil-Copin, Goodman, and Slade [2] and the included citations.

The purpose of this paper is to study elastic manifolds with self-repelling terms, and to estimate the effective radius in the case. Self-repelling elastic manifolds were first introduced by Kantor, Kardar and Nelson in [6] as generalizations of polymer models to higher dimensions, in order to capture the behaviour of sheets of covalently bonded atoms and of polymerized lipid surfaces, among others. See [6, 8, 7, 12] and references therein for additional details. To our knowledge no one has studied these models in the mathematical literature. In this class of models, we define an  $\mathbb{R}^D$ -valued DGFF over  $[-N, N]^d \cap \mathbb{Z}^d$ , and use Neumann boundary conditions since these are most closely tied to a random surface with free boundary conditions. In two dimensions, that is  $d = D = 2$ , we get fairly tight bounds on the effective radius of the manifold  $R_N$ , which is proportional to  $N$  in the upper and lower bounds, up to a logarithmic correction. Our results in higher dimensions are not as sharp, as we derive a lower bound on  $R_N$  that is proportional to  $N$ , but the upper bound is of order  $N^{d/2}$ . This proves however that self-repelling elastic manifolds experience a substantial stretching at any dimension.

We remark that for the case where  $D = d = 1$  we expect  $R_N$  to have the same asymptotic behaviour as a one dimensional self-repelling polymer, that is  $R_N \sim N$ . This result can be derived by the argument of the proof of Theorem 1.1, at least up to a logarithmic constant. We leave the details of the proof to the reader.

The case where  $D < d$  was studied by the authors in a followup paper [11]. It was proved that when the dimension of the domain is  $d = 2$  and the dimension of the range is  $D = 1$ , the effective radius  $R_N$  of the manifold is approximately  $N^{4/3}$ . The results for the case  $d \geq 3$  and  $D < d$  give a lower bound on  $R_N$  of order  $N^{\frac{1}{D}(d - \frac{2(d-D)}{D+2})}$  and an upper bound proportional to  $N^{\frac{d}{2} + \frac{d-D}{D+2}}$ . The results of [11] imply that self-repelling elastic manifolds with a low dimensional range undergo a significantly stronger stretching than in the case  $d = D$ , which is studied in this paper.

The remaining case,  $D > d$ , looks to be much harder. For example, consider the case where the domain of the self-repelling DGFF is  $\{0, \dots, N\}$  and it takes values in  $\mathbb{R}^D$ . For  $D = 2, 3, 4$  the behavior of the effective radius of the self-repelling polymer as  $N \rightarrow \infty$  is still unknown, although we have good information for  $D = 1$  and for  $D > 4$ . See page 400 of Bauerschmidt, Duminil-Copin, Goodman, and Slade [2] and also Bauerschmidt, Slade, and Tomberg, and Wallace [3]. If  $D$  is large enough, then for self-avoiding walks, the lace expansion can be used. For DGFF however, there appears to be no analogue of the lace expansion.

We briefly compare the mathematical results for self-repelling elastic manifolds, which were described above, to conjectures which are available in the physical literature. Most of such conjectures are based on the so called Flory's argument. We refer to Section 7 for the heuristic of derivation of such result for  $D = d = 2$ . In general this heuristic argument suggests that

$$(1.1) \quad R_N \sim N^\nu, \quad \text{with } \nu = \frac{d+2}{D+2}.$$

For the cases where  $D = 1, d = 2$  and  $D = 2, d = 2$  we can confirm (1.1) up to a logarithmic correction by Theorem 1.1 in this paper and by the results in [11]. For the case where  $D = 3$  and  $d = 2$  the physical literature is inconclusive as simulations suggest that  $R_N$  grows linearly in  $N$  in contradiction to (1.1). Not much is known about other cases besides from these heuristic results and some arguably imprecise results using  $\varepsilon$ -expansions (see [8]). We refer to Chapter 10.5.2 of [13] and references therein for additional information about Flory's argument for self-repelling manifolds and for related results on simulations.

**1.2. Setup.** Now we give more precise definitions; the reader can also consult Biskup [4]. Since the standard definition of DGFF involves Dirichlet rather than Neumann boundary conditions, we give details of its construction below. In the following, ordinary letters such as  $x, u$  take values in  $\mathbb{R}$  or  $\mathbb{Z}$ , while boldface letters such as  $\mathbf{x}, \mathbf{u}$  take values in  $\mathbb{R}^d$  for  $d \geq 2$ .

Fix  $d \geq 2, N \geq 1$  and define our parameter set as follows.

$$S_N^d := [-N, N]^d \cap \mathbb{Z}^d.$$

Note that

$$S_N^1 := \{-N, \dots, N\}.$$

Thus  $S_N^d$  is a cube in  $\mathbb{Z}^d$  centered at the origin. Let  $\mathcal{N} = \mathcal{N}_{N,d}$  be the set of unordered nearest neighbor pairs in  $S_N^d$ .

Now we define the discrete Neumann Laplacian  $\Delta = \Delta_{N,d,D}$  as follows. Given functions  $\mathbf{f}, \mathbf{g} : S_N^d \rightarrow \mathbb{R}^D$  we define the energy  $H$  and the inner product  $(\cdot, \cdot) = (\cdot, \cdot)_{N,d,D}$  as follows,

$$\begin{aligned} H(\mathbf{f}, \mathbf{g}) &= \sum_{(\mathbf{x}, \mathbf{y}) \in \mathcal{N}} (\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})) \cdot (\mathbf{g}(\mathbf{x}) - \mathbf{g}(\mathbf{y})), \\ (\mathbf{f}, \mathbf{g}) &= \sum_{\mathbf{x} \in S_N^d} \mathbf{f}(\mathbf{x}) \cdot \mathbf{g}(\mathbf{x}). \end{aligned}$$

We will usually write  $H(\mathbf{f})$  instead of  $H(\mathbf{f}, \mathbf{f})$ . Define the operator  $\Delta = \Delta_{N,d,D}$  on such functions by the requirement

$$H(\mathbf{f}, \mathbf{g}) = -(\mathbf{f}, \Delta \mathbf{g}).$$

If  $D = 1$ , we simply write  $f, g$  instead of  $\mathbf{f}, \mathbf{g}$ .

We claim that  $\Delta$  is a self-adjoint operator defined pointwise as follows. Given  $\mathbf{f} : S_N^d \rightarrow \mathbb{R}^D$ , we first extend the domain of  $\mathbf{f}$  to  $S_{N+1}^d$ . If  $\mathbf{x}, \mathbf{y}$  are nearest neighbors in  $\mathbb{Z}^d$  with  $\mathbf{x} \in S_N^d$  and  $\mathbf{y} \notin S_N^d$ , define  $\mathbf{f}(\mathbf{y}) := \mathbf{f}(\mathbf{x})$ . If  $\mathbf{y} \in S_{N+1}^d \setminus S_N^d$  but  $\mathbf{y}$  is not a nearest neighbor of any point in  $S_N^d$ , let  $\mathbf{f}(\mathbf{y}) := 0$ . We leave it to the reader to use summation by parts to verify that for  $\mathbf{x} \in S_N^d$ ,

$$(1.2) \quad \Delta \mathbf{f}(\mathbf{x}) = \sum_{\mathbf{y} \sim \mathbf{x}} [\mathbf{f}(\mathbf{y}) - 2d \cdot \mathbf{f}(\mathbf{x})],$$

where  $\mathbf{x} \sim \mathbf{y}$  means that  $\mathbf{x}, \mathbf{y}$  are nearest neighbors.

In the case  $D = 1$ , since  $\Delta$  is a self-adjoint operator on a finite-dimensional space, there exists a finite index set  $\mathbb{I} = \mathbb{I}_{N,d}$  to be defined later, and an orthonormal basis of eigenfunctions  $(\varphi_{\mathbf{k}})_{\mathbf{k} \in \mathbb{I}}$  with corresponding eigenvalues  $(\lambda_{\mathbf{k}})_{\mathbf{k} \in \mathbb{I}}$ . We can assume without loss of generality that there is a distinguished index  $\mathbf{0}$  such that  $\varphi_{\mathbf{0}}$  is constant and that  $\lambda_{\mathbf{0}} = 0$ .

Throughout, we fix a parameter  $\beta > 0$ , which has a physical interpretation as the inverse temperature. Let  $(X_{\mathbf{k}}^{(i)})_{\mathbf{k} \in \mathbb{I} \setminus \{\mathbf{0}\}, i=1, \dots, D}$  be a collection of i.i.d. random variables defined on a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  such that

$$X_{\mathbf{k}}^{(i)} \sim N(0, (2\beta\lambda_{\mathbf{k}})^{-1}).$$

For each  $i = 1, \dots, D$  define

$$(1.3) \quad u^{(i)} = \sum_{\mathbf{k} \in \mathbb{I} \setminus \{\mathbf{0}\}} X_{\mathbf{k}}^{(i)} \varphi_{\mathbf{k}},$$

and define DGFF as

$$(1.4) \quad \mathbf{u} = (u^{(1)}, \dots, u^{(D)}).$$

Since the definition of  $\mathbf{u}$  in (1.3) may appear unmotivated, we remark that it is equivalent to stating that  $(\mathbf{u}(\mathbf{x}))_{\mathbf{x} \in S_N^d}$  is a collection of  $\mathbb{R}^D$ -valued centered jointly Gaussian variables with joint density

$$(1.5) \quad \frac{1}{C_{N,\beta,d,D}} \exp(-\beta H(\mathbf{u})) = \frac{1}{C_{N,\beta,d,D}} \exp(\beta(\mathbf{u}, \Delta \mathbf{u}))$$

when  $\mathbf{u}$  is restricted to those  $\mathbf{u}$  for which

$$\sum_{\mathbf{x} \in S_N^d} \mathbf{u}(\mathbf{x}) = 0.$$

Here  $C_{N,\beta,d,D}$  is a normalizing constant which ensures that we have a probability density. We will elaborate on the equivalence of these two definitions in Section 2.2.

We define the local time of the field  $\mathbf{u}$  at level  $\mathbf{z} \in \mathbb{R}^D$  as

$$(1.6) \quad \begin{aligned} \ell_N(\mathbf{z}, \mathbf{u}) &= \#\{\mathbf{x} \in S_N^d : \mathbf{u}(\mathbf{x}) \in [\mathbf{z} - \mathbf{1}/2, \mathbf{z} + \mathbf{1}/2]\} \\ &= \sum_{\mathbf{x} \in S_N^d} \mathbf{1}_{\{\mathbf{u}(\mathbf{x}) \in [\mathbf{z} - \mathbf{1}/2, \mathbf{z} + \mathbf{1}/2]\}}, \end{aligned}$$

where  $\frac{1}{2} = (1/2, \dots, 1/2) \in \mathbb{R}^D$ .

Now we define a weakly self-avoiding Gaussian free field. Throughout, we fix a parameter  $\gamma > 0$ . If  $P_{N,d,D,\beta}$  denotes the original probability measure of  $(\mathbf{u}(\mathbf{x}))_{\mathbf{x} \in S_N^d}$ , we define the probability  $Q_{N,d,D,\beta,\gamma}$  as follows. For ease of notation, we write  $E$  for the expectation with respect to  $P_{N,d,D,\beta}$ . Let

$$(1.7) \quad \begin{aligned} \mathcal{E}_{N,d,D,\gamma} &= \exp\left(-\gamma \int_{\mathbb{R}^D} \ell_N(\mathbf{y}, \mathbf{u})^2 d\mathbf{y}\right), \\ Z_{N,d,D,\beta,\gamma} &= E[\mathcal{E}_{N,d,D,\gamma}] = E^{P_{N,d,D,\beta}}[\mathcal{E}_{N,d,D,\gamma}]. \end{aligned}$$

Then we define for any set  $A \in \mathcal{F}$ ,

$$(1.8) \quad Q_{N,d,D,\beta,\gamma}(A) = \frac{1}{Z_{N,d,D,\beta,\gamma}} E[\mathcal{E}_{N,d,D,\gamma} \mathbf{1}_A].$$

For ease of notation, we will usually drop the subscripts except for  $N$  and write

$$P_N = P_{N,d,D,\beta}, \quad Q_N = Q_{N,d,D,\beta,\gamma}, \quad \mathcal{E}_N = \mathcal{E}_{N,d,D,\gamma}, \quad Z_N = Z_{N,d,D,\beta,\gamma}.$$

Finally, we define the effective radius of the field  $\mathbf{u}$  as

$$R_N = \max_{\mathbf{w}, \mathbf{z} \in S_N^d} \|\mathbf{u}(\mathbf{z}) - \mathbf{u}(\mathbf{w})\|,$$

where  $\|\cdot\|$  denotes the Euclidian norm.

**1.3. Statement of the main result.** Note that in our main theorem below, we assume that  $D = d$ .

**Theorem 1.1.** *Let  $\mathbf{u}$  be the weakly self-avoiding DGFF on  $S_N^d$  taking values in  $\mathbb{R}^d$ . There are constants  $\varepsilon_0, K_0 > 0$  not depending on  $\beta, \gamma, N$  such that*

(i) *For  $d = 2$ ,*

$$\lim_{N \rightarrow \infty} Q_N \left[ \varepsilon_0 \left( \frac{\gamma}{\beta + \gamma} \right)^{1/2} N (\log N)^{-1/2} \leq R_N \leq K_0 \left( \frac{\beta + \gamma}{\beta} \right)^{1/2} N (\log N)^{3/2} \right] = 1.$$

(ii) *For  $d \geq 3$ ,*

$$\lim_{N \rightarrow \infty} Q_N \left[ \varepsilon_0 \left( \frac{\gamma}{\beta + \gamma} \right)^{1/d} N \leq R_N \leq K_0 \left( \frac{\beta + \gamma}{\beta} \right)^{1/2} N^{d/2} \right] = 1.$$

**Remark 1.2.** *Here is the reason we restrict ourselves to the case of  $\mathbf{u}$  taking values in  $\mathbb{R}^D$  with  $D = d$ . As mentioned in the introduction, there is good information about the radius of the self-repelling random polymer taking values in  $\mathbb{R}^D$  for  $D = 1$ , but not for  $D = 2, 3, 4$ . This is because we can guess that for  $D = 1$ , a self-avoiding polymer has ballistic behavior, i.e.  $u(x) \approx Cx$  roughly speaking. In higher dimensions, it is hard to guess what shape the polymer might take. There are results for  $d > 4$  using the lace expansion (see den Hollander [5], Chapter 4), but this method seems hard to adapt to elastic manifolds. However, if  $D = d$ , then we can guess that the self-repelling elastic manifold should stretch itself by dilation in all directions so that  $\mathbf{u}(\mathbf{x}) \approx C\mathbf{x}$ . This guess allows us to carry out the analysis.*

**Remark 1.3.** *Note that when  $\gamma = 0$ , i.e. the self-repelling penalization in (1.8) is set zero, the following statement holds for any  $d \geq 2$ . There exists  $K > 0$  large enough such that,*

$$\lim_{N \rightarrow \infty} P_{N,d,D,\beta} (R_N \leq K \beta^{1/2} \log N) = 1.$$

*This result follows from Proposition 3.1 and by repeating the same steps as in the proof of the upper bound in Theorem 2.1 of [4]. Theorem 1.1 therefore suggests that self-repelling elastic manifolds undergo a substantial stretching at any dimension.*

**Remark 1.4.** *Theorem 1.1 verifies the conjecture by Kantor, Kardar and Nelson in [8] for the case where  $d = 2$  and  $D = 2$ , up to a logarithmic correction. Although in the model that was presented in [8] the DGFF is defined on the triangular lattice, the heuristics that yields*

their result is based on Flory's argument which also applies for the rectangular lattice.

**1.4. Outline of the proof.** We describe the outline for the case  $d = 2$  as the proof for  $d \geq 3$  follows similar lines. Define the following events.

$$(1.9) \quad \begin{aligned} A_N^{(1)} &= \left\{ R_N > K_0 \left( \frac{\beta + \gamma}{\beta} \right)^{1/2} N (\log N)^{3/2} \right\}, \\ A_N^{(2)} &= \left\{ R_N < \varepsilon_0 \left( \frac{\gamma}{\beta + \gamma} \right)^{1/2} N (\log N)^{-1/2} \right\}. \end{aligned}$$

It suffices to show that for  $i = 1, 2$  we have

$$\lim_{N \rightarrow \infty} Q_N(A_N^{(i)}) = 0.$$

From (1.8) we see that it is enough to find:

- (1) a lower bound on  $Z_N$ , derived in Section 2,
- (2) and an upper bound on  $E^{P_N}[\mathcal{E}_N \mathbf{1}_{A_N^{(i)}}]$  for  $i = 1, 2$ , obtained in Sections 5 and 6, respectively.

Finally, the upper bounds divided by the lower bound should vanish as  $N \rightarrow \infty$ .

## 2. LOWER BOUND ON THE PARTITION FUNCTION

In this section we derive the following lower bound on  $Z_N$ .

**Proposition 2.1.** *Let  $\beta > 0$ . Then there exists a constant  $C > 0$  not depending on  $N$ ,  $\beta$  and  $\gamma$  such that*

(i) *for  $d = 2$ ,*

$$\log Z_N \geq -C(\beta + \gamma)N^2 \log N.$$

(ii) *for  $d \geq 3$ ,*

$$\log Z_N \geq -C(\beta + \gamma)N^d.$$

In order to prove Proposition 2.1 we will introduce some additional definitions and auxiliary lemmas.

**2.1. The probability density and associated random walk.** Note that (1.2) implies that  $\Delta_{N,d,1}$  is the generator of a continuous-time simple random walk  $\mathbf{X} = (\mathbf{X}_t)_{t \geq 0}$  which is reflected at the boundary of  $S_N^d$ . More precisely, if  $\mathbf{X}$  attempts to leave  $S_N^d$ , then it stays where it is. Later we will use  $\mathbf{X}$  to obtain the lower bounds in Proposition 2.1.

**2.2. The probability density of the GFF.** Now we study the probability density formula (1.5) in more detail. We focus on  $D = 1$ , and note that  $\Delta = \Delta_{N,d,1}$  is nonpositive definite. As in (1.3), we need to take into account the 0 eigenvalue of  $\Delta$ , which has a constant eigenfunction  $\varphi_{\mathbf{0}}$  with  $\mathbf{0} = (0, \dots, 0) \in S_N^d$ . In order to do that we let  $\mathbf{V}^+ = \mathbf{V}_{N,d}^+$  be the vector space with the standard basis  $\{\mathbf{e}_{\mathbf{x}} : \mathbf{x} \in S_N^d\}$ . We denote vectors in  $\mathbf{V}^+$  as having one component  $v_{\mathbf{x}}$  for each position  $\mathbf{x} \in S_N^d$ , so that  $\mathbf{v} = \sum_{\mathbf{x} \in S_N^d} v_{\mathbf{x}} \mathbf{e}_{\mathbf{x}}$ . Let

$$(2.1) \quad \mathbf{V} = \mathbf{V}_{N,d} = \left\{ \mathbf{v} \in \mathbf{V}^+ : \sum_{\mathbf{x} \in S_N^d} v_{\mathbf{x}} = 0 \right\}.$$

Next, we wish to put a natural measure on  $\mathbf{V}$ . Clearly  $\mathbf{V}$  is a subspace of  $\mathbf{V}^+$ , so we can choose an orthonormal basis  $\{\mathbf{b}_k\}_{k=1}^{(2N+1)^d-1}$  such that each basis element is perpendicular to the constant vector  $\sum_{\mathbf{x} \in S_N^d} \mathbf{e}_{\mathbf{x}}$ . Using such a basis, we can construct Lebesgue measure on  $\mathbf{V}$  in the usual way, to be translation invariant. Also, any such orthonormal basis gives rise to the same measure, which we denote as  $\mu = \mu_{N,d}$ .

Then we can define DGFF on  $\mathbf{V}$  having density with respect to  $\mu$  given by (1.5) with  $D = 1$ . Note that if we were to use  $\mathbf{V}^+$ , then (1.5) (without the normalizing constant) would integrate to  $\infty$ . The extension of (1.5) to  $D \geq 2$  is done by taking the product of the  $D = 1$  densities of  $(u^{(i)})_{i=1}^D$ , which are i.i.d, and then using (1.2).

**2.3. The orthonormal function basis.** We first specify the orthonormal basis  $\{\varphi_{\mathbf{k}}\}$  in (1.3) of eigenfunctions of  $\Delta_{N,d,1}$  on  $[-N, N]^d \cap \mathbb{Z}^d$  taking values in  $\mathbb{R}$ . We note that each basis function  $\varphi_{\mathbf{k}}$  can be represented as a product of  $d$  functions  $\phi_j : \mathbb{R}^d \rightarrow \mathbb{R}$ , as follows

$$(2.2) \quad \varphi_{\mathbf{k}}(\mathbf{x}) = \phi_{k_1}(x_1) \dots \phi_{k_d}(x_d),$$

where  $\mathbf{x} = (x_1, \dots, x_d)$  and  $\mathbf{k} = (k_1, \dots, k_d)$ ,  $-N \leq k_i \leq N$ , and  $1 \leq i \leq d$ . Here  $\{\phi_j\}_{j=-N}^N$  is an orthonormal basis of eigenfunctions of  $\Delta_{N,1,1}$ , the Laplacian with Neumann boundary conditions on  $S_N^1 = \{-N, \dots, N\}$ . Note that if  $\lambda_{\mathbf{k}}$  is the eigenvalue of  $\varphi_{\mathbf{k}}$  and  $\lambda_k$  is the eigenvalue corresponding to  $\phi_k$ , then satisfies

$$(2.3) \quad \lambda_{\mathbf{k}} = \sum_{i=1}^d \lambda_{k_i}.$$

We can explicitly define these eigenfunctions and eigenvalues as follows. Let

$$\phi_0(x) = (2N+1)^{-1/2},$$

and for  $k = 1, \dots, N$  denote

$$(2.4) \quad \phi_k(n) = \frac{1}{a_{k,N}} \sin \left( \frac{(2k-1)\pi}{2N+1} n \right),$$

where  $a_{k,N}$  is a normalizing constant so that

$$\sum_{n=-N}^N \phi_k^2(n) = 1.$$

A calculation (perhaps using Wolfram Alpha) yields

$$a_{k,N} = \left( \frac{1}{2} \csc \left( \frac{(2k-1)\pi}{2N+1} \right) \sin(2k\pi) + N + \frac{1}{2} \right)^{1/2} = \left( N + \frac{1}{2} \right)^{1/2},$$

since  $k$  is an integer.

For  $k = 1, \dots, N$  we further define

$$\phi_{-k}(n) = \frac{1}{b_{k,N}} \cos \left( \frac{2k\pi}{2N+1} n \right),$$

where as in the previous case,  $b_{k,N}$  is a normalizing constant such that

$$\sum_{n=-N}^N \phi_{-k}^2(n) = 1.$$

As before, we can compute

$$b_{k,N} = \left( \frac{1}{2} \csc \left( \frac{2k\pi}{2N+1} \right) \sin(2k\pi) + N + \frac{1}{2} \right)^{1/2} = \left( N + \frac{1}{2} \right)^{1/2},$$

since  $k$  is an integer.

Our basis comprises all such combinations as in (2.2), excluding the constant eigenfunction

$$\varphi_{(0,\dots,0)}(\mathbf{x}) = \phi_0(x_1) \dots \phi_0(x_d).$$

We denote by  $N(d)$  the number of function in our basis,

$$N(d) = |S_N^d| - 1 = (2N+1)^d - 1.$$

**2.4. Incorporating drift.** From (1.5) it follows that  $C_{N,\beta,d,D}$  is given by

$$(2.5) \quad \begin{aligned} C_{N,\beta,d,D} &= \int_{\mathbb{R}^{N(d)}} \exp \left( - \sum_{i=1}^d \sum_{k=1}^{N(d)} \frac{(x_k^{(i)})^2}{2(2\beta\lambda_k)^{-1}} \right) \prod_{i=1}^d \prod_{k=1}^{N(d)} dx_k^{(i)} \\ &= \frac{1}{(2\beta)^{dN(d)/2}} \prod_{k=1}^{N(d)} \frac{1}{\lambda_k^{d/2}}. \end{aligned}$$

Next we incorporate a linear drift into each of the components  $u^{(i)}$  of  $\mathbf{u}$ , calling the resulting component  $u_a^{(i)}$ .

$$u_a^{(i)}(\mathbf{x}) = \sum_{k=1}^{N(d)} X_{\mathbf{k}}^{(i)} \varphi_{\mathbf{k}}(\mathbf{x}) + ax_i, \quad i = 1, \dots, d$$

where  $a > 0$  is a constant to be determined later.

Using (1.3) and (2.3), we get

$$(2.6) \quad u_a^{(i)}(\mathbf{x}) = \sum_{(k_1, \dots, k_d) \in S_N^d \setminus \{\mathbf{0}\}} X_{k_1, \dots, k_d}^{(i)} \phi_{k_1}(x_1) \dots \phi_{k_d}(x_d) + ax_i.$$

where  $\mathbf{0} = (0, \dots, 0) \in \mathbb{Z}^d$ .

Regarding  $x_i$  as a function on  $S_N^1 = \{-N, \dots, N\}$  and expanding it in terms of our eigenfunctions, we see that there are coefficients  $\alpha_j^{(i)}$  such that

$$(2.7) \quad x_i = (\phi_0)^{d-1} \sum_{j \in S_N^1 \setminus \{0\}} \phi_j(x_i) \alpha_j^{(i)},$$

where we have included  $(\phi_0)^{(1-d)}$  for convenience in later calculations. Recall that

$$\phi_0 = \phi_0(x) = (2N+1)^{-1/2}.$$

In (2.7), we do not include  $j = 0$  because  $x_i$  is orthogonal to the constant function  $\phi_0$ .

Next we find  $\alpha_j^{(i)}$  in (2.7). Since  $\alpha_j^{(i)}$  are used to expand the function  $f(x) = x$  for each coordinate  $i$ , we can omit the superscript  $i$  and write just  $\alpha_j$  in what follows. Since  $\{\phi_j\}_{j=-N}^N$  forms an orthonormal basis, we get

$$(2.8) \quad \alpha_j = \phi_0^{(1-d)} \sum_{n=-N}^N n \phi_j(n), \quad j \neq 0, \quad \text{and} \quad \alpha_0 = 0.$$

From (2.6) and (2.7) we have

$$\begin{aligned} u_a^{(i)}(\mathbf{x}) &= \sum_{(k_1, \dots, k_d) \in S_N^d \setminus \{\mathbf{0}\}} X_{k_1, \dots, k_d}^{(i)} \phi_{k_1}(x_1) \dots \phi_{k_d}(x_d) \\ &\quad + a \phi_0^{d-1} \sum_{j \in S_N^1 \setminus \{0\}} \phi_j(x_i) \alpha_j. \end{aligned}$$

We can represent  $u_a^{(i)}$  as follows:

$$(2.9) \quad u_a^{(i)}(x) = \sum_{(k_1, \dots, k_d) \in S_N^d \setminus \{j\mathbf{e}_i: j \in S_N^1\}} X_{k_1, \dots, k_d}^{(i)} \phi_{k_1}(x_1) \cdots \phi_{k_d}(x_d) \\ + \phi_0^{d-1} \sum_{j \in S_N^1 \setminus \{0\}} (X_{j\mathbf{e}_i}^{(i)} + a\alpha_j^{(i)}) \phi_j(x_i).$$

For  $i = 1, \dots, d$  let  $\mathbf{x}^{(i)} \in \mathbf{V}$  and define

$$F(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(d)}) = \sum_{i=1}^d \sum_{(k_1, \dots, k_d) \in S_N^d \setminus \{\mathbf{0}\}} \frac{(x_{k_1, \dots, k_d}^{(i)})^2}{2(2\beta\lambda_{k_1, \dots, k_d})^{-1}}.$$

We rewrite  $Z_N$  in (1.7) as follows. We should emphasize that the local time  $\ell_N$  is random and hence a function of the random variables  $(X_{\mathbf{k}}^{(i)})$ , hence we can write

$$g_N(\mathbf{y}, (X_{\mathbf{k}}^{(i)})) := \ell_N(\mathbf{y}, \mathbf{u})$$

where for readability we have omitted the specification that  $i = 1, \dots, d$  and  $\mathbf{k} \in S_N^d \setminus \{\mathbf{0}\}$ . We get that,

$$(2.10) \quad Z_N = \int_{\mathbb{R}^{N(d)}} \exp \left( -F(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(d)}) - \gamma \int_{\mathbb{R}^d} g_N^2(\mathbf{y}, (\mathbf{x}^{(i)})) d\mathbf{y} \right) \\ \times \prod_{i=1}^d \prod_{(k_1, \dots, k_d) \in S_N^d \setminus \{\mathbf{0}\}} dx_{k_1, \dots, k_d}^{(i)}.$$

In order to find the Radon-Nikodym derivative that allows the drift addition in (2.9) we note that,

$$(2.11) \quad \sum_{i=1}^d \sum_{(k_1, \dots, k_d) \in S_N^d \setminus \{\mathbf{0}\}} \frac{(x_{k_1, \dots, k_d}^{(i)})^2}{2(2\beta\lambda_{k_1, \dots, k_d})^{-1}} \\ = \sum_{i=1}^d \left( \sum_{(k_1, \dots, k_d) \in S_N^d \setminus \{j\mathbf{e}_i: j \in S_N^1\}} \frac{(x_{k_1, \dots, k_d}^{(i)})^2}{2(2\beta\lambda_{k_1, \dots, k_d})^{-1}} \right. \\ \left. + \sum_{j \in S_N^1 \setminus \{0\}} \frac{(x_{j\mathbf{e}_i}^{(i)} + a\alpha_j)^2}{2(2\beta\lambda_{j\mathbf{e}_i})^{-1}} - \sum_{j \in S_N^1 \setminus \{0\}} \frac{2a\alpha_j x_{j\mathbf{e}_i}^{(i)} + (a\alpha_j)^2}{2(2\beta\lambda_{j\mathbf{e}_i})^{-1}} \right).$$

We therefore define  $\hat{P}^{(a)}$  (resp.  $\hat{E}^{(a)}$ ) be the measure (expectation) under which  $u$  is shifted as in (2.9). Then (2.10) and (2.11) imply

$$(2.12) \quad \frac{d\hat{P}^{(a)}}{dP} = \exp \left( - \sum_{i=1}^d \sum_{j \in S_N^1 \setminus \{0\}} \frac{2a\alpha_j x_{j\mathbf{e}_i}^{(i)} + (a\alpha_j)^2}{2(2\beta\lambda_{j\mathbf{e}_i})^{-1}} \right).$$

We can therefore rewrite  $Z_N$  in (2.10) as follows,

$$(2.13) \quad Z_N = \hat{E}^{(a)} \left[ \exp \left( \sum_{i=1}^d \sum_{j \in S_N^1 \setminus \{0\}} \frac{2a\alpha_j X_{j\mathbf{e}_i}^{(i)} + (a\alpha_j)^2}{2(2\beta\lambda_{j\mathbf{e}_i})^{-1}} - \gamma \int_{\mathbb{R}^d} \ell_N^2(\mathbf{y}, \mathbf{u}) d\mathbf{y} \right) \right].$$

We define

$$(2.14) \quad Y_{j\mathbf{e}_i}^{(i)} = \frac{2a\alpha_j X_{j\mathbf{e}_i}^{(i)} + (a\alpha_j)^2}{2(2\beta\lambda_{j\mathbf{e}_i})^{-1}}.$$

Using Jensen's inequality, we get that

$$(2.15) \quad \begin{aligned} \log Z_N &\geq \hat{E}^{(a)} \left[ -\gamma \int_{\mathbb{R}^d} \ell_N^2(\mathbf{y}, \mathbf{u}) d\mathbf{y} \right] - \hat{E}^{(a)} \left[ - \sum_{i=1}^d \sum_{j \in S_N^1 \setminus \{0\}} Y_{j\mathbf{e}_i}^{(i)} \right] \\ &=: -(I_1 + I_2). \end{aligned}$$

The following proposition gives some essential bounds on  $I_i$ ,  $i = 1, 2$ .

**Proposition 2.2.** *Let  $\beta, \gamma > 0$ . Then there exists a constant  $C > 0$  not depending on  $N, \beta, \gamma$  such that*

(i) for  $d = 2$ ,

$$I_1 \leq C\gamma((a^{-2}(\log N)^2) \vee 1)N^2,$$

(ii) for  $d \geq 3$ ,

$$I_1 \leq C\gamma(a^{-2} \vee 1)N^d,$$

(iii) for any  $d \geq 2$ ,

$$I_2 \leq C\beta a^2 N^d.$$

The proof of Proposition (2.2)(i) and (ii) is postponed to Section 3. The proof of Proposition (2.2)(iii) is given in Section 4.

### 2.5. Proof of Proposition 2.1.

*Proof of Proposition 2.1.* From (2.15) and Proposition 2.2(i) and (iii) it follows that for  $d = 2$ ,

$$(2.16) \quad \log \hat{Z}_N \geq -(I_1 + I_2) \geq -C [\gamma((a^{-2}(\log N)^2) \vee 1)N^2 + \beta N^2 a^2].$$

Taking  $a^2 = \log N$  we have

$$\log \hat{Z}_N \geq -C(\beta + \gamma)N^2 \log N.$$

The proof for  $d \geq 3$  follows the same lines with the only modification that we are using Proposition 2.2(ii) and choosing  $a = 1$ .  $\square$

### 3. PROOF OF PROPOSITION (2.2)(I) AND (II)

*Proof of Proposition 2.2(i) and (ii).* We can write

$$\begin{aligned} \tilde{I}_1 &:= \hat{E}^{(a)} \left[ \int_{\mathbb{R}^d} \ell_N(\mathbf{y}, \mathbf{u})^2 d\mathbf{y} \right] \\ &= \hat{E}^{(a)} \left[ \int_{\mathbb{R}^d} \left( \sum_{\mathbf{z} \in S_N^d} \mathbf{1}_{[\mathbf{y}-1/2, \mathbf{y}+1/2]}(\mathbf{u}(\mathbf{z})) \right)^2 d\mathbf{y} \right] \\ (3.1) \quad &= \sum_{\mathbf{z} \in S_N^d} \hat{E}^{(a)} \left[ \int_{\mathbb{R}^d} \mathbf{1}_{[\mathbf{y}-1/2, \mathbf{y}+1/2]}(\mathbf{u}(\mathbf{z})) d\mathbf{y} \right] \\ &\quad + \sum_{\mathbf{z}, \mathbf{w} \in S_N^d, \mathbf{z} \neq \mathbf{w}} \hat{E}^{(a)} \left[ \int_{\mathbb{R}^d} \mathbf{1}_{\mathbf{u}(\mathbf{z}), \mathbf{u}(\mathbf{w}) \in [\mathbf{y}-1/2, \mathbf{y}+1/2]} d\mathbf{y} \right] \\ &= (2N+1)^d + \sum_{\mathbf{z}, \mathbf{w} \in S_N^d, \mathbf{z} \neq \mathbf{w}} \int_{|\mathbf{y}| \leq 1} \hat{p}_{\mathbf{z}, \mathbf{w}}(\mathbf{y}) d\mathbf{y}, \end{aligned}$$

where  $\hat{p}_{\mathbf{z}, \mathbf{w}}$  is the density of  $\mathbf{u}(\mathbf{z}) - \mathbf{u}(\mathbf{w})$  under  $\hat{P}^{(a)}$ .

We will need the following proposition which will be proved in Section 7.

**Proposition 3.1.** *There exist constants  $C_1, C_2 > 0$  such that,*

- (i) *for  $d = 2$ , for all  $\mathbf{w}, \mathbf{z} \in S_N^2$  with  $\mathbf{w} \neq \mathbf{z}$ , and for  $i = 1, 2$  we have*

$$C_1 \beta^{-1} \leq \text{Var}(u^{(i)}(\mathbf{z}) - u^{(i)}(\mathbf{w})) \leq C_2 \beta^{-1} (\log N)^2,$$

- (ii) *for  $d \geq 3$ , for all  $\mathbf{w}, \mathbf{z} \in S_N^d$  with  $\mathbf{w} \neq \mathbf{z}$ , and for  $i = 1, \dots, d$  we have*

$$C_1 \beta^{-1} \leq \text{Var}(u^{(i)}(\mathbf{z}) - u^{(i)}(\mathbf{w})) \leq C_2 \beta^{-1}.$$

Note that from (2.6) we have

$$\hat{E}^{(a)}[u^{(i)}(\mathbf{z}) - u^{(i)}(\mathbf{w})] = a(z_i - w_i), \quad \text{for } i = 1, \dots, d.$$

Since  $(u^{(i)})_{i=1, \dots, d}$  are independent, we have for any  $\mathbf{y} \in \mathbb{R}^d$

$$(3.2) \quad \hat{p}_{\mathbf{z}, \mathbf{w}}(\mathbf{y}) := \prod_{i=1}^d \hat{p}_{\mathbf{z}, \mathbf{w}}^{(i)}(y_i)$$

and therefore

$$(3.3) \quad \int_{\|\mathbf{y}\| \leq 1} \hat{p}_{\mathbf{z}, \mathbf{w}}(\mathbf{y}) d\mathbf{y} \leq \prod_{i=1}^d \int_{-1}^1 \hat{p}_{\mathbf{z}, \mathbf{w}}^{(i)}(y_i) dy_i,$$

where  $\hat{p}_{\mathbf{z}, \mathbf{w}}^{(i)}$  is the density of  $u^{(i)}(\mathbf{z}) - u^{(i)}(\mathbf{w})$  under  $\hat{P}^{(a)}$ .

We distinguish between the following two cases.

**Case 1:**  $d = 2$ . Then from Proposition 3.1(i) we have

$$(3.4) \quad \hat{p}_{\mathbf{z}, \mathbf{w}}^{(i)}(y_i) \leq \frac{1}{\sqrt{2\pi C_1 \beta^{-1}}} \exp\left(-\frac{a^2(y_i - (z_i - w_i))^2}{2\beta^{-1} C_2 (\log N)^2}\right).$$

From (3.3) and (3.4) we therefore get

$$(3.5) \quad \begin{aligned} & \sum_{\mathbf{z}, \mathbf{w} \in S_N^2, \mathbf{z} \neq \mathbf{w}} \int_{\|\mathbf{y}\| \leq 1} \hat{p}_{\mathbf{z}, \mathbf{w}}(\mathbf{y}) d\mathbf{y} \\ & \leq \sum_{\mathbf{z}, \mathbf{w} \in S_N^2, \mathbf{z} \neq \mathbf{w}} \int_{\mathbf{y} \in [-1, 1]^2} \frac{1}{2\pi C_1 \beta^{-1}} \exp\left(-\frac{a^2 \|\mathbf{z} - \mathbf{w} - \mathbf{y}\|^2}{2C_2 \beta^{-1} (\log N)^2}\right) d\mathbf{y} \\ & \leq \int_{\mathbf{y} \in [-1, 1]^2} J(\mathbf{y}) d\mathbf{y}. \end{aligned}$$

where

$$J(\mathbf{y}) := \sum_{\mathbf{z}, \mathbf{w} \in S_N^2} \frac{1}{2\pi C_1 \beta^{-1}} \exp\left(-\frac{a^2 \|\mathbf{z} - \mathbf{w} - \mathbf{y}\|^2}{2C_2 \beta^{-1} (\log N)^2}\right).$$

We will use the following lemma, which will be proved in the end of this section.

**Lemma 3.2.** *Let  $\kappa > 0$ . Then for all  $y \in [-1, 1]$  and  $w \in S_N^1$  we have*

$$\sum_{z \in S_N^1} \exp\left(-\kappa(z - w - y)^2\right) \leq 3 + \int_{-\infty}^{\infty} \exp\left(-\kappa(z - w - y)^2\right) dy.$$

Using Lemma 3.2 and integrating over the Gaussian density gives,

$$\begin{aligned}
J(\mathbf{y}) &= \frac{1}{2\pi C_1 \beta^{-1}} \left( \sum_{w_1 \in S_N^1} \sum_{z_1 \in S_N^1} \exp \left( - \frac{a^2(z_1 - w_1 - y_1)^2}{2C_2 \beta^{-1} (\log N)^2} \right) \right) \\
&\quad \times \left( \sum_{w_2 \in S_N^1} \sum_{z_2 \in S_N^1} \exp \left( - \frac{a^2(z_2 - w_2 - y_2)^2}{2C_2 \beta^{-1} (\log N)^2} \right) \right) \\
&\leq C a^{-2} (\log N)^2 \\
&\quad \times \sum_{w_1 \in S_N^1} \left( 3 + \frac{1}{\sqrt{2\pi C_2 \beta^{-1} a^{-2} (\log N)^2}} \right. \\
(3.6) \quad &\quad \times \int_{-\infty}^{\infty} \exp \left( - \frac{(z_1 - w_1 - y_1)^2}{2C_2 \beta^{-1} a^{-2} (\log N)^2} \right) dz_1 \Bigg) \\
&\quad \times \sum_{w_2 \in S_N^1} \left( 3 + \frac{1}{\sqrt{2\pi C_2 \beta^{-1} a^{-2} (\log N)^2}} \right. \\
&\quad \times \int_{-\infty}^{\infty} \exp \left( - \frac{(z_2 - w_2 - y_2)^2}{2C_2 \beta^{-1} a^{-2} (\log N)^2} \right) dz_2 \Bigg) \\
&\leq C a^{-2} (\log N)^2 \sum_{w_1 \in S_N^1} \sum_{w_2 \in S_N^1} 16 \\
&\leq C a^{-2} N^2 (\log N)^2.
\end{aligned}$$

The number 16 in the next to last line above comes from evaluating the standard Gaussian integrals in the previous two lines.

Substituting (3.6) into (3.5) gives

$$\sum_{\mathbf{z}, \mathbf{w} \in S_N^2, \mathbf{z} \neq \mathbf{w}} \int_{\|\mathbf{y}\| \leq 1} \hat{p}_{\mathbf{z}, \mathbf{w}}(\mathbf{y}) d\mathbf{y} \leq C a^{-2} N^2 (\log N)^2.$$

Together with (3.1) we get that

$$\begin{aligned}
\tilde{I}_1 &\leq (2N + 1)^2 + C a^{-2} N^2 (\log N)^2 \\
(3.7) \quad &\leq \tilde{C} (a^{-2} (\log N)^2 \vee 1) N^2.
\end{aligned}$$

**Case 2:**  $d \geq 3$ . From Proposition 3.1(ii) we have

$$(3.8) \quad \hat{p}_{\mathbf{z}, \mathbf{w}}^{(i)}(y_i) \leq \frac{1}{\sqrt{2\pi C_1 \beta^{-1}}} \exp \left( - \frac{a^2(z_i - w_i - y_i)^2}{2C_2 \beta^{-1}} \right).$$

From (3.3) and (3.8) we therefore get

$$\begin{aligned} & \sum_{\mathbf{z}, \mathbf{w} \in S_N^d, \mathbf{z} \neq \mathbf{w}} \int_{\|\mathbf{y}\| \leq 1} \hat{p}_{\mathbf{z}, \mathbf{w}}(\mathbf{y}) \\ & \leq C \sum_{\mathbf{z}, \mathbf{w} \in S_N^d, \mathbf{z} \neq \mathbf{w}} \int_{\mathbf{y} \in [-1, 1]^d} \frac{1}{(2\pi C_1 \beta^{-1})^{d/2}} \exp\left(-\frac{a^2 \|\mathbf{z} - \mathbf{w} - \mathbf{y}\|^2}{2C_2 \beta^{-1}}\right) d\mathbf{y}. \end{aligned}$$

The rest of the proof is similar to Case 1, with the only difference that we do not have the dependence in  $\log N$  in the Gaussian density. This leads to,

$$\tilde{I}_1 \leq \tilde{C}(a^{-2} \vee 1)N^d.$$

Since from (2.15) we have that

$$I_1 = \gamma \tilde{I}_1,$$

this completes the proof of Proposition 2.2 parts (i) and (ii).  $\square$

*Proof of Lemma 3.2.* Using the fact that  $f(z) := \exp\left(-\kappa(z-w-y)^2\right)$  is monotone decreasing for  $z > w+y$  and since  $y \in [-1, 1]$  we get,

$$(3.9) \quad \sum_{z \geq w+2} \exp\left(-\kappa(z-w-y)^2\right) \leq \int_{w+1}^{\infty} \exp\left(-\kappa(z-w-y)^2\right) dz.$$

Similarly using the fact that  $f(z)$  is monotone increasing for  $z < w+y$  we get

$$(3.10) \quad \sum_{z \leq w-2} \exp\left(-\kappa(z-w-y)^2\right) \leq \int_{-\infty}^{w-1} \exp\left(-\kappa(z-w-y)^2\right) dz.$$

Finally, we have,

$$(3.11) \quad \sum_{z=w-1}^{w+1} \exp\left(-\kappa(z-w-y)^2\right) \leq 3.$$

From (3.9)–(3.11) we get

$$\sum_{z \in S_N^1} \exp\left(-\kappa(z-w-y)^2\right) \leq 3 + \int_{-\infty}^{\infty} \exp\left(-\kappa(z-w-y)^2\right) dz.$$

$\square$

## 4. PROOF OF PROPOSITION (2.2) (III)

*Proof of Proposition (2.2) (iii).* Recall that  $I_2$  was defined in (2.15),

$$(4.1) \quad I_2 = - \sum_{i=1}^d \sum_{\ell \in S_N^1 \setminus \{0\}} \hat{E}^{(a)} \left[ Y_{\ell \mathbf{e}_i}^{(i)} \right],$$

where  $Y_{\ell \mathbf{e}_i}^{(i)}$  was defined in (2.14). We further introduce the following notation:

$$z_{k_1, \dots, k_d}^{(i)} = \frac{(x_{k_1, \dots, k_d}^{(i)})^2}{2(2\beta\lambda_{k_1, \dots, k_d})^{-1}}, \quad w_{j \mathbf{e}_i}^{(i)} = \frac{(x_{j \mathbf{e}_i}^{(i)} + a\alpha_j)^2}{2(2\beta\lambda_{j \mathbf{e}_i})^{-1}},$$

and

$$y_{\ell \mathbf{e}_i}^{(i)} = \frac{2a\alpha_\ell x_{\ell \mathbf{e}_i}^{(i)} + (a\alpha_\ell)^2}{2(2\beta\lambda_{\ell \mathbf{e}_i})^{-1}}.$$

Then from (2.11) and (2.12) we have

$$(4.2) \quad \begin{aligned} & \hat{E}^{(a)} \left[ Y_{\ell \mathbf{e}_i}^{(i)} \right] \\ &= \frac{1}{C_{N, \beta, d, D}} \int y_{\ell \mathbf{e}_i}^{(i)} \exp \left( - \sum_{m=1}^d \left( \sum_{(k_1, \dots, k_d) \in S_N^d \setminus \{j \mathbf{e}_m : j \in S_N^1\}} z_{k_1, \dots, k_d}^{(m)} \right. \right. \\ & \quad \left. \left. + \sum_{l \in S_N^1 \setminus \{0\}} w_{l \mathbf{e}_m}^{(i)} \right) \right) \prod_{r=1}^d \prod_{k=1}^{N(d)} dx_k^{(r)}, \end{aligned}$$

where  $C_{N, \beta, d, D}$  was defined in (2.5).

Since the expected value in (4.2) is symmetric with respect to  $i$ , we can use  $i = 1$  in what follows in order to ease the notation.

$$(4.3) \quad \begin{aligned} & \hat{E}^{(a)} \left[ Y_{\ell \mathbf{e}_1}^{(1)} \right] = \frac{1}{C_{N, \beta, d, D}} \int y_{\ell \mathbf{e}_1}^{(1)} \exp(-w_{\ell \mathbf{e}_1}^{(1)}) \\ & \quad \times \int \exp \left( - \left( \sum_{i=1}^d \sum_{(k_1, \dots, k_d) \in S_N^d \setminus \{j \mathbf{e}_i : j \in S_N^1\}} z_{k_1, \dots, k_d}^{(i)} \right. \right. \\ & \quad \left. \left. + \sum_{i=2}^d \sum_{j \in \{-N, \dots, N\} \setminus \{0\}} w_{j \mathbf{e}_i}^{(i)} + \sum_{j \in S_N^1 \setminus \{0, \ell\}} w_{j \mathbf{e}_1}^{(1)} \right) \right) \prod_{r=1}^d \prod_{k=1}^{N(d)} dx_k^{(r)}. \end{aligned}$$

We notice that we have three types of integrals above, which can be evaluated as follows. We have  $d((2N+1)^d - (2N+1))$  integrals of the

form

$$\begin{aligned} \int_{\mathbb{R}} \exp(-z_{k_1, \dots, k_d}^{(i)}) dx_{k_1, \dots, k_d}^{(i)} &= \int_{\mathbb{R}} \exp\left(-\frac{(x_{k_1, \dots, k_d}^{(i)})^2}{2(2\beta\lambda_{k_1, \dots, k_d})^{-1}}\right) dx_{k_1, \dots, k_d}^{(i)} \\ &= \sqrt{2\pi}(2\beta\lambda_{k_1, \dots, k_d})^{-1/2}. \end{aligned}$$

We have  $2N(d-1) + 2N - 1$  integrals of the form

$$\begin{aligned} \int_{\mathbb{R}} \exp(-w_{j\mathbf{e}_i}^{(i)}) dx_{j\mathbf{e}_i}^{(i)} &= \int_{\mathbb{R}} \exp\left(-\frac{(x_{\mathbf{0}_i(j)}^{(i)} + a\alpha_j)^2}{2(2\beta\lambda_{j\mathbf{e}_i})^{-1}}\right) dx_{j\mathbf{e}_i}^{(i)} \\ &= \sqrt{2\pi}(2\beta\lambda_{j\mathbf{e}_i})^{-1/2}, \end{aligned}$$

and one integral as follows

$$\begin{aligned} &\int y_{\ell\mathbf{e}_1} \exp(-w_{\ell\mathbf{e}_1}^{(1)}) dx_{\ell\mathbf{e}_1}^{(1)} \\ &= \int_{\mathbb{R}} \frac{2a\alpha_\ell x_{j\mathbf{e}_1}^{(1)} + (a\alpha_\ell)^2}{2(2\beta\lambda_{\ell\mathbf{e}_1})^{-1}} \exp\left(-\frac{(x_{\ell\mathbf{e}_1}^{(1)} + a\alpha_\ell)^2}{2(2\beta\lambda_{\ell\mathbf{e}_1})^{-1}}\right) dx_{\ell\mathbf{e}_1}^{(1)} \\ &= -\sqrt{2\pi} \frac{(a\alpha_\ell)^2}{2(2\beta\lambda_{\ell\mathbf{e}_1})^{-1/2}}. \end{aligned}$$

Plugging in all the above integrals into (4.3) gives

$$\begin{aligned} &\hat{E}^{(a)} \left[ Y_{\ell\mathbf{e}_1}^{(1)} \right] \\ &= -\frac{1}{C_{N,\beta,d,D}} \sqrt{2\pi} \frac{(a\alpha_\ell)^2}{2(2\beta\lambda_{\ell\mathbf{e}_1})^{-1/2}} \\ &\quad \times \prod_{i=1}^d \prod_{(k_1, \dots, k_d) \in S_N^d \setminus \{j\mathbf{e}_i: j \in S_N^1\}} \sqrt{2\pi}(2\beta\lambda_{k_1, \dots, k_d})^{-1/2} \\ &\quad \times \prod_{i=2}^d \prod_{j \in S_N^1 \setminus \{0\}} \sqrt{2\pi}(2\beta\lambda_{j\mathbf{e}_i})^{-1/2} \prod_{j \in S_N^1 \setminus \{0, \ell\}} \sqrt{2\pi}(2\beta\lambda_{j\mathbf{e}_1})^{-1/2} \\ &= -\frac{1}{C_{N,\beta,d,D}} \frac{(2\pi)^{((2N+1)^d-1)d/2} (a\alpha_\ell)^2 \lambda_{\ell\mathbf{e}_1}}{2(2\beta)^{((2N+1)^d-1)d/2-1}} \prod_{i=1}^d \prod_{(k_1, \dots, k_d) \in S_N^d \setminus \{\mathbf{0}\}} \frac{1}{\lambda_{k_1, \dots, k_d}^{1/2}}. \end{aligned}$$

Together with (2.5) we get

$$(4.4) \quad \hat{E}^{(a)} \left[ Y_{\ell\mathbf{e}_1}^{(1)} \right] = -\beta(a\alpha_\ell)^2 \lambda_{\ell\mathbf{e}_1}.$$

Plugging (4.4) into (4.1) gives

$$\begin{aligned}
 I_2 &= \beta \sum_{i=1}^d \sum_{\ell \in S_N^1 \setminus \{0\}} (a\alpha_\ell)^2 \lambda_{\ell \mathbf{e}_i} \\
 &= d\beta a^2 \sum_{\ell \in S_N^1 \setminus \{0\}} \alpha_\ell^2 \lambda_{\ell \mathbf{e}_1},
 \end{aligned}
 \tag{4.5}$$

where we used the fact that  $\lambda_{\ell \mathbf{e}_i} = \lambda_{\ell \mathbf{e}_1}$ , by the symmetry of the eigenvalues (see (2.3)).

In order to complete the proof we introduce the following lemma, which will be proved at the end of this section.

**Lemma 4.1.** *There exists a constant  $C > 0$  not depending on  $N$  and  $\beta$  such that,*

$$\sum_{j \in S_N^1 \setminus \{0\}} \alpha_j^2 \lambda_{j \mathbf{e}_1} \leq CN^d.$$

From Lemma 4.1 and (4.5) we conclude that

$$I_2 \leq C\beta a^2 N^2, \tag{4.6}$$

which completes the proof of Proposition 2.2 part (iii).  $\square$

*Proof of Lemma 4.1.* From (2.8) we have

$$n = \phi_0^{d-1} \sum_{j \in S_N^1 \setminus \{0\}} \phi_j(n) \alpha_j,$$

where we recall that

$$\phi_0 = \phi_0(x) = (2N+1)^{-1/2}. \tag{4.7}$$

Since  $\{\phi_j\}_{-N}^N$  are orthonormal we get

$$\alpha_j = \sum_{n=-N}^N \frac{n}{\phi_0^{d-1}} \phi_j(n), \quad j \neq 0, \quad \text{and} \quad \alpha_0 = 0.$$

We recall the eigenfunctions from Section 2.3. For the cosine eigenfunctions we have

$$\sum_{n=-N}^N n \cos\left(\frac{2k\pi}{2N+1}n\right) = 0,$$

and therefore for all  $k = 1, \dots, N$ ,

$$\alpha_{-k} = 0. \tag{4.8}$$

So together with (2.4) we have

$$\begin{aligned}
\alpha_k &= \sum_{n=-N}^N \frac{n}{\phi_0^{d-1}} \phi_k(n) \\
&= (2N+1)^{(d-1)/2} \sum_{n=-N}^N n \phi_k(n) \\
&= (2N+1)^{(d-1)/2} \frac{1}{(N+1/2)^{1/2}} \sum_{n=-N}^N n \sin\left(\frac{(2k-1)\pi}{2N+1} n\right) \\
&= \sqrt{2}(2N+1)^{(d-2)/2} \sum_{n=-N}^N n \sin\left(\frac{(2k-1)\pi}{2N+1} n\right).
\end{aligned}$$

Let

$$Y := \frac{(2k-1)\pi}{2N+1}.$$

Then we have,

$$\begin{aligned}
\frac{\alpha_k}{\sqrt{2}(2N+1)^{(d-2)/2}} &= \sum_{n=-N}^N n \sin(Yn) \\
&= \frac{1}{2} \csc^2\left(\frac{Y}{2}\right) \left[ (N+1) \sin(NY) - N \sin((N+1)Y) \right] \\
&= \frac{(N+1) \sin(NY) - N \sin((N+1)Y)}{2 \sin^2\left(\frac{Y}{2}\right)} \\
&= \frac{\sin(NY) + 2N \sin\left(-\frac{Y}{2}\right) \cos\left((N+\frac{1}{2})Y\right)}{2 \sin^2\left(\frac{Y}{2}\right)}.
\end{aligned}$$

Note that

$$\left(N + \frac{1}{2}\right) Y = \frac{1}{2}(2N+1)Y = \left(k - \frac{1}{2}\right) \pi,$$

and so for any integer  $k$ ,

$$\cos\left(\left(N + \frac{1}{2}\right) Y\right) = \cos\left(\left(k - \frac{1}{2}\right) \pi\right) = 0.$$

Therefore

$$\frac{\alpha_k}{\sqrt{2}(2N+1)^{(d-2)/2}} = \frac{\sin(NY)}{2 \sin^2\left(\frac{Y}{2}\right)}.$$

Note that since  $k \in \{1, 2, \dots, N\}$  we have

$$0 < \frac{Y}{2} \leq \frac{\pi}{2}$$

and so

$$\left| \sin \left( \frac{Y}{2} \right) \right| \leq \frac{Y}{2} \leq C \frac{k}{N}.$$

It follows that,

$$\begin{aligned} |\alpha_k| &\leq C(2N+1)^{(d-2)/2} \left| \frac{\sin(NY)}{\sin^2 \left( \frac{Y}{2} \right)} \right| \\ (4.9) \quad &\leq CN^{(d-2)/2} \left| \frac{1}{\sin^2 \left( \frac{Y}{2} \right)} \right| \\ &\leq CN^{(d+2)/2} \left( \frac{1}{k} \right)^2. \end{aligned}$$

Recall that  $\lambda_k$  is eigenvalue of  $\phi_k$ . Since  $\lambda_0 = 0$ , from (2.3) and (2.4) we have

$$(4.10) \quad \lambda_{j\mathbf{e}_1} = \lambda_j = \left( \frac{(2k-1)\pi}{2N+1} \right)^2.$$

Using (4.8), (4.9) and (4.10) we get

$$\begin{aligned} \sum_{j \in S_N^1 \setminus \{0\}} \alpha_j^2 \lambda_{j\mathbf{e}_1} &\leq C \sum_{j=1}^N N^{d+2} \left( \frac{1}{j} \right)^4 \left( \frac{(2j-1)\pi}{2N+1} \right)^2 \\ &\leq CN^d \sum_{j=1}^N \frac{1}{j^2} \\ &\leq CN^d. \end{aligned}$$

□

## 5. LARGE DISTANCE TAIL ESTIMATES

We define

$$\bar{u}^{(i)} = \max_{\mathbf{w}, \mathbf{z} \in S_N} |u^{(i)}(\mathbf{z}) - u^{(i)}(\mathbf{w})|.$$

Assume first that  $d = 2$ . Let  $\alpha > 0$ , then from (1.8) we have

$$(5.1) \quad \log Q_N(R_N > \alpha N(\log N)^{3/2}) \leq \log P(R_N > \alpha N(\log N)^{3/2}) - \log Z_N.$$

Note that

$$(5.2) \quad \begin{aligned} P(R_N > \alpha N(\log N)^{3/2}) &\leq 2P\left(\cup_{i=1}^2 \{\bar{u}^{(i)} > \alpha N(\log N)^{3/2}/2\}\right) \\ &\leq 4P(\bar{u}^{(1)} > \alpha N(\log N)^{3/2}/2), \end{aligned}$$

where we used the fact that  $\bar{u}^{(i)}$  are i.i.d..

We will use the following standard bound on the tail distribution of the Gaussian random variable  $W \sim N(0, \sigma^2)$ ,

$$P(W > a) \leq \frac{\sigma}{\sigma + a} \exp\left(\frac{-a^2}{2\sigma^2}\right).$$

Using this bound and Proposition 3.1(i) we have for any  $K \geq 1$ ,

$$\begin{aligned} P(|u^{(i)}(\mathbf{z}) - u^{(i)}(\mathbf{w})| > \beta^{-1/2} K N (\log N)^{3/2}) \\ \leq C \exp\left\{-\frac{K^2 N^2 (\log N)^3}{2c_1 (\log N)^2}\right\} \\ \leq C \exp\{-\tilde{c}_1 K^2 N^2 \log N\}. \end{aligned}$$

Note that this bound is uniform in  $\mathbf{z}, \mathbf{w} \in S_N^2$ .

It follows that

$$\begin{aligned} P(\bar{u}^{(i)} > \beta^{-1/2} K N (\log N)^{3/2}/2) \\ \leq \sum_{\mathbf{w}, \mathbf{z} \in S_N^2} P(|u^{(1)}(\mathbf{z}) - u^{(1)}(\mathbf{w})| > \beta^{-1/2} K N (\log N)^{3/2}/2) \\ \leq C(2N+1)^4 \exp\left\{-\frac{\tilde{c}_1}{4} K^2 N^2 \log N\right\} \\ \leq C \exp\left\{-\frac{\tilde{c}_1}{8} K^2 N^2 \log N\right\}. \end{aligned}$$

Together with (5.2) we have

$$P(R_N > \alpha \beta^{-1/2} N (\log N)^{3/2}) \leq C \exp\left\{-\frac{\tilde{c}_1 \alpha^2 N^2 \log N}{8}\right\}.$$

Using this bound together with Proposition 2.1(i) and (5.1) we get for  $d = 2$ , and  $\alpha \geq 1$ ,

$$\begin{aligned} \log Q_N(R_N > \alpha \beta^{-1/2} (\beta + \gamma)^{1/2} N (\log N)^{3/2}) \\ \leq \log P(R_N > \alpha \beta^{-1/2} (\beta + \gamma)^{1/2} N \log N) - \log Z_N \\ \leq -(\beta + \gamma) N^2 \log N (c_3 \alpha^2 - c_4). \end{aligned}$$

We then can choose  $\alpha$  to be large enough to get the large distance tail estimate in Theorem 1.1. The proof for  $d \geq 3$  follows similar lines, only now we use Proposition 2.1(ii) and Proposition 3.1(ii).

## 6. SMALL DISTANCE TAIL ESTIMATES

Let  $\varepsilon > 0$ , then from (1.8) we have the following:

$$(6.1) \quad \log Q_N(R_N < \varepsilon N) \leq \log E \left[ \exp \left\{ -\gamma \int_{\mathbb{R}^d} \ell_N(\mathbf{y}, \mathbf{u})^2 d\mathbf{y} \right\} \mathbf{1}_{\{R < \varepsilon N\}} \right] - \log Z_N.$$

Let

$$(6.2) \quad \tilde{J} := E \left[ \exp \left\{ -\gamma \int_{\mathbb{R}^d} \ell_N(\mathbf{y}, \mathbf{u})^2 d\mathbf{y} \mathbf{1}_{\{R < \varepsilon N\}} \right\} \right].$$

Note that on  $\{R < \varepsilon N\}$  we have

$$(6.3) \quad \begin{aligned} \int_{\mathbb{R}^d} \ell_N(\mathbf{y}, \mathbf{u})^2 d\mathbf{y} &= 2^d N^d \varepsilon^d \int_{-\varepsilon N}^{\varepsilon N} \cdots \int_{-\varepsilon N}^{\varepsilon N} \ell_N(\mathbf{y}, \mathbf{u})^2 \frac{1}{2^d N^d \varepsilon^d} d\mathbf{y} \\ &\geq 2^d N^d \varepsilon^d \left( \int_{-\varepsilon N}^{\varepsilon N} \cdots \int_{-\varepsilon N}^{\varepsilon N} \ell_N(\mathbf{y}, \mathbf{u}) \frac{1}{2^d N^d \varepsilon^d} d\mathbf{y} \right)^2 \\ &= \frac{1}{2^d N^d \varepsilon^d} \left( \int_{-\varepsilon N}^{\varepsilon N} \cdots \int_{-\varepsilon N}^{\varepsilon N} \ell_N(\mathbf{y}, \mathbf{u}) d\mathbf{y} \right)^2, \end{aligned}$$

where we used Jensen's inequality. Since on  $\{R_N < \varepsilon N\}$  we have

$$\int_{-\varepsilon N}^{\varepsilon N} \cdots \int_{-\varepsilon N}^{\varepsilon N} \ell_N(\mathbf{y}, \mathbf{u}) d\mathbf{y} = |S_N| = (2N + 1)^d,$$

and together with (6.3) we get that

$$(6.4) \quad \int_{\mathbb{R}^d} \ell_N(\mathbf{y}, \mathbf{u})^2 d\mathbf{y} \geq \frac{2^d N^d}{\varepsilon^d}.$$

From (6.2) and (6.4) we have

$$(6.5) \quad \tilde{J} \leq e^{-\gamma \frac{2^d N^d}{\varepsilon^d}}.$$

Together with (6.1), (6.2), (6.5) with  $\varepsilon(\log N)^{-1/2}$  instead of  $\varepsilon$ , and Proposition 2.1(i) we get for  $d = 2$ ,

$$\begin{aligned} \log Q_N(R_N < \varepsilon \gamma^{1/2} (\beta + \gamma)^{-1/2} N (\log N)^{-1/2}) \\ \leq -(\beta + \gamma) \left( \frac{4N^2 \log N}{\varepsilon^2} - CN^2 \log N \right). \end{aligned}$$

By choosing  $\varepsilon > 0$  small enough it follows that

$$\lim_{N \rightarrow \infty} \log Q_N(R_N < \varepsilon \gamma^{1/2} (\beta + \gamma)^{-1/2} N (\log N)^{-1/2}) = 0.$$

Repeating the same steps as in the case where  $d = 2$  gives the following result for  $d \geq 3$ ,

$$\log Q_N (R_N < \varepsilon \gamma^{1/d} (\beta + \gamma)^{-1/d} N) \leq -(\beta + \gamma) \left( \frac{N^d}{\varepsilon^2} - CN^d \right).$$

Then choosing  $\varepsilon > 0$  sufficiently small and taking the limit where  $N \rightarrow \infty$  completes the proof of Theorem 1.1.

## 7. PROOF OF PROPOSITION 3.1

*Proof of Proposition 3.1.* We prove the result for the case where  $\beta = 1$ . The extension for any  $\beta > 0$  follows from (1.3) by scaling. We start with the proof of the lower bound. We want to show that for  $d \geq 2$ , there exists a constant  $C = C(d)$  such that for all  $\mathbf{w}, \mathbf{z} \in S_N^d$ ,  $\mathbf{w} \neq \mathbf{z}$ ,  $i = 1, \dots, d$  we have

$$C(d) \leq \text{Var}(u^{(i)}(\mathbf{z}) - u^{(i)}(\mathbf{w})).$$

Since  $(u^{(i)})_{i=1, \dots, d}$  are i.i.d., we will omit the superscript  $i$  for ease of notation. Let  $\mathcal{F}_{\mathbf{z}}$  be the  $\sigma$ -field generated by  $\{u(\mathbf{v}) : \mathbf{v} \in S_N^d \setminus \{\mathbf{z}\}\}$  and define

$$\hat{u}(\mathbf{z}) = E[u(\mathbf{z}) | \mathcal{F}_{\mathbf{z}}].$$

By using the conditional expectation projection theorem and then conditioning on  $\mathcal{F}_{\mathbf{z}}$  we get

$$\begin{aligned} \text{Var}(u(\mathbf{z}) - u(\mathbf{w})) &= E[(u(\mathbf{z}) - u(\mathbf{w}))^2] \\ &\geq E[(u(\mathbf{z}) - \hat{u}(\mathbf{z}))^2] \\ &= E[E[(u(\mathbf{z}) - \hat{u}(\mathbf{z}))^2 | \mathcal{F}_{\mathbf{z}}]]. \end{aligned}$$

Hence it is enough to show that there exists a constant  $C(d) > 0$  not depending on  $\mathbf{z}$  such that

$$(7.1) \quad \text{Var}[u(\mathbf{z}) | \mathcal{F}_{\mathbf{z}}] = E[(u(\mathbf{z}) - \hat{u}(\mathbf{z}))^2 | \mathcal{F}_{\mathbf{z}}] \geq C(d).$$

We consider the nearest neighbor values of  $u(\mathbf{z})$ , which we denote by  $\{u(\mathbf{y}) : \mathbf{y} \sim \mathbf{z}\}$ . These values are fixed once we condition on  $\mathcal{F}_{\mathbf{z}}$ . We further denote by  $\mathcal{N}(\mathbf{z})$  the number neighboring sites of  $\mathbf{z}$ . Note that for any  $d \geq 2$  we have  $d \leq \mathcal{N}(\mathbf{z}) \leq 2d$ . Then the part of the exponent of (1.5) which is relevant to (7.1) is

$$\sum_{\mathbf{y} \sim \mathbf{z}} (u(\mathbf{z}) - u(\mathbf{y}))^2 = \mathcal{N}(\mathbf{z}) \cdot u(\mathbf{z})^2 + 2u(\mathbf{z}) \sum_{\mathbf{y} \sim \mathbf{z}} u(\mathbf{y}) + C,$$

where  $C$  is  $\mathcal{F}_{\mathbf{z}}$ -measurable. By completing the squares we get

$$(7.2) \quad \sum_{\mathbf{y} \sim \mathbf{z}} (u(\mathbf{z}) - u(\mathbf{y}))^2 = \mathcal{N}(\mathbf{z}) \left( u(\mathbf{z}) - \frac{1}{\mathcal{N}(\mathbf{z})} \sum_{\mathbf{y} \sim \mathbf{z}} u(\mathbf{y}) \right)^2 + C' \\ = \mathcal{N}(\mathbf{z}) (u(\mathbf{z}) - C'')^2 + C',$$

where again  $C'$  and  $C''$  are  $\mathcal{F}_{\mathbf{z}}$ -measurable. It follows from the equality in (7.1) and from (7.2) that conditioned on  $\mathcal{F}_{\mathbf{z}}$ ,  $u(\mathbf{z})$  is a Gaussian random variable with variance  $(2\mathcal{N}(\mathbf{z}))^{-1} > (4d)^{-1}$  and therefore (7.1) follows.

Next we prove the upper bound. Recall that  $\mathbf{V}$  was defined in (2.1). Let  $\mathbf{e}_{\mathbf{x}} - \mathbf{e}_{\mathbf{y}} \in \mathbf{V}$ . Recall that  $u$  has density function given by (1.5). Let  $\mathbf{X} = \{\mathbf{X}_t\}_{t \geq 0}$  be the continuous time Markov chain associated with  $\Delta$  and  $\{\mathcal{P}_t\}_{t \geq 0}$  the corresponding probability transition function.

It follows that the variance of  $u(\mathbf{x}) - u(\mathbf{y})$  is given in terms of  $\Delta^{-1}$  as follows:

$$(7.3) \quad \text{Var}(u(\mathbf{x}) - u(\mathbf{y})) = \langle (\mathbf{e}_{\mathbf{x}} - \mathbf{e}_{\mathbf{y}}), \Delta^{-1}(\mathbf{e}_{\mathbf{x}} - \mathbf{e}_{\mathbf{y}}) \rangle \\ = \int_0^\infty \langle (\mathbf{e}_{\mathbf{x}} - \mathbf{e}_{\mathbf{y}}), \mathcal{P}_t(\mathbf{e}_{\mathbf{x}} - \mathbf{e}_{\mathbf{y}}) \rangle dt.$$

**7.1. Estimation of the integrand in (7.3) for small  $t$ .** We consider the case where  $t \leq K_0 N^2 \log N$  for some constant  $K_0 > 0$  to be determined. Note that we can extend  $\mathcal{P}_t$  to a semigroup on all of  $\mathbf{V}^+$ , corresponding to the same random process  $\mathbf{X}_t$ . We write  $P_{\mathbf{x}}$  to indicate the starting point  $\mathbf{X}_0 = \mathbf{x}$ .

Next, we note that the components  $X^{(k)}$ ,  $k = 1, \dots, d$  of  $\mathbf{X}$  are independent, since their jump times are independent Poisson processes which determine the jumps of  $\mathbf{X}$ . Let  $\mathbf{x} = (x_1, \dots, x_d) \in S_N^d$ , then

$$(7.4) \quad \langle \mathbf{e}_{\mathbf{x}}, \mathcal{P}_t \mathbf{e}_{\mathbf{x}} \rangle = P_{\mathbf{x}}[\mathbf{X}_t = \mathbf{x}] = \prod_{k=1}^d P_{x_k}(X_t^{(k)} = x_k).$$

In what follows we focus on the marginal distribution of  $X^{(k)}$ . We write  $z = x_k$ , and  $Z_t = X_t^{(k)}$  and get

$$(7.5) \quad P_z(Z_t = z) = \sum_{n=0}^{\infty} P(T_t = n) P_z(S_n = z),$$

where  $T_t$  Poisson process with intensity  $1/d$  and  $S_n$  is a discrete-time nearest-neighbor one-dimensional simple random walk with reflection at the boundary  $\pm N$ .

Let  $Y$  be a Poisson random variable with mean  $\lambda$  then using Markov inequality we get for all  $\theta > 0$  and  $y > \lambda$ ,

$$(7.6) \quad \begin{aligned} P(Y > y) &\leq \frac{E[e^{\theta Y}]}{e^{\theta y}} \\ &\leq \frac{(e\lambda)^y e^{-\lambda}}{y^y}, \end{aligned}$$

where we choose  $\theta = \log(y/\lambda) > 0$  in the second inequality.

Let  $K_1 > eK_0$ . Recall that  $t \leq K_0 N^2 \log N$  and that  $T_t$  is a Poisson random variable with mean  $(2d)^{-1}t$ . From (7.6) we get the following bound on the tail distribution of  $T_t$ ,

$$(7.7) \quad \begin{aligned} P(T_t > K_1 N^2 \log N) &\leq \frac{e^{-(2d)^{-1}t} ((2d)^{-1}te)^{K_1 N^2 \log N}}{(K_1 N^2 \log N)^{K_1 N^2 \log N}} \\ &\leq e^{-cN^2 \log N}, \end{aligned}$$

for some constant  $c > 0$ .

Using the reflection principle we note that

$$(7.8) \quad P_z(S_n = z) = \sum_{k \in \mathbb{Z}} P_z(W_n = 2Nk + (-1)^k z),$$

where  $\{W_n\}_{n \geq 1}$  is a discrete-time simple random walk on  $\mathbb{Z}$ .

Since  $P_z(W_n = z) = 0$  if  $n$  is odd, we need to take into account only even number of steps, so we have

$$P_z(W_{2m} = z) = \binom{2m}{m} 2^{-2m}.$$

Note that the transition probability from  $z$  to all other points  $2Nk + (-1)^k z$  in the right-hand side of (7.8) can be computed according to the same binomial distribution, since these transitions also require an even number of steps. Moreover the maximum of the above binomial distribution (i.e.  $\text{Bin}(2m, 1/2)$ ) is attained at  $m$  so we must have

$$(7.9) \quad P_z(W_{2m} = 2Nk + (-1)^k z) \leq P_z(W_{2m} = z), \quad \text{for all } k \in \mathbb{Z}.$$

Let  $\ell(N)$  be the number of points from the set  $\{2Nk + (-1)^k z : z \in \mathbb{Z}\}$  visited by  $W_n$  up to  $n = \lfloor K_1 N^2 \log N \rfloor$ . We will use a special case of Corollary A.2.7 in [9] which states that there exist constants  $C_1, c_2 > 0$  such that

$$P\left(\max_{i=0, \dots, n} |W_n| > s\sqrt{n}\right) \leq C_1 e^{-c_2 s^2}, \quad \text{for all } n \geq 0, s > 0,$$

in order to bound the tail probability of  $\ell(N)$  as follows. Let  $K_3 > 0$  then we have

$$\begin{aligned}
 (7.10) \quad P(\ell(N) > K_3 \log N) &\leq P\left(\sup_{0 \leq i \leq K_1 N^2 \log N} |W_i| > \frac{K_3}{4} N \log N\right) \\
 &\leq C_1 e^{-c_2(K_3^2/(16K_2)) \log N} \\
 &\leq C_1 N^{-c_2 K_3^2/(16K_2)}.
 \end{aligned}$$

We would like to bound (7.5) on the event  $T_t \leq K_1 N^2 \log N$ . From (7.8) we have

$$\begin{aligned}
 (7.11) \quad &\sum_{m=0}^{K_1 N^2 \log N} P(T_t = 2m) P_z(S_{2m} = z) \\
 &= \sum_{m=0}^{K_1 N^2 \log N} P(T_t = 2m) \sum_{k \in \mathbb{Z}} P_z(W_{2m} = 2Nk + (-1)^k z) \\
 &= \sum_{m=0}^{K_1 N^2 \log N} P(T_t = 2m) \sum_{|k| \leq 2^{-1} K_3 \log N} P_z(W_{2m} = 2Nk + (-1)^k z) \\
 &\quad + \sum_{m=0}^{K_1 N^2 \log N} P(T_t = 2m) \sum_{|k| > 2^{-1} K_3 \log N} P_z(W_{2m} = 2Nk + (-1)^k z) \\
 &:= J_1 + J_2.
 \end{aligned}$$

From (7.10) it follows that  $J_2$  is bounded by

$$\begin{aligned}
 (7.12) \quad J_2 &\leq \sum_{m=0}^{K_1 N^2 \log N} P(T_t = 2m) P(\ell(N) > K_3 \log N) \\
 &\leq C_1 N^{-c_2 K_3^2/(16K_2)} \sum_{m=0}^{\infty} P(T_t = 2m) \\
 &\leq C N^{-r}.
 \end{aligned}$$

for  $r > 0$  to be determined. Note that the last inequality follows by choosing  $K_3$  large enough.

Using (7.9) and we get for  $J_1$  that

$$(7.13) \quad J_1 \leq K_3 \log N \sum_{m=0}^{K_1 N^2 \log N} P(T_t = 2m) P_z(W_{2m} = z).$$

We estimate the above sum as follows,

$$\begin{aligned}
 (7.14) \quad \sum_{m=0}^{\infty} P(T_t = 2m) P_z(W_{2m} = z) &= \sum_{m=0}^{\infty} \frac{(t/d)^{2m}}{(2m)!} e^{-t/d} \cdot \binom{2m}{m} 2^{-2m} \\
 &= e^{-t/d} \sum_{m=0}^{\infty} \frac{(t/(2d))^{2m}}{(m!)^2}.
 \end{aligned}$$

Recall that

$$I_0(2z) = \sum_{m=0}^{\infty} \frac{z^{2m}}{(m!)^2},$$

where  $I_0$  is a modified Bessel function of the first kind, so for large values of  $y$ ,  $I_0$  has the following asymptotics,

$$I_0(y) \sim \frac{e^y}{\sqrt{2\pi y}}$$

and so by taking  $y = 2z = 2t/(2d) = t/d$ , it follows that there exists a constant  $C > 0$  not depending on  $z$  such that

$$\begin{aligned}
 (7.15) \quad \sum_{m=0}^{\infty} P(T_t = 2m) P_z(W_{2m} = z) &\leq C e^{-t/d} \frac{e^{t/d}}{\sqrt{2\pi t/d}} \\
 &\leq C t^{-1/2}, \quad \text{for all } t \geq 1.
 \end{aligned}$$

From (7.11)–(7.13) and (7.15) it follows that

$$\sum_{m=0}^{K_1 N^2 \log N} P(T_t = 2m) P_z(S_{2m} = z) \leq C \log N t^{-1/2} + N^{-r}.$$

Using this bound together with (7.5) and (7.7) we have

$$\begin{aligned}
 (7.16) \quad P_z(Z_t = z) &\leq C_1 \log N t^{-1/2} + C_2 N^{-r} + e^{-cN^2 \log N} \\
 &\leq C(\log N t^{-1/2} + N^{-r}).
 \end{aligned}$$

From (7.4) and (7.16) we get

$$\begin{aligned}
 (7.17) \quad &| \langle (\mathbf{e}_x - \mathbf{e}_y), \mathcal{P}_t(\mathbf{e}_x - \mathbf{e}_y) \rangle | \\
 &\leq 2 \langle \mathbf{e}_x, \mathcal{P}_t \mathbf{e}_x \rangle + 2 \langle \mathbf{e}_y, \mathcal{P}_t \mathbf{e}_y \rangle \\
 &\leq C(d)(\log N t^{-d/2} + N^{-rd}), \quad \text{for all } 0 \leq t \leq K_0 N^2 \log N.
 \end{aligned}$$

From (7.17) and by choosing  $r$  sufficiently large we finally get,

$$\begin{aligned}
 (7.18) \quad &\int_1^{K_0 N^2 \log N} | \langle (\mathbf{e}_x - \mathbf{e}_y), \mathcal{P}_t(\mathbf{e}_x - \mathbf{e}_y) \rangle | dt \leq \begin{cases} C(\log N)^2 & \text{if } d = 2, \\ C & \text{if } d \geq 3. \end{cases}
 \end{aligned}$$

**7.2. Estimation of the integrand in (7.3) for large  $t$ .** Here we consider the case where  $t \geq KN^2 \log N$  with  $K$  sufficiently large. We again consider the one-dimensional projection of the continuous time random walk  $Z$  reflected at  $\pm N$ , as in (7.5). We expect that in this case there is a large probability of  $Z$  reaching the boundary on or before time  $t$ , hence we expect the semigroup  $\mathcal{P}_t$  to even out any given function. Thus we keep both  $\mathbf{e}_x, \mathbf{e}_y$  in (7.3), use the fact that  $\langle \mathbf{e}_x - \mathbf{e}_y, \mathcal{P}_t(\mathbf{e}_x - \mathbf{e}_y) \rangle$  mostly cancels out. We therefore show that  $P_x(Z_t = x) - P_y(Z_t = x)$  is small for large  $t$  uniformly in  $x, y$ .

We will use the coupling method in order to bound the difference in the probabilities above. To this end, we construct two i.i.d copies  $Z^{(1)}, Z^{(2)}$  of the process  $Z$  with  $Z_0^{(1)} = x$  and  $Z_0^{(2)} = y$  on the same probability space. We seek a (random) coupling time  $\tau$  such that if  $t > \tau$  then  $Z_t^{(1)} = Z_t^{(2)}$ . In that case we would have

$$(7.19) \quad |P_x(Z_t = x) - P_y(Z_t = x)| \leq P(\tau > t).$$

Next we derive an upper on  $P(\tau > t)$ . For any process  $Y$ , and for  $z \in S_N^1$  define

$$\tau_z^Y = \inf\{t \geq 0 : Y_t = z\}.$$

Without loss of generality assume that  $x > y$ . We observe that in this case we have for  $Z^{(1)}$  and  $Z^{(2)}$  as above that  $\tau \leq \tau_{-N}^{Z^{(1)}}$ . Using reflection and translation invariance we get

$$(7.20) \quad \begin{aligned} P(\tau > t) &\leq P_x(\tau_{-N}^{Z^{(1)}} > t) \\ &\leq P_{N-x}(\tau_{2N}^Y > t) \\ &\leq P_0(\tau_{2N}^Y > t), \end{aligned}$$

where  $Y = \{Y_t\}_{t \geq 0}$  is a continuous time simple random walk on  $\mathbb{Z}$ , reflected at 0 with jumps rate similar to  $Z^{(1)}$ .

Recall that  $W$  as simple random walk on  $\mathbb{Z}$  and  $\{T_t\}_{t \geq 0}$  is a Poisson process with intensity  $1/d$ . By Proposition 2.4.5 in [9] we get that there exists constants  $C_1, C_2 > 0$  such that for all integer  $n > 0$  and any  $r > 0$  we have

$$P_0(\tau_n^{|W|} > rn^2) \leq C_1 e^{-C_2 r}.$$

We therefore get

$$\begin{aligned}
P_0(\tau_{2N}^Y > t) &= \sum_{m=0}^{\infty} P(T_t = m) P_0(\tau_{2N}^{|W|} > m) \\
&\leq C_1 \sum_{m=0}^{\infty} e^{-t/d} \frac{(t/d)^m}{(m)!} e^{-C_2 m N^{-2}} \\
(7.21) \quad &= C_1 E[e^{-C_2 N^{-2} T_t}] \\
&= C_1 \exp \left\{ \frac{t}{d} (e^{-C_2 N^{-2}} - 1) \right\} \\
&\leq \tilde{C}_1 e^{-\tilde{C}_2 t N^{-2}},
\end{aligned}$$

where we have used the expression for the characteristic function of  $T_t$ . Combining (7.19), (7.20) and (7.21) gives

$$\begin{aligned}
(7.22) \quad |P_x(Z_t = x) - P_y(Z_t = x)| &\leq \tilde{C}_1 e^{-\tilde{C}_2 t N^{-2}}, \\
&\text{for all } x, y \in [-N, N], \ t > 0.
\end{aligned}$$

From (7.4) we get,

$$\begin{aligned}
(7.23) \quad |\langle \mathbf{e}_x, \mathcal{P}_t \mathbf{e}_x \rangle - \langle \mathbf{e}_y, \mathcal{P}_t \mathbf{e}_x \rangle| &= |P_x[\mathbf{X}_t = \mathbf{x}] - P_y[\mathbf{X}_t = \mathbf{x}]| \\
&= \left| \prod_{k=1}^d P_{x_k}(X_t^{(k)} = x_k) - \prod_{k=1}^d P_{y_k}(X_t^{(k)} = x_k) \right| \\
&\leq C(d) e^{-\tilde{C}_2 t N^{-2}},
\end{aligned}$$

where we used (7.22) and the triangle inequality in the last inequality.

Then by choosing  $K_0$  large enough, we get from (7.23) that

$$\begin{aligned}
(7.24) \quad \int_{K_0 N^2 \log N}^{\infty} |\langle \mathbf{e}_x - \mathbf{e}_y, \mathcal{P}_t(\mathbf{e}_x - \mathbf{e}_y) \rangle| dt &\leq C(d) \int_{K_0 N^2 \log N}^{\infty} e^{-\tilde{C}_2 t N^{-2}} dt \\
&\leq C N^2 N^{-\tilde{C}_2 K_0} \\
&\leq \tilde{C}.
\end{aligned}$$

Finally from (7.18), (7.24) and by noting that the integral on the right-hand side of (7.3) is bounded trivially by a constant on the integration region  $[0, 1]$  we get the upper bound in Proposition 3.1.  $\square$

## APPENDIX A. SOME HEURISTIC IDEAS

We introduce some heuristic ideas to shed light on our main results. We consider parameters  $i, j \in \{-N, \dots, N\}$ , and let  $\mathbf{X}_{i,j} \in \mathbb{R}^2$ . We

define the energy of for the field  $\mathbf{X}$  as

$$E_N(\mathbf{X}) = \frac{1}{2} \sum_{|(i,j)-(i',j')|=1} |\mathbf{X}_{i,j} - \mathbf{X}_{i',j'}|^2$$

Then we define  $P_N$  to be a constant times  $\exp(-E_N(\mathbf{X}))$ .

For  $\mathbf{x} \in \mathbb{R}^2$ , let  $\ell_N(\mathbf{x})$  be the number of points  $\mathbf{X}_{i,j}$  within distance  $1/2$  of  $\mathbf{x}$ . We weight the probability  $P_N$  by  $\exp(-\gamma \mathcal{E}_N(\mathbf{X}))$  where

$$\mathcal{E}_N(\mathbf{X}) = \int_{\mathbb{R}^2} \ell_N(\mathbf{x})^2 d\mathbf{x}.$$

Now we argue heuristically, and suppose that the  $\mathbf{X}_{i,j}$  are evenly spread in a ball of radius  $R$ . One way to spread them evenly is to let

$$\mathbf{X}_{i,j} \approx \frac{R}{N} \cdot (i, j)$$

In that case,

$$E_N(\mathbf{X}) \approx C \left( \frac{R}{N} \right)^2 N^2 = CR^2.$$

Also, with the hypothesis of even spreading, we have that either  $\ell(\mathbf{x}) = 0$  or

$$\ell_N(\mathbf{x}) \approx C \frac{N^2}{R^2}$$

and then

$$\mathcal{E}_N(\mathbf{X}) \approx C \left( \frac{N^2}{R^2} \right)^2 R^2 = C \frac{N^4}{R^2}$$

Equating  $E_N(\mathbf{X})$  and  $\mathcal{E}_N(\mathbf{X})$ , we get

$$R^2 = \frac{N^4}{R^2}$$

and so

$$R = N.$$

#### ACKNOWLEDGMENTS

We are very grateful to the Associate Editor and to the anonymous referee for careful reading of the manuscript and for a number of useful comments and suggestions that significantly improved this paper.

**Funding.** The work of Carl Mueller is partially supported by the Simons grant 513424.

**Availability of data and material.** Not applicable.

**Compliance with ethical standards.** The authors have no conflicts of interest to declare that are relevant to the content of this article.

**Code availability.** Not applicable.

## REFERENCES

- [1] L. Balents and D. S. Fisher. Large- $N$  expansion of  $(4 - \epsilon)$ -dimensional oriented manifolds in random media. *Phys. Rev. B*, 48(9):5949, 1993.
- [2] R. Bauerschmidt, H. Duminil-Copin, J. Goodman, and G. Slade. Lectures on self-avoiding walks. In *Probability and statistical physics in two and more dimensions*, volume 15 of *Clay Math. Proc.*, pages 395–467. Amer. Math. Soc., Providence, RI, 2012.
- [3] R. Bauerschmidt, G. Slade, A. Tomberg, and B. C. Wallace. Finite-order correlation length for four-dimensional weakly self-avoiding walk and  $|\varphi|^4$  spins. *Ann. Henri Poincaré*, 18(2):375–402, 2017.
- [4] Marek Biskup. Extrema of the two-dimensional discrete Gaussian free field. In *Random graphs, phase transitions, and the Gaussian free field*, volume 304 of *Springer Proc. Math. Stat.*, pages 163–407. Springer, Cham, [2020] ©2020.
- [5] F. den Hollander. *Random polymers*, volume 1974 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2009. Lectures from the 37th Probability Summer School held in Saint-Flour, 2007.
- [6] Y. Kantor, M. Kardar, and D. R. Nelson. Statistical mechanics of tethered surfaces. *Phys. Rev. Lett.*, 57:791–794, Aug 1986.
- [7] Y. Kantor, M. Kardar, and D. R. Nelson. Tethered surfaces: Statics and dynamics. *Phys. Rev. A*, 35:3056–3071, Apr 1987.
- [8] M. Kardar and D. R. Nelson.  $\epsilon$  expansions for crumpled manifolds. *Phys. Rev. Lett.*, 58:1289–1292, Mar 1987.
- [9] G. F. Lawler and V. Limic. *Random walk: a modern introduction*. Cambridge University Press, 2010.
- [10] M. Mezard and G. Parisi. Manifolds in random media: two extreme cases. *J. de Phys. I France*, 2(12):2231–2242, 1992.
- [11] C. Mueller and E. Neuman. Self-Repelling Elastic Manifolds with Low Dimensional Range. *Journal of Stochastic Analysis*, 3(2):Article 1, 2022.
- [12] D. Nelson, T. Piran, and S. Weinberg. *Statistical Mechanics of Membranes and Surfaces*. WORLD SCIENTIFIC, 2nd edition, 2004.
- [13] M. Plischke and B. Bergersen. *Equilibrium statistical physics*. World Scientific Publishing Co., Inc., Teaneck, NJ, third edition, 2006.

CARL MUELLER, DEPARTMENT OF MATHEMATICS, UNIVERSITY OF ROCHESTER, ROCHESTER, NY 14627

*Email address:* `carl.e.mueller@rochester.edu`

EYAL NEUMAN, DEPARTMENT OF MATHEMATICS, IMPERIAL COLLEGE, LONDON, UK

*Email address:* `e.neumann@imperial.ac.uk`