



# The random walk penalised by its range in dimensions $d \geq 3$

Nathanael Berestycki, Raphaël Cerf

## ► To cite this version:

Nathanael Berestycki, Raphaël Cerf. The random walk penalised by its range in dimensions  $d \geq 3$ . 2020. [⟨hal-02497146⟩](#)

**HAL Id: hal-02497146**

**<https://hal.science/hal-02497146v1>**

Preprint submitted on 3 Mar 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

# The random walk penalised by its range in dimensions $d \geq 3$

Nathanaël Berestycki  
Universität Wien<sup>‡</sup>

Raphaël Cerf<sup>\*†</sup>  
École Normale Supérieure

March 3, 2020

## Abstract

We study a self-attractive random walk such that each trajectory of length  $N$  is penalised by a factor proportional to  $\exp(-|R_N|)$ , where  $R_N$  is the set of sites visited by the walk. We show that the range of such a walk is close to a solid Euclidean ball of radius approximately  $\rho_d N^{1/(d+2)}$ , for some explicit constant  $\rho_d > 0$ . This proves a conjecture of Bolthausen [6] who obtained this result in the case  $d = 2$ .

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
1.1	Main results . . . . .	2
1.2	Heuristics . . . . .	5
1.3	Relation with other works . . . . .	6
<b>2</b>	<b>Further results and organisation of the paper</b>	<b>7</b>
2.1	Preliminary lower bound on the partition function . . . . .	7
2.2	Upper bound on the numerator . . . . .	8
2.3	Organisation of paper . . . . .	8
<b>3</b>	<b>Martingale estimates</b>	<b>11</b>
3.1	The classical martingale . . . . .	11
3.2	The fundamental inequality . . . . .	12
3.3	Estimate for a fixed profile . . . . .	12
3.4	Donsker–Varadhan estimate for the random walk in the full space . . . . .	14
3.5	Correction for the origin . . . . .	20

---

<sup>\*</sup>DMA, Ecole Normale Supérieure, CNRS, PSL Research University, 75005 Paris.

<sup>†</sup>Laboratoire de Mathématiques d'Orsay, Université Paris-Sud, CNRS, Université Paris-Saclay, 91405 Orsay.

<sup>‡</sup>On leave from the University of Cambridge

<b>4</b>	<b>A priori estimates</b>	<b>23</b>
4.1	Control of the range . . . . .	23
4.2	Control of the Dirichlet energy . . . . .	23
4.3	High density blocks . . . . .	24
4.4	Local averaging . . . . .	25
<b>5</b>	<b>Upper bound: proof of Theorem 2.6</b>	<b>26</b>
5.1	The coarse grained profile . . . . .	27
5.2	Continuation of proof of Theorem 2.6 . . . . .	29
5.3	Truncation . . . . .	32
5.4	End of proof of Theorem 2.6 . . . . .	34
<b>6</b>	<b>Continuous version of Theorem 2.6</b>	<b>39</b>
6.1	Linear interpolation . . . . .	39
6.2	Rescaling . . . . .	41
<b>7</b>	<b>Application of the quantitative Faber–Krahn inequality</b>	<b>47</b>
7.1	Statement . . . . .	47
7.2	Application of Faber–Krahn: proof of Theorem 1.5 . . . . .	53
<b>8</b>	<b>Filling the ball and proof of Theorem 1.7</b>	<b>55</b>
8.1	Time spent in mesoscopic balls . . . . .	55
8.2	Bridges in mesoscopic balls . . . . .	56
8.3	Conditioning and summation . . . . .	62
8.4	Transition probabilities for bridges . . . . .	63
8.5	End of proof of Theorem 8.1 . . . . .	68
<b>A</b>	<b>Functional inequalities on the rescaled lattice</b>	<b>70</b>
<b>B</b>	<b>Inequalities on the original lattice</b>	<b>71</b>

# 1 Introduction

## 1.1 Main results

Let  $P$  be the law of the discrete-time simple random walk  $(S_n)_{n \in \mathbb{N}}$  on  $\mathbb{Z}^d$ ,  $d \geq 1$ , starting from the origin. Let  $N$  be a positive integer. In [6], Bolthausen proposed the following model for a self-attractive random walk: let us denote by  $R_N$  the set of the points visited by the random walk until time  $N$  and by  $|R_N|$  its cardinality. We define a new probability on the set of the  $N$ -steps trajectories by setting

$$\frac{d\tilde{P}_N}{dP} = \frac{1}{Z_N} \exp(-|R_N|), \quad (1.1)$$

where the normalization factor (or partition function)  $Z_N$  is given by

$$Z_N = E(\exp(-|R_N|)). \quad (1.2)$$

Clearly,  $\tilde{P}_N$  favours configurations where the trajectory is localised on a small number of points. Bolthausen asked what can be said about a typical realisation of  $\tilde{P}_N$ . The question is particularly natural from the point of view of large deviations theory. Indeed one of the early successes of the theory, due to Donsker and Varadhan [9], was a determination of the first order asymptotics of the partition function  $Z_N$ :

$$Z_N = \exp \left( -(1 + o(1)) \chi_d N^{d/d+2} \right), \quad (1.3)$$

for some  $\chi_d > 0$  depending only on the dimension. Bolthausen was able to show that in dimension  $d = 2$ , under  $\tilde{P}_N$  a typical trajectory localises on a Euclidean ball of radius approximately  $\rho_2 N^{1/4}$  for some constant  $\rho_2 > 0$ . His analysis strongly suggests that in general dimensions  $d \geq 3$ , a similar result holds except that the walk now localises on a ball of radius approximately  $\rho_d N^{1/(d+2)}$ , where  $\rho_d > 0$  is a specific constant depending only on the ambient dimension  $d$ .

The main goal of this paper is to verify Bolthausen's conjecture. Bolthausen actually provided support for his conjecture by showing that two (admittedly crucial) estimates implied the conjecture in general dimension  $d \geq 2$ ; these two estimates were in turn proved for  $d = 2$ . The two theorems below provide a proof of these two estimates in the general case  $d \geq 3$ , thereby proving as a corollary of Bolthausen's paper [6] that his conjecture is true and hence completing his programme. We now state below these two results from which Bolthausen's conjecture follows.

We suppose without loss of generality that

$$n = N^{1/(d+2)}$$

is an integer. We define the local time  $L_N$  as

$$\forall x \in \mathbb{Z}^d \quad L_N(x) = \sum_{k=0}^{N-1} 1_{\{S_k=x\}}.$$

We define the continuous rescaled version  $\ell_N$  of  $L_N$  by

$$\forall x \in \mathbb{R}^d \quad \ell_N(x) = \frac{n^d}{N} L_N(\lfloor nx \rfloor). \quad (1.4)$$

For  $x \in \mathbb{R}^d$ , we let  $\phi_x$  be the principal eigenfunction (normalised so that  $\|\phi_x\|_2^2 := \int_{\mathbb{R}^d} \phi_x^2 = 1$ ) of  $-\Delta$  in  $B(x, \rho_d)$ , with Dirichlet boundary conditions. We denote by  $L^1(\mathbb{R}^d)$  the set of the integrable Borel functions on  $\mathbb{R}^d$  and we use the standard norm:

$$\forall f \in L^1(\mathbb{R}^d) \quad \|f\|_1 = \int_{\mathbb{R}^d} |f(x)| dx.$$

The first result below is a quantitative shape theorem in the  $L^1$  sense for the local time profile at time  $N$ .

**Theorem 1.5** Let  $\mathcal{L}_n$  be the set of functions defined by

$$\mathcal{L}_n = \left\{ \ell \in L^1(\mathbb{R}^d) : \|\ell\|_1 = 1, \ell \geq 0, \inf_{x \in \mathbb{R}^d} \|\ell - (\phi_x)^2\|_1 \geq 1/n^{1/800} \right\}.$$

For  $n$  large enough, we have

$$E(e^{-|R_N|}; \ell_N \in \mathcal{L}_n) \leq \exp \left( -n^d \chi_d - n^{d-\frac{1}{17}} \right) \quad (1.6)$$

where  $\chi_d$  is as in (1.3).

The upper bound obtained in formula (1.6) is in fact negligible compared to the partition function  $Z_N$ . This is a consequence of another result of [6] which is recalled in proposition 2.1.

The second main result says that if  $x$  is such that  $\ell_N$  is close to  $(\phi_x)^2$  in the  $L^1$  sense above, then actually almost all of the ball of radius  $\rho_d n$  around  $x$  has been filled by the range of the walk. More precisely, set

$$\mathcal{G}_{loc,x} = \left\{ \|\ell_N - (\phi_x)^2\|_1 \leq \frac{1}{n^{1/800}} \right\}.$$

For  $\kappa > 0$  and  $x \in \mathbb{R}^d$ , we define

$$\mathcal{R}_{\kappa,x} = \left\{ \forall z \in B(x, \rho_d n(1 - n^{-\kappa})), \quad \ell_N(z) > 0 \right\}.$$

This is the event that the ball of radius  $\rho_d n(1 - n^{-\kappa})$  around  $x$  is filled.

**Theorem 1.7** There exists  $\kappa > 0$  such that, for any  $a > 0$ ,

$$\frac{1}{Z_N} E(e^{-|R_N|}; \mathcal{G}_{loc,x}; (\mathcal{R}_{\kappa,x})^c) = o(N^{-a}).$$

Together with Bolthausen's results [6] (see p.877, immediately below Conjecture 1.3), Theorems 1.5 and 1.7 immediately imply the following, which is the main conclusion of our paper.

**Corollary 1.8** Let us denote by  $B(x, r)$  the  $d$  dimensional Euclidean ball centered at  $x$  of radius  $r$ . There exists a positive constant  $\rho_d$ , which depends only on the dimension  $d$ , such that, for any  $\varepsilon > 0$ , as  $N \rightarrow \infty$ ,

$$\tilde{P}_N \left( \exists x \in \mathbb{R}^d \quad B(x, \rho_d(1 - \varepsilon)N^{\frac{1}{d+2}}) \cap \mathbb{Z}^d \subset R_N \subset B(x, \rho_d(1 + \varepsilon)N^{\frac{1}{d+2}}) \right) \longrightarrow 1.$$

We mention here a result obtained along the way, which we feel is interesting in its own right. This is a Donsker–Varadhan large deviation estimate which is valid for the random walk in the full space  $\mathbb{Z}^d$ , and so bypasses the assumption of compactness for the state space which underlies [9]. Let  $D$  be an arbitrary finite subset of  $\mathbb{Z}^d$ . For  $t \in \mathbb{N}$ , we define

$$\tau(D, t) = \inf \left\{ k \geq 1 : L_k(D) = t \right\}$$

and we set

$$\forall x \in D \quad L_t^D(x) = L_{\tau(D,t)} = \sum_{k=0}^{\tau(D,t)-1} 1_{\{S_k=x\}}.$$

The function  $L_t^D$  is a function from  $D$  to  $\mathbb{N}$ . We shall work in the functional space  $\ell^1(D)$  equipped with the norm

$$\forall f \in \ell^1(D) \quad \|f\|_{1,D} = \sum_{x \in D} |f(x)|.$$

For a function  $f : \mathbb{Z}^d \rightarrow \mathbb{R}$ , we define

$$\mathcal{E}(f, D) = \frac{1}{2d} \sum_{y, z \in D; |y-z|=1} (f(y) - f(z))^2. \quad (1.9)$$

We define  $\overline{D}$  as

$$\overline{D} = D \cup \{x \in \mathbb{Z}^d : \exists y \in D \quad |x - y| = 1\}.$$

When  $D = \mathbb{Z}^d$  we will simply write  $\|f\|_1$  and  $\mathcal{E}(f)$ .

**Theorem 1.10** Let  $C$  be a closed convex subset of  $\ell^1(D)$ . For any  $t \geq 1$ , we have

$$\inf_{x \in D} P_x \left( \frac{1}{t} L_t^D \in C, \tau(D, t) < \infty \right) \leq \exp \left( -t \inf_{h \in C} \frac{1}{2} \mathcal{E}(\sqrt{h}, D) \right).$$

## 1.2 Heuristics

We begin a discussion of the above results (such as Corollary 1.8) with a rough heuristics explaining where the limit shape comes from. Note that a random walk will stay in a box of diameter  $n$  for duration  $N$  with probability approximately  $\exp(-O(N/n^2))$ , since it may leave this box with positive probability every  $n^2$  units of time. On the other hand, the energetic contribution to (1.2) of such configurations is of order  $\exp(-n^d)$ . Balancing entropy and energy, we expect that the trajectories that contribute most to (1.2) are such that  $n \asymp N^{1/(d+2)}$ , which explains the scaling in Corollary 1.8. We now discuss this in a bit more detail, but still ignoring many technical details.

If  $U \subset \mathbb{R}^d$  is an open bounded subset, and  $n = N^{1/(d+2)}$  is as above, then the probability for the random walk to remain in  $nU$  for a long time  $N$  is approximately  $\exp(-\lambda_U N/(2n^2))$ , where  $\lambda_U$  is the principal eigenvalue of  $-\Delta$  in  $U$  with Dirichlet boundary conditions on  $\partial U$ . Hence the contribution to (1.2) coming from trajectories staying in  $nU$  should be well approximated by  $\exp(-n^d(\lambda_U/2 + |U|))$ , where  $|U|$  is the Lebesgue measure of  $U$ . Using the **Faber–Krahn inequality**, it can be seen that  $\inf_U \{\lambda_U/2 + |U|\}$  is attained when  $U$  is a Euclidean ball of radius  $r$ , say. The radius  $r$  can then be determined by noting that principal eigenvalues obey diffusive scaling, i.e.,  $\lambda_{B(r)} = \lambda/r^2$ , where  $\lambda$  is the principal eigenvalue in the unit ball. Hence, if  $\omega_d$  is the volume of the unit ball, we deduce that  $r = \rho_d$  is obtained as the minimiser of the following functional:

$$\rho_d = \arg \min \left\{ \frac{\lambda}{2r^2} + \omega_d r^d : r > 0 \right\} = \left( \frac{\lambda}{d\omega_d} \right)^{1/(d+2)}.$$

The constant  $\rho_d$  is the one which appears in the theorem. As already mentioned, Bolthausen proved the corresponding result in two dimensions [6]. The main problem to extend Bolthausen’s proof to dimensions 3 and higher was to extend Lemma 3.1 and Proposition 4.1 from his paper [6]. The rest of his proof is written for arbitrary dimensions  $d \geq 3$  and it is solely the statement of these two lemmas which depend on  $d$  being equal to 2 in his proof.

Lemma 3.1 in [6] can be seen as a quantitative Faber–Krahn inequality on the torus, saying that if a set  $U$  is not far from minimising  $\lambda_U/2 + |U|$  then  $U$  itself is not far from a Euclidean ball. Unfortunately, such an inequality is not available yet in dimensions three and higher. Even an analogue of the quantitative isoperimetric inequality on the torus has not been proved so far. Therefore we cannot use the standard projection of the random walk on the torus, as in Bolthausen’s proof. This creates a serious difficulty. Indeed, as far as the probabilistic estimates on the local time are concerned, it is very convenient to work on the torus: the state space of the walk becomes compact and one can readily use the classical large deviations estimates of Donsker and Varadhan [9]. The good news is that a quantitative Faber–Krahn inequality has been proved recently in  $\mathbb{R}^d$  by Brasco, De Philippis and Velichkov [7]. Ultimately, we are to use this inequality. Therefore we have to deal with the random walk in the full space and we cannot afford the luxury of projecting its trajectories on the torus. A key point to carry out this program is to develop the relevant large deviations estimates. Indeed, the random walk being transient, the classical Donsker–Varadhan theory cannot be applied directly.

**Warning.** In the probability literature, one usually works with the half-Laplacian  $\Delta/2$ , which is the infinitesimal generator of the Brownian motion. In Bolthausen’s paper, the notation  $\lambda(G)$  corresponds to half of the quantity defined above. We choose here to stick to the convention employed in the papers on the Faber–Krahn inequality.

### 1.3 Relation with other works

A Brownian analogue of Corollary 1.8 was proved in dimension  $d = 2$  by Sznitman [18] using the method of enlargement of obstacles. Very briefly, the starting point of this method (adapted to the discrete setting of this paper) consists in viewing the weighted probability measure  $\hat{P}_N$  as the annealed probability measure corresponding to a random walk in a random medium in which there is an obstacle at every site with probability  $1 - e^{-1}$ . Note that  $e^{-|R_N|}$  then corresponds to the annealed probability that the walk has not encountered any obstacle for time  $N$ .

A refinement of this method enabled Povel [16] to establish the same result in dimension  $d \geq 3$ . It is important to note however that in the continuum, one cannot of course hope that the range of Brownian motion will fill a ball completely – there will always be small holes. For this reason, both results in [18] and [16] are restricted to a statement of the so-called *confinement property*, i.e., a statement that the range is contained in a ball of the appropriate radius (corresponding to  $R = \rho_d n(1 + \varepsilon)$  in our setup). The question of whether the range will visit any macroscopic ball within this ball of radius  $R$  is only addressed tangentially, see e.g. Theorem 4.3 in [18] and the discussion at the end of Section 1 in [16]. Needless to say, the method of enlargement of obstacles is very different from the strategy employed by Bolthausen in [6]. Curiously, neither [18] nor [16] discuss what their results imply for the discrete case, though both briefly mention the paper [6].

At the time we were finishing this paper, we learnt of the independent and nearly simultaneous work of Ding, Fukushima, Sun and Xu [8], who obtained an alternative proof of Corollary 1.8. In fact, their result implies a more precise control on the size of the boundary  $\partial R_N$  under  $\hat{P}_N$ , showing that with high probability,  $|\partial R_N| \leq (\log n)^c n^{d-1}$  for some  $c > 0$  and all for all  $N$  large enough. Their starting point is the paper of Povel [16], whose results are used freely in the discrete setting rather than in the continuum. (As pointed out in [8], a translation of Sznitman’s method of enlargement of obstacles to a discrete setup was undertaken previously in [1] – interestingly this predates [16]). Given this, what remains to be proved is that the range of the random walk covers all of the ball of radius  $\rho_d n(1 - \varepsilon)$ , i.e., our Theorem 1.7. Hence the overlap with our paper is reduced to the proof of this theorem, which occupies Section 8 of this paper. The major differences with the approach of [8] can be summarised as follows:

- once we have proved Theorem 1.5, we know a bit more than the confinement property, since we know that the local time profile is close in the  $L^1$  sense to the eigenfunction. This implies in particular that mesoscopic balls are visited frequently, a step which is therefore easy for us (Lemma 8.2) but which requires an argument in [8] (more precisely, the authors of [8] argue that if a ball is not frequently visited it must be close to the boundary).

- We have found a way to control uniformly the probability that a given set of  $k$  points is avoided by the random walk in such a mesoscopic ball. Surprisingly, the control we get here is good enough that it works for *any* configuration of points, no matter what its geometry, and depends only on its cardinality. This is a major technical difference with [8] where the bound given depends on how many points in the set are far from one another. The key additional idea which allows us to do this here is to partition the points in the set according to their distance to the boundary

and only consider points at a given distance from the boundary, where this distance is chosen to maximise the number of such points. This results in an arguably simpler line of reasoning from the conceptual point of view.

We also mention the related work [2] which was the initial motivation of our investigation. In this paper, the penalisation by the range  $e^{-|R_N|}$  is replaced by the size of the boundary of the range  $e^{-|\partial R_N|}$ . This turns the random walk into a polymer interface model. A conjecture in [2] states that a shape theorem takes place on the scale  $n = N^{1/(d+1)}$  instead of  $n = N^{1/(d+2)}$ . This shape could then be thought of as a Wulff crystal shape for the random walk. Despite partial results in [2], this conjecture remains wide open at positive temperature. However, in the limit of zero temperature (for a more general model where the boundary size is measured through i.i.d. random variables attached to edges), Biskup and Procaccia [5, 4] were able to prove this conjecture. Observe that the random media representation which is the starting point of the method of enlargement of obstacles is not available for such a model.

Finally, for another approach to large deviations without compactness, see [15].

**Acknowledgements.** We thank the authors of [8] for useful discussions regarding their work. Part of this research was carried out when NB was a guest at Ecole Normale Supérieure, Paris, whose support and hospitality is gratefully acknowledged. This project started in 2013, while R. Cerf was visiting Cambridge, thanks to the support of the IUF. NB's research was partly supported by EPSRC grants EP/L018896/1 and EP/I03372X/1.

## 2 Further results and organisation of the paper

### 2.1 Preliminary lower bound on the partition function

To explain some further details about our approach (including intermediate theorems of interest in their own right, see below), it will be useful to start by recalling the following lower bound due to Bolthausen on the partition function which is a quantitative improvement on the result of Donsker and Varadhan.

Let  $n$  be an integer such that  $n^{d+2} = N$ . Without loss of generality, we can assume that  $N$  is such an integer power, and we do so throughout the paper.

**Proposition 2.1** There exists a constant  $c \in \mathbb{R}$ , which depends on the dimension  $d$  only, such that, for  $N$  large enough, we have

$$Z_N \geq \exp \left( -\chi_d n^d - cn^{d-1} \right)$$

where

$$\chi_d = \frac{\lambda}{\rho_d^2} + \omega_d \rho_d^d$$

is the same constant which appears in (1.3).

**Proof.** See Proposition 2.1 in [6] (note that the proof is valid in any dimension). □



## 2.2 Upper bound on the numerator

Once we have a lower bound on the normalizing constant, the main problem is to obtain an adequate upper bound on the integral of  $\exp(-|R_N|)$  over an arbitrary event  $\mathcal{A}$ . We can then rule out those events  $\mathcal{A}$  for which we can obtain an upper bound which is negligible compared to the previous lower bound. Throughout the computations, we use the following convention. For  $\mathcal{A}$  an event, we write

$$E\left(\exp(-|R_N|); \mathcal{A}\right) = E\left(\exp(-|R_N|)1_{\mathcal{A}}\right). \quad (2.2)$$

The central object in our study is the local time of the random walk, defined as

$$\forall x \in \mathbb{Z}^d \quad L_N(x) = \sum_{k=0}^{N-1} 1_{\{S_k=x\}}. \quad (2.3)$$

Our estimates will involve its square root, which we denote by  $f_N$ :

$$\forall x \in \mathbb{Z}^d \quad f_N(x) = \sqrt{L_N(x)}. \quad (2.4)$$

In order to compare functions to one another, we will make use of various  $\ell^p$  norms, which, unlike Bolthausen [6], we take to be unscaled. Thus we define, for a function  $f : \mathbb{Z}^d \rightarrow \mathbb{R}$ ,

$$\|f\|_p = \left( \sum_{x \in \mathbb{Z}^d} |f(x)|^p \right)^{1/p}. \quad (2.5)$$

A fundamental step in the proof of Theorem 1.5, which takes up a substantial portion of this paper, is the following quantitative result which allows us to get an upper bound on (2.2).

**Theorem 2.6** Let  $\mathcal{F}$  be a collection of functions from  $\mathbb{Z}^d$  to  $\mathbb{R}^+$ . For any  $\kappa \geq 1$ , we have, for  $n$  large enough,

$$\begin{aligned} E(e^{-|R_N|}; f_N \in \mathcal{F}) &\leq \exp(-\kappa n^d) + \\ &\exp\left(n^{d-1/8} - \inf \left\{ \left| \left\{ x \in \mathbb{Z}^d : h(x) > 0 \right\} \right| + \frac{N}{2}(1 - n^{-1/4}) \left( \max\left(\sqrt{\mathcal{E}(\sqrt{h})} - \frac{1}{n^{9/8}}, 0\right) \right)^2 \right. \right. \\ &\quad \left. \left. : h \in \ell^1(\mathbb{Z}^d), h \geq 0, \exists f \in \mathcal{F} \quad \left\| h - \frac{1}{N} f^2 \right\|_1 \leq \frac{1}{n^{1/16}} \right\} \right). \end{aligned}$$

## 2.3 Organisation of paper

Sections 3, 4, and 5 are devoted to the proof of Theorem 2.6. Section 6 and 7 explain how to deduce Theorem 1.5 from Theorem 2.6. Section 8 deals with a proof of Theorem 1.7.

As the proofs are rather lengthy, let us sum up the main steps of the proofs.

### Main steps in proof of Theorem 2.6.

(i) (Section 3) We estimate probabilities of the form  $P(\|L_N - g^2\|_{1,D} < \Gamma)$  for some small  $\Gamma$ , where  $\|f\|_{1,D}$  is the norm of the function  $f$  restricted to  $D$ . To this end, we develop a new type of large deviations estimates for the random walk (Theorem 1.10, see Section 3.4 for its proof). In fact,

that result involves the infimum over a set of admissible starting points. In order to apply it to the random walk starting from the origin, we have to introduce a correcting factor, which however does not destroy the leading term in the estimates (Section 3.5).

(ii) (Section 4.1) We show that, up to events whose  $\tilde{P}_N$  probability is negligible, for some  $c > 0$ , we have  $|R_N| \leq cn^d$ . This is a direct consequence of the definition of  $\tilde{P}_N$ .

(iii) (Section 4.2) We show that, up to events whose  $\tilde{P}_N$  probability is negligible, for some  $\kappa \geq 1$ , we have  $\mathcal{E}(f_N) \leq \kappa n^d \ln n$ , where  $\mathcal{E}$  is the Dirichlet energy. First we estimate the probability that  $f_N$  is equal to a fixed function  $f$ . This estimate relies on the classical martingale used by Donsker and Varadhan. We then bound the number of functions satisfying the constraint  $\text{supp } f \leq cn^d$ .

From now onwards, we need only to consider trajectories satisfying the points (ii) and (iii). Having a control on the Dirichlet energy yields automatically a control on the norm in  $\ell^{2^*}(\mathbb{Z}^d)$ , where  $2^* = 2d/(d-2)$ , via a discrete Sobolev–Poincaré inequality. We will use the three bounds

$$\mathcal{E}(f_N) \leq \kappa n^d \ln n, \quad |R_N| \leq cn^d, \quad \|f_N\|_{2^*} \leq c_{PS} \sqrt{\kappa n^d \ln n},$$

to develop an adequate coarse-grained image of the local time.

(iv) (Section 4.3) We partition the space into blocks of side length  $n$ . We focus on the blocks  $B$  such that  $\|f_N\|_{2^*, B} \geq \delta n^{d/2^*+1}$ . The exponent  $d/2^* + 1$  corresponds to the typical situation for a block actively visited by the random walk until time  $N$ : the number of visits per site should be of order  $n^2$ , so  $f_N$  is of order  $n$  throughout the block. We keep record of the indices of these blocks. More precisely, we denote by  $X$  the set of the centers of these blocks; of course the set  $X$  depends on  $f_N$  and is random. We denote by  $E$  the union of these blocks and by  $D$  the region  $E$  enlarged with all the blocks on the frontier. We control the norm of  $f_N$  outside the region  $E$ . This is done with the help of a discrete Poincaré–Sobolev inequality and the control of the  $\ell^{2^*}$  norm:

$$\sum_{x \in \mathbb{Z}^d \setminus E} (f_N(x))^{2^*} \leq c_d (\delta^{2^*} n^{d+2^*})^{1-2/2^*} (n^d + \kappa n^d \ln n).$$

(v) (Section 4.4) We introduce a length scale  $M$ . We partition the space into blocks  $B'(\underline{x})$  of side length  $M$ . We perform a local average of  $f_N$  on each such block and we get a function  $\overline{f_N}^M$ . We control the norm of the difference  $f_N - \overline{f_N}^M$  with the help of a discrete Poincaré–Wirtinger inequality.

(vi) (Section 5.1) We discretise next the values of the functions  $\overline{f_N}^M$  over the blocks  $B(\underline{x})$  which are included in  $E$  with a discretisation step  $\eta > 0$ . This way we obtain a function  $\eta \overline{f_N}^M$ , which is the coarse grained profile.

At this point, we can write

$$E(e^{-|R_N|}; f_N \in \mathcal{F}) \leq \sum_{X, g} E(e^{-|R_N|}; f_N \in \mathcal{F}, \eta \overline{f_N}^M = g),$$

where the summation extends over the admissible set of blocks  $X$  and profiles  $g$ . Whenever  $\eta \overline{f_N}^M = g$ , we can show that  $\|f_N - g\|_{2, D} < \Gamma_0$ , and  $\|L_N - g^2\|_{1, D} < \Gamma_1$ , where  $\Gamma_0, \Gamma_1$  are explicit functions of  $N$ .

(viii) (Section 5.2) We apply our large deviations inequality to bound the expectation

$$E(e^{-|R_N|}; \|L_N - g^2\| < \Gamma_1).$$

To do so, we approximate the cardinality of the range  $|R_N|$  by the cardinality of the points where the coarse grained profile is quite large. The resulting upper bound depends on an infimum over a set of functions  $h$  defined on the domain  $D$ , and more specifically on their Dirichlet energy inside  $D$ .

(ix) (Section 5.3) With the help of a truncation and the control of the  $\ell^{2*}$  norm outside  $E$ , we relate the Dirichlet energy in the full space to the Dirichlet energy restricted to  $D$ . This involves essentially a discrete integration by parts.

(x) (Section 5.4) We use the inequality proved in step (ix). This way we get an upper bound involving the Dirichlet energy in the full space, however we have to work further to get rid of the truncation operator. After some tedious computations, we obtain an upper bound depending only on the collection  $\mathcal{F}$ . We plug this upper bound in the previous sums. It remains only to count the number of terms in the sums. We choose finally the parameters  $M, \delta, \eta$  adequately to get the desired upper bound.

**Main steps in the proof of Theorem 1.7.** We fix a mesoscopic scale  $m = n^{1-2\kappa}$  where  $\kappa > 0$  is a small parameter. Suppose  $\mathcal{G}_{loc,x}$  holds and take  $x = 0$  without loss of generality. We aim to show that the ball of radius  $m$  around a point  $z$  in the desired range  $B(0, \rho_d n(1 - n^{-\kappa}))$  is entirely visited.

(i) (Section 8.1) We first note that deterministically on  $\mathcal{G}_{loc,0}$ , the walk spends a lot of time in  $B(z, m)$ . This is because otherwise there would be a polynomial error in the  $L^1$  distance between  $\ell_N$  and  $(\phi_x)^2$ .

(ii) (Section 8.2) We show that under  $\tilde{P}_N$  there are many disjoint portions of the walk of duration  $m^2$  where the walk starts and ends inside the bulk of the ball (say,  $B(z, m/2)$ ) and never leaves  $B(z, m)$  throughout this interval of time. We call such a portion a bridge. To do so we use a change of measure argument whose cost (i.e., the value of the Radon–Nikodym derivative of this change of measure) is comparable to the entropic term in the partition function  $Z_N$ : that is, of order  $\exp(-(\lambda/2\rho_d^2)n^d)$ . Since on  $\mathcal{G}_{loc,0}$  the size of the range  $|R_N|$  is also essentially deterministically lower bounded up to a small error, the energetic term  $e^{-|R_N|}$  together with the cost of the change of measure is of order at most  $Z_N$ , which allows us to compare effectively  $\tilde{P}_N$  to this new measure. A technical difficulty is that the change of measure technique is better implemented in continuous time rather than discrete time. Once we work under this change of measure, it is easy to check that the number of bridges is as desired: indeed, every time the walk is in the bulk of the ball, there is a decent chance that the next  $m^2$  units of time will result in a bridge. Moreover, by Step 1 we control the number of trials, so ultimately the desired result follows from standard large deviations for Binomial random variables.

(iii) (Section 8.3) We fix  $k \geq 1$  and a set  $\mathcal{X}$  of  $k$  points in  $B(z, m)$ , and try to estimate the  $\tilde{P}_N$  probability that  $\mathcal{X}$  is avoided by the walk. We can condition on everything that happens outside of  $B(z, m)$ ; on the event that the range avoids exactly the set  $\mathcal{X}$  the size of the range is then deterministic. Furthermore, if we condition on the number and endpoints of the bridges the probability that all the bridges avoid  $\mathcal{X}$  is exactly the product for individual bridges to avoid  $\mathcal{X}$ .

Hence the probability that the range restricted to  $B(z, m)$  is strictly smaller than this ball can be written as a sum over  $k \geq 1$ , and over all subsets  $\mathcal{X}$  of size  $k$  of the product of probabilities that bridges avoid  $\mathcal{X}$ .

(iv) (Section 8.4) To estimate the latter we need to control transition probabilities for bridges that are uniform in the starting and end points of the bridge. It is here that it is useful to have taken the starting and end points of the bridge in the bulk of the ball  $B(z, m/2)$  and not near the boundary. This shows in particular that the probability for a bridge to find itself at a specific point at distance  $r$  from the boundary at some specific time which is neither close to the start or the end of the bridge, is proportional to  $(r/m)^2$ . This follows essentially from a gambler's ruin probability argument.

(v) (Section 8.5) By the previous step it suffices to estimate the probability that a given bridge avoids  $\mathcal{X}$ . We aim to find a bound that is uniform on the geometry of  $\mathcal{X}$  and depends only on the number of points  $k$  in  $\mathcal{X}$ . Intuitively, the easiest configuration to avoid is when  $\mathcal{X}$  is clumped together as a solid ball of radius  $R = k^{1/d}$ , so this should provide a lower bound on the desired probability. In that case our estimates on the transition probabilities from the previous step show via a moment computation that the probability for a bridge to hit  $\mathcal{X}$  should be at least  $km^{2-d}/R^2 = k^{1-2/d}m^{2-d}$ , independently of the geometry of  $\mathcal{X}$  (ignoring boundary effects). This turns out to be true and can be deduced relatively easily from the fact that the Green function of the random walk in  $\mathbb{Z}^d$ ,  $d \geq 3$ , is essentially monotone in the distance. (Such arguments can be used to prove isoperimetric inequalities for the capacity of a set, but we did not include this here for the sake of brevity).

In order to deal with boundary effects, in a way that is still uniform in the geometry of  $\mathcal{X}$ , we divide the ball  $B(z, m)$  into concentric annuli  $A_j$  at distance  $2^j$  from the boundary of the ball ( $j \geq 1$ ). If all the points of  $\mathcal{X}$  were in the annulus  $A_j$  it would be possible to control the boundary effects in a uniform way. Indeed the expected time spent in  $\mathcal{X}$  would be, by the gambler's ruin estimate from Step (iv), proportional to  $r^2$  where  $r = 2^j$  is the distance to the boundary. However the expected time spent in  $\mathcal{X}$  starting from a point in  $\mathcal{X}$  would also be bounded by a factor proportional to  $r^2$  as well using Step (iv) again. Hence the  $r^2$  terms cancel in the moment computation, and we could use the above bound. When  $\mathcal{X}$  is not contained in a single annulus  $A_j$ , we can instead consider  $\mathcal{X}_j = \mathcal{X} \cap A_j$ , where  $A_j$  is chosen so that it contains most points of  $\mathcal{X}$ . Then since it suffices to hit  $\mathcal{X}_j$ , we can apply the above lower bound with  $k$  replaced by  $|\mathcal{X}_j|$ . By the choice of  $j$ , this is at least  $k/\log m$  instead of  $k$ . It turns out that plugging this extra logarithmic factor does not substantially alter the conclusion.

## 3 Martingale estimates

### 3.1 The classical martingale

The crucial ingredient to derive the relevant probabilistic estimates on the random walk is the family of martingales used by Donsker and Varadhan, which we define thereafter. Let  $u$  be a positive function defined on  $\mathbb{Z}^d$ . To the function  $u$  we associate the function  $V$  defined on  $\mathbb{Z}^d$  by

$$\forall x \in \mathbb{Z}^d \quad V(x) = \frac{1}{2d} \sum_{y \in \mathbb{Z}^d, |x-y|=1} u(y).$$

For  $n \geq 0$ , we set

$$M_n = \left( \prod_{k=0}^{n-1} \frac{u(S_k)}{V(S_k)} \right) u(S_n).$$

We claim that the process  $(M_n)_{n \in \mathbb{N}}$  is a martingale. Indeed, for any  $n \geq 0$ ,

$$\begin{aligned} E(M_n \mid S_0, \dots, S_{n-1}) &= \left( \prod_{k=0}^{n-1} \frac{u(S_k)}{V(S_k)} \right) E(u(S_n) \mid S_{n-1}) \\ &= \left( \prod_{k=0}^{n-1} \frac{u(S_k)}{V(S_k)} \right) V(S_{n-1}) = M_{n-1}. \end{aligned}$$

In the same way, if the random walk  $(S_n)_{n \in \mathbb{N}}$  starts from an arbitrary point  $x \in \mathbb{Z}^d$ , and if we denote by  $P_x$  and  $E_x$  the associated probability and expectation, then, under  $P_x$ , the process  $(M_n)_{n \in \mathbb{N}}$  is again a martingale.

### 3.2 The fundamental inequality

Since  $(M_n)_{n \in \mathbb{N}}$  is a martingale, then

$$E_x(M_N) = E_x(M_0) = u(x). \quad (3.1)$$

Let us express  $M_N$  with the help of the local time  $L_N$ :

$$\begin{aligned} M_N &= \exp \left( \sum_{0 \leq k < N} \ln \frac{u(S_k)}{V(S_k)} \right) u(S_N) \\ &= \exp \left( \sum_{y \in \mathbb{Z}^d} \ln \frac{u(y)}{V(y)} L_N(y) \right) u(S_N). \end{aligned}$$

Since  $|S_N - S_0| \leq N$ , then

$$u(S_N) \geq \inf \{ u(y) : |y - x| \leq N \}.$$

Reporting this inequality in the martingale equality (3.1), we get the following fundamental inequality.

**Lemma 3.2** For any  $N \geq 1$ , any  $x \in \mathbb{Z}^d$  and any positive function  $u$  defined on  $\mathbb{Z}^d$ , we have

$$E_x \left( \exp \left( \sum_{y \in \mathbb{Z}^d} \ln \frac{u(y)}{V(y)} L_N(y) \right) \right) \leq \frac{u(x)}{\inf \{ u(y) : |y - x| \leq N \}}.$$

### 3.3 Estimate for a fixed profile

For  $f$  a function from  $\mathbb{Z}^d$  to  $\mathbb{R}$ , we define its discrete Dirichlet energy  $\mathcal{E}(f)$  by

$$\mathcal{E}(f) = \frac{1}{2d} \sum_{\substack{y, z \in \mathbb{Z}^d \\ |y - z| = 1}} (f(y) - f(z))^2.$$

In this section, we shall prove the following estimate for the probability that the square root of the local time is equal to a fixed profile.

**Proposition 3.3** Let  $\phi$  be a function from  $\mathbb{Z}^d$  to  $[0, +\infty[$  such that

$$\sum_{y \in \mathbb{Z}^d} \phi(y)^2 = N.$$

For any  $N \geq 1$ , any  $\alpha \in ]0, 1[$ , we have

$$P(f_N = \phi) \leq \frac{\phi(0)}{\alpha} \exp \left( -\frac{1}{2} \mathcal{E}(\phi) + \alpha \sqrt{N |\text{supp } \phi|} \right).$$

**Proof.** To bound the probability  $P(f_N = \phi)$ , we proceed as follows. Let  $\alpha \in ]0, 1[$  and let  $u : \mathbb{Z}^d \rightarrow ]0, +\infty[$  be the positive function defined on  $\mathbb{Z}^d$  by

$$\forall y \in \mathbb{Z}^d \quad u(y) = \max(\phi(y), \alpha).$$

Obviously, we have

$$E \left( \exp \left( \sum_{y \in \mathbb{Z}^d} \ln \frac{u(y)}{V(y)} L_N(y) \right) \right) \geq P(f_N = \phi) \exp \left( \sum_{y \in \mathbb{Z}^d} \ln \frac{u(y)}{V(y)} \phi(y)^2 \right).$$

Since the random walk starts from 0, then  $f_N(0) \geq 1$ , so we need only to consider functions  $\phi$  such that  $\phi(0) \geq 1$ . In this case, we have  $u(0) = \phi(0) \geq \alpha$ . Applying the fundamental estimate of lemma 3.2, we get

$$P(f_N = \phi) \leq \frac{\phi(0)}{\alpha} \exp \left( - \sum_{y \in \mathbb{Z}^d} \ln \frac{u(y)}{V(y)} \phi(y)^2 \right).$$

Next, we have, for  $y \in \mathbb{Z}^d$ ,

$$- \ln \frac{u(y)}{V(y)} = \ln \left( 1 + \frac{V(y) - u(y)}{u(y)} \right) \leq \frac{\Delta_1 u(y)}{u(y)},$$

where  $\Delta_1$  is the discrete Laplacian operator, defined by

$$\Delta_1 u(y) = V(y) - u(y).$$

Reporting in the previous inequality, we arrive at

$$P(f_N = \phi) \leq \frac{\phi(0)}{\alpha} \exp \left( \sum_{y \in \mathbb{Z}^d} \frac{\Delta_1 u(y)}{u(y)} \phi(y)^2 \right).$$

We evaluate next the sum in the exponential. For  $y \in \text{supp } \phi$ , we have  $\phi(y) \geq 1 > \alpha$ , whence  $\phi(y) = u(y)$ , therefore

$$\begin{aligned}
\sum_{y \in \mathbb{Z}^d} \frac{\Delta_1 u(y)}{u(y)} \phi(y)^2 &= \sum_{y \in \text{supp } \phi} \Delta_1 u(y) \phi(y) \\
&= \sum_{y \in \text{supp } \phi} \frac{1}{2d} \sum_{z: |z-y|=1} (u(z) - u(y)) \phi(y) \\
&\leq \sum_{y \in \text{supp } \phi} \frac{1}{2d} \sum_{z: |z-y|=1} (\phi(z) - \phi(y) + \alpha) \phi(y) \\
&= \frac{1}{2d} \sum_{\substack{y, z \in \text{supp } \phi \\ |y-z|=1}} \phi(y) \phi(z) + \sum_{y \in \text{supp } \phi} (\alpha \phi(y) - \phi(y)^2) \\
&= -\frac{1}{4d} \sum_{\substack{y, z \in \text{supp } \phi \\ |y-z|=1}} (\phi(y) - \phi(z))^2 + \alpha \sum_{y \in \text{supp } \phi} \phi(y).
\end{aligned}$$

Moreover

$$\left( \sum_{y \in \text{supp } \phi} \phi(y) \right)^2 \leq \left( \sum_{y \in \text{supp } \phi} \phi(y)^2 \right) |\text{supp } \phi| = N |\text{supp } \phi|.$$

Putting together the previous inequalities, we obtain the inequality stated in the proposition.  $\square$

### 3.4 Donsker–Varadhan estimate for the random walk in the full space

In order to estimate the probability that  $L_N$  belongs to a ball centered at a fixed function  $g^2$ , we develop here a deviation inequality, valid for any value of  $N$  and for any convex set of functions. This deviation inequality is stated in Theorem 1.10 of the introduction and we prove it here. We use the notation introduced just before the statement of Theorem 1.10.

**Proof.** Let  $u$  be a positive function defined on  $\mathbb{Z}^d$  and let  $x \in \mathbb{Z}^d$ . Let  $(M_n)_{n \in \mathbb{N}}$  be the martingale constructed with the function  $u$  and the random walk  $(S_n)_{n \in \mathbb{N}}$ , defined in section 3.1. We first remark that  $\tau(D, t)$  is a stopping time for  $(M_n)_{n \in \mathbb{N}}$ . Indeed, the event  $\{\tau(D, t) = n\}$  is measurable with respect to  $L_n(D)$ , hence also to  $S_1, \dots, S_{n-1}$ . We apply next the optional stopping theorem. Let  $t, n \geq 1$ , we have

$$E_x(M_{n \wedge \tau(D, t)}) = E_x(M_0) = u(x).$$

We note simply  $\tau$  instead of  $\tau(D, t)$  and we bound from below the lefthand member:

$$\begin{aligned}
E_x(M_{n \wedge \tau}) &\geq E_x(M_{n \wedge \tau} 1_{\tau < \infty}) \\
&= E_x \left( \exp \left( \sum_{y \in \mathbb{Z}^d} \left( \ln \frac{u(y)}{V(y)} \right) L_{n \wedge \tau}(y) \right) u(S_{n \wedge \tau}) 1_{\tau < \infty} \right). \tag{3.4}
\end{aligned}$$

From now onwards, we suppose that the function  $u$  is superharmonic on  $\mathbb{Z}^d \setminus D$ , i.e., we suppose that  $u$  is such that

$$\forall y \in \mathbb{Z}^d \setminus D \quad u(y) \geq V(y) = \frac{1}{2d} \sum_{\substack{z \in \mathbb{Z}^d \\ |y-z|=1}} u(z). \tag{3.5}$$

We denote by  $\mathcal{S}^+$  the collection of the positive functions  $u$  defined on  $\mathbb{Z}^d$  which satisfy (3.5). The superharmonicity of  $u$  implies that

$$\sum_{y \in \mathbb{Z}^d} \left( \ln \frac{u(y)}{V(y)} \right) L_{n \wedge \tau}(y) \geq \sum_{y \in D} \left( \ln \frac{u(y)}{V(y)} \right) L_{n \wedge \tau}(y). \quad (3.6)$$

Reporting inequality (3.6) in inequality (3.4), we get

$$u(x) \geq E_x \left( \exp \left( \sum_{y \in D} \left( \ln \frac{u(y)}{V(y)} \right) L_{n \wedge \tau}(y) \right) u(S_{n \wedge \tau}) 1_{\tau < \infty} \right).$$

On the event  $\tau < \infty$ , we have

$$\begin{aligned} \forall y \in D \quad \lim_{n \rightarrow \infty} L_{n \wedge \tau}(y) &= L_\tau(y) = L_t^D(y), \\ \lim_{n \rightarrow \infty} S_{n \wedge \tau} &= S_\tau. \end{aligned}$$

By Fatou's lemma, we have

$$u(x) \geq E_x \left( \exp \left( \sum_{y \in D} \left( \ln \frac{u(y)}{V(y)} \right) L_t^D(y) \right) u(S_\tau) 1_{\tau < \infty} \right).$$

Yet  $S_\tau$  belongs to  $\overline{D}$ , thus

$$u(x) \geq \left( \inf_{y \in \overline{D}} u(y) \right) E_x \left( \exp \left( \sum_{y \in D} \left( \ln \frac{u(y)}{V(y)} \right) L_t^D(y) \right) 1_{\tau < \infty} \right). \quad (3.7)$$

Recall that

$$\overline{D} = D \cup \{x \in \mathbb{Z}^d : \exists y \in D \quad |x - y| = 1\}.$$

Taking now the infimum over  $x \in \overline{D}$ , we obtain

$$\inf_{x \in \overline{D}} E_x \left( \exp \left( \sum_{y \in D} \left( \ln \frac{u(y)}{V(y)} \right) L_t^D(y) \right) 1_{\tau < \infty} \right) \leq 1. \quad (3.8)$$

We proceed by bounding from below the lefthand member of (3.8) as follows: for any  $x \in \overline{D}$ ,

$$\begin{aligned} & E_x \left( \exp \left( \sum_{y \in D} \left( \ln \frac{u(y)}{V(y)} \right) L_t^D(y) \right) 1_{\tau < \infty} \right) \\ & \geq E_x \left( \exp \left( \sum_{y \in D} \left( \ln \frac{u(y)}{V(y)} \right) L_t^D(y) \right); \tau < \infty, \frac{1}{t} L_t^D \in C \right) \\ & \geq \exp \left( \inf_{g \in C} \sum_{y \in D} \left( \ln \frac{u(y)}{V(y)} \right) t g(y) \right) P_x \left( \frac{1}{t} L_t^D \in C, \tau < \infty \right). \end{aligned} \quad (3.9)$$



Let us define

$$\phi_t(C) = \inf_{x \in D} P_x \left( \frac{1}{t} L_t^D \in C, \tau < \infty \right). \quad (3.10)$$

Whenever  $\tau < \infty$ , the function  $\frac{1}{t} L_t^D$  belongs to the set  $M_1(D)$  defined by

$$M_1(D) = \left\{ \phi \in \ell^1(D) : \sum_{x \in D} \phi(x) = 1, \forall x \in D \quad 0 \leq \phi(x) \leq 1 \right\}.$$

Therefore we can replace the set  $C$  by its intersection with  $M_1(D)$  in formula (3.9) and the previous inequalities yield that

$$\phi_t(C) \leq \exp \left( -t \inf_{g \in C \cap M_1(D)} \sum_{y \in D} \left( \ln \frac{u(y)}{V(y)} \right) g(y) \right). \quad (3.11)$$

This inequality holds for any function  $u$  in  $\mathcal{S}^+$ . In order to get a functional which is convex, we perform a change of functions and we set  $\phi = \ln u$ . We denote by  $\mathcal{T}$  the image of  $\mathcal{S}^+$  under this change of functions, i.e.,

$$\mathcal{T} = \{ \ln u : u \in \mathcal{S}^+ \}.$$

We rewrite inequality (3.11) as follows: for any  $\phi \in \mathcal{T}$ ,

$$\phi_t(C) \leq \exp \left( -t \inf_{g \in C \cap M_1(D)} \sum_{y \in D} -\ln \left( \frac{1}{2d} \sum_{\substack{z \in \mathbb{Z}^d \\ |y-z|=1}} \exp(\phi(z) - \phi(y)) \right) g(y) \right). \quad (3.12)$$

We define a map  $\Phi$  on  $\mathcal{T} \times M_1(D)$  by

$$\Phi(\phi, g) = \sum_{y \in D} -\ln \left( \frac{1}{2d} \sum_{\substack{z \in \mathbb{Z}^d \\ |y-z|=1}} \exp(\phi(z) - \phi(y)) \right) g(y).$$

Optimizing the previous inequality (3.12) over the function  $\phi$ , we get

$$\phi_t(C) \leq \exp \left( -t \sup_{\phi \in \mathcal{T}} \inf_{g \in C \cap M_1(D)} \Phi(\phi, g) \right). \quad (3.13)$$

The map  $\Phi$  is linear in  $g$ . We shall next prove that it is convex in  $\phi$ . In fact, the convexity in  $\phi$  is a consequence of the convexity of the functions  $t \in \mathbb{R} \mapsto \exp(t)$  and  $t \in \mathbb{R}^+ \mapsto -\ln(t)$ . The delicate point is to check that the domain of definition of  $\Phi$  is convex. This is the purpose of the next lemma.

**Lemma 3.14** The set  $\mathcal{T}$  is convex.

**Proof.** Let  $\phi, \psi$  belong to  $\mathcal{T}$  and let  $\alpha, \beta \in ]0, 1[$  be such that  $\alpha + \beta = 1$ . There exist  $u, v \in \mathcal{S}^+$  such that  $\phi = \ln u, \psi = \ln v$ , whence

$$\alpha\phi + \beta\psi = \ln(u^\alpha v^\beta),$$

and we have to check that  $w = u^\alpha v^\beta$  is in  $\mathcal{S}^+$ . Let  $y$  be a fixed point in  $\mathbb{Z}^d \setminus D$ . We apply the discrete Hölder inequality to the functions  $u^\alpha$ ,  $v^\beta$ , with respect to the measure  $\nu_y$  which is the uniform measure on the neighbours of  $y$ ,

$$\nu_y = \frac{1}{2d} \sum_{\substack{z \in \mathbb{Z}^d \\ |y-z|=1}} \delta_z,$$

and with the exponents  $p = 1/\alpha$ ,  $q = 1/\beta$ . We obtain

$$\int u^\alpha v^\beta d\nu_y \leq \left( \int u d\nu_y \right)^\alpha \left( \int v d\nu_y \right)^\beta.$$

This inequality can be rewritten as

$$\frac{1}{2d} \sum_{\substack{z \in \mathbb{Z}^d \\ |y-z|=1}} u^\alpha(z) v^\beta(z) \leq \left( \frac{1}{2d} \sum_{\substack{z \in \mathbb{Z}^d \\ |y-z|=1}} u(z) \right)^\alpha \left( \frac{1}{2d} \sum_{\substack{z \in \mathbb{Z}^d \\ |y-z|=1}} v(z) \right)^\beta$$

We finally use the fact that  $u, v$  are superharmonic to conclude.  $\square$

Since the set  $D$  is finite, the set  $\ell^1(D)$  is finite dimensional, and the map  $\Phi$  is continuous with respect to  $g$  and any norm on  $\ell^1(D)$ . Similarly, for any  $\phi \in \mathcal{T}$  and  $g \in M_1(D)$ , the quantity  $\Phi(\phi, g)$  depends only on the values of  $\phi$  on the set  $\overline{D}$ , which is finite, thus the map  $\Phi$  is also continuous with respect to  $\phi$  and the  $\ell^1$  norm. Moreover the set  $C \cap M_1(D)$  is compact and convex (these are essential conditions in order to apply the minimax theorem). Therefore, by the famous minimax theorem (see [12]),

$$\sup_{\phi \in \mathcal{T}} \inf_{g \in C \cap M_1(D)} \Phi(\phi, g) = \inf_{g \in C \cap M_1(D)} \sup_{\phi \in \mathcal{T}} \Phi(\phi, g). \quad (3.15)$$

Let us fix  $g \in C \cap M_1(D)$  and let us bound from below  $\sup_{\phi \in \mathcal{T}} \Phi(\phi, g)$ . Let us fix  $\phi \in \mathcal{T}$ , we have, by convexity of  $-\ln$ ,

$$\Phi(\phi, g) \geq -\ln \left( \sum_{y \in D} \left( \frac{1}{2d} \sum_{\substack{z \in \mathbb{Z}^d \\ |y-z|=1}} \exp(\phi(z) - \phi(y)) \right) g(y) \right).$$

Using the inequality  $-\ln(t) \geq 1 - t$  for  $t > 0$ , we get

$$\Phi(\phi, g) \geq 1 - \sum_{y \in D} \left( \frac{1}{2d} \sum_{\substack{z \in \mathbb{Z}^d \\ |y-z|=1}} \exp(\phi(z) - \phi(y)) \right) g(y) \quad (3.16)$$

Let  $\varepsilon$  be such that

$$0 < \varepsilon < \min \{ g(x) : x \in D, g(x) > 0 \}.$$

We define next an adequate function  $u_\varepsilon$ . The entrance time  $T_0$  of  $D$  is

$$T_0 = \inf \{ n \geq 0 : S_n \in D \}.$$

We define

$$\forall x \in \mathbb{Z}^d \quad u_\varepsilon(x) = E_x \left( \sqrt{\max(g(S_{T_0}), \varepsilon)} \mid T_0 < \infty \right).$$

With this definition, we have that

$$\forall x \in D \quad u_\varepsilon(x) = \sqrt{\max(g(x), \varepsilon)}.$$

We claim that this function  $u_\varepsilon$  belongs to  $\mathcal{S}^+$ . Obviously it is strictly positive everywhere. The function  $u_\varepsilon$  is harmonic on  $\mathbb{Z}^d \setminus D$ , hence it is also superharmonic. We apply inequality (3.16) with the function  $\phi_\varepsilon(x) = \ln u_\varepsilon$  and we get

$$\begin{aligned} \sup_{\phi \in \mathcal{T}} \Phi(\phi, g) &\geq \Phi(\phi_\varepsilon, g) = 1 - \sum_{y \in D, g(y) > 0} \frac{1}{2d} \sum_{\substack{z \in \mathbb{Z}^d \\ |y-z|=1}} \frac{u_\varepsilon(z)}{u_\varepsilon(y)} g(y) \\ &= 1 - \sum_{y \in D, g(y) > 0} \frac{1}{2d} \sum_{\substack{z \in \mathbb{Z}^d \\ |y-z|=1}} \frac{u_\varepsilon(z)}{\sqrt{\max(g(y), \varepsilon)}} g(y) \\ &= 1 - \sum_{y \in D} \frac{1}{2d} \sum_{\substack{z \in \mathbb{Z}^d \\ |y-z|=1}} u_\varepsilon(z) \sqrt{g(y)}. \end{aligned} \tag{3.17}$$

This inequality holds for  $\varepsilon > 0$  sufficiently small. We send next  $\varepsilon$  to 0. We have

$$\forall x \in \mathbb{Z}^d \quad \lim_{\varepsilon \rightarrow 0} u_\varepsilon(x) = E_x \left( \sqrt{g(S_{T_0})} \mid T_0 < \infty \right).$$

Passing to the limit as  $\varepsilon$  goes to 0 in inequality (3.17), we get

$$\sup_{\phi \in \mathcal{T}} \Phi(\phi, g) \geq 1 - \sum_{y \in D} \frac{1}{2d} \sum_{\substack{z \in \mathbb{Z}^d \\ |y-z|=1}} E_z \left( \sqrt{g(S_{T_0})} \mid T_0 < \infty \right) \sqrt{g(y)}. \tag{3.18}$$

Let us introduce the hitting time  $T_1$  of  $D$  is

$$T_1 = \inf \{ n \geq 1 : S_n \in D \}.$$

Notice that the entrance and hitting time are equal whenever the starting point is in  $\mathbb{Z}^d \setminus D$ . However, if the starting point belongs to  $D$ , then  $T_0 = 0$  and  $T_1 \geq 1$ . With the help of  $T_1$ , we rewrite formula (3.18) as

$$\begin{aligned} \sup_{\phi \in \mathcal{T}} \Phi(\phi, g) &\geq 1 - \sum_{y \in D} E_y \left( \sqrt{g(S_{T_1})} \mid T_0 < \infty \right) \sqrt{g(y)} \\ &= 1 - \sum_{y \in D} \sum_{z \in D} P_y(S_{T_1} = z \mid T_0 < \infty) \sqrt{g(z)} \sqrt{g(y)}. \end{aligned} \tag{3.19}$$

Using the time reversibility of the random walk  $(S_n)_{n \in \mathbb{N}}$ , we have

$$\forall y, z \in D \quad P_y(S_{T_1} = z \mid T_0 < \infty) = P_z(S_{T_1} = y \mid T_0 < \infty), \tag{3.20}$$

whence

$$\sum_{y \in D} \sum_{z \in D} P_y(S_{T_1} = z \mid T_0 < \infty) g(z) = \sum_{z \in D} \sum_{y \in D} P_z(S_{T_1} = y \mid T_0 < \infty) g(z) = \sum_{z \in D} g(z) = 1. \quad (3.21)$$

The identities (3.21) yield that

$$1 - \sum_{y \in D} \sum_{z \in D} P_y(S_{T_1} = z \mid T_0 < \infty) \sqrt{g(z)} \sqrt{g(y)} = \frac{1}{2} \sum_{y \in D} \sum_{z \in D} P_y(S_{T_1} = z \mid T_0 < \infty) (\sqrt{g(z)} - \sqrt{g(y)})^2. \quad (3.22)$$

Reporting (3.22) in (3.19), we get

$$\sup_{\phi \in \mathcal{T}} \Phi(\phi, g) \geq \frac{1}{2} \sum_{y \in D} \sum_{z \in D} P_y(S_{T_1} = z \mid T_0 < \infty) (\sqrt{g(z)} - \sqrt{g(y)})^2. \quad (3.23)$$

Now, for any  $y, z \in D$  such that  $|y - z| = 1$ , we have

$$P_y(S_{T_1} = z \mid T_0 < \infty) \geq \frac{1}{2d},$$

thus we have the lower bound

$$\sup_{\phi \in \mathcal{T}} \Phi(\phi, g) \geq \sum_{\substack{y, z \in D \\ |y - z| = 1}} \frac{1}{2d} (\sqrt{g(z)} - \sqrt{g(y)})^2 = \frac{1}{2} \mathcal{E}(\sqrt{g}, D). \quad (3.24)$$

Taking the infimum with respect to  $g \in C \cap M_1(D)$  and coming back to inequality (3.13), we obtain the desired result.  $\square$

**Remark 3.25** To clarify the proof of Theorem 1.10 it may be useful to consider the *finite* Markov chain on  $D$  obtained from the simple random walk on  $\mathbb{Z}^d$  by restricting it to the times when it visits  $D$ . This is the Markov chain where the transition probabilities are given by

$$q(y, z) = P_y(S_{T_1} = z).$$

In particular, this chain coincides with the simple random walk on the vertices  $y$  for which all neighbours are in  $D$ . However all the vertices on the boundary of  $D$  are also connected to one another, and so this chain can jump from any boundary vertex to any other boundary vertex (corresponding to an excursion of the simple random walk away from  $D$ ).

Then note that the right hand side of (3.23) is nothing but the Dirichlet energy for this Markov chain. On the other hand, (3.24) shows that the Dirichlet energy for this Markov chain of a function  $g : D \rightarrow \mathbb{R}$  is always lower bounded by the Dirichlet energy  $\mathcal{E}(g; D)$  since we can simply restrict to the transitions on  $\mathbb{Z}^d$  and ignore the additional connections along the boundary.

Effectively, this argument reduces the simple random walk on the infinite state space  $\mathbb{Z}^d$  to a finite state space, in a way that is conceptually reminiscent of the compactification arguments to use Donsker–Varadhan large deviation estimates in Bolthausen’s work [6] and many other works on large deviations. However, the advantage of this approach is that it does not alter significantly the geometry of the ambient space and so eventually lets us use functional inequalities that are more readily available in  $\mathbb{R}^d$  than on a torus.

### 3.5 Correction for the origin

The problem with the inequality of Theorem 1.10 is the presence of the infimum over  $x \in \overline{D}$ . In order to go around it, we shall take advantage of the fact that our trajectories are constrained to have a range of cardinality less than  $cn^d$ . We first bound from below the probability to travel between an arbitrary point of the range and the origin. To this end, we shall use a standard estimate on multinomial coefficients, that we recall next.

**Lemma 3.26** For any  $k \geq 1$ ,  $r \geq 2$ , any  $k_1, \dots, k_r \in \{0, \dots, k\}$  such that  $k_1 + \dots + k_r = k$ , we have

$$\left| \frac{1}{k} \ln \frac{k!}{k_1! \dots k_r!} + \sum_{i=1}^r \frac{k_i}{k} \ln \frac{k_i}{k} \right| \leq \frac{r}{k} (\ln k + 2).$$

**Proof.** The proof of this estimate is standard (see for instance [11]). Setting, for  $k \in \mathbb{N}$ ,  $f(k) = \ln k! - k \ln k + k$ , we have

$$\begin{aligned} \ln \frac{k!}{k_1! \dots k_r!} &= \ln k! - \sum_{i=1}^r \ln k_i! \\ &= k \ln k - k + f(k) - \sum_{i=1}^r (k_i \ln k_i - k_i + f(k_i)) \\ &= - \sum_{i=1}^r k_i \ln \frac{k_i}{k} + f(k) - \sum_{i=1}^r f(k_i). \end{aligned}$$

Comparing the discrete sum

$$\ln k! = \sum_{1 \leq i \leq k} \ln i$$

to the integral  $\int_1^k \ln x dx$ , we see that  $1 \leq f(k) \leq \ln k + 2$  for all  $k \geq 1$ . On one hand,

$$f(k) - \sum_{i=1}^r f(k_i) \leq \ln k + 2 - r \leq r(\ln k + 2),$$

on the other hand,

$$f(k) - \sum_{i=1}^r f(k_i) \geq 1 - \sum_{i=1}^r (\ln k_i + 2) \geq 1 - 2r - r \ln k \geq -r(\ln k + 2)$$

and we have the desired inequalities. □

The box of side length  $r > 0$  centered at the origin is the set

$$\Lambda(r) = \{ (x_1, \dots, x_d) \in \mathbb{R}^d : \forall i \in \{1, \dots, d\} \quad -r/2 < x_i \leq r/2 \}.$$

With the help of Lemma 3.26, we obtain the following lower bound for the symmetric random walk.

**Lemma 3.27** Let  $c > 0$ . There exists  $c' > 0$  such that, for  $n$  large enough,

$$\forall k \in \{n^{d+1}, \dots, n^{d+2}\}, \quad \forall x \in \Lambda(cn^d), \quad P_x(S_k = 0) + P_x(S_{k-1} = 0) \geq \exp(-c'n^{2d}/k).$$

**Proof.** Let  $x = (x_1, \dots, x_d) \in \Lambda(cn^d)$  and let  $k \in \{n^{d+1}, \dots, n^{d+2}\}$ . We have

$$P_x(S_k = 0) = P_0(S_k = x) = \sum_{\substack{i_1, \dots, i_d \\ j_1, \dots, j_d}} \frac{1}{(2d)^k} \frac{k!}{i_1! j_1! \dots i_d! j_d!},$$

where the sum runs over the indices  $i_1, j_1, \dots, i_d, j_d$  such that

$$i_1 - j_1 = x_1, \dots, i_d - j_d = x_d, \quad i_1 + j_1 + \dots + i_d + j_d = k.$$

The index  $i_1$  corresponds to the number of moves associated to the vector  $(1, 0, \dots, 0)$ , the index  $j_1$  to the number of moves associated to the vector  $(-1, 0, \dots, 0)$ , and so on. This set of indices is empty in the case where  $k$  and  $x_1 + \dots + x_d$  don't have the same parity, that is why we have to consider the sum  $P_x(S_k = 0) + P_x(S_{k-1} = 0)$ . To simplify the discussion, we assume that the term  $P_x(S_k = 0)$  is non-zero. We will get the desired lower bound by considering only one term in the sum, the term corresponding to

$$i_1 = \frac{k}{2d} + \frac{x_1}{2}, j_1 = \frac{k}{2d} - \frac{x_1}{2}, \dots, i_d = \frac{k}{2d} + \frac{x_d}{2}, j_d = \frac{k}{2d} - \frac{x_d}{2}.$$

To be precise, we should take integer parts in the above formula, but this would become a bit messy, so we do as if all the fractions were integers. We get that

$$P_x(S_k = 0) \geq \frac{1}{(2d)^k} \frac{k!}{\left(\frac{k}{2d} + \frac{x_1}{2}\right)! \left(\frac{k}{2d} - \frac{x_1}{2}\right)! \dots \left(\frac{k}{2d} + \frac{x_d}{2}\right)! \left(\frac{k}{2d} - \frac{x_d}{2}\right)!}.$$

Taking the  $\ln$  and using the inequality of Lemma 3.26, we obtain

$$\begin{aligned} \ln P_x(S_k = 0) &\geq -k \ln(2d) \\ &\quad - \sum_{i=1}^d \left( \frac{k}{2d} + \frac{x_i}{2} \right) \ln \left( \frac{1}{2d} + \frac{x_i}{2k} \right) + \left( \frac{k}{2d} - \frac{x_i}{2} \right) \ln \left( \frac{1}{2d} - \frac{x_i}{2k} \right) - 2d(\ln k + 2) \\ &\geq - \sum_{i=1}^d \left( \frac{k}{2d} + \frac{x_i}{2} \right) \ln \left( 1 + \frac{dx_i}{k} \right) + \left( \frac{k}{2d} - \frac{x_i}{2} \right) \ln \left( 1 - \frac{dx_i}{k} \right) - 2d(\ln k + 2) \\ &\geq - \sum_{i=1}^d \left( \frac{k}{2d} + \frac{x_i}{2} \right) \left( \frac{dx_i}{k} \right) - \left( \frac{k}{2d} - \frac{x_i}{2} \right) \left( \frac{dx_i}{k} \right) - 2d(\ln k + 2) \\ &\geq - \sum_{i=1}^d \left( \frac{d}{k} \right) x_i^2 - 2d(\ln k + 2) \geq -d \left( \frac{d}{k} \right) (cn^d)^2 - 2d(\ln k + 2). \end{aligned}$$

Using the hypothesis on  $k$ , we conclude that

$$\ln P_x(S_k = 0) \geq -d^2 c^2 \left( \frac{n^{2d}}{k} \right) - 2d((d+2) \ln n + 2).$$

Since  $n^{2d}/k \geq n^{d-2} \geq n$ , then for  $n$  large enough, we obtain the desired estimate.  $\square$

We will use Lemma 3.27 together with the next proposition in order to have a deviation inequality for the trajectories starting from the origin. The idea is to let the random walk evolve naturally

over a certain time interval, so that it has a chance to visit all the points of  $\overline{D}$ , and in particular the point of  $\overline{D}$  realizing the infimum in the inequality of Theorem 1.10. Of course, we are forced to introduce a correcting factor, but we will adjust our parameters so that this correcting factor is not disturbing.

**Proposition 3.28** Let  $D$  be an arbitrary subset of  $\mathbb{Z}^d$  and let  $g$  be a function in  $\ell^1(D)$ . Let  $k, t$  be positive integers with  $k < t$ . Let  $r > 0$ . For any  $x \in \overline{D}$ , we have

$$P\left(\left\|\frac{1}{t}L_t^D - g^2\right\|_{1,D} \leq r, \tau < \infty\right) \leq \frac{2P_x\left(\left\|\frac{1}{t}L_t^D - g^2\right\|_{1,D} \leq r + 4\frac{k}{t-k}, \tau < \infty\right)}{P_x(S_k = 0) + P_x(S_{k-1} = 0)}.$$

**Proof.** Let  $s, t$  be such that  $0 < s < t$ . We have

$$\begin{aligned} \left\|\frac{1}{t}L_t^D - \frac{1}{s}L_s^D\right\|_{1,D} &\leq \left\|\frac{1}{t}L_t^D - \frac{1}{s}L_t^D\right\|_{1,D} + \left\|\frac{1}{s}L_t^D - \frac{1}{s}L_s^D\right\|_{1,D} \\ &\leq \left|\frac{1}{t} - \frac{1}{s}\right|t + \frac{t-s}{s} \leq 2\frac{t-s}{s}. \end{aligned} \quad (3.29)$$

Let  $x \in \mathbb{Z}^d$  be an arbitrary starting point. Let also  $r > 0$  and let  $g$  be an arbitrary function from  $D$  to  $\mathbb{R}$ . We condition on the state visited by the random walk at time  $t-s$ , we apply the strong Markov property and we use twice the inequality (3.29) to get:

$$\begin{aligned} P_x\left(\left\|\frac{1}{t}L_t^D - g^2\right\|_{1,D} \leq r, \tau < \infty\right) &\geq \sum_{y \in \mathbb{Z}^d} P_y\left(S_{t-s} = y, \left\|\frac{1}{s}(L_t^D - L_s^D) - g^2\right\|_{1,D} \leq r - 2\frac{t-s}{s}, \tau < \infty\right) \\ &\geq \sum_{y \in \mathbb{Z}^d} P_x(S_{t-s} = y) P_y\left(\left\|\frac{1}{s}L_s^D - g^2\right\|_{1,D} \leq r - 2\frac{t-s}{s}, \tau < \infty\right) \\ &\geq P_x(S_{t-s} = 0) P\left(\left\|\frac{1}{t}L_t^D - g^2\right\|_{1,D} \leq r - 4\frac{t-s}{s}, \tau < \infty\right). \end{aligned} \quad (3.30)$$

It might happen that  $P_x(S_{t-s} = 0)$  vanishes if  $|x|_1$  and  $t-s$  don't have the same parity. To avoid this nasty detail, we apply the inequality with  $s+1$  instead of  $s$ . Noting that

$$\frac{t-s-1}{s+1} \leq \frac{t-s}{s},$$

we get

$$P_x\left(\left\|\frac{1}{t}L_t^D - g^2\right\|_{1,D} \leq r, \tau < \infty\right) \geq P_x(S_{t-s-1} = 0) P\left(\left\|\frac{1}{t}L_t^D - g^2\right\|_{1,D} \leq r - 4\frac{t-s}{s}, \tau < \infty\right).$$

Summing this inequality and inequality (3.30), and setting  $k = t-s$ , we obtain, for any  $0 < k < t$ ,

$$\begin{aligned} 2P_x\left(\left\|\frac{1}{t}L_t^D - g^2\right\|_{1,D} \leq r, \tau < \infty\right) &\geq \\ &(P_x(S_k = 0) + P_x(S_{k-1} = 0)) P\left(\left\|\frac{1}{t}L_t^D - g^2\right\|_{1,D} \leq r - 4\frac{k}{t-k}, \tau < \infty\right). \end{aligned}$$

We finally make the change of variable  $r' = r - 4\frac{k}{t-k}$  and we obtain the desired inequality.  $\square$

## 4 A priori estimates

### 4.1 Control of the range

We shall obtain a control on the size of  $|R_N|$ . Let  $c > 0$ . We write

$$\tilde{P}_N(|R_N| > cn^d) \leq \frac{\exp(-cn^d)}{E(\exp(-|R_N|))}.$$

From theorem 1 of [10], there exists a positive constant  $k(1, d)$  such that

$$\lim_{N \rightarrow \infty} \frac{1}{n^d} \ln E(\exp(-|R_N|)) = -k(1, d).$$

We conclude that, for any  $c > 0$  and for  $N$  large enough, we have

$$\tilde{P}_N(|R_N| > cn^d) \leq \exp(-(c - 2k(1, d))n^d). \quad (4.1)$$

### 4.2 Control of the Dirichlet energy

We shall obtain a bound on  $\mathcal{E}(f_N)$  conditionally on the size of  $|R_N|$ .

**Proposition 4.2** Let  $c \geq 1$  and  $\kappa \geq 1$  be such that  $\kappa/2 - 4dc > 0$ . For  $n$  large enough, we have

$$P(\mathcal{E}(f_N) \geq \kappa n^d \ln n, |R_N| \leq cn^d) \leq \exp\left(-\left(\frac{\kappa}{2} - 4dc\right)n^d \ln n\right).$$

**Proof.** Let  $\lambda, c > 0$ . We write

$$P(\mathcal{E}(f_N) \geq \lambda n^d, |R_N| \leq cn^d) = \sum_{\phi} P(f_N = \phi),$$

where the summation extends over the profiles  $\phi : \mathbb{Z}^d \rightarrow [0, +\infty[$  such that

$$\phi(0) \geq 1, \quad \sum_{y \in \mathbb{Z}^d} \phi(y)^2 = N, \quad |\text{supp } \phi| \leq cn^d, \quad \mathcal{E}(\phi) \geq \lambda n^d.$$

For such a profile  $\phi$ , we have, thanks to the inequality of proposition 3.3:

$$P(f_N = \phi) \leq \frac{N}{\alpha} \exp\left(-\frac{\lambda}{2}n^d + \alpha\sqrt{Ncn^d}\right).$$

For the probability  $P(f_N = \phi)$  to be positive, it is necessary that  $\text{supp } \phi$  is a connected subset of  $\mathbb{Z}^d$  containing the origin, and that  $1 \leq \phi(y)^2 \leq N$  for all  $y \in \text{supp } \phi$ . The number of possible choices for the support of  $\phi$  is bounded by  $(c_d)^{cn^d}$ , where  $c_d$  is a constant depending on the dimension  $d$  only. Once the support is fixed, the number of choices for the profile  $\phi$  is bounded by  $N^{cn^d}$ . In the end, the total number of profiles  $\phi$  satisfying the previous constraints is bounded by  $(c_d N)^{cn^d}$ . Therefore

$$P(\mathcal{E}(f_N) \geq \lambda n^d, |R_N| \leq cn^d) \leq (c_d N)^{cn^d} \frac{N}{\alpha} \exp\left(-\frac{\lambda}{2}n^d + \alpha\sqrt{Ncn^d}\right).$$



We choose  $\alpha$  and  $\lambda$  of the form

$$\alpha = \frac{\ln n}{n}, \quad \lambda = \kappa \ln n.$$

Recalling that  $n^{d+2} = N$ , the previous inequality can be rewritten as

$$\begin{aligned} P(\mathcal{E}(f_N) \geq \kappa n^d \ln n, |R_N| \leq cn^d) &\leq \frac{n^{d+3}}{\ln n} \exp\left(cn^d(\ln c_d + (d+2)\ln n) - \frac{\kappa}{2}n^d \ln n + \sqrt{cn^d \ln n}\right) \\ &\leq \exp\left(-\left(\frac{\kappa}{2} - 4dc\right)n^d \ln n\right) \end{aligned}$$

where the last inequality holds for  $n$  large enough.  $\square$

### 4.3 High density blocks

We divide  $\mathbb{Z}^d$  into boxes called blocks in the following way. The box of side length  $r > 0$  centered at the origin is the set

$$\Lambda(r) = \{(x_1, \dots, x_d) \in \mathbb{R}^d : \forall i \in \{1, \dots, d\} \quad -r/2 < x_i \leq r/2\}.$$

Let  $n$  be a positive integer. For  $\underline{x} \in \mathbb{Z}^d$ , we define the block indexed by  $\underline{x}$  as

$$B(\underline{x}) = n\underline{x} + \Lambda(n).$$

Note that the blocks partition  $\mathbb{R}^d$ . We perform here a deterministic construction on any function  $f : \mathbb{Z}^d \rightarrow \mathbb{R}^+$  in order to record the blocks where the function  $f$  has a high density. This construction will be applied to build the coarse grained profile of the function  $f_N$ . Let  $f$  be a function from  $\mathbb{Z}^d$  to  $\mathbb{R}^+$ . Let  $n \geq 1, \delta > 0$ . We recall the notation  $2^* = 2d/(d-2)$  and we define

$$\underline{X}(f, \delta) = \left\{ \underline{x} \in \mathbb{Z}^d : \sum_{y \in B(\underline{x})} (f(y))^{2^*} \geq \delta^{2^*} n^{d+2^*} \right\}.$$

The exponent  $d+2^*$  corresponds to the typical situation for a block actively visited by the random walk until time  $N$ : the number of visits per site should be of order  $n^2$ , so  $f_N$  is of order  $n$  throughout the block. The blocks corresponding to vertices outside of  $\underline{X}(f, \delta)$  are low density blocks. Our goal is to control the total contribution of these blocks to the  $2^*$ -norm of  $f$ .

**Lemma 4.3** There exists a constant  $c_d$  depending on the dimension  $d$  only such that, for any function  $f$  from  $\mathbb{Z}^d$  to  $\mathbb{R}^+$ , we have

$$\sum_{\underline{x} \in \mathbb{Z}^d \setminus \underline{X}(f, \delta)} \sum_{y \in B(\underline{x})} (f(y))^{2^*} \leq c_d (\delta^{2^*} n^{d+2^*})^{1-2/2^*} \left( \frac{1}{n^2} (\|f\|_2)^2 + \mathcal{E}(f) \right).$$

**Proof.** Let  $f$  be a function from  $\mathbb{Z}^d$  to  $\mathbb{R}^+$ . Let  $\underline{x} \in \mathbb{Z}^d$ . We write

$$\sum_{y \in B(\underline{x})} (f(y))^{2^*} = \left( \sum_{y \in B(\underline{x})} (f(y))^{2^*} \right)^{1-2/2^*} \left( \sum_{y \in B(\underline{x})} (f(y))^{2^*} \right)^{2/2^*}. \quad (4.4)$$

To control the last factor, we apply the discrete Poincaré–Sobolev inequality stated in corollary B.1:

$$\begin{aligned} \left( \sum_{y \in B(\underline{x})} (f(y))^{2^*} \right)^{2/2^*} &\leq (c_{PS})^2 \left( \frac{1}{n} \|f\|_{2, B(\underline{x})} + \sqrt{\mathcal{E}(f, B(\underline{x}))} \right)^2 \\ &\leq 2(c_{PS})^2 \left( \frac{1}{n^2} (\|f\|_{2, B(\underline{x})})^2 + \mathcal{E}(f, B(\underline{x})) \right). \end{aligned} \quad (4.5)$$

If  $\underline{x}$  does not belong to  $\underline{X}(f, \delta)$ , then we have

$$\sum_{y \in B(\underline{x})} (f(y))^{2^*} < \delta^{2^*} n^{d+2^*}. \quad (4.6)$$

Plugging inequalities (4.5) and (4.6) into (4.4), we obtain

$$\sum_{y \in B(\underline{x})} (f(y))^{2^*} \leq (\delta^{2^*} n^{d+2^*})^{1-2/2^*} 2(c_{PS})^2 \left( \frac{1}{n^2} (\|f\|_{2, B(\underline{x})})^2 + \mathcal{E}(f, B(\underline{x})) \right).$$

Summing this inequality over the blocks outside of  $\underline{X}(f, \delta)$ , we get the estimate stated in the lemma.  $\square$

#### 4.4 Local averaging

In order to build a coarse grained image of the local time, we shall perform a local averaging. The averaging operation is deterministic, so we define it here for any function  $f : \mathbb{Z}^d \rightarrow \mathbb{R}$ . Let  $M$  be an integer which is a divisor of  $n$ . Let  $f$  be a function from  $\mathbb{Z}^d$  to  $\mathbb{R}$ . To  $f$  we associate a function  $\bar{f}^M$  obtained by performing a local average of  $f$  over boxes of side length  $M$ . Let us define the function  $\bar{f}^M$  precisely. For  $\underline{x} \in \mathbb{Z}^d$ , we define the block indexed by  $\underline{x}$  of side  $M$  as

$$B'(\underline{x}) = M\underline{x} + \Lambda(M).$$

We define a function  $\bar{f}^M$  from  $\mathbb{Z}^d$  to  $\mathbb{R}$  which is constant on the blocks  $B'(\underline{x})$ ,  $\underline{x} \in \mathbb{Z}^d$ , by setting

$$\forall \underline{x} \in \mathbb{Z}^d \quad \forall y \in B'(\underline{x}) \quad \bar{f}^M(y) = \frac{1}{|B'(\underline{x})|} \sum_{z \in B'(\underline{x})} f(z).$$

We have

$$\begin{aligned} \|\bar{f}^M\|_2^2 &= \sum_{x \in \mathbb{Z}^d} (\bar{f}^M(x))^2 \\ &= \sum_{\underline{x} \in \mathbb{Z}^d} |B'(\underline{x})| \left( \frac{1}{|B'(\underline{x})|} \sum_{z \in B'(\underline{x})} f(z) \right)^2 \\ &\leq \sum_{\underline{x} \in \mathbb{Z}^d} \sum_{z \in B'(\underline{x})} (f(z))^2 = \|f\|_2^2. \end{aligned} \quad (4.7)$$

We shall next bound the  $\ell^2$  norm of the difference between  $f$  and  $\bar{f}^M$ . Let  $\underline{x} \in \mathbb{Z}^d$ . We apply the discrete Poincaré–Wirtinger inequality stated in corollary B.2 to the function  $f$  and the block  $B'(\underline{x})$ :

$$(\|f - \bar{f}^M\|_{2, B'(\underline{x})})^2 \leq (M c_{PW})^2 \mathcal{E}(f, B'(\underline{x})).$$

We sum over  $\underline{x} \in \mathbb{Z}^d$ :

$$(\|f - \bar{f}^M\|_2)^2 \leq (M c_{PW})^2 \sum_{\underline{x} \in \mathbb{Z}^d} \mathcal{E}(f, B'(\underline{x})) \leq (M c_{PW})^2 \mathcal{E}(f).$$

Taking the square root, we conclude that

$$\|f - \bar{f}^M\|_2 \leq M c_{PW} \sqrt{\mathcal{E}(f)}. \quad (4.8)$$

## 5 Upper bound: proof of Theorem 2.6

We start here the proof of Theorem 2.6. Let  $c \geq 1$  and  $\kappa \geq 1$ . We consider the event

$$\mathcal{A} = \{ \mathcal{E}(f_N) \leq \kappa n^d \ln n, |R_N| \leq cn^d \}.$$

Let  $\mathcal{F}$  be a collection of functions from  $\mathbb{Z}^d$  to  $\mathbb{R}^+$ . The estimate on the range (4.1) and proposition 4.2 yield that, for  $n$  large enough,

$$\begin{aligned} E(e^{-|R_N|}; f_N \in \mathcal{F}) &\leq E(e^{-|R_N|}; f_N \in \mathcal{F}, \mathcal{A}) \\ &\quad + \exp(-(c - 2k(1, d))n^d) + \exp\left(-\left(\frac{\kappa}{2} - 4dc\right)n^d \ln n\right). \end{aligned} \quad (5.1)$$

From now onwards, we focus on estimating the term  $E(e^{-|R_N|}; f_N \in \mathcal{F}, \mathcal{A})$ . So we suppose that the event  $\mathcal{A}$  occurs and we deal only with trajectories of the random walk belonging to  $\mathcal{A}$ . Applying corollary B.3, we have

$$\|f_N\|_{2^*} \leq c_{PS} \sqrt{2d\mathcal{E}(f_N)} \leq c_{PS} \sqrt{2d\kappa n^d \ln n}. \quad (5.2)$$

We apply next the deterministic construction of section 4.3 to the function  $f_N$ . The corresponding set  $\underline{X}(f_N, \delta)$  satisfies:

$$|\underline{X}(f_N, \delta)| \delta^{2^*} n^{d+2^*} \leq (\|f_N\|_{2^*})^{2^*}. \quad (5.3)$$

Therefore, combining inequalities (5.2) and (5.3),

$$|\underline{X}(f_N, \delta)| \leq \frac{1}{\delta^{2^*} n^{d+2^*}} \left( c_{PS} \sqrt{2d\kappa n^d \ln n} \right)^{2^*} \leq \frac{1}{\delta^{2^*}} (c_{PS})^{2^*} (2d\kappa \ln n)^{2^*/2}. \quad (5.4)$$

In addition, since the range  $R_N$  is connected and has cardinality at most  $cn^d$ , then certainly the set  $\underline{X}(f_N, \delta)$  is included in the box  $\Lambda(3cn^{d-1})$ ; otherwise, the range  $R_N$  would contain a vertex which is outside the box  $\Lambda(3cn^d)$  (when coming back to the original lattice, the scale is multiplied by  $n$ ), and a connected set containing 0 which exits the box  $\Lambda(3cn^d)$  must contain a path of length at least  $3cn^d/2$ . We denote by  $\mathcal{X}(N, \delta)$  the collection of the subsets  $X$  of  $\mathbb{Z}^d$  satisfying these constraints. There exists a constant  $c_d$  depending on the dimension  $d$  only such that, for  $n$  large enough,

$$|\mathcal{X}(N, \delta)| \leq \exp(c_d C_0), \quad (5.5)$$

where

$$C_0 = \frac{1}{\delta^{2^*}} \kappa^4 (\ln n)^5. \quad (5.6)$$

We decompose the expectation according to the value  $X$  of  $\underline{X}$  in  $\mathcal{X}(N, \delta)$ :

$$E(e^{-|R_N|}; f_N \in \mathcal{F}, \mathcal{A}) = \sum_{X \in \mathcal{X}(N, \delta)} E(e^{-|R_N|}; f_N \in \mathcal{F}, \mathcal{A}, \underline{X}(f_N, \delta) = X). \quad (5.7)$$

We fix next  $X \in \mathcal{X}(N, \delta)$  and we shall estimate the probability appearing in the sum. We perform the local averaging on  $f_N$ , thereby getting the function  $\overline{f_N}^M$ . Using inequality (4.8), we have

$$\|f_N - \overline{f_N}^M\|_2 \leq M c_{PW} \sqrt{\mathcal{E}(f_N)} \leq M c_{PW} \sqrt{\kappa n^d \ln n}. \quad (5.8)$$

### 5.1 The coarse grained profile

We build here the coarse grained image of the local time. Since  $M$  is a divisor of  $n$ , then each block  $B(\underline{x})$ , for  $\underline{x} \in \mathbb{Z}^d$ , is the disjoint union of the blocks  $B'(\underline{y})$ ,  $\underline{y} \in \mathbb{Z}^d$ , which are included in it. Let us make this statement more precise. For  $\underline{x} \in \mathbb{Z}^d$ , we denote by  $Y(\underline{x})$  the subset of  $\mathbb{Z}^d$  defined by:

$$Y(\underline{x}) = \{ \underline{y} \in \mathbb{Z}^d : B'(\underline{y}) \subset B(\underline{x}) \}.$$

With this definition, we have

$$\forall \underline{x} \in \mathbb{Z}^d \quad B(\underline{x}) = \bigcup_{\underline{y} \in Y(\underline{x})} B'(\underline{y}).$$

We recall that the norm  $|\cdot|_\infty$  is defined by

$$\forall (x_1, \dots, x_d) \in \mathbb{R}^d \quad |(x_1, \dots, x_d)|_\infty = \max_{1 \leq i \leq d} |x_i|.$$

For  $X$  a subset of  $\mathbb{Z}^d$ , we define

$$\widehat{X} = X \cup \{ x \in \mathbb{Z}^d \setminus X : \exists y \in X \quad |x - y|_\infty = 1 \}.$$

We have the simple bound

$$|\widehat{X}| \leq 3^d |X|. \quad (5.9)$$

We define also

$$Y(X) = \bigcup_{\underline{x} \in X} Y(\underline{x}).$$

Since the blocks  $B'(\underline{x})$ ,  $\underline{x} \in \mathbb{Z}^d$ , are pairwise disjoint, we have

$$|Y(\widehat{X})| \times |\Lambda(M) \cap \mathbb{Z}^d| \leq |\widehat{X}| \times |\Lambda(n) \cap \mathbb{Z}^d|,$$

whence, using inequality (5.9),

$$|Y(\widehat{X})| \leq \frac{n^d}{M^d} |\widehat{X}| \leq \frac{3^d n^d}{M^d} |X|. \quad (5.10)$$

We take now  $X = \underline{X}(f_N, \delta)$ . The function  $\overline{f_N}^M$  is constant on each block  $B'(\underline{y})$ ,  $\underline{y} \in Y(\widehat{X})$ , and

$$\forall x \in \mathbb{Z}^d \quad 0 \leq \overline{f_N}^M(x) \leq \sup_{y \in \mathbb{Z}^d} f_N(y) \leq \sqrt{N}.$$

We shall work in the domain

$$D = \bigcup_{\underline{x} \in \widehat{X}} B(\underline{x}) = \bigcup_{\underline{y} \in Y(\widehat{X})} B'(\underline{y}). \quad (5.11)$$

We define also the set

$$E = \bigcup_{\underline{x} \in X} B(\underline{x}). \quad (5.12)$$

We apply Lemma 4.3 to  $f_N$ . Recalling that

$$(\|f_N\|_2)^2 = N = n^{d+2}, \quad \mathcal{E}(f_N) \leq \kappa n^d \ln n,$$

we obtain

$$\sum_{x \in \mathbb{Z}^d \setminus E} (f_N(x))^{2^*} \leq c_d (\delta^{2^*} n^{d+2^*})^{1-2/2^*} (n^d + \kappa n^d \ln n). \quad (5.13)$$

We discretize next the values of the functions  $\overline{f_N^M}$ . Let  $\eta > 0$ . We define the function  $\eta \overline{f_N^M}$  by setting

$$\forall x \in \mathbb{Z}^d \quad \eta \overline{f_N^M}(x) = \eta \left\lfloor \frac{1}{\eta} \overline{f_N^M}(x) \right\rfloor.$$

The function  $\eta \overline{f_N^M}$  restricted to  $D$ , denoted by  $\eta \overline{f_N^M}|_D$ , is the coarse-grained image of the function  $f_N$ . By construction, we have

$$\forall x \in \mathbb{Z}^d \quad |\overline{f_N^M}(x) - \eta \overline{f_N^M}(x)| \leq \eta,$$

whence, using inequalities (5.4) and (5.9) (recall that our  $\ell^p$  norms are unscaled, see the definition in formula (2.5)),

$$\begin{aligned} (\|\overline{f_N^M} - \eta \overline{f_N^M}\|_{2,D})^2 &\leq \eta^2 |\widehat{X}| n^d \\ &\leq \eta^2 3^d \frac{1}{\delta^{2^*}} (c_{PS})^{2^*} (\kappa \ln n)^{2^*/2} n^d. \end{aligned} \quad (5.14)$$

Putting together inequalities (5.8) and (5.14), we see that, on the event  $\underline{X}(f_N, \delta) = X$ , we have

$$\|f_N - \eta \overline{f_N^M}\|_{2,D} < \Gamma_0, \quad (5.15)$$

where

$$\Gamma_0 = M c_{PW} \sqrt{\kappa n^d \ln n} + \left( \eta^2 3^d \frac{1}{\delta^{2^*}} (c_{PS})^{2^*} (\kappa \ln n)^{2^*/2} n^d \right)^{1/2}. \quad (5.16)$$

The function  $\eta \overline{f_N^M}|_D$  belongs to the collection  $\mathcal{G}$  of the functions which are constant on each block  $B'(\underline{y})$ ,  $\underline{y} \in Y(\widehat{X})$ , and with values in the set

$$\left\{ k\eta : k \in \mathbb{N}, 0 \leq k < \frac{\sqrt{N}}{\eta} \right\}.$$

Using inequality (5.10), a simple upper bound on the cardinality of  $\mathcal{G}$  is given by

$$|\mathcal{G}| \leq \left( \frac{\sqrt{N}}{\eta} \right)^{|Y(\widehat{X})|} \leq \left( \frac{\sqrt{N}}{\eta} \right)^{\frac{3^d n^d}{M^d}} |X|. \quad (5.17)$$

Notice that the collection  $\mathcal{G}$  depends on  $f_N$  only through the set of blocks  $\underline{X}(f_N, \delta)$ . More precisely, once we know that  $\underline{X}(f_N, \delta) = X$ , then  $\mathcal{G}$  depends only on  $X$  and the parameters  $M, \eta$ , so we write  $\mathcal{G} = \mathcal{G}(X, M, \eta)$ . We come back to equation (5.7) and we decompose further the expectation as follows:

$$E(e^{-|R_N|}; f_N \in \mathcal{F}, \mathcal{A}, \underline{X}(f_N, \delta) = X) \leq \sum_{g \in \mathcal{G}(X, M, \eta)} E(e^{-|R_N|}; f_N \in \mathcal{F}, \mathcal{A}, \underline{X}(f_N, \delta) = X, \eta \overline{f_N}^M|_D = g). \quad (5.18)$$

Our large deviations inequality will involve the local time  $L_N$ , so we try to estimate  $\|L_N - g^2\|_{1,D}$  once we know that  $\eta \overline{f_N}^M|_D = g$ . Let  $g \in \mathcal{G}(X, M, \eta)$ . By the Cauchy-Schwarz inequality,

$$\|L_N - g^2\|_{1,D} = \|f_N^2 - g^2\|_{1,D} \leq \|f_N - g\|_{2,D} \|f_N + g\|_{2,D}.$$

Suppose that  $\eta \overline{f_N}^M|_D = g$ . Thanks to inequality (5.15), we have

$$\|g\|_{2,D} \leq \|f_N\|_{2,D} + \|g - f_N\|_{2,D} \leq \sqrt{N} + \Gamma_0. \quad (5.19)$$

Setting  $\Gamma_1 = \Gamma_0(2\sqrt{N} + \Gamma_0)$ , we deduce from the previous inequalities that

$$\|L_N - g^2\|_{1,D} \leq \Gamma_1. \quad (5.20)$$

Inequality (5.18) and the above inequalities yield that

$$E(e^{-|R_N|}; f_N \in \mathcal{F}, \mathcal{A}, \underline{X}(f_N, \delta) = X) \leq \sum_{\substack{g \in \mathcal{G}(X, M, \eta) \\ \|g\|_{2,D} \leq \sqrt{N} + \Gamma_0}} E(e^{-|R_N|}; f_N \in \mathcal{F}, \mathcal{A}, \underline{X}(f_N, \delta) = X, \|L_N - g^2\|_{1,D} < \Gamma_1). \quad (5.21)$$

## 5.2 Continuation of proof of Theorem 2.6

We still fix  $g \in \mathcal{G}(X, M, \eta)$ . We come back to equation (5.21) and we estimate the expectation

$$E(e^{-|R_N|}; f_N \in \mathcal{F}, \mathcal{A}, \underline{X}(f_N, \delta) = X, \|L_N - g^2\|_{1,D} < \Gamma_1).$$

We need first to bound from below  $|R_N|$  when  $L_N$  is close to  $g^2$ . Let  $\lambda > 0$ . If  $x \in D$  is such that  $g^2(x) \geq \lambda$  and  $|L_N(x) - g^2(x)| < \lambda/2$ , then certainly  $L_N(x) > 0$ . Therefore

$$\begin{aligned} |R_N| &\geq |\{x \in D : g^2(x) \geq \lambda, |L_N(x) - g^2(x)| < \lambda/2\}| \\ &\geq |\{x \in D : g^2(x) \geq \lambda\}| - |\{x \in D : |L_N(x) - g^2(x)| \geq \lambda/2\}|. \end{aligned} \quad (5.22)$$

Moreover, by Markov's inequality,

$$|\{x \in D : |L_N(x) - g^2(x)| \geq \lambda/2\}| \leq \frac{2}{\lambda} \|L_N - g^2\|_{1,D}.$$

We conclude that

$$\begin{aligned} E(e^{-|R_N|}; f_N \in \mathcal{F}, \mathcal{A}, \underline{X}(f_N, \delta) = X, \|L_N - g^2\|_{1,D} < \Gamma_1) &\leq \\ &\exp\left(-|\{x \in D : g^2(x) \geq \lambda\}| + \frac{2\Gamma_1}{\lambda}\right) \times P(\mathcal{B}), \end{aligned} \quad (5.23)$$

where  $\mathcal{B}$  is the event defined by

$$\mathcal{B} = \left\{ f_N \in \mathcal{F}, \mathcal{A}, \underline{X}(f_N, \delta) = X, \|L_N - g^2\|_{1,D} < \Gamma_1 \right\}. \quad (5.24)$$

It remains to estimate the probability  $P(\mathcal{B})$ , for which we ultimately wish to use Theorem 1.10. Recall the set  $E$  from (5.12). Then we have

$$\|L_N\|_{1,D} \geq \|L_N\|_{1,E} = \|L_N\|_1 - \|L_N\|_{1,\mathbb{Z}^d \setminus E}. \quad (5.25)$$

We estimate  $\|L_N\|_{1,\mathbb{Z}^d \setminus E}$  in the next lemma.

**Lemma 5.26** For  $n$  large enough, we have  $\|L_N\|_{1,\mathbb{Z}^d \setminus E} \leq \Gamma_2$  where

$$\Gamma_2 = c \delta^{4/d} n^{d+2} (c_d \kappa \ln n)^{2/2^*}.$$

**Proof.** We use Hölder's inequality with the exponents  $2^*/2$  and  $d/2$  to write

$$\|L_N\|_{1,\mathbb{Z}^d \setminus E} \leq \sum_{\underline{x} \in \mathbb{Z}^d \setminus X} \sum_{y \in B(\underline{x})} (f_N(y))^2 \leq \left( \sum_{\underline{x} \in \mathbb{Z}^d \setminus X} \sum_{y \in B(\underline{x})} (f_N(y))^{2^*} \right)^{2/2^*} |\text{supp } f_N \setminus E|^{1-2/2^*}.$$

Moreover, on the event  $\mathcal{A}$ , we have

$$|\text{supp } f_N \setminus E| \leq |\text{supp } f_N| \leq cn^d.$$

Using also inequality (5.13), we conclude that  $\|L_N\|_{1,\mathbb{Z}^d \setminus E} \leq \Gamma_2$ .  $\square$

It follows from inequality (5.25) and Lemma 5.26 that

$$P(\mathcal{B}) = P(\mathcal{B}, \|L_N\|_{1,D} \geq N - \Gamma_2) = \sum_{N - \Gamma_2 \leq t \leq N} P(\mathcal{B}, \|L_N\|_{1,D} = t).$$

Now, if  $\|L_N\|_{1,D} = t$ , then  $\tau(D, t) < \infty$  and  $L_N(y) = L_t^D(y)$  for any  $y \in D$ , thus

$$\|L_N - g^2\|_{1,D} = \|L_t^D - g^2\|_{1,D},$$

and we obtain the bound

$$P(\mathcal{B}) \leq \sum_{N - \Gamma_2 \leq t \leq N} P(f_N \in \mathcal{F}, \mathcal{A}, \underline{X}(f_N, \delta) = X, \|L_t^D - g^2\|_{1,D} < \Gamma_1, \tau < \infty). \quad (5.27)$$

Let us fix  $t$  such that

$$N - \Gamma_2 \leq t \leq N. \quad (5.28)$$

The time has now come to make specific choices for the parameters  $\delta, M, \eta, \lambda$  introduced in the course of the proof (recall that  $\lambda$  was introduced in (5.22)). We suppose that

$$\delta = \frac{1}{n^\alpha}, \quad M = n^\beta, \quad \eta = \frac{1}{n^\gamma}, \quad \lambda = n^\rho, \quad (5.29)$$

where  $\alpha, \beta, \gamma, \rho$  are positive exponents, which satisfy furthermore

$$\begin{aligned} \beta < 1, \quad \frac{2^*}{2}\alpha - \gamma < \beta, \quad 2^*\alpha < d\beta, \\ 1 + \beta < \rho, \quad 2 - 4\alpha/d < \rho, \quad \rho < 2. \end{aligned} \quad (5.30)$$

These conditions imply that, as  $N$  or  $n$  goes to  $\infty$ ,

$$\Gamma_0 \stackrel{\ln}{\sim}_{n \rightarrow \infty} n^{d/2+\beta}, \quad \Gamma_1 \stackrel{\ln}{\sim}_{n \rightarrow \infty} n^{d+1+\beta}, \quad \Gamma_2 \stackrel{\ln}{\sim}_{n \rightarrow \infty} n^{d+2-4\alpha/d}, \quad (5.31)$$

where  $\sim_{n \rightarrow \infty}^{\ln}$  means that the logarithms are equivalent. In particular, we have

$$\frac{\Gamma_0}{\sqrt{N}} \rightarrow 0, \quad \frac{\Gamma_1}{N} \rightarrow 0, \quad \frac{\Gamma_2}{N} \rightarrow 0, \quad \frac{n^{d+1}}{\Gamma_1} \rightarrow 0, \quad (5.32)$$

whence for  $n$  large enough

$$\Gamma_0 < \sqrt{N}, \quad \Gamma_1 < 3\Gamma_0\sqrt{N}, \quad n^{d+1} < \Gamma_1 < n^{d+2}, \quad \Gamma_2 < N/2. \quad (5.33)$$

We apply Proposition 3.28 with  $r = \Gamma_1/t$ , and  $k = \lfloor \Gamma_1 \rfloor$ : for any  $x \in \bar{D}$ , we have

$$\begin{aligned} P(\|L_t^D - g^2\|_{1,D} < \Gamma_1, \tau < \infty) &= P\left(\left\|\frac{1}{t}L_t^D - \frac{1}{t}g^2\right\|_{1,D} < \frac{\Gamma_1}{t}, \tau < \infty\right) \\ &\leq \frac{2P_x\left(\left\|\frac{1}{t}L_t^D - \frac{1}{t}g^2\right\|_{1,D} \leq \frac{\Gamma_1}{t} + 4\frac{\Gamma_1}{t-\Gamma_1}, \tau < \infty\right)}{P_x(S_{\lfloor \Gamma_1 \rfloor} = 0) + P_x(S_{\lfloor \Gamma_1 \rfloor - 1} = 0)}. \end{aligned}$$

Moreover, using (5.28) and (5.32), for  $n$  large enough,

$$\frac{\Gamma_1}{t} + 4\frac{\Gamma_1}{t-\Gamma_1} \leq \frac{5\Gamma_1}{t-\Gamma_1} \leq \frac{7\Gamma_1}{t}.$$

Since  $|R_N| \leq cn^d$ , then the range  $R_N$  is included in the box  $\Lambda(2cn^d)$  and its trace on the renormalized lattice is included in the box  $\Lambda(2cn^{d-1})$ . The set  $\bar{D}$  is obtained by enlarging slightly this trace and then coming back to the original lattice, hence it is included in the box  $\Lambda(3cn^d)$ . We bound the denominator with the help of Lemma 3.27 (notice that  $n^{d+1} < \Gamma_1 < n^{d+2}$ ) and we take the infimum over  $x \in \bar{D}$ :

$$P(\|L_t^D - g^2\|_{1,D} < \Gamma_1, \tau < \infty) \leq 2 \exp\left(c' \frac{n^{2d}}{\lfloor \Gamma_1 \rfloor}\right) \inf_{x \in \bar{D}} P_x\left(\left\|\frac{1}{t}L_t^D - \frac{1}{t}g^2\right\|_{1,D} < \frac{7\Gamma_1}{t}, \tau < \infty\right). \quad (5.34)$$

For  $k \in \mathbb{N}$  and  $r > 0$ , we define the closed convex set

$$C(g, k, r) = \left\{ h \in \ell^1(D) : \left\| h - \frac{1}{k}g^2 \right\|_{1,D} \leq r \right\}.$$

We apply the deviation inequality of Theorem 1.10 to the set  $C(g, t, 7\Gamma_1/t)$ :

$$\inf_{x \in \bar{D}} P_x\left(\left\|\frac{1}{t}L_t^D - \frac{1}{t}g^2\right\|_{1,D} \leq \frac{7\Gamma_1}{t}, \tau < \infty\right) \leq \exp\left(-t \inf_{h \in C(g, t, 7\Gamma_1/t)} \frac{1}{2} \mathcal{E}(\sqrt{h}, D)\right). \quad (5.35)$$

The previous inequalities (5.34) and (5.35) yield that

$$P(\|L_t^D - g^2\|_{1,D} < \Gamma_1, \tau < \infty) \leq 2 \exp\left(c' \frac{n^{2d}}{\lfloor \Gamma_1 \rfloor} - t \inf_{h \in C(g, t, 7\Gamma_1/t)} \frac{1}{2} \mathcal{E}(\sqrt{h}, D)\right). \quad (5.36)$$



We deal next with the infimum in the exponential. Our first goal is to obtain a bound which depends on  $N$  (and not on  $t$ ). Let  $h \in C(g, t, 7\Gamma_1/t)$ . We have

$$\begin{aligned} \left\| h - \frac{1}{N}g^2 \right\|_{1,D} &\leq \left\| h - \frac{1}{t}g^2 \right\|_{1,D} + \left\| \frac{1}{t}g^2 - \frac{1}{N}g^2 \right\|_{1,D} \\ &\leq \frac{7\Gamma_1}{t} + \left| \frac{1}{t} - \frac{1}{N} \right| \|g^2\|_{1,D}. \end{aligned}$$

Now, from (5.28), we have

$$\left| \frac{1}{t} - \frac{1}{N} \right| = \frac{N-t}{Nt} \leq \frac{\Gamma_2}{N(N-\Gamma_2)}.$$

Therefore, using the inequalities (5.19), (5.28) and (5.33), we get

$$\begin{aligned} \left\| h - \frac{1}{N}g^2 \right\|_{1,D} &\leq \frac{7\Gamma_1}{t} + \frac{\Gamma_2}{N(N-\Gamma_2)}(\Gamma_0 + \sqrt{N})^2 \\ &\leq \frac{7\Gamma_1}{N-\Gamma_2} + \frac{4\Gamma_2}{N-\Gamma_2} \leq \frac{14\Gamma_1}{N} + \frac{8\Gamma_2}{N}. \end{aligned} \quad (5.37)$$

Setting

$$r(N) = \frac{14\Gamma_1}{N} + \frac{8\Gamma_2}{N},$$

we conclude that

$$C\left(g, t, \frac{7\Gamma_1}{t}\right) \subset C(g, N, r(N)).$$

Recalling that  $t \geq N - \Gamma_2$ , inequality (5.36) now implies

$$\begin{aligned} P(\|L_t^D - g^2\|_{1,D} < \Gamma_1, \tau < \infty) &\leq \\ &2 \exp\left(c' \frac{n^{2d}}{[\Gamma_1]} - (N - \Gamma_2) \inf \left\{ \frac{1}{2} \mathcal{E}(\sqrt{h}, D) : h \in C(g, N, r(N)) \right\}\right). \end{aligned}$$

The good point is that this upper bound does not depend any more on  $t$ . Plugging this upper bound in equations (5.23) and (5.27), we obtain

$$\begin{aligned} E(e^{-|R_N|}; \mathcal{B}) &\leq 2(\Gamma_2 + 1) \exp\left(-|\{x \in D : g^2(x) \geq \lambda\}| + \frac{2\Gamma_1}{\lambda}\right. \\ &\quad \left.+ c' \frac{n^{2d}}{[\Gamma_1]} - (N - \Gamma_2) \inf \left\{ \frac{1}{2} \mathcal{E}(\sqrt{h}, D) : h \in C(g, N, r(N)) \right\}\right). \end{aligned} \quad (5.38)$$

Our next goal is to remove the dependence on  $g$  in the upper bound. We wish to obtain an upper bound which depends only on the set  $\mathcal{F}$ . To do so, we shall control the Dirichlet energy  $\mathcal{E}(\sqrt{h}, D)$  restricted to  $D$  with the help of the Dirichlet energy  $\mathcal{E}(\sqrt{h})$  in the whole space. Of course, this creates a correcting factor, which we study in the next section.

### 5.3 Truncation

We first define a truncation operator associated to the set  $X$ . Let  $\phi : \mathbb{R}^d \rightarrow [0, 1]$  be a function such that

- $\text{supp } \phi \subset \Lambda(7/4)$ ,

- $\phi$  is piecewise affine on  $\Lambda(7/4) \setminus \Lambda(5/4)$ ,
- $\phi$  is equal to 1 on  $\Lambda(5/4)$ ,
- the gradient of  $\phi$  has Euclidean norm less than 4 on  $\Lambda(7/4) \setminus \Lambda(5/4)$ .

To the set  $X$ , we associate the function  $\phi_X$  defined by

$$\forall y \in \mathbb{Z}^d \quad \phi_X(y) = \max \{ \phi(y/n - x) : x \in X \}.$$

By construction, we have

$$\text{supp } \phi_X \subset \bigcup_{x \in X} \Lambda(nx, 2n) \subset D, \quad (5.39)$$

where the domain  $D$  was defined in (5.11). Let  $f$  be any function in  $\ell^1(\mathbb{Z}^d)$ . We define the function  $\phi_X f$  by setting

$$\forall y \in \mathbb{Z}^d \quad (\phi_X f)(y) = \phi_X(y) f(y).$$

If  $f$  is a function in  $\ell^1(D)$ , we extend  $f$  outside  $D$  by setting  $f(x) = 0$  for  $x \in \mathbb{Z}^d \setminus D$  and the previous definition still makes sense. We recall that the set  $E$  is defined in (5.12).

**Proposition 5.40** Let  $f$  be any function in  $\ell^1(D)$ . We have the inequality

$$\mathcal{E}(f, D) \geq \left( \max \left( \sqrt{\mathcal{E}(\phi_X f)} - \frac{4}{n} \|f\|_{2, D \setminus E}, 0 \right) \right)^2.$$

**Proof.** Since the support of  $\phi_X f$  is included in  $D$  (see the inclusion (5.39)), then we have

$$\begin{aligned} \mathcal{E}(\phi_X f) &= \mathcal{E}(\phi_X f, D) = \frac{1}{2d} \sum_{\substack{y, z \in D \\ |y-z|=1}} \left( \phi_X(y) f(y) - \phi_X(z) f(z) \right)^2 \\ &= \frac{1}{2d} \sum_{\substack{y, z \in D \\ |y-z|=1}} \left( \phi_X(y) (f(y) - f(z)) - (\phi_X(z) - \phi_X(y)) f(z) \right)^2 \\ &= \frac{1}{2d} \sum_{\substack{y, z \in D \\ |y-z|=1}} \left( (\phi_X(y))^2 (f(y) - f(z))^2 + (\phi_X(z) - \phi_X(y))^2 (f(z))^2 \right. \\ &\quad \left. - 2\phi_X(y) (f(y) - f(z)) (\phi_X(z) - \phi_X(y)) f(z) \right). \end{aligned}$$

From the definition of  $\phi_X$ , it follows that if  $|y - z| = 1$ , then  $|\phi_X(z) - \phi_X(y)| \leq 4/n$ , and  $\phi_X(z) = \phi_X(y)$  unless both  $y, z$  are in  $D \setminus E$ . Therefore we have the following bound:

$$\mathcal{E}(\phi_X f) \leq \mathcal{E}(f, D) + \left( \frac{4}{n} \right)^2 \sum_{z \in D \setminus E} f(z)^2 + \frac{4}{nd} \sum_{\substack{y, z \in D \setminus E \\ |y-z|=1}} |f(y) - f(z)| |f(z)|.$$

Using the Cauchy–Schwarz inequality, we get

$$\begin{aligned} \sum_{\substack{y, z \in D \setminus E \\ |y-z|=1}} |f(y) - f(z)| |f(z)| &\leq \left( \sum_{\substack{y, z \in D \setminus E \\ |y-z|=1}} (f(y) - f(z))^2 \sum_{\substack{y, z \in D \setminus E \\ |y-z|=1}} (f(z))^2 \right)^{\frac{1}{2}} \\ &\leq \sqrt{2d \mathcal{E}(f, D \setminus E)} \sqrt{2d} \|f\|_{2, D \setminus E} = 2d \sqrt{\mathcal{E}(f, D)} \|f\|_{2, D \setminus E}. \end{aligned}$$

Reporting in the previous inequality, we have

$$\begin{aligned}\mathcal{E}(\phi_X f) &\leq \mathcal{E}(f, D) + \frac{16}{n^2} (\|f\|_{2, D \setminus E})^2 + \frac{8}{n} \sqrt{\mathcal{E}(f, D)} \|f\|_{2, D \setminus E} \\ &= \left( \sqrt{\mathcal{E}(f, D)} + \frac{4}{n} \|f\|_{2, D \setminus E} \right)^2,\end{aligned}$$

from which we deduce easily the inequality stated in the proposition.  $\square$

#### 5.4 End of proof of Theorem 2.6

We come back to inequality (5.38). We wish to obtain an upper bound which depends on the set  $\mathcal{F}$  and not on  $g$ . We know that  $f_N$  belongs to  $\mathcal{F}$ , so we pick a function  $h$  in  $C(g, N, r(N))$  and we try to control the distance between  $\frac{1}{N} L_N$  and  $\phi_X^2 h$ . We write

$$\left\| \frac{1}{N} L_N - \phi_X^2 h \right\|_1 \leq \left\| \frac{1}{N} L_N - \frac{1}{N} \phi_X^2 L_N \right\|_1 + \left\| \frac{1}{N} \phi_X^2 L_N - \phi_X^2 h \right\|_1$$

and we control separately each term. For the first term, we use the fact that  $\phi_X$  is equal to 1 on  $E$  and lemma 5.26 to get

$$\left\| \frac{1}{N} L_N - \frac{1}{N} \phi_X^2 L_N \right\|_1 \leq \left\| \frac{1}{N} L_N \right\|_{1, \mathbb{Z}^d \setminus E} \leq \frac{\Gamma_2}{N}. \quad (5.41)$$

For the second term, we have, recalling that  $\text{supp } \phi_X \subset D$  and the definition of  $\mathcal{B}$  in (5.24), and using (5.37),

$$\begin{aligned}\left\| \frac{1}{N} \phi_X^2 L_N - \phi_X^2 h \right\|_1 &\leq \left\| \frac{1}{N} L_N - h \right\|_{1, D} \\ &\leq \left\| \frac{1}{N} L_N - \frac{1}{N} g^2 \right\|_{1, D} + \left\| \frac{1}{N} g^2 - h \right\|_{1, D} \\ &\leq \frac{\Gamma_1}{N} + r(N).\end{aligned} \quad (5.42)$$

Inequalities (5.41) and (5.42) together yield

$$\left\| \frac{1}{N} L_N - \phi_X^2 h \right\|_1 \leq \frac{\Gamma_2}{N} + \frac{\Gamma_1}{N} + r(N). \quad (5.43)$$

Furthermore, we have, by inequalities (5.20) and (5.43),

$$\begin{aligned}\left\| \frac{1}{N} g^2 - \phi_X^2 h \right\|_{1, D} &\leq \left\| \frac{1}{N} g^2 - \frac{1}{N} L_N \right\|_{1, D} + \left\| \frac{1}{N} L_N - \phi_X^2 h \right\|_{1, D} \\ &\leq \frac{\Gamma_1}{N} + \frac{\Gamma_2}{N} + \frac{\Gamma_1}{N} + r(N).\end{aligned} \quad (5.44)$$

Let us set

$$\Gamma_3 = \frac{\Gamma_2}{N} + 2 \frac{\Gamma_1}{N} + r(N) \leq \frac{16\Gamma_1}{N} + \frac{9\Gamma_2}{N}. \quad (5.45)$$

We work next on the first term in the exponential appearing in (5.38). For any  $f \in \ell^1(D)$ , we have

$$\begin{aligned}\left| \left\{ x \in D : g^2(x) \geq \lambda \right\} \right| &\geq \left| \left\{ x \in D : f(x) \geq \frac{2\lambda}{N}, |f(x) - \frac{g^2(x)}{N}| < \frac{\lambda}{N} \right\} \right| \\ &\geq \left| \left\{ x \in D : f(x) \geq \frac{2\lambda}{N} \right\} \right| - \left| \left\{ x \in D : |f(x) - \frac{1}{N} g^2(x)| \geq \frac{\lambda}{N} \right\} \right|.\end{aligned}$$

Moreover, by Markov's inequality,

$$\left| \left\{ x \in D : \left| f(x) - \frac{1}{N} g^2(x) \right| \geq \frac{\lambda}{N} \right\} \right| \leq \frac{N}{\lambda} \left\| f - \frac{1}{N} g^2 \right\|_{1,D},$$

whence

$$\left| \left\{ x \in D : g^2(x) \geq \lambda \right\} \right| \geq \left| \left\{ x \in D : f(x) \geq \frac{2\lambda}{N} \right\} \right| - \frac{N}{\lambda} \left\| f - \frac{1}{N} g^2 \right\|_{1,D}.$$

Let  $h \in C(g, N, r(N))$  and let us apply this inequality with  $f = \phi_X^2 h$ . Together with inequality (5.44), we obtain

$$\left| \left\{ x \in D : g^2(x) \geq \lambda \right\} \right| \geq \left| \left\{ x \in D : (\phi_X^2 h)(x) \geq \frac{2\lambda}{N} \right\} \right| - \frac{N}{\lambda} \Gamma_3.$$

Taking the supremum over  $h$  in  $C(g, N, r(N))$ , we get

$$\left| \left\{ x \in D : g^2(x) \geq \lambda \right\} \right| \geq \sup \left\{ \left| \left\{ x \in D : (\phi_X^2 h)(x) \geq \frac{2\lambda}{N} \right\} \right| : h \in C(g, N, r(N)) \right\} - \frac{N}{\lambda} \Gamma_3.$$

Plugging this into inequality (5.38), we arrive at

$$\begin{aligned} E(e^{-|R_N|}; \mathcal{B}) &\leq 2(\Gamma_2 + 1) \exp \left( \frac{N}{\lambda} \Gamma_3 + \frac{2\Gamma_1}{\lambda} + c' \frac{n^{2d}}{[\Gamma_1]} \right. \\ &\quad \left. - \sup \left\{ \left| \left\{ x \in D : (\phi_X^2 h)(x) \geq \frac{2\lambda}{N} \right\} \right| : h \in C(g, N, r(N)) \right\} \right. \\ &\quad \left. - (N - \Gamma_2) \inf \left\{ \frac{1}{2} \mathcal{E}(\sqrt{h}, D) : h \in C(g, N, r(N)) \right\} \right) \\ &\leq 2(\Gamma_2 + 1) \exp \left( \frac{N}{\lambda} \Gamma_3 + \frac{2\Gamma_1}{\lambda} + c' \frac{n^{2d}}{[\Gamma_1]} \right. \\ &\quad \left. - \inf \left\{ \left| \left\{ x \in D : (\phi_X^2 h)(x) \geq \frac{2\lambda}{N} \right\} \right| + \frac{N - \Gamma_2}{2} \mathcal{E}(\sqrt{h}, D) : h \in C(g, N, r(N)) \right\} \right). \end{aligned}$$

Let again  $h \in C(g, N, r(N))$ . We apply Proposition 5.40 to the function  $f = \sqrt{h}$ :

$$\mathcal{E}(\sqrt{h}, D) \geq \left( \max \left( \sqrt{\mathcal{E}(\phi_X \sqrt{h})} - \frac{4}{n} \|\sqrt{h}\|_{2, D \setminus E}, 0 \right) \right)^2.$$

Now, using inequalities (5.41), (5.42) and (5.45), we get

$$\left( \|\sqrt{h}\|_{2, D \setminus E} \right)^2 = \|h\|_{1, D \setminus E} \leq \left\| h - \frac{1}{N} L_N \right\|_{1, D \setminus E} + \left\| \frac{1}{N} L_N \right\|_{1, \mathbb{Z}^d \setminus E} \leq \frac{\Gamma_1}{N} + r(N) + \frac{\Gamma_2}{N} \leq \Gamma_3. \quad (5.46)$$

We conclude that

$$\begin{aligned} E(e^{-|R_N|}; \mathcal{B}) &\leq 2(\Gamma_2 + 1) \exp \left( \frac{N}{\lambda} \Gamma_3 + \frac{2\Gamma_1}{\lambda} + c' \frac{n^{2d}}{[\Gamma_1]} \right. \\ &\quad \left. - \inf \left\{ \left| \left\{ x \in D : (\phi_X^2 h)(x) \geq \frac{2\lambda}{N} \right\} \right| + \right. \right. \\ &\quad \left. \left. \frac{N - \Gamma_2}{2} \left( \max \left( \sqrt{\mathcal{E}(\phi_X \sqrt{h})} - \frac{4\sqrt{\Gamma_3}}{n}, 0 \right) \right)^2 : h \in C(g, N, r(N)) \right\} \right). \quad (5.47) \end{aligned}$$

Our next goal is to remove the term  $2\lambda/N$  appearing in the infimum. Let us fix again  $h \in C(g, N, r(N))$ . We define an auxiliary function  $\tilde{h}$  by setting

$$\forall x \in \mathbb{Z}^d \quad \tilde{h}(x) = \left( \max \left( \phi_X \sqrt{h}(x) - \sqrt{\frac{2\lambda}{N}}, 0 \right) \right)^2.$$

Let also define

$$A = \left\{ x \in \mathbb{Z}^d : \phi_X \sqrt{h}(x) > \sqrt{\frac{2\lambda}{N}} \right\} = \left\{ x \in \mathbb{Z}^d : \tilde{h}(x) > 0 \right\}.$$

Obviously, we have  $A \subset D$  (since  $\text{supp } \phi_X \subset D$ ) and

$$|A| \leq \left| \left\{ x \in D : \phi_X^2 h(x) \geq \frac{2\lambda}{N} \right\} \right|.$$

We have then

$$\begin{aligned} \mathcal{E}(\sqrt{\tilde{h}}) &= \frac{1}{2d} \sum_{\substack{y, z \in \mathbb{Z}^d \\ |y-z|=1}} \left( \sqrt{\tilde{h}(y)} - \sqrt{\tilde{h}(z)} \right)^2 \\ &= \frac{1}{2d} \sum_{\substack{y, z \in A \\ |y-z|=1}} \left( \phi_X \sqrt{h}(y) - \phi_X \sqrt{h}(z) \right)^2 + \frac{1}{d} \sum_{\substack{y \in A, z \notin A \\ |y-z|=1}} \left( \phi_X \sqrt{h}(y) - \sqrt{\frac{2\lambda}{N}} \right)^2 \\ &\leq \mathcal{E}(\phi_X \sqrt{h}). \end{aligned}$$

We conclude that

$$\begin{aligned} &\left| \left\{ x \in D : (\phi_X^2 h)(x) \geq \frac{2\lambda}{N} \right\} \right| + \frac{N - \Gamma_2}{2} \left( \max \left( \sqrt{\mathcal{E}(\phi_X \sqrt{h})} - \frac{4\sqrt{\Gamma_3}}{n}, 0 \right) \right)^2 \\ &\geq \left| \left\{ x \in \mathbb{Z}^d : \tilde{h}(x) > 0 \right\} \right| + \frac{N - \Gamma_2}{2} \left( \max \left( \sqrt{\mathcal{E}(\sqrt{\tilde{h}})} - \frac{4\sqrt{\Gamma_3}}{n}, 0 \right) \right)^2. \end{aligned} \quad (5.48)$$

We evaluate next the distance between  $\tilde{h}$  and  $L_N/N$ . To that end, we introduce the function

$$\forall x \in \mathbb{Z}^d \quad \widetilde{\frac{1}{N} L_N}(x) = \left( \max \left( \frac{1}{\sqrt{N}} f_N(x) - \sqrt{\frac{2\lambda}{N}}, 0 \right) \right)^2,$$

and we write

$$\left\| \tilde{h} - \frac{1}{N} L_N \right\|_1 \leq \left\| \tilde{h} - \widetilde{\frac{1}{N} L_N} \right\|_1 + \left\| \widetilde{\frac{1}{N} L_N} - \frac{1}{N} L_N \right\|_1. \quad (5.49)$$

We shall control the first term with the help of the following lemma.

**Lemma 5.50** For any  $s, t > 0$ , any  $a > 0$ , we have

$$\left| \left( \max(\sqrt{t} - a, 0) \right)^2 - \left( \max(\sqrt{s} - a, 0) \right)^2 \right| \leq |t - s|.$$

**Proof.** If  $\sqrt{t} \geq \sqrt{s} \geq a$ , then

$$\begin{aligned} \left( \max(\sqrt{t} - a, 0) \right)^2 - \left( \max(\sqrt{s} - a, 0) \right)^2 &= (\sqrt{t} - a)^2 - (\sqrt{s} - a)^2 \\ &= t - s - 2a(\sqrt{t} - \sqrt{s}) \leq t - s. \end{aligned}$$

If  $\sqrt{t} \geq a \geq \sqrt{s}$ , then

$$\begin{aligned} \left( \max(\sqrt{t} - a, 0) \right)^2 - \left( \max(\sqrt{s} - a, 0) \right)^2 &= (\sqrt{t} - a)^2 \leq (\sqrt{t} - \sqrt{s})^2 \\ &= \left( \frac{t - s}{\sqrt{t} + \sqrt{s}} \right)^2 \leq (t - s) \frac{t + s}{(\sqrt{t} + \sqrt{s})^2} \leq t - s. \end{aligned}$$

In each case, we obtain the desired inequality.  $\square$

Now, thanks to Lemma 5.50 applied with

$$a = \sqrt{2\lambda/N}, \quad t = \phi_X^2 h(x), \quad s = \frac{1}{N} L_N(x),$$

we have

$$\left\| \widetilde{h - \frac{1}{N} L_N} \right\|_1 \leq \sum_{x \in \mathbb{Z}^d} \left| \phi_X^2 h(x) - \frac{1}{N} L_N(x) \right| = \left\| \phi_X^2 h - \frac{1}{N} L_N \right\|_1 \leq \Gamma_3,$$

where the last inequality is a consequence of inequalities (5.43) and (5.45). We deal finally with the second term of formula (5.49). We set  $a = \sqrt{2\lambda/N}$  and we define

$$B = \left\{ x \in \mathbb{Z}^d : \frac{1}{N} L_N(x) \geq a^2 \right\}.$$

We have

$$\begin{aligned} \left\| \widetilde{\frac{1}{N} L_N} - \frac{1}{N} L_N \right\|_1 &= \sum_{x \in B} \left| \left( \frac{1}{\sqrt{N}} f_N(x) - a \right)^2 - \frac{1}{N} L_N(x) \right| + \sum_{x \in \mathbb{Z}^d \setminus B} \frac{1}{N} L_N(x) \\ &\leq \sum_{x \in B} \left| a^2 - 2a \frac{1}{\sqrt{N}} f_N(x) \right| + a^2 |\text{supp } f_N| \\ &\leq 3a \sum_{x \in B} \left| \frac{1}{\sqrt{N}} f_N(x) \right| + a^2 |\text{supp } f_N| \\ &\leq 3a |\text{supp } f_N|^{1/2} \left\| \frac{1}{N} L_N \right\|_1^{1/2} + a^2 |\text{supp } f_N| \\ &\leq 3\sqrt{\frac{2\lambda}{N} cn^d} + \frac{2\lambda}{N} cn^d \leq 4\sqrt{\frac{2c\lambda}{n^2}}, \end{aligned}$$

where the last inequalities come from the fact that, on the event  $\mathcal{A}$ , the range of the random walk has cardinality at most  $cn^d$ , and that  $\lambda/n^2 = n^{\rho-2}$  goes to 0 as  $n$  goes to  $\infty$ . Plugging the previous inequalities in inequality (5.49), we obtain that

$$\left\| \widetilde{h - \frac{1}{N} L_N} \right\|_1 \leq \Gamma_3 + 4\sqrt{\frac{2c\lambda}{n^2}}.$$

We conclude that, if  $h \in C(g, N, r(N))$  and  $f_N \in \mathcal{F}$ , then  $\tilde{h} \in V(\mathcal{F}, \Gamma_4, N)$ , where

$$\Gamma_4 = \Gamma_3 + 4\sqrt{\frac{2c\lambda}{n^2}},$$

and

$$V(\mathcal{F}, \Gamma, N) = \left\{ h \in \ell^1(\mathbb{Z}^d) : \exists f \in \mathcal{F} \quad \left\| h - \frac{1}{N} f^2 \right\|_1 \leq \Gamma \right\}.$$

Together with inequalities (5.47) and (5.48), we obtain finally that

$$E(e^{-|R_N|}; \mathcal{B}) \leq 2(\Gamma_2 + 1) \exp \left( \frac{N}{\lambda} \Gamma_3 + \frac{2\Gamma_1}{\lambda} + c' \frac{n^{2d}}{[\Gamma_1]} \right. \\ \left. - \inf \left\{ \left| \{ x \in \mathbb{Z}^d : \tilde{h}(x) > 0 \} \right| + \frac{N - \Gamma_2}{2} \left( \max \left( \sqrt{\mathcal{E}(\sqrt{\tilde{h}})} - \frac{4\sqrt{\Gamma_3}}{n}, 0 \right) \right)^2 : \tilde{h} \in V(\mathcal{F}, \Gamma_4, N) \right\} \right).$$

We have reached our goal, indeed, this last upper bound depends only on the set  $\mathcal{F}$  and it is uniform over  $g \in \mathcal{G}(X, M, \eta)$ . We report this upper bound in the successive decompositions in sums presented in formulas (5.7) and (5.21). These two formulas imply that

$$E(e^{-|R_N|}; f_N \in \mathcal{F}, \mathcal{A}) \leq |\mathcal{X}(N, \delta)| \times |\mathcal{G}(X, M, \eta)| \times \sup E(e^{-|R_N|}; \mathcal{B}),$$

where the supremum is over  $X \in \mathcal{X}(N, \delta)$  and  $g \in \mathcal{G}(X, M, \eta)$ . Using the combinatorial bounds (5.4), (5.5), (5.17), and recalling the definition of  $C_0$  given in (5.6), we have, for some constant  $c_d$  depending on the dimension only,

$$|\mathcal{X}(N, \delta)| \leq \exp(c_d C_0), \quad |\mathcal{G}| \leq \left( \frac{\sqrt{N}}{\eta} \right) \frac{n^d}{M^d} c_d C_0.$$

Together with our upper bound on the last expectation, this yields

$$\ln E(e^{-|R_N|}; f_N \in \mathcal{F}, \mathcal{A}) \leq C_1 - \inf \left\{ \left| \{ x \in \mathbb{Z}^d : \tilde{h}(x) > 0 \} \right| + \right. \\ \left. \frac{N - \Gamma_2}{2} \left( \max \left( \sqrt{\mathcal{E}(\sqrt{\tilde{h}})} - \frac{4\sqrt{\Gamma_3}}{n}, 0 \right) \right)^2 : \tilde{h} \in V(\mathcal{F}, \Gamma_4, N) \right\}, \quad (5.51)$$

where

$$C_1 = c_d C_0 + \frac{n^d}{M^d} c_d C_0 \ln \left( \frac{\sqrt{N}}{\eta} \right) + \ln(2(\Gamma_2 + 1)) + \frac{N}{\lambda} \Gamma_3 + \frac{2\Gamma_1}{\lambda} + c' \frac{n^{2d}}{[\Gamma_1]}.$$

We compute the asymptotic expansion of these different terms in powers of  $n$ . Recalling the definitions of  $C_0$  and  $\Gamma_3$  given in (5.6) and (5.45), we obtain, using (5.29),

$$C_0 \underset{n \rightarrow \infty}{\sim} n^{\alpha 2^*}, \quad \Gamma_3 \underset{n \rightarrow \infty}{\sim} n^{\beta-1} + n^{-4\alpha/d}. \quad (5.52)$$

Together with the relations (5.31), this yields

$$C_1 \underset{n \rightarrow \infty}{\sim} n^{d(1-\beta)+\alpha 2^*} + n^{d+1+\beta-\rho} + n^{d+2-4\alpha/d-\rho}. \quad (5.53)$$

In addition, we have

$$\Gamma_4 \underset{n \rightarrow \infty}{\sim} n^{\beta-1} + n^{-4\alpha/d} + n^{\rho/2-1}.$$

We make next a specific choice for the values of the exponents, which satisfies the constraints stated before. However we do not try to look for the best possible exponents. So, we choose

$$\alpha = \frac{d}{12}, \quad \beta = \frac{3}{4}, \quad \gamma = \frac{d^2}{12(d-2)}, \quad \rho = \frac{15}{8}. \quad (5.54)$$

With this choice, we obtain that

$$C_1 \underset{n \rightarrow \infty}{\sim} n^{d-1/8}, \quad \Gamma_2 \underset{n \rightarrow \infty}{\sim} n^{d+5/3}, \quad \Gamma_3 \underset{n \rightarrow \infty}{\sim} n^{-1/4}, \quad \Gamma_4 \underset{n \rightarrow \infty}{\sim} n^{-1/16}.$$

The equivalents above are logarithmic, however, by perturbing slightly the values of the exponents  $\alpha, \beta, \gamma, \rho$ , we can ensure that, for  $n$  large enough,

$$C_1 \leq n^{d-1/8}, \quad \Gamma_2 \leq Nn^{-1/4}, \quad 4\sqrt{\Gamma_3} \leq n^{-1/8}, \quad \Gamma_4 \leq n^{-1/16}.$$

Plugging these inequalities in (5.51), we obtain finally the statement of Theorem 2.6.

## 6 Continuous version of Theorem 2.6

### 6.1 Linear interpolation

We need a version of Theorem 2.6 in which the discrete Dirichlet energy is replaced by the continuous one, so that we can use quantitative versions of the Faber–Krahn inequality. The first step is to transform the local time into a function defined continuously everywhere on  $\mathbb{R}^d$ . We will state such a version in Theorem 6.6. We start preparing for this result.

Let  $f$  be a function defined on  $\mathbb{Z}^d$  with values in  $\mathbb{R}^+$ . We define a function  $\tilde{f}$  on  $\mathbb{R}^d$  by interpolating linearly  $f$  successively in the  $d$  directions of the axis. More precisely, we set  $f_0 = f$  and for all  $(x_1, \dots, x_d) \in \mathbb{Z}^d$ ,

$$\forall \alpha \in [0, 1] \quad f_1(\alpha x_1 + (1-\alpha)(x_1+1), x_2, \dots, x_d) = \alpha f_0(x_1, \dots, x_d) + (1-\alpha)f_0(x_1+1, \dots, x_d).$$

We define iteratively, for  $2 \leq k \leq d$ , for all  $(x_1, \dots, x_{k-1}) \in \mathbb{R}^{k-1}$ , for all  $(x_k, \dots, x_d) \in \mathbb{Z}^{d-k+1}$ ,

$$\begin{aligned} \forall \alpha \in [0, 1] \quad f_k(x_1, \dots, x_{k-1}, \alpha x_k + (1-\alpha)(x_k+1), x_{k+1}, \dots, x_d) = \\ \alpha f_{k-1}(x_1, \dots, x_{k-1}, x_k, x_{k+1}, \dots, x_d) + (1-\alpha)f_{k-1}(x_1, \dots, x_{k-1}, x_k+1, x_{k+1}, \dots, x_d) \end{aligned}$$

and finally  $\tilde{f} = f_d$ . Let  $C$  be the unit cube

$$C = \{x = (x_1, \dots, x_d) \in \mathbb{R}^d : 0 \leq x_i \leq 1, 1 \leq i \leq d\}.$$

Let us denote by  $D^d$  the union of all the lines parallel to the axis which go through the points of  $\mathbb{Z}^d$ . We make the following observations relating  $f$  to  $\tilde{f}$ .

**Lemma 6.1** The function  $\tilde{f}$  is continuous on  $\mathbb{R}^d$  and  $C^\infty$  on  $\mathbb{R}^d \setminus D^d$ . We have

$$\int_{\mathbb{R}^d} \tilde{f}(x) dx = \sum_{x \in \mathbb{Z}^d} f(x), \quad (6.2)$$

$$\int_{\mathbb{R}^d} \tilde{f}(x)^2 dx \leq \sum_{x \in \mathbb{Z}^d} f(x)^2, \quad (6.3)$$

$$\int_{\mathbb{R}^d} |\nabla \tilde{f}(x)|^2 dx \leq d\mathcal{E}(f). \quad (6.4)$$



**Proof.** With the help of a standard induction, we get the following formula for  $\tilde{f}$ :

$$\begin{aligned} \forall (y_1, \dots, y_d) \in \mathbb{Z}^d \quad \forall (x_1, \dots, x_d) \in (y_1, \dots, y_d) + C \\ \tilde{f}(x_1, \dots, x_d) = \\ \sum_{\varepsilon_1, \dots, \varepsilon_d \in \{0,1\}} f(y_1 + \varepsilon_1, \dots, y_d + \varepsilon_d) \prod_{1 \leq k \leq d} \left( (1 - \varepsilon_k)(1 - (x_k - y_k)) + \varepsilon_k(x_k - y_k) \right). \end{aligned}$$

Let us compute the integral of  $\tilde{f}$  over  $\mathbb{R}^d$ . We have

$$\int_C \tilde{f}(x) dx = \int_{0 \leq x_1 \leq 1} \left( \dots \left( \int_{0 \leq x_d \leq 1} \tilde{f}(x) dx_d \right) \dots \right) dx_1.$$

Let us fix  $(x_1, \dots, x_{d-1}) \in [0, 1]^{d-1}$  and let us compute

$$\begin{aligned} \int_{0 \leq x_d \leq 1} \tilde{f}(x_1, \dots, x_d) dx_d &= \int_{0 \leq x_d \leq 1} \left( (1 - x_d)f_{d-1}(x_1, \dots, x_{d-1}, 0) + x_d f_{d-1}(x_1, \dots, x_{d-1}, 1) \right) dx_d \\ &= \frac{1}{2} \left( f_{d-1}(x_1, \dots, x_{d-1}, 0) + f_{d-1}(x_1, \dots, x_{d-1}, 1) \right). \end{aligned}$$

Iterating this computation, we obtain

$$\int_C \tilde{f}(x) dx = \frac{1}{2^d} \sum_{x_1, \dots, x_d \in \{0,1\}} f(x_1, \dots, x_d).$$

Each point of  $\mathbb{Z}^d$  belongs to  $2^d$  integer translates of  $C$ , therefore

$$\int_{\mathbb{R}^d} \tilde{f}(x) dx = \sum_{y \in \mathbb{Z}^d} \int_{y+C} \tilde{f}(x) dx = \sum_{x \in \mathbb{Z}^d} f(x)$$

which proves (6.2). Inequality (6.3) is proved in the same way, by using  $d$  times the convexity of the function  $t \rightarrow t^2$  in the linear interpolations. Our next goal is to compute the integral of  $(\nabla \tilde{f})^2$  over  $\mathbb{R}^d$ . Let us first compute its partial derivatives. For simplicity, we deal with the derivative with respect to the last variable  $x_d$ , and we consider only the points  $x$  in the unit cube  $C$ :

$$\frac{\partial \tilde{f}}{\partial x_d}(x) = \sum_{\varepsilon_1, \dots, \varepsilon_{d-1} \in \{0,1\}} D_d f(\varepsilon_1, \dots, \varepsilon_{d-1}) \prod_{k=1}^{d-1} \left( (1 - \varepsilon_k)(1 - x_k) + \varepsilon_k x_k \right),$$

where we define, for  $\varepsilon_1, \dots, \varepsilon_{d-1} \in \{0, 1\}$ ,

$$D_d f(\varepsilon_1, \dots, \varepsilon_{d-1}) = f(\varepsilon_1, \dots, \varepsilon_{d-1}, 1) - f(\varepsilon_1, \dots, \varepsilon_{d-1}, 0).$$

We apply Fubini's theorem to write

$$\begin{aligned} \int_C \left( \frac{\partial \tilde{f}}{\partial x_d}(x) \right)^2 dx &= \sum_{\substack{\varepsilon_1, \dots, \varepsilon_{d-1} \in \{0,1\} \\ \varepsilon'_1, \dots, \varepsilon'_{d-1} \in \{0,1\}}} D_d f(\varepsilon_1, \dots, \varepsilon_{d-1}) D_d f(\varepsilon'_1, \dots, \varepsilon'_{d-1}) \\ &\quad \prod_{k=1}^{d-1} \left( \int_0^1 \left( (1 - \varepsilon_k)(1 - x_k) + \varepsilon_k x_k \right) \left( (1 - \varepsilon'_k)(1 - x_k) + \varepsilon'_k x_k \right) dx_k \right). \end{aligned}$$

We compute the value of the integrals. For  $\varepsilon, \varepsilon' \in \{0, 1\}$ , we have

$$\int_0^1 \left( (1 - \varepsilon)(1 - x) + \varepsilon x \right) \left( (1 - \varepsilon')(1 - x) + \varepsilon' x \right) dx = \frac{1}{6} (1 + \delta(\varepsilon, \varepsilon')),$$

where  $\delta$  is the Kronecker symbol. Next, we have the bound

$$\begin{aligned} \int_C \left( \frac{\partial \tilde{f}}{\partial x_d}(x) \right)^2 dx &\leq \sum_{\substack{\varepsilon_1, \dots, \varepsilon_{d-1} \in \{0, 1\} \\ \varepsilon'_1, \dots, \varepsilon'_{d-1} \in \{0, 1\}}} \left( D_d f(\varepsilon_1, \dots, \varepsilon_{d-1}) \right)^2 \prod_{k=1}^{d-1} \frac{1}{6} (1 + \delta(\varepsilon_k, \varepsilon'_k)) \\ &= \frac{1}{2^{d-1}} \sum_{\varepsilon_1, \dots, \varepsilon_{d-1} \in \{0, 1\}} \left( D_d f(\varepsilon_1, \dots, \varepsilon_{d-1}) \right)^2. \end{aligned}$$

Next, we sum the previous inequality over all integer translates of the unit cube  $C$ . We obtain

$$\begin{aligned} \int_{\mathbb{R}^d} \left( \frac{\partial \tilde{f}}{\partial x_d}(x) \right)^2 dx &= \sum_{y \in \mathbb{Z}^d} \int_{y+C} \left( \frac{\partial \tilde{f}}{\partial x_d}(x) \right)^2 dx \\ &\leq \sum_{y \in \mathbb{Z}^d} \frac{1}{2^{d-1}} \sum_{\varepsilon_1, \dots, \varepsilon_{d-1} \in \{0, 1\}} \left( f(y_1 + \varepsilon_1, \dots, y_{d-1} + \varepsilon_{d-1}, y_d + 1) - f(y_1 + \varepsilon_1, \dots, y_{d-1} + \varepsilon_{d-1}, y_d) \right)^2 \\ &= \sum_{y \in \mathbb{Z}^d} \left( f(y_1, \dots, y_{d-1}, y_d + 1) - f(y_1, \dots, y_{d-1}, y_d) \right)^2. \end{aligned}$$

Summing finally the integrals associated to each partial derivatives, we conclude that

$$\int_{\mathbb{R}^d} |\nabla \tilde{f}|^2 dx \leq d \mathcal{E}(f).$$

Notice that, in the definition (1.9) of  $\mathcal{E}$ , each edge of the lattice appears twice in the summation. This proves (6.4), which finishes the proof of the lemma.  $\square$

## 6.2 Rescaling

We denote by  $\mathbb{Z}_n^d$  the lattice  $\mathbb{Z}^d$  rescaled by a factor  $n$ :  $\mathbb{Z}_n^d = \frac{1}{n} \mathbb{Z}^d$ . We shall simultaneously rescale the space by a factor  $n$  and the values of  $L_N$  by a factor  $n^d/N = 1/n^2$ . Starting with the local time  $L_N$ , which is a function defined on  $\mathbb{Z}^d$ , we define the function  $\ell_N$  on  $\mathbb{R}^d$  by setting

$$\forall x \in \mathbb{R}^d \quad \ell_N(x) = \frac{n^d}{N} L_N(\lfloor nx \rfloor). \quad (6.5)$$

For  $g$  a continuous function from  $\mathbb{R}^d$  to  $\mathbb{R}$ , we define its support as

$$\text{supp } g = \left\{ x \in \mathbb{R}^d : g(x) \neq 0 \right\},$$

and we denote by  $|\text{supp } g|$  its Lebesgue measure.

Here is a continuous analogue of Theorem 2.6.

**Theorem 6.6** Let us denote by  $D^d$  the union of all the lines parallel to the axis which go through the points of  $\mathbb{Z}^d$ . Let  $\mathcal{L}$  be a collection of functions from  $\mathbb{R}^d$  to  $\mathbb{R}^+$  such that  $\|\ell\|_2 = 1$  for  $\ell \in \mathcal{L}$ . For any  $\kappa \geq 1$ , we have, for  $n$  large enough,

$$E(e^{-|R_N|}; \ell_N \in \mathcal{L}) \leq 3 \exp(-\kappa n^d) + \exp \left( 2n^{d-1/8} - n^d \inf \left\{ |\text{supp } g| + \frac{1}{2d}(1 - n^{-1/4}) \left( \max \left( \sqrt{\int_{\mathbb{R}^d} |\nabla g|^2 dx} - \frac{\sqrt{d}}{n^{1/8}}, 0 \right) \right)^2 \right. \right. \\ \left. \left. : g \text{ is continuous } \mathbb{R}^d \rightarrow \mathbb{R}^+ \text{ and } C^\infty \text{ on } \mathbb{R}^d \setminus \frac{1}{n}D^d, \right. \right. \\ \left. \left. \exists \ell \in \mathcal{L} \quad \left\| g^2 - \ell \right\|_{L^1(\mathbb{R}^d)} \leq \frac{3}{n^{1/16}} \right\} \right).$$

**Proof.** Let us define formally the operator which transforms the function  $f_N$  into  $\ell_N$  as in (6.5). To a function  $f : \mathbb{Z}^d \rightarrow \mathbb{R}^+$ , we associate a function  $\Phi_n(f) : \mathbb{R}^d \rightarrow \mathbb{R}^+$  by setting

$$\forall x \in \mathbb{R}^d \quad \Phi_n(f)(x) = \frac{n^d}{N} \left( f(\lfloor nx \rfloor) \right)^2. \quad (6.7)$$

With this definition, we check that  $\ell_n = \Phi_n(f_N)$ . Therefore

$$E(e^{-|R_N|}; \ell_N \in \mathcal{L}) = E(e^{-|R_N|}; f_N \in \mathcal{F}),$$

where  $\mathcal{F}$  is the collection of functions defined by

$$\mathcal{F} = \Phi_n^{-1}(\mathcal{L}) = \{ f : \mathbb{Z}^d \rightarrow \mathbb{R}^+, \Phi_n(f) \in \mathcal{L} \}.$$

We apply the upper bound of theorem 2.6. For  $h$  a function from  $\mathbb{Z}^d$  to  $\mathbb{R}^+$ , we denote its support by

$$\text{supp } h = \left\{ x \in \mathbb{Z}^d : h(x) \neq 0 \right\}.$$

Let  $\kappa \geq 1$ . For  $n$  large enough, we have

$$E(e^{-|R_N|}; \ell_N \in \mathcal{L}) \leq \exp(-\kappa n^d) + \exp \left( n^{d-1/8} - \inf \left\{ |\text{supp } h| + \frac{N}{2}(1 - n^{-1/4}) \left( \max \left( \sqrt{\mathcal{E}(\sqrt{h})} - \frac{1}{n^{9/8}}, 0 \right) \right)^2 \right. \right. \\ \left. \left. : h \in \ell^1(\mathbb{Z}^d), \exists f \in \mathcal{F} \quad \left\| h - \frac{1}{N} f^2 \right\|_1 \leq \frac{1}{n^{1/16}} \right\} \right).$$

In the infimum appearing above, we can limit ourselves to functions  $h$  for which the infimum is of order  $n^d$ , otherwise the exponential becomes negligible compared to the first term  $\exp(-\kappa n^d)$ . The relevant functions  $h$  should be such that

$$|\text{supp } h| \leq 2\kappa n^d, \quad \mathcal{E}(\sqrt{h}) \leq \frac{3\kappa}{n^2}. \quad (6.8)$$

We make also the change of function  $h \rightarrow h^2$ . We obtain that, for  $n$  large enough,

$$E(e^{-|R_N|}; \ell_N \in \mathcal{L}) \leq 3 \exp(-\kappa n^d) + \exp \left( n^{d-1/8} - \inf \left\{ |\text{supp } h| + \frac{N}{2}(1 - n^{-1/4}) \left( \max \left( \sqrt{\mathcal{E}(h)} - \frac{1}{n^{9/8}}, 0 \right) \right)^2 : h \in \ell^2(\mathbb{Z}^d), h \geq 0, |\text{supp } h| \leq 2\kappa n^d, \mathcal{E}(h) \leq \frac{3\kappa}{n^2}, \exists f \in \mathcal{F} \quad \left\| h^2 - \frac{1}{N} f^2 \right\|_1 \leq \frac{1}{n^{1/16}} \right\} \right). \quad (6.9)$$

We apply now the linear interpolation procedure described in section 6.1 to the function  $h$  appearing in the above infimum. Since  $\|f^2\|_1 = N$  for any  $f \in \mathcal{F}$ , the constraint on the function  $h$  implies that

$$\left| \|h^2\|_1 - 1 \right| \leq \frac{1}{n^{1/16}}. \quad (6.10)$$

Starting from a function  $h$  in  $\ell^2(\mathbb{Z}^d)$ , we apply the interpolation procedure and we obtain a function  $\tilde{h}$  in  $L^2(\mathbb{R}^d)$  which satisfies, according to inequalities (6.3) and (6.4),

$$\|\tilde{h}\|_{L^2(\mathbb{R}^d)} \leq \|h\|_{\ell^2(\mathbb{Z}^d)}, \quad \int_{\mathbb{R}^d} |\nabla \tilde{h}|^2 dx \leq d \mathcal{E}(h). \quad (6.11)$$

We wish to obtain an infimum involving only continuous functions, so we have to get completely rid of the discrete function  $h$ . Therefore we should control  $|\text{supp } h|$  and the distance between  $h$  and the set  $\mathcal{L}$ . Let us start with  $|\text{supp } h|$ . The linear interpolation  $\tilde{h}$  of  $h$  is non-zero only in the unit cubes having at least one vertex in the support of  $h$ , therefore

$$|\{x \in \mathbb{R}^d : \tilde{h}(x) > 0\}| \leq |\text{supp } h| + |\{x \in \mathbb{Z}^d \setminus \text{supp } h : \exists y \in \text{supp } h \quad |x - y|_\infty \leq 1\}|. \quad (6.12)$$

Notice that we use  $|\cdot|$  to denote the Lebesgue measure for a continuous set and the cardinality for a discrete set. Each point  $x$  of  $\mathbb{Z}^d$  admits  $3^d$  points  $y \in \mathbb{Z}^d$  such that  $|x - y|_\infty = 1$ , therefore

$$|\{x \in \mathbb{R}^d : \tilde{h}(x) > 0\}| \leq (3^d + 1) |\text{supp } h|. \quad (6.13)$$

This inequality will be useful, however we need a better lower bound for  $|\text{supp } h|$ , therefore we need a better upper bound on the last term in inequality (6.12). This is a delicate matter. Our strategy is to use the bound on the discrete Dirichlet energy to control this boundary term, and to do so, we truncate the function at a fixed level  $\lambda > 0$ . Since  $\tilde{h}(x)$  is a linear interpolation between the  $2^d$  values of  $h$  at the vertices of the unit cube containing  $x$ , we have

$$|\{x \in \mathbb{R}^d : \tilde{h}(x) > \lambda\}| \leq |\text{supp } h| + \left| \left\{ x \in \mathbb{Z}^d \setminus \text{supp } h : \exists y \in \text{supp } h \quad |x - y|_\infty \leq 1, \quad h(y) \geq \lambda \right\} \right|. \quad (6.14)$$

Let  $C$  be a unit cube with vertices in  $\mathbb{Z}^d$  such that one vertex  $x$  of  $C$  is not in  $\text{supp } h$  and another vertex  $y$  of  $C$  satisfies  $h(y) \geq \lambda$ . Then there exist two vertices  $x', y'$  of  $C$  which are nearest neighbours and which satisfy

$$h(x') - h(y') \geq \frac{\lambda}{2^d},$$

and the contribution to the discrete Dirichlet energy  $\mathcal{E}(h)$  of the edges belonging to the boundary of  $C$  is larger or equal than

$$\frac{1}{d} (h(x') - h(y'))^2 \geq \frac{1}{d} \frac{\lambda^2}{2^{2d}}.$$

An edge belongs to at most  $2^{d-1}$  unit cubes. The number  $M$  of cubes making such a contribution satisfies therefore

$$M \frac{1}{d} \frac{\lambda^2}{2^{2d}} \frac{1}{2^{d-1}} \leq \mathcal{E}(h).$$

Moreover the last term in inequality (6.14) is bounded from above by  $M2^{2d}$ . Taking into account the bound on  $\mathcal{E}(h)$  given in (6.8), we conclude that

$$|\text{supp } h| \geq |\{x \in \mathbb{R}^d : \tilde{h}(x) > \lambda\}| - M2^{2d} \geq |\{x \in \mathbb{R}^d : \tilde{h}(x) > \lambda\}| - \frac{d2^{5d} 3\kappa}{\lambda^2 n^2}. \quad (6.15)$$

In order to delay the rescaling of the space, we define an intermediate operator  $\Phi$  acting on the functions as follows. To a function  $f : \mathbb{Z}^d \rightarrow \mathbb{R}^+$ , we associate a function  $\Phi(f) : \mathbb{R}^d \rightarrow \mathbb{R}^+$  by setting

$$\forall x \in \mathbb{R}^d \quad \Phi(f)(x) = \left( f(\lfloor x \rfloor) \right)^2.$$

We rewrite (6.7) with the help of  $\Phi$  as follows:

$$\forall x \in \mathbb{R}^d \quad \Phi_n(f)(x) = \frac{n^d}{N} \Phi(f)(nx). \quad (6.16)$$

Let now  $f$  be an element of  $\mathcal{F}$  such that

$$\left\| h^2 - \frac{1}{N} f^2 \right\|_1 \leq \frac{1}{n^{1/16}}.$$

By definition of  $\mathcal{F}$ , we have that  $\Phi_n(f) \in \mathcal{L}$ . On the one hand, we have

$$\left\| \Phi(h) - \frac{1}{N} \Phi(f) \right\|_{L^1(\mathbb{R}^d)} = \left\| h^2 - \frac{1}{N} f^2 \right\|_{\ell^1(\mathbb{Z}^d)} \leq \frac{1}{n^{1/16}}. \quad (6.17)$$

On the other hand, we have

$$\left\| \tilde{h}^2 - \Phi(h) \right\|_{L^1(\mathbb{R}^d)} = \int_{\mathbb{R}^d} |\tilde{h}(x)^2 - h(\lfloor x \rfloor)^2| dx. \quad (6.18)$$

Let us define

$$\forall x \in \mathbb{Z}^d \quad D(x) = \{ y \in \mathbb{R}^d : \lfloor y \rfloor = x \}.$$

We can rewrite the previous integral as

$$\begin{aligned}
\left\| \tilde{h}^2 - \Phi(h) \right\|_{L^1(\mathbb{R}^d)} &= \sum_{x \in \mathbb{Z}^d} \int_{D(x)} |\tilde{h}(y)^2 - h(\lfloor y \rfloor)^2| dy \\
&\leq \sum_{x \in \mathbb{Z}^d} \sup_{y \in D(x)} |\tilde{h}(y)^2 - h(x)^2| \\
&\leq \sum_{x \in \mathbb{Z}^d} 2^d \max \{ |h(y)^2 - h(z)^2| : y, z \in \overline{D(x)} \cap \mathbb{Z}^d, |y - z| = 1 \} \\
&\leq \sum_{x \in \mathbb{Z}^d} 2^d \sum_{\substack{y, z \in \overline{D(x)} \cap \mathbb{Z}^d \\ |y - z| = 1}} |h(y)^2 - h(z)^2| \\
&\leq 2^{2d} \sum_{\substack{y, z \in \mathbb{Z}^d \\ |y - z| = 1}} |h(y) - h(z)| \times |h(y) + h(z)| \\
&\leq 2^{2d} \left( \sum_{\substack{y, z \in \mathbb{Z}^d \\ |y - z| = 1}} |h(y) - h(z)|^2 \right)^{1/2} \left( \sum_{\substack{y, z \in \mathbb{Z}^d \\ |y - z| = 1}} |h(y) + h(z)|^2 \right)^{1/2} \\
&\leq 2^{2d} (2d\mathcal{E}(h))^{1/2} 8d \|h\|_{\ell^2(\mathbb{Z}^d)}. \tag{6.19}
\end{aligned}$$

Using inequalities (6.8) and (6.10) (remember that we changed  $h$  into  $h^2$  just afterwards), we obtain that

$$\left\| \tilde{h}^2 - \Phi(h) \right\|_{L^1(\mathbb{R}^d)} \leq 32d^2 2^{2d} \frac{\sqrt{3\kappa}}{n}. \tag{6.20}$$

Combining this inequality and inequality (6.17), we conclude that, for  $n$  large enough,

$$\left\| \tilde{h}^2 - \frac{1}{N} \Phi(f) \right\|_{L^1(\mathbb{R}^d)} \leq 32d^2 2^{2d} \frac{\sqrt{3\kappa}}{n} + \frac{1}{n^{1/16}} \leq \frac{2}{n^{1/16}}. \tag{6.21}$$

In addition, the function  $\tilde{h}$  has bounded support and it is continuous on  $\mathbb{R}^d$ . Let us denote by  $D^d$  the union of all the lines parallel to the axis which go through the points of  $\mathbb{Z}^d$ . The function  $\tilde{h}$  is also  $C^\infty$  on  $\mathbb{R}^d \setminus D^d$ . Thus we can take the infimum over the set of functions  $\tilde{h}$  having these properties. We are now ready to substitute  $\tilde{h}$  to  $h$  in inequality (6.9). We bound from below the infimum in the exponential with the help of inequalities (6.11), (6.13), (6.15) and (6.21):

$$\begin{aligned}
E(e^{-|R_N|}; \ell_N \in \mathcal{L}) &\leq 3 \exp(-\kappa n^d) + \exp \left( n^{d-1/8} - \right. \\
&\inf \left\{ \left| \{x \in \mathbb{R}^d : \tilde{h}(x) > \lambda\} \right| - \frac{d2^{5d} 3\kappa}{\lambda^2 n^2} + \right. \\
&\quad \left. \frac{N}{2} (1 - n^{-1/4}) \left( \max \left( \sqrt{\frac{1}{d} \int_{\mathbb{R}^d} |\nabla \tilde{h}|^2 dx} - \frac{1}{n^{9/8}}, 0 \right) \right)^2 \right. \\
&\quad \left. : \tilde{h} \text{ is continuous } \mathbb{R}^d \rightarrow \mathbb{R}^+ \text{ and } C^\infty \text{ on } \mathbb{R}^d \setminus D^d, \quad |\text{supp } \tilde{h}| \leq (3^d + 1) 2\kappa n^d, \right. \\
&\quad \left. \exists f \in \mathcal{F} \quad \left\| \tilde{h}^2 - \frac{1}{N} \Phi(f) \right\|_{L^1(\mathbb{R}^d)} \leq \frac{2}{n^{1/16}} \right\}. \tag{6.22}
\end{aligned}$$

We are almost done. We make a change of scale in order to replace  $\mathcal{F}$  by  $\mathcal{L}$  in the infimum. To the function  $\tilde{h} : \mathbb{R}^d \rightarrow \mathbb{R}^+$ , we associate the function  $\tilde{h}_n : \mathbb{R}^d \rightarrow \mathbb{R}^+$  obtained by rescaling the space by a factor  $n$  and the values by  $n^{d/2}$ :

$$\forall x \in \mathbb{R}^d \quad \tilde{h}_n(x) = n^{d/2} \tilde{h}(nx).$$

We have then

$$\begin{aligned} |\operatorname{supp} \tilde{h}_n| &= n^d |\operatorname{supp} \tilde{h}|, \\ |\{x \in \mathbb{R}^d : \tilde{h}(x) > \lambda\}| &= n^d |\{x \in \mathbb{R}^d : \tilde{h}_n(x) > n^{d/2} \lambda\}|, \\ \int_{\mathbb{R}^d} |\nabla \tilde{h}_n(x)|^2 dx &= n^d \int_{\mathbb{R}^d} |n \nabla \tilde{h}(nx)|^2 dx = n^2 \int_{\mathbb{R}^d} |\nabla \tilde{h}(x)|^2 dx. \end{aligned} \quad (6.23)$$

Moreover, using the identity (6.16), we see that

$$\left\| \tilde{h}_n^2 - \Phi_n(f) \right\|_{L^1(\mathbb{R}^d)} = \int_{\mathbb{R}^d} n^d \left| \tilde{h}(nx)^2 - \frac{1}{N} \Phi(f)(nx) \right| dx = \left\| \tilde{h}^2 - \frac{1}{N} \Phi(f) \right\|_{L^1(\mathbb{R}^d)}. \quad (6.24)$$

Remember that  $\mathcal{F} = \Phi_n^{-1}(\mathcal{L})$ , thus  $\Phi_n(f) \in \mathcal{L}$  whenever  $f \in \mathcal{F}$ . The identities (6.23) and (6.24) allow to rewrite (6.22) as follows:

$$\begin{aligned} E(e^{-|R_N|}; \ell_N \in \mathcal{L}) &\leq 3 \exp(-\kappa n^d) + \exp \left( n^{d-1/8} - \inf \left\{ n^d |\{x \in \mathbb{R}^d : \tilde{h}_n(x) > n^{d/2} \lambda\}| \right. \right. \\ &\quad \left. \left. - \frac{d 2^{5d} 3\kappa}{\lambda^2 n^2} + \frac{N}{2} (1 - n^{-1/4}) \left( \max \left( \sqrt{\frac{1}{dn^2} \int_{\mathbb{R}^d} |\nabla \tilde{h}_n|^2 dx} - \frac{1}{n^{9/8}}, 0 \right) \right)^2 \right. \right. \\ &\quad \left. \left. : \tilde{h}_n \text{ is continuous } \mathbb{R}^d \rightarrow \mathbb{R}^+ \text{ and } C^\infty \text{ on } \mathbb{R}^d \setminus \frac{1}{n} D^d, \quad |\operatorname{supp} \tilde{h}_n| \leq (3^d + 1) 2\kappa, \right. \right. \\ &\quad \left. \left. \exists \ell \in \mathcal{L} \quad \left\| \tilde{h}_n^2 - \ell \right\|_{L^1(\mathbb{R}^d)} \leq \frac{2}{n^{1/16}} \right\} \right). \end{aligned} \quad (6.25)$$

We choose  $\lambda = n^{-\frac{d+1}{2}}$  and we set

$$\forall x \in \mathbb{R}^d \quad g(x) = \max \left( \tilde{h}_n(x) - \frac{1}{\sqrt{n}}, 0 \right).$$

This function  $g$  satisfies

$$\begin{aligned} |\{x \in \mathbb{R}^d : \tilde{h}_n(x) > n^{d/2} \lambda\}| &= |\{x \in \mathbb{R}^d : g(x) > 0\}|, \\ 1 + \frac{2}{n^{1/16}} &\geq \int_{\mathbb{R}^d} |\tilde{h}_n|^2 dx \geq \int_{\mathbb{R}^d} |g|^2 dx, \\ \int_{\mathbb{R}^d} |\nabla \tilde{h}_n|^2 dx &\geq \int_{\mathbb{R}^d} |\nabla g|^2 dx. \end{aligned} \quad (6.26)$$

Moreover we have

$$\begin{aligned}
\|g^2 - \tilde{h}_n^2\|_{L^1(\mathbb{R}^d)} &= \int_{\mathbb{R}^d} |g(x)^2 - \tilde{h}_n(x)^2| dx \\
&\leq \frac{1}{\sqrt{n}} \int_{\mathbb{R}^d} |g(x) + \tilde{h}_n(x)| dx \\
&\leq \frac{1}{\sqrt{n}} \left( \|g\|_{L^2(\mathbb{R}^d)} (|\text{supp } g|)^{1/2} + \|\tilde{h}_n\|_{L^2(\mathbb{R}^d)} (|\text{supp } \tilde{h}_n|)^{1/2} \right) \\
&\leq \frac{2}{\sqrt{n}} \|\tilde{h}_n\|_{L^2(\mathbb{R}^d)} (|\text{supp } \tilde{h}_n|)^{1/2} \\
&\leq \frac{4}{\sqrt{n}} ((3^d + 1)2\kappa)^{1/2}.
\end{aligned} \tag{6.27}$$

Plugging the inequalities (6.26) and (6.27) in the infimum of (6.25), we get, for  $n$  large enough,

$$\begin{aligned}
E(e^{-|R_N|}; \ell_N \in \mathcal{L}) &\leq 3 \exp(-\kappa n^d) + \exp \left( n^{d-1/8} - \right. \\
&\quad \inf \left\{ n^d |\text{supp } g| - d 2^{5d} 3\kappa n^{d-1} + \frac{N}{2} (1 - n^{-1/4}) \left( \max \left( \sqrt{\frac{1}{dn^2} \int_{\mathbb{R}^d} |\nabla g|^2 dx} - \frac{1}{n^{9/8}}, 0 \right) \right)^2 \right. \\
&\quad \left. : g \text{ is continuous } \mathbb{R}^d \rightarrow \mathbb{R}^+ \text{ and } C^\infty \text{ on } \mathbb{R}^d \setminus \frac{1}{n} D^d, \right. \\
&\quad \left. \left. \exists \ell \in \mathcal{L} \quad \|g^2 - \ell\|_{L^1(\mathbb{R}^d)} \leq \frac{3}{n^{1/16}} \right\} \right).
\end{aligned}$$

For  $n$  large enough, this inequality can be rewritten as in the statement of theorem (6.6).  $\square$

## 7 Application of the quantitative Faber–Krahn inequality

### 7.1 Statement

A crucial ingredient to extend Bolthausen’s result to dimensions  $d \geq 3$  is the quantitative Faber–Krahn inequality. Currently, the best version of this inequality is due to Brasco, De Philippis and Velichkov [7]. A weaker version was proved before by Fusco, Maggi and Pratelli [13], and we could very well rely on this weaker version to achieve our goal. Let  $G$  be an open subset of  $\mathbb{R}^d$  having finite Lebesgue measure. Let  $\lambda(G)$  be the first eigenvalue of the Dirichlet–Laplacian of  $G$ , defined by

$$\lambda(G) = \inf \left\{ \int_G |\nabla u|^2 dx : \|u\|_{2,G} = 1 \right\}.$$

Recall that our notation differs from Bolthausen’s by a factor  $1/2$ : the notation  $\lambda(G)$  in [6] is half of the quantity defined above.

To control the distance of  $G$  to a ball, we define the Fraenkel asymmetry

$$\mathcal{A}(G) = \inf \left\{ \frac{|G \Delta B|}{|B|} : B \text{ ball such that } |B| = |G| \right\},$$

where  $G \Delta B$  is the symmetric difference between the sets  $G$  and  $B$ .



**Theorem 7.1** There exists a positive constant  $\sigma$ , which depends on the dimension  $d$  only, such that, for any an open subset  $G$  of  $\mathbb{R}^d$  having finite Lebesgue measure, we have, for any  $d$  dimensional ball  $B$ ,

$$|G|^{2/d} \lambda(G) - |B|^{2/d} \lambda(B) \geq \sigma \mathcal{A}(G)^2.$$

This result is proved by Brasco, De Philippis and Velichkov [7]. With the help of this quantitative Faber-Krahn inequality, we shall prove the crucial Lemma 7.5, which is the  $d$  dimensional counterpart of lemma A.1 in Bolthausen's paper [6]. Let  $\lambda_d$  be the principal eigenvalue of  $-\Delta$  in the unit ball  $B(0, 1)$  of  $\mathbb{R}^d$ , let  $\omega_d$  be the volume of  $B(0, 1)$ , and let us define

$$\rho_d = \left( \frac{\lambda_d}{d^2 \omega_d} \right)^{\frac{1}{d+2}}.$$

Let  $G$  be an open subset of  $\mathbb{R}^d$  having finite Lebesgue measure and let  $r > 0$  be such that  $|G| = r^d \omega_d$ . The classical Faber-Krahn inequality states that

$$\lambda(G) \geq \frac{\lambda_d}{r^2}, \quad (7.2)$$

while the inequality of theorem 7.1 can be rewritten as

$$\lambda(G) \geq \frac{1}{r^2} \left( \lambda_d + \frac{\sigma \mathcal{A}(G)^2}{(\omega_d)^{2/d}} \right). \quad (7.3)$$

By the classical Faber-Krahn inequality, we have

$$|G| + \frac{1}{2d} \lambda(G) \geq r^d \omega_d + \frac{\lambda_d}{2dr^2}.$$

Let us define

$$\forall r > 0 \quad \psi(r) = r^d \omega_d + \frac{\lambda_d}{2dr^2}. \quad (7.4)$$

The function  $\psi$  admits a unique minimum on  $\mathbb{R}^+$  at  $r = \rho_d$ . The constant  $\chi_d$  is defined through the variational formula

$$\chi_d = \inf \left\{ |G| + \frac{1}{2d} \lambda(G) : G \text{ open subset of } \mathbb{R}^d \right\}.$$

We conclude from the previous inequalities that

$$\chi_d = \frac{d+2}{2} \left( \frac{\lambda_d}{d^2} \right)^{\frac{d}{d+2}} (\omega_d)^{\frac{2}{d+2}}.$$

Let  $\phi$  be the eigenfunction of  $-\Delta$  in  $B(0, \rho_d)$  with Dirichlet boundary conditions associated to  $\lambda_d$  and normalized so that  $\|\phi\|_2 = 1$ ,  $\phi \geq 0$ . We extend  $\phi$  to  $\mathbb{R}^d$  by setting it equal to 0 outside  $B(0, \rho_d)$ . For  $x \in \mathbb{R}^d$ , we denote by  $\phi_x$  the translate of  $\phi$  defined by

$$\forall y \in \mathbb{R}^d \quad \phi_x(y) = \phi(y - x).$$

We state next the counterpart of Lemma A.1 of Bolthausen's paper [6]. The difference is that we work in dimensions  $d \geq 3$  and in the full space rather than in the torus. The spirit of the proof is exactly the same as in Bolthausen's case, the major new input is the quantitative Faber-Krahn inequality. Equipped with this powerful inequality, the proof becomes more transparent than Bolthausen's proof, which contains somehow a two-dimensional version of a quantitative isoperimetric inequality.

**Lemma 7.5** If  $g : \mathbb{R}^d \rightarrow \mathbb{R}^+$  is  $C^\infty$  and such that  $\|g\|_2 = 1$ ,  $g \geq 0$ , and if

$$\varepsilon = \inf_{x \in \mathbb{R}^d} \|g - \phi_x\|_2 > 0$$

is small enough, then

$$|\{g > 0\}| + \frac{1}{2d} \int_{\mathbb{R}^d} |\nabla g|^2 dx \geq \chi_d + \varepsilon^{48}.$$

**Proof.** There are many constants involved throughout this proof. So we denote by  $c$  a generic positive constant which depends only on the dimension  $d$ , and we warn that of course the value of  $c$  changes from one formula to another! Let  $g$  be a function as in the statement of the lemma and let us set

$$G = \{x \in \mathbb{R}^d : g(x) > 0\}.$$

We need only to consider the case where  $|G| < \infty$  and the function  $g$  is such that

$$|G| + \frac{1}{2d} \int_{\mathbb{R}^d} |\nabla g|^2 dx \leq 2\chi_d.$$

The Sobolev inequality implies then that there exists a constant  $c_S(d)$  depending on the dimension only such that

$$\|g\|_4^2 \leq c_S(d) \int_{\mathbb{R}^d} (g^2 + |\nabla g|^2) dx \leq c_S(d)(1 + 4d\chi_d). \quad (7.6)$$

Let  $r > 0$  be such that  $|G| = r^d \omega_d$ . Applying inequality (7.3) to the set  $G$ , we obtain

$$\begin{aligned} |\{g > 0\}| + \frac{1}{2d} \int_{\mathbb{R}^d} |\nabla g|^2 dx &\geq r^d \omega_d + \frac{1}{2d} \lambda(G) \\ &\geq r^d \omega_d + \frac{1}{2dr^2} \left( \lambda_d + \frac{\sigma \mathcal{A}(G)^2}{(\omega_d)^{2/d}} \right) = \psi(r) + \frac{\sigma \mathcal{A}(G)^2}{2dr^2(\omega_d)^{2/d}}, \end{aligned} \quad (7.7)$$

where  $\psi(r)$  is the function defined in (7.4). We recall that the function  $\psi$  admits a unique minimum at  $r = \rho_d$  and moreover  $\psi''(\rho_d) > 0$ . Therefore there exist two positive constants  $\eta, c$  such that  $\eta < \rho_d/2$  and

$$\forall r \in ]\rho_d - \eta, \rho_d + \eta[ \quad \psi(r) \geq \psi(\rho_d) + c(r - \rho_d)^2. \quad (7.8)$$

For  $\varepsilon_0 > 0$  small enough, we can take  $\eta$  sufficiently small to ensure that we have in addition

$$\forall r \in \mathbb{R}^+ \setminus ]\rho_d - \eta/2, \rho_d + \eta/2[ \quad \psi(r) \geq \psi(\rho_d) + (\varepsilon_0)^{48}.$$

In particular, the inequality stated in the lemma holds if  $\varepsilon < \varepsilon_0$  and if  $r$  does not belong to the interval  $]\rho_d - \eta/2, \rho_d + \eta/2[$ . From now onwards, we suppose that  $\varepsilon < \varepsilon_0$  and we consider only the cases of sets  $G$  such that  $\rho_d - \eta/2 < r < \rho_d + \eta/2$ . From inequality (7.8), we see that we need only to consider the case where

$$c(r - \rho_d)^2 \leq \varepsilon^{48}. \quad (7.9)$$

Recalling that  $\psi(r) \geq \chi_d$ , it follows from (7.7) that if

$$\frac{\sigma \mathcal{A}(G)^2}{8d(\rho_d)^2(\omega_d)^{2/d}} \geq \varepsilon^{48},$$

then the inequality of the lemma is satisfied. From now onwards, we consider only the cases of sets  $G$  satisfying

$$\mathcal{A}(G) < \sqrt{\frac{1}{\sigma} 8d(\rho_d)^2(\omega_d)^{2/d}\varepsilon^{48}} = c\varepsilon^{24}. \quad (7.10)$$

Therefore there exists  $x_0 \in \mathbb{R}^d$  and  $r_0 > 0$  such that

$$|G \Delta B(x_0, r_0)| \leq c(2\rho_d)^d \omega_d \varepsilon^{24} \leq c\varepsilon^{24}. \quad (7.11)$$

This inequality, together with inequality (7.9), imply that

$$|r_0 - \rho_d| \leq c\varepsilon^{24}. \quad (7.12)$$

Let  $\delta > 0$ . We shall compare  $\int_{\mathbb{R}^d} |\nabla g|^2 dx$  with  $\lambda(B(x_0, r_0 + \delta))$ . To this end, we truncate smoothly the function  $g$  as follows. Let  $h$  be a  $C^\infty$  function from  $\mathbb{R}^d$  to  $[0, 1]$  satisfying

$$\forall x \in \mathbb{R}^d \quad h(x) = \begin{cases} 1 & \text{if } x \in B(x_0, r_0) \\ 0 & \text{if } x \notin B(x_0, r_0 + \delta) \end{cases},$$

as well as the following bound on its gradient:

$$\forall x \in \mathbb{R}^d \quad |\nabla h(x)| \leq \frac{c}{\delta}.$$

We set  $\tilde{g} = hg$  and we estimate the Dirichlet energy of  $\tilde{g}$  as follows. Let

$$A = B(x_0, r_0 + \delta) \setminus B(x_0, r_0),$$

we have

$$\begin{aligned} \int_{\mathbb{R}^d} |\nabla \tilde{g}|^2 dx &= \int_{B(x_0, r_0)} |\nabla g|^2 dx + \int_A |\nabla(hg)|^2 dx \\ &\leq \int_{\mathbb{R}^d} |\nabla g|^2 dx + \int_A |\nabla h|^2 g^2 dx + \int_A 2hg |\nabla h| |\nabla g| dx. \end{aligned} \quad (7.13)$$

Moreover

$$\int_A |\nabla h|^2 g^2 dx \leq \left(\frac{c}{\delta}\right)^2 \int_{A \cap G} g^2 dx.$$

Yet, using (7.11), we have

$$|A \cap G| \leq |G \setminus B(x_0, r_0)| \leq c\varepsilon^{24}, \quad (7.14)$$

and, by Hölder's inequality,

$$\int_{A \cap G} g^2 dx \leq \int_{G \setminus B(x_0, r_0)} g^2 dx \leq \left( \int_{\mathbb{R}^d} g^4 dx \right)^{\frac{1}{2}} |G \setminus B(x_0, r_0)|^{\frac{1}{2}} \leq \|g\|_4^2 \sqrt{c\varepsilon^{24}}. \quad (7.15)$$

We control the last integral of (7.13) as follows:

$$\int_A 2hg |\nabla h| |\nabla g| dx \leq 2\frac{c}{\delta} \int_{A \cap G} |g \nabla g| dx \leq 2\frac{c}{\delta} \left( \int_{\mathbb{R}^d} |\nabla g|^2 dx \right)^{\frac{1}{2}} \left( \int_{A \cap G} g^2 dx \right)^{\frac{1}{2}}. \quad (7.16)$$

We plug these inequalities in (7.13). Together with inequality (7.6), we obtain

$$\int_{\mathbb{R}^d} |\nabla \tilde{g}|^2 dx \leq \int_{\mathbb{R}^d} |\nabla g|^2 dx + c \frac{\varepsilon^{12}}{\delta^2} + c \frac{\varepsilon^6}{\delta}. \quad (7.17)$$

The end of the argument is the same as in lemma A.1 of [6]. We denote by  $\widehat{\phi}_0$  the normalized eigenfunction in  $B(x_0, r_0 + \delta)$ , that is,

$$\forall x \in B(x_0, r_0 + \delta) \quad \widehat{\phi}_0(x) = \left( \frac{\rho_d}{r_0 + \delta} \right)^{d/2} \phi \left( \frac{\rho_d}{r_0 + \delta} (x - x_0) \right).$$

With a change of variables, we obtain that

$$\int_{\mathbb{R}^d} (\widehat{\phi}_0)^2 dx = 1$$

and

$$\int_{\mathbb{R}^d} |\nabla \widehat{\phi}_0|^2 dx = \left( \frac{\rho_d}{r_0 + \delta} \right)^2 \int_{\mathbb{R}^d} |\nabla \phi|^2 dx.$$

We have then, using (7.14),

$$\begin{aligned} |\{g > 0\}| + \frac{1}{2d} \int_{\mathbb{R}^d} |\nabla g|^2 dx &\geq |B(x_0, r_0)| - c\varepsilon^{24} + \frac{1}{2d} \int_{\mathbb{R}^d} |\nabla \tilde{g}|^2 dx - c \frac{\varepsilon^{12}}{\delta^2} - c \frac{\varepsilon^6}{\delta} \\ &\geq |B(x_0, r_0 + \delta)| + \frac{1}{2d} \int_{\mathbb{R}^d} |\nabla \widehat{\phi}_0|^2 dx - c\delta - c\varepsilon^{24} + \frac{1}{2d} \int_{\mathbb{R}^d} (|\nabla \tilde{g}|^2 - |\nabla \widehat{\phi}_0|^2) dx - c \frac{\varepsilon^{12}}{\delta^2} - c \frac{\varepsilon^6}{\delta}. \end{aligned} \quad (7.18)$$

By inequality (7.12), we have

$$|r_0 + \delta - \rho_d| \leq c\varepsilon^{24} + \delta. \quad (7.19)$$

Thus, for  $\varepsilon$  and  $\delta$  small enough, the value  $r_0 + \delta$  belongs to the interval  $]\rho_d - \eta, \rho_d + \eta[$  and we can apply (7.8) to get

$$|B(x_0, r_0 + \delta)| + \frac{1}{2d} \int_{\mathbb{R}^d} |\nabla \widehat{\phi}_0|^2 dx \geq \psi(r_0 + \delta) \geq \psi(\rho_d) + c(r_0 + \delta - \rho_d)^2. \quad (7.20)$$

To estimate the second integral in (7.18), we proceed as in lemma A.1 of [6]. We denote by  $\delta_{12}$  the difference between the first and the second eigenvalues of the Laplacian in  $B(x_0, r_0 + \delta)$ . We have

$$\frac{1}{2d} \int_{\mathbb{R}^d} (|\nabla \tilde{g}|^2 - |\nabla \widehat{\phi}_0|^2) dx \geq \delta_{12} \|\tilde{g} - \widehat{\phi}_0\|_2^2.$$

Again, since  $r_0 + \delta$  belongs to the interval  $]\rho_d - \eta, \rho_d + \eta[$ , which is included in  $]\rho_d/2, 3\rho_d/2[$ , there exists a constant  $c > 0$  depending on the dimension  $d$  only such that  $\delta_{12} \geq c$ . Moreover we have

$$\varepsilon \leq \|g - \phi_{x_0}\|_2 \leq \|g - \tilde{g}\|_2 + \|\tilde{g} - \widehat{\phi}_0\|_2 + \|\widehat{\phi}_0 - \phi_{x_0}\|_2.$$

Now, thanks to inequalities (7.6) and (7.15),

$$\|g - \tilde{g}\|_2^2 = \int_{\mathbb{R}^d} ((1 - h)g)^2 dx \leq \int_{G \setminus B(x_0, r_0)} g^2 dx \leq c\varepsilon^{12}.$$

Using the fact that the eigenfunction  $\phi$  has bounded support and is Lipschitz, we have

$$\begin{aligned} \|\widehat{\phi}_0 - \phi_{x_0}\|_2^2 &= \int_{\mathbb{R}^d} \left( \left( \frac{\rho_d}{r_0 + \delta} \right)^{d/2} \phi \left( \frac{\rho_d}{r_0 + \delta} (x - x_0) \right) - \phi(x - x_0) \right)^2 dx \\ &\leq c \left( \left( \frac{\rho_d}{r_0 + \delta} \right)^{d/2} - 1 \right)^2 + c \left( \left( \frac{\rho_d}{r_0 + \delta} \right) - 1 \right)^2 \\ &\leq c(r_0 + \delta - \rho_d)^2. \end{aligned}$$

The previous inequalities yield

$$\frac{1}{2d} \int_{\mathbb{R}^d} (|\nabla \widetilde{g}|^2 - |\nabla \widehat{\phi}_0|^2) dx \geq c \left( \varepsilon - c\varepsilon^{12} - c(r_0 + \delta - \rho_d)^2 \right)^2.$$

Reporting in the inequality (7.18), and using inequalities (7.19) and (7.20), we get

$$|\{g > 0\}| + \frac{1}{2d} \int_{\mathbb{R}^d} |\nabla g|^2 dx \geq \psi(\rho_d) + c \left( c\varepsilon - c\varepsilon^{12} - c(c\varepsilon^{24} + \delta)^2 \right)^2 - c\delta - c\varepsilon^{24} - c\frac{\varepsilon^{12}}{\delta^2} - c\frac{\varepsilon^6}{\delta}.$$

Choosing  $\delta = \varepsilon^3$ , we conclude that, for  $\varepsilon$  small enough,

$$|\{g > 0\}| + \frac{1}{2d} \int_{\mathbb{R}^d} |\nabla g|^2 dx \geq \chi_d + c\varepsilon^2.$$

Therefore the inequality stated in the lemma is satisfied for  $\varepsilon$  small enough whenever the conditions (7.9) and (7.10) are fulfilled.  $\square$

We shall extend slightly Lemma 7.5, in two ways. First, we will relax the condition  $\|g\|_2 = 1$ , second, we will consider functions which are not  $C^\infty$  on the whole space, but on the complement of a countable union of lines. We could probably reach the Sobolev space  $W^{1,2}(\mathbb{R}^d)$ , however this won't be necessary for our purpose.

**Corollary 7.21** Let  $D$  be a subset of  $\mathbb{R}^d$  which is a countable union of lines. Let  $\varepsilon > 0$ . If  $g : \mathbb{R}^d \rightarrow \mathbb{R}^+$  is continuous and  $C^\infty$  on  $\mathbb{R}^d \setminus D$ , and if

$$\begin{aligned} 1 - \varepsilon^{49} &\leq \|g\|_2 \leq 1 + \varepsilon^{49}, \\ \forall x \in \mathbb{R}^d \quad \|g - \phi_x\|_2 &\geq 2\varepsilon, \end{aligned}$$

then, for  $\varepsilon$  small enough,

$$|\{g > 0\}| + \frac{1}{2d} \int_{\mathbb{R}^d} |\nabla g|^2 dx \geq \chi_d + \varepsilon^{49}.$$

**Proof.** The condition on the regularity is not problematic, in fact the proof of Lemma 7.5 can be used to deal directly with these functions. The only thing we need to do is to rescale the function  $g$  in order to have a function of  $L^2$  norm one, for which we can use lemma 7.5. So let  $g$  be a function satisfying the hypothesis of corollary 7.21 and let us set

$$h = \frac{1}{\|g\|_2} g.$$

We have obviously  $\|h\|_2 = 1$ . Moreover, for any  $x \in \mathbb{R}^d$ ,

$$\begin{aligned} 2\varepsilon &\leq \|g - \phi_x\|_2 \leq \|g - h\|_2 + \|h - \phi_x\|_2 \\ &\leq \left|1 - \frac{1}{\|g\|_2}\right| \|g\|_2 + \|h - \phi_x\|_2 \leq \varepsilon^{49} + \|h - \phi_x\|_2. \end{aligned}$$

We can thus apply the inequality of Lemma 7.5 to  $h$ . Moreover, we have

$$\begin{aligned} |\{h > 0\}| + \frac{1}{2d} \int_{\mathbb{R}^d} |\nabla h|^2 dx &= |\{g > 0\}| + \frac{1}{2d(\|g\|_2)^2} \int_{\mathbb{R}^d} |\nabla g|^2 dx \\ &\leq \frac{1}{(1 - \varepsilon^{49})^2} \left( |\{g > 0\}| + \frac{1}{2d} \int_{\mathbb{R}^d} |\nabla g|^2 dx \right). \end{aligned}$$

In the end, we get that, for  $\varepsilon$  small enough,

$$|\{g > 0\}| + \frac{1}{2d} \int_{\mathbb{R}^d} |\nabla g|^2 dx \geq (1 - \varepsilon^{49})^2 (\chi_d + \varepsilon^{48}),$$

and the last term is larger than  $\chi_d + \varepsilon^{49}$  for  $\varepsilon$  sufficiently small.  $\square$

## 7.2 Application of Faber–Krahn: proof of Theorem 1.5

The time has come to apply the quantitative Faber–Krahn inequality to the random walk in order to prove Theorem 1.5. Note that Theorem 1.5 is the analogue of Proposition 3.1 in [6].

We shall apply theorem 6.6 to the set  $\mathcal{L}_n$ . To this end, we consider a function  $g$  satisfying the constraints of the infimum appearing in theorem 6.6. So, let  $g$  be a continuous function  $\mathbb{R}^d \rightarrow \mathbb{R}^+$  such that  $g$  is  $C^\infty$  on  $\mathbb{R}^d \setminus \frac{1}{n}D^d$ , and there exists  $\ell \in \mathcal{L}_n$  satisfying

$$\|g^2 - \ell\|_{L^1(\mathbb{R}^d)} \leq \frac{3}{n^{1/16}}.$$

This implies in particular that

$$1 - \frac{3}{n^{1/16}} \leq \|g\|_2 \leq 1 + \frac{3}{n^{1/16}}.$$

Since  $\ell$  belongs to  $\mathcal{L}_n$ , then we have, for any  $x \in \mathbb{R}^d$ , for  $n$  large enough,

$$\begin{aligned} \frac{1}{n^{1/800}} &\leq \|\ell - (\phi_x)^2\|_1 \leq \|\ell - g^2\|_1 + \|g^2 - (\phi_x)^2\|_1 \\ &\leq \frac{3}{n^{1/16}} + \int_{\mathbb{R}^d} |g - \phi_x| \times |g + \phi_x| dx \\ &\leq \frac{3}{n^{1/16}} + \|g - \phi_x\|_2 \times \|g + \phi_x\|_2 \\ &\leq \frac{3}{n^{1/16}} + 3\|g - \phi_x\|_2. \end{aligned}$$

It follows that, for  $n$  large enough,

$$\forall x \in \mathbb{R}^d \quad \|g - \phi_x\|_2 \geq \frac{1}{4n^{1/800}}.$$

We apply the inequality of corollary 7.21 to  $g$  with  $\varepsilon^{49} = 3/n^{1/16}$ . For  $n$  large enough, we have

$$|\{g > 0\}| + \frac{1}{2d} \int_{\mathbb{R}^d} |\nabla g|^2 dx \geq \chi_d + \frac{1}{n^{1/16}}. \quad (7.22)$$

In order to exploit the inequality of theorem 6.6, we will first restrict the set of the functions  $g$  which are relevant in the infimum. Let  $\kappa \geq 1$  and let  $\delta > 0$  be such that

$$\omega_d \left( \frac{\lambda_d}{2\delta} \right)^{d/2} \geq 2\kappa.$$

Let us set

$$G = \{x \in \mathbb{R}^d : g(x) > 0\}.$$

Let  $r > 0$  be such that  $|G| = r^d \omega_d$ . By the classical Faber–Krahn inequality 7.2, we have

$$\int_{\mathbb{R}^d} |\nabla g|^2 dx \geq \lambda(G) (\|g\|_2)^2 \geq \frac{\lambda_d}{r^2} \left(1 - \frac{3}{n^{1/16}}\right)^2.$$

Suppose that

$$\int_{\mathbb{R}^d} |\nabla g|^2 dx \leq \delta.$$

We would then have, for  $n$  large enough,

$$r^2 \geq \frac{\lambda_d}{2\delta}, \quad |G| = \omega_d r^d \geq \omega_d \left( \frac{\lambda_d}{2\delta} \right)^{d/2} \geq 2\kappa,$$

and for such a function  $g$ , the functional in the infimum is larger or equal than  $2\kappa$ . This will also be the case if the Dirichlet energy of  $g$  is too large, say larger than  $4d\kappa$ . So, up to terms which are negligible compared to  $\exp(-\kappa n^d)$ , we can restrict ourselves to functions  $g$  such that

$$\delta \leq \int_{\mathbb{R}^d} |\nabla g|^2 dx \leq 4d\kappa.$$

We have then, for  $n$  large enough,

$$\left( \max \left( \sqrt{\int_{\mathbb{R}^d} |\nabla g|^2 dx} - \frac{\sqrt{d}}{n^{1/8}}, 0 \right) \right)^2 \geq \int_{\mathbb{R}^d} |\nabla g|^2 dx - \frac{2\sqrt{d}}{n^{1/8}} \sqrt{4d\kappa} \geq \int_{\mathbb{R}^d} |\nabla g|^2 dx - \frac{1}{n^{1/9}}. \quad (7.23)$$

We apply theorem 6.6 and we use inequality (7.23) to simplify the infimum. We have, for  $n$  large enough,

$$\begin{aligned} E(e^{-|R_N|}; \ell_N \in \mathcal{L}_n) &\leq 4 \exp(-\kappa n^d) + \\ &\exp \left( 2n^{d-1/8} - n^d (1 - n^{-1/4}) \inf \left\{ |\text{supp } g| + \frac{1}{2d} \int_{\mathbb{R}^d} |\nabla g|^2 dx - \frac{1}{2dn^{1/9}} \right. \right. \\ &\quad \left. \left. : g \text{ is continuous } \mathbb{R}^d \rightarrow \mathbb{R}^+ \text{ and } C^\infty \text{ on } \mathbb{R}^d \setminus \frac{1}{n} D^d, \quad \exists \ell \in \mathcal{L} \quad \left\| g^2 - \ell \right\|_{L^1(\mathbb{R}^d)} \leq \frac{3}{n^{1/16}} \right\} \right). \end{aligned}$$

We can finally use inequality (7.22)! We obtain

$$E(e^{-|R_N|}; \ell_N \in \mathcal{L}_n) \leq 4 \exp(-\kappa n^d) + \exp \left( 2n^{d-1/8} - n^d (1 - n^{-1/4}) \left( \chi_d + \frac{1}{n^{1/16}} - \frac{1}{2dn^{1/9}} \right) \right).$$

By choosing  $\kappa > \chi_d$ , for  $n$  large enough, we obtain the statement of theorem 1.5.

## 8 Filling the ball and proof of Theorem 1.7

We now come to the proof of Theorem 1.7. We will show that with high probability under  $\tilde{P}_N$  a ball of approximately the right radius is entirely filled. The analogue of this result in [6] (Proposition 4.1), is however only valid for  $d = 2$ , whereas the proof below holds for any  $d \geq 2$ . Fix  $x \in \mathbb{R}^d$ , and let

$$\mathcal{G}_{loc,x} = \{\|\ell_N - (\phi_x)^2\|_1 \leq 1/n^s\}$$

with  $s = 1/800$  is as in Theorem 1.5 and we recall that the rescaled local time profile  $\ell_N$  is defined in (6.5). For  $\kappa > 0$  and  $x \in \mathbb{R}^d$ ,

$$\mathcal{R}_{\kappa,x} = \{\forall z \in B(x, \rho_d(1 - n^{-\kappa})) \cap \mathbb{Z}_n^d, \ell_N(z) > 0\}.$$

This is the event that the Euclidean ball of radius  $\rho_d n(1 - n^{-\kappa})$  around  $nx$ , restricted to the unscaled lattice  $\mathbb{Z}^d$ , is filled by the walk. The main result of this section, which is the analogue of Proposition 4.1 in [6], is the following.

**Theorem 8.1** There exists  $\kappa > 0$  such that for all  $n$  (or equivalently  $N$ ) large enough,

$$\frac{1}{Z_N} E(e^{-|R_N|}; \mathcal{G}_{loc,x}; (\mathcal{R}_{\kappa,x})^c) \leq \exp(-n^\kappa).$$

### 8.1 Time spent in mesoscopic balls

We recall that constants  $C, c > 0$  denote constants depending only on the dimension, whose precise numerical value is allowed to change from line to line. We will use Landau's notations  $O(\cdot), o(\cdot)$  where the implicit constants are allowed to depend on the dimension only. The notation  $\lesssim, \gtrsim, \asymp$  denote inequalities and equality up to constant respectively: thus  $a_n \lesssim b_n$  means  $a_n \leq Cb_n$  for some constant  $C$  (depending only on the dimension).  $\kappa$  will be a small parameter eventually chosen in a way that depends only on the dimension, so at the end of the proof we will be able to absorb its value in a generic constant  $c > 0$ , but we will refrain from doing so during the course of the proof.

Without loss of generality we take  $x = 0$  in the rest of the proof of Theorem 8.1; write  $\mathcal{G}_{loc}$  and  $\mathcal{R}_\kappa$  for  $\mathcal{G}_{loc,x}$  and  $\mathcal{R}_{\kappa,x}$ . Fix  $z \in B(0, \rho_d n(1 - n^{-\kappa}))$ . We let  $m = \rho_d n^{1-2\kappa}$  where  $\kappa$  is sufficiently small;  $m$  is a mesoscopic scale (quite close to  $n$ ) and our first task will be to control the amount of time spent in a ball of that scale around the point  $z$ . Let

$$B = B(z, m) \cap \mathbb{Z}^d$$

be the discretised ball of radius  $m$  around  $z$ . Note that  $B \subset B(0, \rho_d n(1 - n^{-\kappa}/2))$ . The idea will be to condition on some information outside  $B$  including the local time of the random walk on every site outside of  $B$ . Let  $B^\circ$  denote the ball of radius  $m/2$  around  $x$ , i.e.,  $B^\circ = B(x, m/2) \cap \mathbb{Z}^d$ .

We make the following simple deterministic observations (recall that we expect  $L_N(x)$  to be typically of order  $n^2$  at any point in the bulk of  $R_N$ ).

**Lemma 8.2** If  $\mathcal{G}_{loc}$  holds, then necessarily, for some sufficiently small but fixed  $\kappa$  and  $\delta > 0$  depending only on the dimension  $d$ , we have

$$L_N(B^\circ) \geq \delta m^d n^{2-2\kappa}. \quad (8.3)$$



Furthermore,

$$|R_N| \geq \omega_d(\rho_d n)^d - cn^{d-\varepsilon} \quad (8.4)$$

where  $\varepsilon$  depends only  $d$ .

**Proof.** Let  $\varphi = \phi^2$ . Note that as  $z \rightarrow \partial B(0, \rho_d)$ , for some constant  $C > 0$  depending only on the dimension,

$$\varphi(z) \sim Cd(z)^2, \text{ where } d(z) = \text{dist}_{\mathbb{R}^d}(z, \partial B(0, \rho_d)). \quad (8.5)$$

Indeed recall that  $\phi$ , the first eigenfunction in  $B(0, \rho_d)$ , has nonzero normal derivative on  $\partial B(0, \rho_d)$ .

Let  $y \in B$ . Then  $\text{dist}(y, \partial B(0, \rho_d n)) \geq \rho_d n^{1-\kappa}/2$  and so  $d(y/n) \geq \rho_d/(2n^\kappa)$ . Now suppose for contradiction that  $L_N(B^\circ) \leq \delta m^d n^{2-2\kappa}$ . Then necessarily, by definition of  $\ell_N$  as a function on  $\mathbb{R}^d$  in (6.5),

$$\int_{(B^\circ/n)} \ell_N(y) dy \leq \sum_{y \in \mathbb{Z}^d \cap B^\circ} \frac{1}{n^{d+2}} L_N(y) = \frac{L_N(B^\circ)}{n^{d+2}} \leq \delta(m/n)^d n^{-2\kappa} = \delta \rho_d^d n^{-2(d+1)\kappa}$$

whereas

$$\int_{(B^\circ/n)} \varphi(y) dy \geq \omega_d \left(\frac{m/2}{n}\right)^d C \rho_d^2 n^{-2\kappa} / 4 = \left(\frac{C \rho_d^{d+2} \omega_d}{2^{2+d}}\right) n^{-2(d+1)\kappa}.$$

Thus if we take  $\delta = C \rho_d^{d+2} \omega_d / 2^{3+d}$ ,

$$\|\ell_N - \varphi\|_1 \geq \left| \int_{(B^\circ/n)} \ell_N(y) dy - \int_{(B^\circ/n)} \varphi(y) dy \right| \geq \left(\frac{C \rho_d^{d+2} \omega_d}{2^{3+d}}\right) n^{-2(d+1)\kappa}.$$

Hence if  $\kappa$  is small enough that  $2(d+1)\kappa \leq s$  we see that this cannot hold at the same time as  $\mathcal{G}_{loc}$ . This shows (8.3) with this choice of parameters.

The proof of (8.4) follows a similar argument.  $\square$

## 8.2 Bridges in mesoscopic balls

Let us call an interval of time  $[t_1, t_2]$  a **bridge** if  $X[t_1, t_2] \subset B$  and  $X_{t_1}, X_{t_2} \in B^\circ$ . (We warn the reader that this differs from the more standard notion of bridge). We call  $a = X_{t_1}$  and  $b = X_{t_2}$  its endpoints and  $t_2 - t_1$  its length; note that we require that  $a, b \in B$  be far away from  $\partial B$ ; in fact we require them to be in  $B^\circ$ . We will consider bridges of length at least  $m^2$  and our first goal will be to show that there are sufficiently many such disjoint bridges.

Let  $\mathcal{G}_{br} = \mathcal{G}_{br, \delta}$  be the event that there are at least  $a \delta n^{2-2\kappa} m^{d-2}$  bridges of length  $m^2$ , where  $\delta$  is as in Lemma 8.2, and where  $a$  is a suitable constant depending also only on the dimension, which will be chosen below.

**Lemma 8.6** We can choose  $\kappa$  sufficiently small and  $a > 0$ , depending only on the dimension, so that

$$E(e^{-|R_N|}; \mathcal{G}_{loc}; (\mathcal{G}_{br})^c) \leq Z_N \exp(-cn^{d-6\kappa}),$$

for some constant  $c > 0$  depending only on  $d$ .

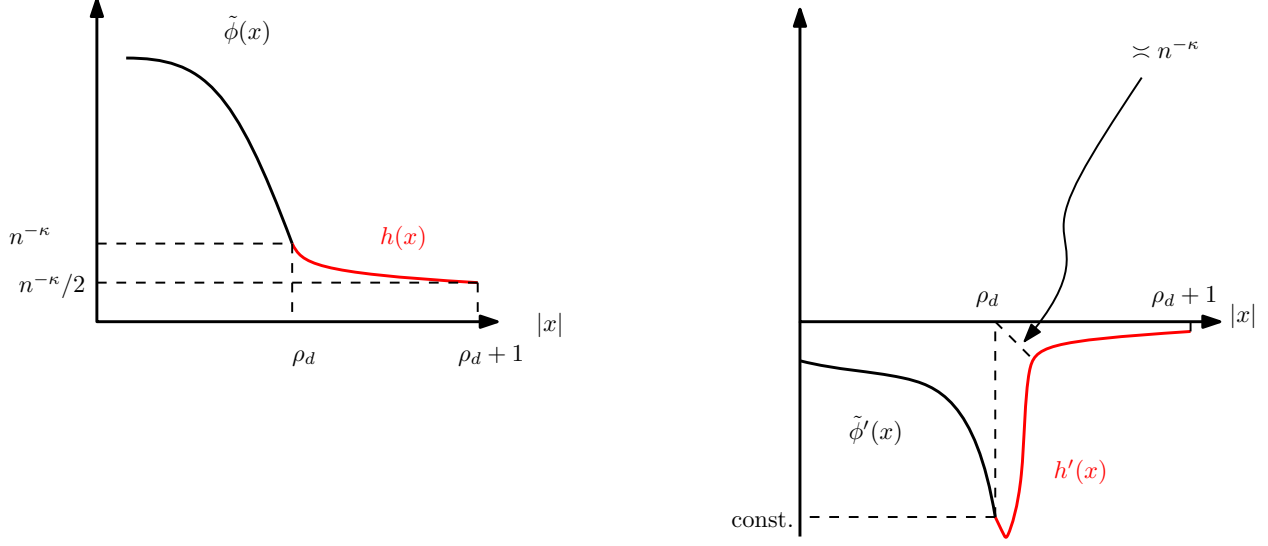


Figure 1: Choice of function  $h$ : as desired,  $h$  decreases from  $n^{-\kappa}$  to  $n^{-\kappa}/2$  over  $[\rho, \rho + 1]$ . This requires the area above the red function on the right hand side (its derivative) to be less than  $n^{-\kappa}/2$ . The value of the constant in the right hand side is the normal (radial) derivative of  $\phi$ . To achieve this, the second derivative will be at most  $\lesssim n^\kappa$ , as claimed in (8.7).

**Proof.** We already know that the walk spends at least  $\delta n^{2-2\kappa} m^d$  units of time in  $B^\circ$ . Roughly speaking, every time the walk is in  $B^\circ$  there is a positive probability that during the next  $m^2$  units of time, the walk stays in  $B$  and its position at the end of this interval is again in  $B^\circ$ , thereby completing a bridge; independently of the past. So we wish to use standard Chernoff bounds for deviations of binomial random variables. In order to implement this strategy, we must however take care that we are working under the weighted probability measure  $\tilde{P}_N$  and not  $P$ . We will deal with this complication by performing a suitable change of measure (which as it turns out is essentially the same as the one used by Bolthausen in his proof of Proposition 4.1 in [6]). The first step will be to work in continuous time rather than discrete time. By Lemma 8.2 (and more precisely (8.4)) we write

$$E(e^{-|R_N|}; \mathcal{G}_{loc}; (\mathcal{G}_{br})^c) \leq \exp(-\omega_d(\rho_d n)^d + cn^{d-\varepsilon}) P(\mathcal{G}_{loc}; \mathcal{G}_{br}^c)$$

and we interpret the event in the right hand side of the above inequality as an event for the jump chain of a continuous Markov chain whose jump rates from  $x$  to  $y$  is  $1/(2d)$  if  $x$  and  $y$  are neighbours and zero else. Let  $(\tilde{X}_t, t \geq 0)$  be this process, let  $\tilde{P}$  denote its law and let  $J(\tilde{X})$  be the jump chain of  $\tilde{X}$ . Set  $\tilde{\mathcal{G}}_{loc} = \{J(\tilde{X}) \in \mathcal{G}_{loc}\}$  and, similarly,  $\tilde{\mathcal{G}}_{br} = \{J(\tilde{X}) \in \mathcal{G}_{br}\}$ . Then  $P(\mathcal{G}_{loc}; (\mathcal{G}_{br})^c) = \tilde{P}(\tilde{\mathcal{G}}_{loc}; (\tilde{\mathcal{G}}_{br})^c)$ .

Define a modification  $\tilde{\phi}$  of  $\phi$  as follows:

$$\tilde{\phi}(x) = \phi(x) + n^{-\kappa} \quad ; \quad x \in B(0, \rho_d)$$

We will also define  $\tilde{\phi}$  outside of  $B(0, \rho_d)$  so that it is positive everywhere and also in such a way that it is reasonably smooth near the boundary of that ball (indeed, if we do not take a positive function we cannot use it to change the underlying probability measure). Of course we could set

$\tilde{\phi}$  to be constant outside of that ball, equal to its value on the boundary, but this turns out to not be sufficiently smooth; in particular the Laplacian on this sphere would be too large.

Instead we define

$$\tilde{\phi}(x) = h(|x|)$$

where

$$h : [\rho_d, \rho_d + 1] \rightarrow [n^{-\kappa}/2, n^{-\kappa}]$$

is a smooth monotone decreasing convex function such that its derivative at  $1 + \rho_d$  is 0 while its derivative at  $\rho_d$  is the radial derivative of  $\phi$  on  $\partial B(0, \rho_d)$ , and such that

$$\sup_{x \in [\rho_d, \rho_d + 1]} \frac{d^2}{dx^2} h \lesssim n^\kappa. \quad (8.7)$$

It is elementary to check that such a function  $h$  exists, so that  $\tilde{\phi}$  is well defined (see Figure 1 for an illustration of the function  $\tilde{\phi}$  and its derivative).

With the help of  $\tilde{\phi}$  we can define a new probability measure  $\tilde{Q}$  to be the law of the Markov chain whose transition rates are given by

$$q(x, y) = \frac{\tilde{\phi}(y/n)}{\tilde{\phi}(x/n)}$$

whenever  $x \neq y$  are neighbours in  $\mathbb{Z}^d$ , and  $q(x, y) = 0$  otherwise. Note that since  $\tilde{\phi}$  is positive,  $\tilde{Q}$  is indeed equivalent to  $\tilde{P}$  and furthermore, letting  $\tilde{t}_N$  be the first time  $\tilde{X}$  has jumped  $N$  times,

$$\left. \frac{d\tilde{P}}{d\tilde{Q}} \right|_{\mathcal{F}_{\tilde{t}_N}} = \frac{\tilde{\phi}(\tilde{X}_0/n)}{\tilde{\phi}(\tilde{X}_{\tilde{t}_N}/n)} \exp \left( \int_0^{\tilde{t}_N} \frac{\Delta_1 \tilde{\phi}}{\tilde{\phi}}(\tilde{X}_s/n) ds \right). \quad (8.8)$$

(See, e.g., e.g. [17], IV, (22.8)). Here recall that  $\Delta_1$  is the discrete Laplacian, i.e.  $2d$  times the generator of the Markov chain  $\tilde{X}$  under  $\tilde{P}$ .

**Step 1.** We will argue that

$$E(e^{-|R_N|}; \mathcal{G}_{loc}; (\mathcal{G}_{br})^c) \leq Z_N e^{-N^{1-5\kappa}} + Z_N e^{3n^{d-\varepsilon}} \tilde{Q}(\tilde{\mathcal{G}}_{loc}; (\tilde{\mathcal{G}}_{br})^c). \quad (8.9)$$

Essentially the proof consists in analysing carefully the integral in (8.8). We follow roughly the arguments in [6] (see equations (4.3) and (4.4)), with additional details. We start by observing that if  $x \in B(0, \rho_d n) \cap \mathbb{Z}^d$  is such that all its  $2d$  neighbours are also in  $B(0, \rho_d n) \cap \mathbb{Z}^d$ , then by a Taylor expansion,

$$\begin{aligned} \Delta_1 \tilde{\phi}(x/n) &= \Delta_1 \phi(x/n) = \frac{1}{2} \Delta \phi(x/n) + O(n^{-3}) \\ &= -\frac{\lambda}{2n^2} \phi(x/n) + O(n^{-3}) \\ &= -\frac{\lambda}{2n^2} \tilde{\phi}(x/n) + O(n^{-3}) + O(n^{-2-\kappa}). \end{aligned}$$

Assume that  $\kappa < 1$  without loss of generality so that  $O(n^{-3}) = O(n^{-2-\kappa})$ . Hence if furthermore  $x \in B(0, \rho_d n(1 - n^{-0.9\kappa}))$ ,

$$\left| \frac{\Delta_1 \tilde{\phi}}{\tilde{\phi}}(x/n) + \frac{\lambda}{2n^2} \right| \lesssim n^{-2-\kappa+0.9\kappa} \leq n^{-2-0.1\kappa} \quad (8.10)$$

using (8.5).

Furthermore, for any other  $x$ ,

$$|\Delta_1 \tilde{\phi}(x/n)| \lesssim n^{-2+\kappa}$$

by (8.7) (where the implicit constant depends only on  $\tilde{\phi}$  and so on the dimension), so that

$$\left| \frac{\Delta_1 \tilde{\phi}}{\tilde{\phi}}(x/n) + \frac{\lambda}{2n^2} \right| \lesssim n^{-2+2\kappa} \quad (8.11)$$

for such  $x$ , and the same remark holds about the implicit constant.

Combining (8.10) and (8.11), it follows that

$$\int_0^{\tilde{t}_N} \frac{\Delta_1 \tilde{\phi}}{\tilde{\phi}}(\tilde{X}_s/n) ds = (-\lambda/2) \frac{\tilde{t}_N}{n^2} + O(\tilde{t}_N n^{-2-0.1\kappa}) + O\left(n^{-2+2\kappa} \tilde{L}_{\tilde{t}_N}(B(0, \rho_d n(1 - n^{-0.9\kappa}))^c)\right), \quad (8.12)$$

where for a set  $A \subset \mathbb{Z}^d$  and a time  $t > 0$ ,  $\tilde{L}_t(A) = \int_0^t 1_{\{\tilde{X}_s \in A\}} ds$  is the local time of  $\tilde{X}$  (in continuous time) in the set  $A$  up to time  $t$ .

Set  $\tilde{\mathcal{A}} = \tilde{\mathcal{G}}_{loc} \cap (\tilde{\mathcal{G}}_{br})^c$ . Set  $\tilde{\mathcal{A}}'$  to be the event

$$\tilde{\mathcal{A}} \quad ; \quad |\tilde{t}_N - N| \leq N^{1-\kappa} \quad ; \quad \tilde{L}_{\tilde{t}_N}(B(0, \rho_d n(1 - n^{-0.9\kappa}))^c) \leq N n^{-2.1\kappa}.$$

Then, still writing  $R_N$  for the range of the (continuous time) walk at the time  $\tilde{t}_N$  of its  $N$ th jump,

$$E_{\tilde{P}}(e^{-|R_N|}; \tilde{\mathcal{A}}) \leq E_{\tilde{P}}(e^{-|R_N|}; \tilde{\mathcal{A}}') + \tilde{P}(\tilde{\mathcal{A}} \setminus \tilde{\mathcal{A}}'). \quad (8.13)$$

We bound separately each of those terms. We start with the first term. On  $\mathcal{A}'$ , we see that

$$\int_0^{\tilde{t}_N} \frac{\Delta_1 \tilde{\phi}}{\tilde{\phi}}(\tilde{X}_s/n) ds = (-\lambda/2) n^d + O(n^{d-(d+2)\kappa}) + O(n^{d-0.1\kappa}) + O(n^{d-0.1\kappa})$$

so that if  $\kappa < \varepsilon/(d+2)$  (which implies  $\kappa < 10\varepsilon$ ), all error terms are  $O(n^{d-(d+2)\kappa}) \leq n^{d-\varepsilon}$  and thus, plugging into (8.8), and using the fact that  $\max_{B(0, \rho_d)} \tilde{\phi} \lesssim 1$ ,  $\min_{B(0, \rho_d)} \tilde{\phi} \gtrsim n^{-\kappa}$ , as well as the already established lower bound of Proposition 2.1 on  $Z_N$ ,

$$\begin{aligned} E_{\tilde{P}}(e^{-|R_N|}; \tilde{\mathcal{A}}') &\lesssim n^\kappa \exp(-\omega_d(\rho_d n)^d - (\lambda/2)n^d + 2n^{d-\varepsilon}) \tilde{Q}(\tilde{\mathcal{A}}') \\ &\leq Z_N e^{2n^{d-\varepsilon} + O(n^{d-1})} \tilde{Q}(\tilde{\mathcal{A}}') \\ &\leq Z_N e^{3n^{d-\varepsilon}} \tilde{Q}(\mathcal{G}_{loc}; (\mathcal{G}_{br})^c). \end{aligned} \quad (8.14)$$

Let us now deal with the second term in (8.13). We have,

$$\tilde{P}(\tilde{\mathcal{A}} \setminus \tilde{\mathcal{A}}') \leq \tilde{P}(|\tilde{t}_N - N| \geq N^{1-\kappa}) + \tilde{P}(\tilde{L}_{\tilde{t}_N}(B(0, \rho_d n(1 - n^{-0.9\kappa}))^c) \geq N n^{-2.1\kappa}; \tilde{\mathcal{G}}_{loc}). \quad (8.15)$$

Now, standard Chernoff estimates for exponential random variables show that

$$P(|\tilde{t}_N - N| \geq N^{1-\kappa}) \leq \exp(-cN^{1-3\kappa}) \quad (8.16)$$

for some constant  $c > 0$ . Furthermore, on  $\mathcal{G}_{loc}$ , deterministically we have

$$L_N(B(0, \rho_d n(1 - n^{-0.9\kappa})^c) \leq Nn^{-2.2\kappa},$$

since otherwise, reasoning as in Lemma 8.2,

$$\int_{B(0, \rho_d(1-n^{-0.9\kappa}))^c} \ell_N(y) dy \geq n^{-2.2\kappa} \text{ while } \int_{B(0, \rho_d(1-n^{-0.9\kappa}))^c} \phi^2(y) dy \leq C\omega_d n^{-2.7\kappa}$$

and we would deduce that for  $n$  large enough,  $\|\ell_N - \phi^2\| > n^{-2.3\kappa}$  which by definition cannot take place on  $\mathcal{G}_{loc}$  if  $\kappa \leq s/2.3$  (with  $s = 1/800$  as in Theorem 1.5. Hence

$$\tilde{P}(\tilde{L}_{\tilde{t}_N}(B(0, \rho_d n(1 - n^{-0.9\kappa}))^c) \geq Nn^{-2.1\kappa}; \tilde{\mathcal{G}}_{loc}) \leq P\left(\sum_{i=1}^{\lfloor Nn^{-2.2\kappa} \rfloor} X_i \geq Nn^{-2.1\kappa}\right) \quad (8.17)$$

where  $X_i$  are independent unit exponential random variables. Hence for all  $n$  (or equivalently  $N$ ) large enough,

$$P\left(\sum_{i=1}^{\lfloor Nn^{-2.2\kappa} \rfloor} X_i \geq Nn^{-2.1\kappa}\right) \leq P\left(\sum_{i=1}^{\lfloor Nn^{-2.2\kappa} \rfloor} X_i \geq 2Nn^{-2.2\kappa}\right) \leq \exp(-cNn^{-2.2\kappa}) \leq \exp(-N^{1-3\kappa}),$$

for some constant  $c > 0$ , by elementary Chernoff estimates for exponential random variables. Hence by (8.17),

$$\tilde{P}(\tilde{L}_{\tilde{t}_N}(B(0, \rho_d n(1 - n^{-0.9\kappa}))^c) \geq Nn^{-2.1\kappa}; \tilde{\mathcal{G}}_{loc}) \leq \exp(-N^{1-3\kappa}).$$

Plugging into (8.15), we deduce that if  $\kappa$  is sufficiently small, for all  $N$  large enough,

$$\tilde{P}(\mathcal{A} \setminus \mathcal{A}') \leq \exp(-N^{1-4\kappa}).$$

Since  $Z_N = \exp(-O(n^d))$ , we deduce that

$$\tilde{P}(\mathcal{A} \setminus \mathcal{A}') \leq Z_N \exp(-N^{1-5\kappa}). \quad (8.18)$$

Combining (8.18), (8.14) and (8.13), we obtain

$$E_{\tilde{P}}(e^{-|R_N|}; \mathcal{G}_{loc}; (\mathcal{G}_{br})^c) \leq Z_N e^{-N^{1-5\kappa}} + Z_N e^{3n^{d-\varepsilon}} \tilde{Q}(\tilde{\mathcal{G}}_{loc}; (\tilde{\mathcal{G}}_{br})^c),$$

as desired in (8.9).

**Step 2.** Now it remains to show that  $\tilde{\mathcal{G}}_{loc} \cap \tilde{\mathcal{G}}_{br}$  is overwhelmingly likely under  $\tilde{Q}$ . More precisely, we will argue that

$$\tilde{Q}((\tilde{\mathcal{G}}_{br})^c; \tilde{\mathcal{G}}_{loc}) \leq \exp(-cn^{2-2\kappa}m^{d-2}). \quad (8.19)$$

We claim that every time the walk is in  $B^\circ$ , there is a probability bounded below by a constant  $p$ , say, depending only on the dimension, such that under  $\tilde{Q}$ , the walk will perform a bridge of duration  $m^2$  (recall that this means the walk remains in  $B$  for the next  $m^2$  jumps and ends up in

$B^\circ$  again after this time), uniformly over the initial position in  $B^\circ$  of the walk. To see this, note that  $\min_{B/n} \tilde{\phi} = (1 - o(1)) \max_{B/n} \tilde{\phi}$ , so that the Radon–Nikodym derivative  $d\tilde{Q}/d\tilde{P}$  during an interval of time consisting of the first  $m^2$  jumps of the chain (call this time  $\tilde{t}_{m^2}$ ) is at least

$$\begin{aligned} \frac{d\tilde{Q}}{d\tilde{P}} &\geq \frac{\min_{B/n} \tilde{\phi}}{\max_{B/n} \tilde{\phi}} \exp(\tilde{t}_{m^2} \inf_{x \in B} \frac{\Delta_1 \tilde{\phi}}{\tilde{\phi}}(x/n)) \\ &\geq (1 + o(1)) \exp(-(\lambda/2 + o(1)) \frac{\tilde{t}_{m^2}}{n^2}). \end{aligned} \quad (8.20)$$

Now, under  $\tilde{P}$ ,  $\tilde{t}_{m^2} \leq 2m^2$  with probability at least  $3/4$ , if  $m$  (equivalently  $n$  or  $N$ ) is large enough. Moreover, the probability of making a bridge of duration  $m^2$ , under  $\tilde{P}$ , is clearly at least  $p$  for some constant  $p > 0$  depending only on  $d$ . Since the latter event depends only on the jump chain and the former event depends only on the time parametrisation, which are independent processes under  $\tilde{P}$ , we conclude from (8.20) that for any  $a \in B^\circ$ ,

$$\begin{aligned} \tilde{Q}_a(X[0, \tilde{t}_{m^2}] \text{ is a bridge}) &\geq E_{\tilde{P}_a} \left( \frac{d\tilde{Q}_a}{d\tilde{P}_a}; X[0, \tilde{t}_{m^2}] \text{ is a bridge} \right) \\ &\geq (1 + o(1)) E_{\tilde{P}_a} \left( \exp(-(\lambda/2 + o(1)) \frac{\tilde{t}_{m^2}}{n^2}); \tilde{t}_{m^2} \leq 2m^2; X[0, \tilde{t}_{m^2}] \text{ is a bridge} \right) \\ &\geq (3/4)(1 + o(1)) \tilde{P}_a(X[0, \tilde{t}_{m^2}] \text{ is a bridge}) \geq (3/4)p(1 + o(1)). \end{aligned}$$

Hence  $\tilde{Q}_a(X[0, \tilde{t}_{m^2}] \text{ is a bridge}) \geq p/2$  for  $n$  large enough; thereby proving what we desired (with  $p$  replaced by  $p/2$ , a distinction which is of no consequence in the rest of the argument).

Since we know by (8.3) that on  $\tilde{\mathcal{G}}_{loc}$ , the total amount of discrete steps in  $B^\circ$  (and hence in  $B$ ) is deterministically at least  $\delta m^d n^{2-2\kappa}$ , and each time the walk is in  $B^\circ$  there is a probability  $p/2$  (under  $\tilde{Q}$ ) to make a bridge over the next  $m^2$  jumps, we deduce that the number of bridges stochastically dominates a binomial random variable of parameters  $(\delta m^{d-2} n^{2-2\kappa}, p/2)$ , under  $\tilde{Q}$ . More precisely, let

$$\sharp(s, t) := \text{number of jumps by } \tilde{X} \text{ during } (s, t],$$

and define the sequence of stopping times:

$$\sigma_1 = \inf\{t > 0 : \tilde{X}_t \in B^\circ\}; \quad \tau_1 = \inf\{t > 0 : \sharp(\sigma_1, t) \geq m^2\} \wedge \inf\{t > \sigma_1 : \tilde{X}_t \notin B^\circ\};$$

then inductively, for  $i \geq 2$ :

$$\sigma_i = \inf\{t > \tau_{i-1} : \tilde{X}_t \in B^\circ\}; \quad \tau_i = \inf\{t > 0 : \sharp(\sigma_i, t) \geq m^2\} \wedge \inf\{t > \sigma_i : \tilde{X}_t \notin B^\circ\}.$$

Let

$$\xi_i = 1_{\{\sharp(\sigma_i, \tau_i) \geq m^2\}}; \quad i \geq 1$$

be the Bernoulli variable which is the indicator of the event that the  $i$ th trial results in a bridge. Then note that by Lemma 8.2, if  $j = \lfloor \delta m^{d-2} n^{2-2\kappa} \rfloor$ , then  $\tau_j \leq t_N$  no matter what on  $\tilde{\mathcal{G}}_{loc}$ . Hence if  $S = \sum_{i=1}^j \xi_i$ , then

$$\{S \geq (\delta p/4) m^{d-2} n^{2-2\kappa}\} \subset \tilde{\mathcal{G}}_{br}$$

where the constant  $a > 0$  defining  $\mathcal{G}_{br}$  is taken to be  $a = p/4$ .

Moreover by the Markov property and the above,

$$\tilde{Q}(\xi_i = 1 | (X_t, t \leq \sigma_i)) \geq p/2.$$

Hence  $S$  dominates stochastically a Binomial random variables with parameters  $j$  and  $p/2$ . By standard Chernoff bounds for binomials, for some constant  $c > 0$ , we deduce that

$$\tilde{Q}((\tilde{\mathcal{G}}_{br})^c; \tilde{\mathcal{G}}_{loc}) \leq \exp(-cn^{2-2\kappa}m^{d-2}),$$

which shows (8.19).

Plugging (8.19) into (8.9), we deduce

$$E(e^{-|R_N|}; \mathcal{G}_{loc}; (\mathcal{G}_{br})^c) \leq Z_N(e^{-N^{1-5\kappa}} + e^{n^{d-\varepsilon}} e^{-cn^{2-2\kappa}m^{d-2}}).$$

Since  $m = \rho_d n^{1-2\kappa}$ ,  $\varepsilon$  is fixed (by Lemma 8.2) in a way that depends only on the dimension, and we are free to choose  $\kappa$  as small as we want, we can choose it so that  $n^{2-2\kappa}m^{d-2} = n^{d-6\kappa}$  is much greater than  $n^{d-\varepsilon}$  (i.e., we assume  $6\kappa < \varepsilon$ ) and then for all  $n$  large enough we have

$$E(e^{-|R_N|}; \mathcal{G}_{loc}; (\mathcal{G}_{br})^c) \leq Z_N(e^{-N^{1-5\kappa}} + e^{-cn^{2-2\kappa}m^{d-2}}) \leq Z_N e^{-cn^{d-6\kappa}}.$$

This completes the proof of Lemma 8.6.  $\square$

### 8.3 Conditioning and summation

Call  $P^*$  the conditional probability given the local time at every site in  $\mathbb{Z}^d \setminus B$ . Let  $k \geq 1$ ,  $\mathcal{X} = \{x_1, \dots, x_k\} \subset B$  be a subset of  $k$  distinct points in  $B$  and let  $\mathcal{B}_{\mathcal{X}}$  be the bad event that the range in  $B$  avoids exactly those  $k$  points, i.e.,  $R_N \cap B = B \setminus \mathcal{X}$ . Note that

$$\{B \not\subset R_N\} \subset (\mathcal{G}_{br})^c \cup \bigcup_{k=1}^{|B|} \bigcup_{\mathcal{X} \subset B; |\mathcal{X}|=k} (\mathcal{B}_{\mathcal{X}} \cap \mathcal{G}_{br}). \quad (8.21)$$

So we are led to try and analyse expectations of the form

$$\frac{1}{Z_N} E \left[ e^{-|R_N \cap B^c|} E^* [e^{-|R_N \cap B|} 1_{\mathcal{B}_{\mathcal{X}} \cap \mathcal{G}_{br}}] \right]$$

where  $\mathcal{X} \subset B$  and  $|\mathcal{X}| = k \geq 1$  is arbitrary between 1 and  $|B|$ . The key will be the following estimate:

**Lemma 8.22** For constants  $C, c$  depending only on the dimension  $d$ ,

$$E^* \left[ e^{-|R_N \cap B|} 1_{\mathcal{B}_{\mathcal{X}} \cap \mathcal{G}_{br}} \right] \leq C e^{-|B|+k} \exp\left(-c \frac{n^{2-2\kappa} k}{m^2 (\log m)^{2+10d}}\right).$$

We defer the proof of Lemma 8.22 and instead show how it implies Theorem 8.1.

**Proof.** [Proof of Theorem 8.1, assuming Lemma 8.22] Using (8.21) and Lemma 8.6, we see that since  $|R_N \cap B| \leq |B|$ ,

$$\begin{aligned}
\frac{1}{Z_N} E(e^{-R_N} 1_{\mathcal{G}_{loc}} 1_{B \not\subset R_N}) &\leq e^{-cn^{d-6\kappa}} + \sum_{k=1}^{|B|} \sum_{\mathcal{X} \subset B, |\mathcal{X}|=k} \frac{1}{Z_N} E \left[ e^{-|R_N \cap B^c|} E^* [e^{-|R_N \cap B|} 1_{\mathcal{B}_{\mathcal{X}}; \mathcal{G}_{br}}] \right] \\
&\leq e^{-cn^{d-6\kappa}} + C \sum_{k=1}^{|B|} \sum_{\mathcal{X} \subset B, |\mathcal{X}|=k} \frac{1}{Z_N} E \left[ e^{-|R_N \cap B^c|} e^{-|B|+k} e^{-c \frac{n^{2-2\kappa} k}{m^2 (\log m)^{2+10d}}} \right] \\
&\leq e^{-cn^{d-6\kappa}} + C \sum_{k=1}^{|B|} \binom{|B|}{k} e^k e^{-c \frac{n^{2-2\kappa} k}{m^2 (\log m)^{2+10d}}} \frac{1}{Z_N} E \left[ e^{-|R_N|} \right] \\
&\leq e^{-cn^{d-6\kappa}} + C \sum_{k=1}^{m^d} \binom{|B|}{k} \exp \left( k \left( 1 - c \frac{n^{2-2\kappa} k}{m^2 (\log m)^{2+10d}} \right) \right).
\end{aligned}$$

Now note that on the one hand, the entropic factor satisfies  $\binom{|B|}{k} \leq (\omega_d m^d)^k = e^{Ckd \log m}$ . On the other hand since  $m = \rho_d n^{1-2\kappa}$  we have that  $n^{2-2\kappa} / (m^2 (\log m)^{2+10d}) \gtrsim n^{2\kappa} / (\log n)^{2+10d}$ . This is of course much greater than the exponential factor in the entropic term of  $C \log m$ , and so altogether the above series is exponentially decaying. Hence we can conclude that

$$\frac{1}{Z_N} E(e^{-R_N} 1_{\mathcal{G}_{loc}} 1_{B \not\subset R_N}) \leq C \exp(-n^\kappa)$$

where  $\kappa > 0$  and  $C$  depend only on the dimension. Summing over all  $O(n^d)$  possible centres of the ball  $B$  and using a union bound we immediately deduce the statement of Theorem 8.1.  $\square$

## 8.4 Transition probabilities for bridges

We now start the proof of Lemma 8.22. The idea is to show that for each bridge of duration at least  $m^2$  there is a good chance of hitting our  $k$  points. For this we will need the following bounds on the heat kernel of bridge; we will now further condition on the endpoints of the bridge. Let  $P^{a \rightarrow b; \tau}$  denote the law of a bridge of duration  $\tau$  starting from  $a \in B$  and ending in  $b$ , i.e., simple random walk starting from  $a$ , conditioned to be in  $b$  at time  $\tau$  and to remain in  $B$  throughout  $[0, \tau]$ . Implicit in this notation is the fact that  $P^a(X_\tau = b) > 0$ , i.e., the parity of  $b - a$  is the same as  $\tau$ .

We start with the following lemma.

**Lemma 8.23** Suppose  $\text{dist}(x, \partial B) = r$  with  $1 \leq r \leq m$ . Then if  $s \asymp m^2$ ,

$$P_x(X[0, s] \subset B) \asymp \frac{r}{m}. \quad (8.24)$$

The right hand side is essentially the familiar gambler's ruin probability in one dimension. Here it is important that we use a curved ball  $B$  and not a box (otherwise if  $x$  is near a corner the probability would be much smaller). This lemma could be deduced from Proposition 6.9.4 in [14] but we choose to include a proof in order to make the presentation as self-contained as possible.



**Proof.** The lower bound follows easily from an eigenvalue estimate and optional stopping: let  $\lambda_B^1$  denote the principal eigenvalue of the (discrete) Laplacian  $-\Delta_1$  in the ball  $B$  with Dirichlet boundary conditions. Let  $\phi_B^1$  is the corresponding eigenfunction, normalised so that  $m^{-d} \sum_{x \in B} \phi(x) = 1$ . Then note that

$$M_s = \frac{1}{(1 - \lambda_B^1)^s} \phi_B^1(X_s), \quad s = 0, 1, \dots$$

is a nonnegative martingale. Apply the optional stopping theorem at the bounded stopping time  $\tau_B \wedge s$  (where  $\tau_B$  is the first time the walk leaves  $B$ ) to see that

$$\phi_B^1(x) = \frac{1}{(1 - \lambda_B^1)^s} E(\phi_B^1(X_s) 1_{\tau_B > s}) \leq \frac{\|\phi_B^1\|_\infty P_x(\tau_B > s)}{(1 - \lambda_B^1)^s}.$$

Now, we have already mentioned that  $\lambda_B^1 \asymp 1/m^2$  (with implied constants depending as usual only on the dimension) so that when  $s \asymp m^2$ ,

$$\phi_B^1(x) \lesssim \|\phi_B^1\|_\infty P_x(\tau_B > s).$$

Moreover, using Lemma 2.1 of [6], and using known properties of the principal eigenfunction in the continuum (namely that there is a radial derivative on the boundary of the ball), we see that

$$\phi_B^1(x) \lesssim \frac{r}{m} + O(1/m) \lesssim \frac{r}{m}.$$

Since furthermore the principal eigenfunction on the unit ball in the continuum is bounded, using again Lemma 2.1 in [6], we deduce the lower bound

$$P_x(\tau_B > s) \gtrsim \frac{r}{m}, \tag{8.25}$$

which gives the desired lower bound.

In the other direction, let  $\pi$  be the hyperplane tangent to the ball  $B$  closest to the point  $x$ . Let  $H$  denote the half space in the complement of  $\pi$  that contains  $x$ , and let  $\vec{n}$  be the normal vector to  $\pi$ . Let  $\tau_H = \inf\{t \geq 0 : X_t \notin H\}$ , and note that since  $B \subset H$ ,

$$P(\tau_B > s) \leq P(\tau_H > s).$$

We will show that

$$P(\tau_H > s) \lesssim r/m. \tag{8.26}$$

Consider first the case where  $r = \text{dist}(x, \partial B) \geq \log m$ . Then, using a KMT approximation (strong coupling with Brownian motion), see Theorem 7.1.1 in [14], there is a  $d$ -dimensional Brownian motion  $(W_t^d, t \geq 0)$  such that if  $s \asymp m^2$  then for some constant  $C > 0$ ,

$$\sup_{t \leq s} |W_t - X_t| \leq C \log m$$

on an event  $E_1$  of probability at least  $1 - m^{-1}$ . Let

$$\text{osc} = \sup_{0 \leq t \leq s} |W_t^d - W_{[t]}^d|$$

and note that on an event  $E_2$  of probability at least  $1 - O(m)e^{-(\log m)^2/2} \geq 1 - 1/m$ ,

$$\text{osc} \leq \log m.$$

Let  $\pi^-$  denote a hyperplane parallel to  $\pi$  at distance  $C \log m$  from  $\pi$  such that  $\pi^-$  is in the half space which does not contain  $B$ . Let  $T_{\pi^-}$  denote the first (continuous) time when  $W^d$  hits  $\pi^-$ . Then on  $E_1 \cap E_2$ ,  $\tau_H > s$  implies  $T_{\pi^-} > s$  and hence

$$P_x(\tau_H > s) \leq P_x(T_{\pi^-} > s) + P_x(E_1^c) + P_x(E_2^c) \leq 2/m + P_x(T^- > s).$$

Using rotational invariance of Brownian motion and projecting onto  $\vec{n}$ , letting  $W = W^1$  be a standard one-dimensional Brownian motion and using the reflection principle,

$$P_x(T_{\pi^-} > s) = P_{r+C \log m}(W_t \geq 0 \text{ for all } 0 \leq t \leq s) \asymp \frac{r + C \log m}{m}.$$

Since we assumed initially that  $r \geq \log m$  we see that the right hand side above is  $\asymp r/m$  and so this proves the upper bound in this case.

Now suppose that  $r \leq \log m$ . Let  $\pi^+$  denote another hyperplane parallel to  $\pi$ , also at distance  $\log m$  from  $\pi$  but such that  $\pi^+$  is contained in the half space containing  $B$  (and in particular,  $\pi^+$  intersects  $B$ ). Let  $S$  be the slab comprised between  $\pi$  and  $\pi^+$ , and let  $\tau_S = \inf\{t \geq 0 : X_t \notin S\}$  denote the first time the walk leaves the slab. Then note that if the walk has left  $S$  by time  $s/2$ , it must do so by hitting  $\pi^+$  before  $\pi$  and must then avoid  $\pi$  for at least  $s/2$  units of time. Of course,  $s/2 \asymp m^2$  hence using the result in the case already proved that  $r \geq \log m$ , since  $\pi^+$  is at distance  $\log m$  from  $\pi$ ,

$$P_x(\tau_H > s) \lesssim P_x(\tau_S > s/2) + P_x(\tau_{\pi^+} < \tau_H) \frac{\log m}{m}. \quad (8.27)$$

Let us bound the first term in the right hand side above. Note that the slab  $S$  has a width equal to  $\log m$  by definition. Hence every  $(\log m)^2$  units of time, the walk has a probability bounded below by  $p > 0$  to exit  $S$ . We deduce that

$$P_x(\tau_S > s/2) \leq (1 - p)^{\lfloor (s/2)/(\log m)^2 \rfloor}$$

and since  $s \asymp m^2$  we see that this decays faster than any polynomial in  $m$  and hence in particular is  $O(1/m)$ . Moreover, we claim that

$$P_x(\tau_{\pi^+} < \tau_H) \asymp \frac{r}{\log m} \quad (8.28)$$

so that, combining with (8.27), we get  $P(\tau_H > s) \lesssim r/m$  as desired. To see (8.28), let  $x'$  be the point at which  $\pi$  is tangent to  $B$  and consider the martingale

$$M'_t = (X_t - x') \cdot \vec{n},$$

that is to say, the (signed) distance to the plane  $\pi$  of the walk  $X_t$ . Note that this is indeed a martingale since  $X$  is a martingale and the projection onto  $\vec{n}$  is a linear operation. Apply the optional stopping theorem at the time  $\tau_S = \tau_H \wedge \tau_{\pi^+}$ . If  $\tau_S = \tau_H$  then  $|M'_{\tau_S}| \leq 1$ , whereas if  $\tau_S = \tau_{\pi^+}$  then  $|M_{\tau_S} - \log m| \leq 1$ . Consequently,

$$r = O(1)(1 - P_x(\tau_{\pi^+} < \tau_H)) + (\log m + O(1))P_x(\tau_{\pi^+} < \tau_H)$$

from which it immediately follows that

$$P_x(\tau_{\pi^+} < \tau_H) \leq \frac{r + O(1)}{\log m + O(1)}$$

which proves (8.28). As explained, the lemma follows.  $\square$

We will also need a slight improvement of this estimate where the end point is specified. (This would also follow from Proposition 6.9.4. in [14] but as above we prefer to provide our own proof).

**Lemma 8.29** In the same setting as Lemma 8.23, We have

$$P_x(X[0, t] \subset B; X_t = z) \lesssim t^{-d/2} \frac{\text{dist}(x, \partial B)}{\sqrt{t}} \frac{\text{dist}(z, \partial B)}{\sqrt{t}}.$$

Furthermore if  $z \in B^\circ$  and  $t \asymp m^2$  is such that  $P_x(X_t = z) > 0$ , the same inequality holds with  $\lesssim$  replaced by  $\gtrsim$ .

**Proof.** This uses a simple time reversal argument as well as the Markov property. Split the interval  $[0, t]$  into three intervals of length  $t/3$  each (for this argument we can assume without loss of generality that  $t/3$  is an integer). Observe that the process  $(X_{t-s} : 0 \leq s \leq t)$  is also by a reversibility a random walk which given  $X_t = z$  will be starting from  $z$ . Hence, using the standard fact that  $P_{x'}(X_t = z') \lesssim t^{-d/2}$  for all  $t \geq 1$  and  $x', z' \in \mathbb{Z}^d$ ,

$$\begin{aligned} P_x(X[0, t] \subset B; X_t = z) &\leq \sum_{x', z' \in B} P_x(X[0, t/3] \subset B; X_{t/3} = x') P_{x'}(X[0, t/3] \subset B; X_{t/3} = z') \times \\ &\quad \times P_{z'}(X_{t/3} = z; X[0, t/3] \subset B) \\ &\leq \sum_{x', z' \in B} P_x(X[0, t/3] \subset B; X_{t/3} = x') P_{x'}(X[0, t/3] \subset B; X_{t/3} = z') \times \\ &\quad \times P_z(X_{t/3} = z'; X[0, t/3] \subset B) \\ &\lesssim \sum_{x', z' \in B} P_x(X[0, t/3] \subset B; X_{t/3} = x') \times \frac{1}{t^{d/2}} \times P_z(X_{t/3} = z'; X[0, t/3] \subset B) \\ &\leq t^{-d/2} P_x(X[0, t/3] \subset B) P_z(X[0, t/3] \subset B) \\ &\lesssim t^{-d/2} \frac{\text{dist}(x, \partial B)}{\sqrt{t}} \frac{\text{dist}(z, \partial B)}{\sqrt{t}} \end{aligned}$$

by (8.26), as desired. When  $z \in B^\circ$  and  $t \asymp m^2$ , the opposite inequality holds using the lower bound in Lemma 8.23, and the fact that a Brownian bridge from a point in  $B(0, 1/2)$  to another point in  $B(0, 1/2)$  has a probability bounded below to stay in  $B(0, 1)$  throughout, uniformly in the endpoints of the trajectory within  $B(0, 1/2)$ .  $\square$

The estimate we will rely on is the following:

**Lemma 8.30** Let  $t \geq 2$  and  $x \in B$  with  $\text{dist}(x, \partial B) = r$  with  $1 \leq r \leq m$ . Then if  $\tau = m^2$ , uniformly in  $a, b \in B^\circ$ , and  $s \in [\tau/4; \tau/2]$ ,

$$P^{a \rightarrow b; \tau}(X_s = x) \gtrsim \frac{1}{m^d} \left( \frac{r}{m} \right)^2 \quad (8.31)$$

if the parity of  $x - a$  is the same as  $s$ . Moreover, in that case, uniformly in  $a, b \in B^\circ$ , and  $s \in [\tau/4; \tau/2]$ ,  $1 \leq t \leq \tau/4$ ,  $1 \leq r \leq m$ , the following holds: let  $A = \{x \in B : r \leq \text{dist}(x, \partial B) < 2r\}$  be the annulus at distance  $r$ . Then for any  $x \in A$ ,

$$P^{a \rightarrow b; \tau}(X_{s+t} \in A | X_s = x) \lesssim r^2 (\log t)^{10d} / t. \quad (8.32)$$

**Proof.** Observe that, uniformly in  $a, b \in B^\circ$ ,

$$P_a(X_\tau = b, X[0, \tau] \subset B) \asymp P_a(X_\tau = b) \asymp \frac{1}{m^d}$$

since a Brownian bridge from one point in a ball to another point in a ball has positive probability to stay in that ball throughout. Thus using Lemma 8.29 and reversibility (and noting that the parity of  $b - x$  is the same as  $\tau - s$  under our assumptions),

$$\begin{aligned} P^{a \rightarrow b; \tau}(X_s = x) &\asymp \frac{P_a(X_s = x; X[0, s] \subset B) \times P_x(X_{\tau-s} = b; X[0, \tau-s] \subset B)}{1/m^d} \\ &\gtrsim \frac{P_x(X_s = a; X[0, s] \subset B) \times \frac{1}{m^d} \frac{r}{m}}{1/m^d} \\ &\gtrsim \frac{\frac{1}{m^d} \frac{r}{m} \times \frac{1}{m^d} \frac{r}{m}}{1/m^d} = \frac{1}{m^d} \left(\frac{r}{m}\right)^2 \end{aligned}$$

which proves (8.31).

For the upper bound (8.32), observe that by the Markov property, if  $\tau' = \tau - s$ ,

$$P^{a \rightarrow b; \tau}(X_{s+t} \in A | X_s = x) = P^{x \rightarrow b; \tau'}(X_t \in A) = \frac{P_x(X_t \in A; X[0, \tau'] \subset B; X_{\tau'} = b)}{P_x(X[0, \tau'] \subset B; X_{\tau'} = b)}.$$

Consider first the denominator. Observe that by Lemma 8.29, since  $b \in B^\circ$  and  $\tau' \asymp m^2$ ,

$$P_x(X[0, \tau'] \subset B; X_{\tau'} = b) \asymp \frac{r}{m} \frac{1}{m^d}.$$

Now consider the numerator and suppose we condition on  $X_t = y \in A$ . Since  $s \in [\tau/4, \tau/2]$  we have  $\tau' = \tau - s \in [\tau/2, 3\tau/4]$ . Hence since we also assume that  $t \leq \tau/4$  we deduce that  $\tau' - t \in [\tau/4, \tau/2]$ , and hence  $\tau' - t \asymp \tau' \asymp \tau$ . From this, it is not hard to see that

$$P_y(X([0, \tau' - t] \subset B; X_{\tau'-t} = b) \asymp \frac{r}{m} \frac{1}{m^d}.$$

Consequently the numerator satisfies

$$P_x(X_t \in A; X[0, \tau'] \subset B; X_{\tau'} = b) \lesssim \frac{r}{m} \frac{1}{m^d} P_x(X_t \in A; X[0, t] \subset B) \lesssim \frac{r}{m} \frac{1}{m^d} \sum_{y \in A} P_x(X_t = y, X[0, t] \subset B).$$

Now, if  $|y - x| \leq t^{1/2} (\log t)^{10}$  then we use Lemma 8.23 to obtain that

$$P_x(X_t = y, X[0, t] \subset B) \lesssim \frac{r^2}{t} t^{-d/2},$$

and since  $\#\{y : |y - x| \leq t^{1/2}(\log t)^{10}\} \lesssim t^{d/2}(\log t)^{10d}$ , the contribution to the sum from such points is at most

$$\sum_{y \in A: |y-x| \leq (\log t)^{10}} P_x(X_t = y, X[0, t] \in B) \lesssim \frac{r^2}{t} (\log t)^{10}.$$

Now if  $|y - x| \leq t^{1/2}(\log t)^{10}$ , then

$$P_x(X_t = y, X[0, t] \in B) \leq P_x(X_t = y)$$

and so summing over all such  $y$ , the contribution is at most

$$P_x(|X_t - x| \geq t^{1/2}(\log t)^{10}) \leq \exp(-(\log t)^{20}/2) = o(1/t) \lesssim r^2/t$$

using elementary Chernoff large deviation bounds for sums of i.i.d. random variables. Hence

$$\sum_{y \in A} P_x(X_t = y, X[0, t] \in B) \lesssim \frac{r^2(\log t)^{10d}}{t}$$

and so the numerator satisfies

$$P_x(X_t \in A; X[0, \tau'] \subset B; X_{\tau'} = b) \lesssim \frac{r}{m} \frac{1}{m^d} \frac{r^2(\log t)^{10d}}{t}.$$

Combining with the bound on the denominator, we deduce

$$P^{a \rightarrow b; \tau}(X_{s+t} \in A | X_s = x) \leq \frac{r^2(\log t)^{10d}}{t}$$

as desired.  $\square$

## 8.5 End of proof of Theorem 8.1

We now start the proof of Lemma 8.22. Decompose the ball  $B = B(0, m)$  into dyadic annuli

$$A_j = \{y \in B : \text{dist}(y, \partial B) \in [2^j, 2^{j+1})\}; \quad 0 \leq j \leq \log_2(m)$$

Given our point configuration  $\mathcal{X}$  of  $k$  disjoint points in  $B$ , let  $\mathcal{X}_j = \mathcal{X} \cap A_j$ ; let  $j$  be such that  $|\mathcal{X}_j|$  is maximal. Then note that

$$|\mathcal{X}_j| \gtrsim \frac{k}{\log m}.$$

We will show that any bridge of duration  $\tau = m^2$  has a probability bounded below uniformly in its endpoints  $a, b$  to visit  $\mathcal{X}_j$ . Given a bridge  $X$  of duration  $\tau$ , let

$$L(\mathcal{X}_j) = \int_{\tau/4}^{\tau/2} 1_{X_s \in \mathcal{X}_j} ds$$

denote its local time spent in  $\mathcal{X}_j$  during its second quarter.

**Lemma 8.33** Uniformly over  $a, b \in B^\circ$ ,

$$P^{a \rightarrow b; \tau}(L(\mathcal{X}_j) > 0) \gtrsim \frac{k}{(\log m)^{2+10d} m^d}.$$

**Proof.** We use the trivial identity for nonnegative random variables:

$$P(X > 0) = \frac{E(X)}{E(X|X > 0)}$$

Hence we need a lower bound on  $E^{a \rightarrow b; \tau}(L(\mathcal{X}_j))$  and an upper bound on  $E^{a \rightarrow b; \tau}(L(\mathcal{X}_j)|L(\mathcal{X}_j) > 0)$ .

We start by the lower bound. Using (8.31),

$$\begin{aligned} E^{a \rightarrow b; \tau}(L(\mathcal{X}_j)) &\gtrsim |\mathcal{X}_j|(\tau/4) \frac{1}{m^d} \left(\frac{r}{m}\right)^2 \\ &\gtrsim \frac{k}{\log m} \frac{1}{m^{d-2}} \left(\frac{r}{m}\right)^2 \end{aligned}$$

where  $r = 2^j$ . On the other hand, for the upper bound we can apply the Markov property at the first hitting of  $\mathcal{X}_j$  and (8.32) to deduce that

$$\begin{aligned} E^{a \rightarrow b; \tau}(L(\mathcal{X}_j)|L(\mathcal{X}_j) > 0) &\leq \sup_{s \in [\tau/4, \tau/2]} \sup_{x \in \mathcal{X}_j} \sum_{y \in \mathcal{X}_j} \int_0^{\tau/4} P^{a \rightarrow b; \tau}(X_{s+t} = y | X_s = x) dt \\ &\lesssim \sup_{s \in [\tau/4, \tau/2]} \sup_{x \in \mathcal{X}_j} \int_0^{\tau/4} P^{a \rightarrow b; \tau}(X_{s+t} \in A_j | X_s = x) dt \\ &\lesssim 1 + \int_2^{\tau/4} \frac{r^2 (\log t)^{10d}}{t} dt \lesssim r^2 (\log m)^{1+10d}. \end{aligned}$$

Taking the quotient of these two terms, we deduce

$$P^{a \rightarrow b; \tau}(L(\mathcal{X}_j) > 0) \gtrsim \frac{k}{(\log m)^{2+10d} m^d}$$

as desired.  $\square$

We are now able to complete the proof of Lemma 8.22.

**Proof.** Proof of Lemma 8.22. On  $\mathcal{G}_{br}$  there are at least  $cn^{2-2\kappa}m^{d-2}$  bridges of length  $m^2$ , by definition of this event, where  $c$  depends only on the dimension. When we condition on the endpoints of the bridges they are independent of each other and of anything else under  $P^*$ . Hence, using Lemma 8.33, and the inequality  $1 - x \leq e^{-x}$  valid for all  $x \in \mathbb{R}$ ,

$$\begin{aligned} P^*(\mathcal{G}_{br}; \mathcal{B}_{\mathcal{X}}) &\leq \left(1 - \frac{k}{(\log m)^{2+10d} m^d}\right)^{cn^{2-2\kappa}m^{d-2}} \\ &\leq \exp\left(-c \frac{k}{m^d (\log m)^{2+10d}} n^{2-2\kappa} m^{d-2}\right) \\ &\leq \exp\left(-c \frac{n^{2-2\kappa}}{m^2 (\log m)^{2+10d}} k\right). \end{aligned}$$

Since on this event  $|R_N \cap B| = |B| - k$ , Lemma 8.22 follows.  $\square$

As explained this also finishes the proof of Theorem 8.1.

## A Functional inequalities on the rescaled lattice

Let  $x \in \mathbb{R}^d$  and  $r > 0$ . The cubic box centered at  $x$  of side length  $r$  is

$$\Lambda(x, r) = \left\{ y = (y_1, \dots, y_d) \in \mathbb{R}^d : -r/2 < y_i - x_i \leq r/2, 1 \leq i \leq d \right\}.$$

Let  $n \geq 1$ . We denote by  $\mathbb{Z}_n^d$  the lattice  $\mathbb{Z}^d$  rescaled by a factor  $n$ :

$$\mathbb{Z}_n^d = \frac{1}{n} \mathbb{Z}^d.$$

Let  $x \in \mathbb{Z}_n^d$  and  $r$  a rational number such that  $nr$  is an odd integer. We define the discrete box  $\Lambda_n(x, r)$  as

$$\Lambda_n(x, r) = \Lambda(x, r) \cap \mathbb{Z}_n^d.$$

We rewrite first the general discrete Poincaré–Sobolev inequality proved by Bessemoulin–Chatard, Chainais–Hillairet and Filbet (see theorem 3 of [3]) in the particular case of a box and a cubic mesh and for the exponent

$$2^* = \frac{2d}{d-2}.$$

**Theorem A.1** [Discrete Poincaré–Sobolev inequality] Let  $f$  be a function from  $\Lambda_n(x, r)$  to  $\mathbb{R}$ . There exists a constant  $c_{PS} = c_{PS}(x, r, d)$  which depends on  $x, r, d$  such that

$$\left( \sum_{y \in \Lambda_n(x, r)} \frac{1}{n^d} |f(y)|^{2^*} \right)^{\frac{1}{2^*}} \leq c_{PS} \left( \sum_{y \in \Lambda_n(x, r)} \frac{1}{n^d} |f(y)|^2 \right)^{\frac{1}{2}} + c_{PS} \left( \sum_{\substack{y, z \in \Lambda_n(x, r) \\ |y-z|=1/n}} \frac{1}{n^{d-2}} (f(y) - f(z))^2 \right)^{\frac{1}{2}}.$$

We rewrite now the discrete Poincaré–Wirtinger inequality proved by Bessemoulin–Chatard, Chainais–Hillairet and Filbet (see theorem 5 of [3]) in the particular case of a box and a cubic mesh and for the exponent 2.

**Theorem A.2** [Discrete Poincaré–Wirtinger inequality] Let  $f$  be a function from  $\Lambda_n(x, r)$  to  $\mathbb{R}$ . There exists a constant  $c_{PW} = c_{PW}(x, r, d)$  which depends on  $x, r, d$  such that

$$\left( \sum_{y \in \Lambda_n(x, r)} \frac{1}{n^d} (f(y) - \bar{f})^2 \right)^{\frac{1}{2}} \leq c_{PW} \left( \sum_{\substack{y, z \in \Lambda_n(x, r) \\ |y-z|=1/n}} \frac{1}{n^{d-2}} (f(y) - f(z))^2 \right)^{\frac{1}{2}},$$

where

$$\bar{f} = \frac{1}{(rn)^d} \sum_{y \in \Lambda_n(x, r)} f(y).$$

In general, the constants  $c_{PS}$ ,  $c_{PW}$  depend on both  $x \in \mathbb{Z}_n^d$  and  $r$ . However, the lattices  $\mathbb{Z}_n^d$  being invariant under a translation by an element of  $\mathbb{Z}_n^d$ , these constants are the same for all points  $x \in \mathbb{Z}_n^d$ . From now onwards, we suppose that  $x = 0$ . Let us examine the dependence of the constants  $c_{PS}$ ,  $c_{PW}$  with respect to  $r$ . Let  $f$  be a function defined on  $\Lambda_n(x, r)$  with values in  $\mathbb{R}$  and let us set

$$\forall x \in \Lambda_{nr}(0, 1) \quad g(x) = f(rx).$$

We apply the inequality stated in theorem A.2 to  $g$ :

$$\left( \sum_{y \in \Lambda_{nr}(0, 1)} \frac{1}{(nr)^d} (g(y) - \bar{g})^2 \right)^{\frac{1}{2}} \leq c_{PW}(0, 1, d) \left( \sum_{\substack{y, z \in \Lambda_{nr}(0, 1) \\ |y-z|=1/(nr)}} \frac{1}{(nr)^{d-2}} (g(y) - g(z))^2 \right)^{\frac{1}{2}},$$

and we rewrite everything in terms of the function  $f$ :

$$\begin{aligned} \bar{g} &= \frac{1}{(nr)^d} \sum_{y \in \Lambda_{nr}(0, 1)} f(ry) = \frac{1}{(nr)^d} \sum_{y \in \Lambda_n(0, r)} f(y) = \bar{f}, \\ \sum_{y \in \Lambda_{nr}(0, 1)} (g(y) - \bar{g})^2 &= \sum_{y \in \Lambda_n(0, r)} (f(y) - \bar{f})^2, \\ \sum_{\substack{y, z \in \Lambda_{nr}(0, 1) \\ |y-z|=1/(nr)}} (g(y) - g(z))^2 &= \sum_{\substack{y, z \in \Lambda_n(0, r) \\ |y-z|=1/n}} (f(y) - f(z))^2. \end{aligned}$$

We obtain the following inequality for the function  $f$ :

$$\left( \sum_{y \in \Lambda_n(0, r)} \frac{1}{n^d} (f(y) - \bar{f})^2 \right)^{\frac{1}{2}} \leq c_{PW}(0, 1, d) r \left( \sum_{\substack{y, z \in \Lambda_n(0, r) \\ |y-z|=1/n}} \frac{1}{n^{d-2}} (f(y) - f(z))^2 \right)^{\frac{1}{2}}.$$

We conclude that  $c_{PW}(0, r, d) = c_{PW}(0, 1, d)r$ .

## B Inequalities on the original lattice

We shall adopt a slightly different viewpoint to apply these inequalities. Instead of rescaling the lattice by a factor  $n$ , we will consider functions defined on the lattice  $\mathbb{Z}^d$  but on boxes of side  $n$ . We shall deduce the relevant inequalities from the previous ones by a simple change of variables  $y \rightarrow ny$ . More precisely, let  $x \in \mathbb{Z}^d$ ,  $n \geq 1$  and let  $f$  be a function from  $\Lambda(x, n) \cap \mathbb{Z}^d$  to  $\mathbb{R}$ . Let  $\Lambda_n(x/n, 1)$  be the box

$$\Lambda_n(x/n, 1) = \Lambda(x/n, 1) \cap \mathbb{Z}_n^d.$$

We define a function  $f_n$  on  $\Lambda_n(x/n, 1)$  by setting

$$\forall y \in \Lambda_n(x/n, 1) \quad f_n(y) = f(ny).$$

We apply then the inequalities stated in theorems A.1, A.2 to the function  $f_n$  and we rewrite everything in terms of  $f$ . We first introduce some notation before stating the inequalities. Let  $f$  be a function defined on a subset  $D$  of  $\mathbb{Z}^d$  with values in  $\mathbb{R}$ . For  $p \geq 1$ , we define its  $p$ -norm over  $D$

$$\|f\|_{p, D} = \left( \sum_{y \in D} |f(y)|^p \right)^{\frac{1}{p}}$$



and its Dirichlet energy over  $D$

$$\mathcal{E}(f, D) = \frac{1}{2d} \sum_{\substack{y, z \in D \\ |y-z|=1}} (f(y) - f(z))^2.$$

We recall that the exponent  $2^*$  is equal to  $2^* = 2d/(d-2)$ . The Poincaré–Sobolev inequality stated in theorem A.1 yields the following inequality in a box of side  $n$ .

**Corollary B.1** Let  $x \in \mathbb{Z}^d$ ,  $n \geq 1$  and let  $f$  be a function from  $\Lambda(x, n) \cap \mathbb{Z}^d$  to  $\mathbb{R}$ . There exists a constant  $c_{PS}$  which depends on the dimension  $d$  only such that

$$\|f\|_{2^*, \Lambda(x, n)} \leq c_{PS} \left( \frac{1}{n} \|f\|_{2, \Lambda(x, n)} + \sqrt{2d \mathcal{E}(f, \Lambda(x, n))} \right).$$

The Poincaré–Wirtinger inequality stated in theorem A.2 yields the following inequality in a box of side  $n$ .

**Corollary B.2** Let  $x \in \mathbb{Z}^d$ ,  $n \geq 1$  and let  $f$  be a function from  $\Lambda(x, n) \cap \mathbb{Z}^d$  to  $\mathbb{R}$ . There exists a constant  $c_{PW}$  which depends on the dimension  $d$  only such that

$$\|f - \bar{f}\|_{2, \Lambda(x, n)} \leq n c_{PW} \sqrt{2d \mathcal{E}(f, \Lambda(x, n))},$$

where

$$\bar{f} = \frac{1}{n^d} \sum_{y \in \Lambda(x, n)} f(y).$$

Finally, if we send  $n$  to  $\infty$  in the inequality of corollary B.1, we get the following result.

**Corollary B.3** Let  $f$  be a function defined on  $\mathbb{Z}^d$  with values in  $\mathbb{R}$  having finite support. There exists a constant  $c_{PS}$  which depends on the dimension  $d$  only such that

$$\|f\|_{2^*} \leq c_{PS} \sqrt{2d \mathcal{E}(f)}.$$

**Acknowledgements.** We warmly thank the Referees for their precise reading and their numerous comments which helped to improve the paper.

## References

- [1] Peter Antal. Enlargement of obstacles for the simple random walk. *The Annals of Probability*, pages 1061–1101, 1995.
- [2] Nathanael Berestycki and Ariel Yadin. Condensation of random walks and the wulff crystal. *Annales de l’Institut Henri Poincaré (B): probability and statistics*, to appear. arXiv preprint arXiv:1305.0139, 2013.
- [3] Marianne Bessemoulin-Chatard, Claire Chainais-Hillairet, and Francis Filbet. On discrete functional inequalities for some finite volume schemes. *IMA J. Numer. Anal.*, 35(3):1125–1149, 2015.

- [4] Marek Biskup and Eviatar B Procaccia. Shapes of drums with lowest base frequency under non-isotropic perimeter constraints. *arXiv preprint arXiv:1603.03871*, 2016.
- [5] Marek Biskup and Eviatar B Procaccia. Eigenvalue versus perimeter in a shape theorem for self-interacting random walks. *The Annals of Applied Probability*, 28(1):340–377, 2018.
- [6] Erwin Bolthausen. Localization of a two-dimensional random walk with an attractive path interaction. *Ann. Probab.*, 22(2):875–918, 1994.
- [7] Lorenzo Brasco, Guido De Philippis, and Bozhidar Velichkov. Faber-Krahn inequalities in sharp quantitative form. *Duke Math. J.*, 164(9):1777–1831, 2015.
- [8] Jian Ding, Ryoki Fukushima, Rongfeng Sun, and Changji Xu. Geometry of the random walk range conditioned on survival among Bernoulli obstacles. *arXiv preprint arXiv:1806.08319*, 2018.
- [9] M. D. Donsker and S. R. S. Varadhan. Asymptotic evaluation of certain Markov process expectations for large time. I. II. *Comm. Pure Appl. Math.*, 28:1–47; *ibid.* 28 (1975), 279–301, 1975.
- [10] M. D. Donsker and S. R. S. Varadhan. On the number of distinct sites visited by a random walk. *Comm. Pure Appl. Math.*, 32(6):721–747, 1979.
- [11] Richard S. Ellis. *Entropy, large deviations, and statistical mechanics*. Classics in Mathematics. Springer-Verlag, Berlin, 2006. Reprint of the 1985 original.
- [12] K. Fan. Minimax theorems. *Proc. Nat. Acad. Sci.*, 39:42–47, 1953.
- [13] Nicola Fusco, Francesco Maggi, and Aldo Pratelli. Stability estimates for certain Faber-Krahn, isocapacitary and Cheeger inequalities. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, 8(1):51–71, 2009.
- [14] Gregory F Lawler and Vlada Limic. *Random walk: a modern introduction*, volume 123. Cambridge University Press, 2010.
- [15] Chiranjib Mukherjee and SRS Varadhan. Brownian occupation measures, compactness and large deviations. *The Annals of Probability*, 44(6):3934–3964, 2016.
- [16] Tobias Povel. Confinement of Brownian motion among Poissonian obstacles in  $\mathbb{R}^d, d \geq 3$ . *Probability theory and related fields*, 114(2):177–205, 1999.
- [17] Chris G Rogers and David Williams. *Diffusions, Markov processes and martingales: Volume 2, Itô calculus*, volume 2. Cambridge university press, 2000.
- [18] Alain-Sol Sznitman. On the confinement property of two-dimensional Brownian motion among Poissonian obstacles. *Communications on pure and applied mathematics*, 44(8-9):1137–1170, 1991.