Condensation of a self-attracting random walk

by Nathanaël Berestycki[*] Ariel Yadin[†]

April 2, 2017

Abstract

We introduce a Gibbs measure on nearest-neighbour paths of length t in the Euclidean d-dimensional lattice, where each path is penalised by a factor proportional to the size of its boundary and an inverse temperature β. We prove that, for all $\beta > 0$, the random walk condensates to a set of diameter $(t/\beta)^{1/3}$ in dimension $d = 2$, up to a multiplicative constant. In all dimensions $d \geq 3$, we also prove that the volume is bounded above by $(t/\beta)^{d/(d+1)}$ and the diameter is bounded below by $(t/\beta)^{1/(d+1)}$. We further speculate that the limiting shape shares the same exponents as the KPZ universality class in the planar case. Similar results hold for a random walk conditioned to have local time greater than β everywhere in its range when β is larger than some explicit constant, which in dimension two is the logarithm of the connective constant.

Keywords: Gibbs measure, condensation, self-attractive random walk, Wulff crystal, large deviations, Donsker–Varadhan principle.

MSC 2010 classification: 60K35, 60J27, 60F10

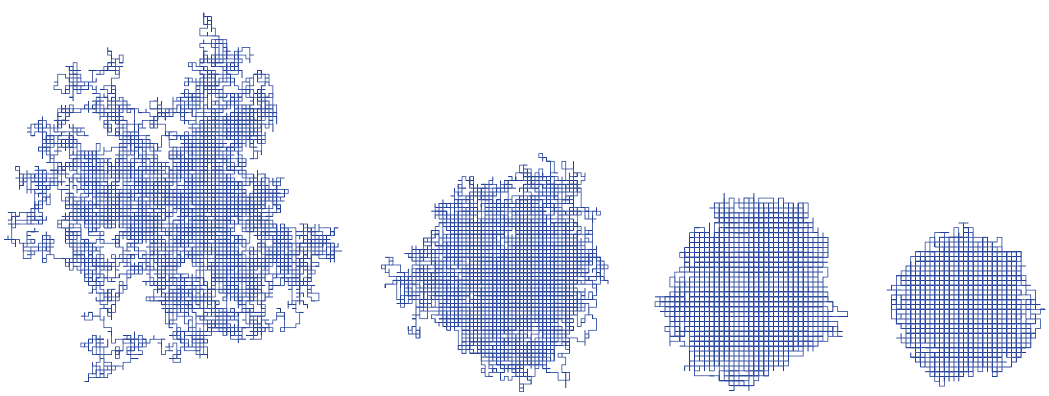

Figure 1: Simulations using a Gibbs sampler algorithm of a random walk cluster with $t = 25{,}000$ steps, corresponding to $\beta = 0.01$, $\beta = 0.1$, $\beta = 1$, and $\beta = 2$.

[*]University of Cambridge. Supported in part by EPSRC grants EP/GO55068/1 and EP/I03372X/1

[†]Ben Gurion University of the Negev.

Contents

1 Introduction

1.1 Statement of the main results

Let $d \geq 2$ and let Ω be the space of nearest neighbour, right continuous, infinite paths $(\omega_t, t \in [0, \infty))$ on $\mathbb{Z}^d$, and let $(X_t(\omega), t \geq 0)$ be the canonical process. For $x \in \mathbb{Z}^d$, let $\mathbb{P}_x$ denote the law of simple random walk on $(\Omega, \mathcal{F})$ in continuous time with rate one started from x, where $\mathcal{F}$ denotes the σ-field generated by X. We call $\mathbb{P} = \mathbb{P}_0$.

Our main result deals with geometric properties of some penalisations of random walks on $\mathbb{Z}^d$ by their boundary. More precisely, we introduce a Gibbs measure μ on random paths defined as follows. Let $R_t = \{v \in \mathbb{Z}^d : X_s = v \text{ for some } s \leq t\}$ denote the *range* of the walk at time t. For a given time t, we consider the Hamiltonian H given by

$$H(\omega) = |\partial R_t|, \tag{1.1}$$

where for a set G, $\partial G = \{x \in G : x \sim y \text{ for some } y \notin G\}$ is the (inner) vertex-boundary of G. (Here $x \sim y$ means that x is a neighbor of y in $\mathbb{Z}^d$). The associated Gibbs measure on random paths $\mu = \mu_t$ is obtained by considering the measure μ defined by

$$\frac{d\mu}{d\,\mathbb{P}}(\omega) = \frac{1}{Z} \exp(-\beta H(\omega)) \tag{1.2}$$

on $\mathcal{F}$. Here $\beta > 0$ is a positive number playing the role of inverse temperature and $Z = Z(t, \beta) = \mathbb{E}(\exp(-\beta |\partial R_t|))$ is a normalising factor called the partition function. In plain words, the Gibbs measure μ penalises every site on the boundary of the range R_t by a fixed amount $e^{-\beta}$. Hence μ favours "highly condensed" configurations. Interpreting the random walk $(X_0, \ldots, X_t)$ as a chain of monomers, the Gibbs measure μ represents the law of a diluted polymer in a poor solvent. With a constraint on the maximal volume and a boundary penalisation, the random set R_t may be seen as a random walk realisation of the Wulff crystal droplet with inverse temperature β. The Wulff construction ([31]) is a method to determine the equilibrium shape of crystals based on surface energy minimisation. Until now this has been studied rigorously in the context of percolation and the Ising model. As far as we are aware, the present paper is the first attempt to study the question through genuinely d-dimensional random walks. Note however that in the 1-dimensional SOS model, namely for a 1+1-dimensional random walk conditioned on describing an atypically large arithmetic area, it was proven by Dobrushin and Hryniv in [14] that a limiting Wulff shape arises.

In the percolation context, a rigorous derivation of the limiting shape was first given by Alexander, Chayes and Chayes [2] in two dimensions, while in the context of the Ising model, this was achieved in a nearly simultaneous way in a celebrated work of Dobrushin, Kotecký and Shlosman [15] (which was derived again by Pfister [27], see also the papers by Ioffe and Schonmann [23] extending the results of [15] to all subcritical temperatures). The three-dimensional case, which is the most delicate, was handled only relatively recently by Cerf [10], with earlier work by Bodineau [7], Cerf and Pisztora [11] as well as Bodineau, Ioffe and Velenik [8]. See [26] for an early reference on the problem of phase separation in the context of the Ising model, [10] for a recent monograph giving a detailed overview of the subject.

We will be interested in describing the geometry of the shape of R_t. Our main result gives precise estimates for the condensation effect that results from the self-interaction in dimension $d \geq 2$. If $G \subset \mathbb{Z}^d$, let $\operatorname{diam} G$ denote the (Euclidean) diameter of G:

$$\operatorname{diam} G = \sup\{|z - w| \ : \ z, w \in G\}.$$

Theorem 1.1. *(i). Let $d = 2$. For any $\beta_0 > 0$ there exist positive constants c_1, c_2 depending only on β_0 such that for all $\beta > \beta_0$,*

$$\mu_t\left(c_1\left(\frac{t}{\beta}\right)^{1/3} \leq \operatorname{diam}(R_t) \leq c_2\left(\frac{t}{\beta}\right)^{1/3}\right) \to 1, \tag{1.3}$$

as $t \to \infty$.

(ii). For all $d \geq 3$ we also have for all β_0 and for all $\beta > \beta_0$,

$$\operatorname{diam}(R_t) \geq c_1\left(\frac{t}{\beta}\right)^{1/(d+1)} \qquad \text{and} \qquad |R_t| \leq c_2\left(\frac{t}{\beta}\right)^{d/(d+1)} \tag{1.4}$$

with μ_t-probability tending to 1 as $t \to \infty$, where the constants c_1 and c_2 depend only on d and β_0.

In principle, the inverse temperature β may even be chosen to depend on t; in which case the same result holds if we also assume $\beta = \beta(t)$ satisfies $\beta(t)/t \to 0$.

Note that the proof gives precise estimates on the probability of these events, as well as estimates on the partition function $Z = Z(t, \beta)$. We refer the reader to Theorem 3.7 for a precise statement, and refer to Open Problem 2 and the following paragraph, at the end of the paper, for a precise conjecture on the geometry of R_t when $d = 2$. In dimensions $d \geq 3$, we conjecture that

$\text{diam}(R_t)$ scales like $(t/\beta)^{1/(d+1)}$, but we have only obtained a lower bound. Our proof supports this conjecture, but the upper bound is elusive (roughly because of topological complications in dimensions ≥ 3). Nevertheless we still manage to get an upper bound on the volume which is consistent with this conjecture. This difficulty is a common feature of all works on Wulff crystal.

A related problem was studied by E. Bolthausen [9] in dimension $d = 2$. In that work, the energy $H(\omega)$ serving to define the Gibbs measure μ in (1.2) is taken to be $\hat{H}(\omega) = |R_t(\omega)|$, the size of the range (as opposed to that of its boundary). Thus $d\hat{\mu} = \hat{Z}_t^{-1} \exp(-\beta \hat{H}(\omega)) d\,\mathbb{P}$. Bolthausen's result is that the random walk condensates to a set of diameter $t^{1/4}$, which is close in the Hausdorff sense to a Euclidean ball of that diameter. In both problems, good bounds on the partition functions Z_t and $\hat{Z}_t$ play a crucial role. In the case where the energy is just the volume, we have $\hat{Z}_t = \mathbb{E}(\exp(-\beta|R_t|))$, and precise asymptotics for this quantity were already obtained by Donsker and Varadhan [18]. This is a considerably easier problem than the one considered here, essentially because $|R_t(\omega)|$ is "almost" a continuous function of its local time profile, viewed as a probability measure on $\mathbb{Z}^d$. In particular, the powerful machinery of large deviations theory provides the right tools to study that question. This goes a long way in explaining the appearance of the Euclidean ball as a limit shape, and explains why the inverse temperature β is not a relevant parameter in that model.

In contrast, here we believe that the limit shape depends on β and is not a rotationally symmetric ball. It may even contain some flat faces: see the simulations shown at the beginning of the paper, and recall that this is also the situation in the the percolation and Ising Wulff crystals. So the microscopic geometry of the lattice is important even to determine the macroscopic shape of the random walk cluster, and thus there is no hope in directly applying the Donsker–Varadhan large deviations machinery to the problem. And indeed, two local time profiles can be macroscopically close while the sizes of boundary of the corresponding ranges can be of widely different orders of magnitude. Instead, we perform a suitable change of measure and obtain quantitative estimates to derive the lower bound on the partition function. This is the main technical part of the paper; the key is to understand how a random walk gets a small (or smooth) boundary. Once this is done, we rely on a discrete isoperimetric inequality to deal efficiently with the entropy coming from the large number of possible shapes of the range.

1.2 Some related problems

Our technique is sufficiently robust that it yields similar results for a number of models which turn out to be quite closely related. One interesting case is the following conditioning problem, initially suggested by Itai Benjamini in private communication with the first author (in fact, it was this question which was initially the focus of the present investigation). For $t \geq 0$, define the event $\mathcal{E}_t$ as follows:

$$\mathcal{E}_t = \{L(t,x) \geq \beta, \forall\, x \in R_t\}. \tag{1.5}$$

Here $L(t,x) = \int_0^t 1_{\{X_s = x\}} ds$ is the amount of time the walk spends at a vertex x. Conditioning on this event gives a uniform lower bound on the density of local time uniformly over the range R_t. This also favours highly condensed configurations. In Benjamini's original question, it was assumed that $\beta = \beta(t) \to \infty$ as $t \to \infty$. The question was to decide whether, conditional on the event $\mathcal{E}_t$, there is a shape theorem for the range R_t.

As we will see later on (see for instance (3.10)), the conditioning also heavily penalises shapes R_t with large boundaries: essentially, every point on the boundary penalises the shape by a factor of order $e^{-\beta}$, and so it is not surprising that a behavior similar to Theorem 1.1 occurs. In particular, the conditioning is already highly nontrivial when $\beta(t) \equiv \beta$ is a fixed constant. This is perhaps counterintuitive initially, since in dimension $d = 2$ for instance, points are typically visited logarithmically many times, so the constraint $\mathcal{E}_t$ does not seem "very" singular.

Unlike in Theorem 1.1, we will need an assumption that $\beta > \beta_0$, where β_0 is an explicit constant: $\beta_0 = \log \alpha$, where α is the connective constant of $\mathbb{Z}^2$. (In dimension $d \geq 3$, that constant takes a different value related to a notion of self-avoiding surfaces, see (3.11) for the definition).

Theorem 1.2. *(i). Let $d = 2$. Let $\beta_0 = \log \alpha$, where α is the connective constant. For all $\beta > \beta_1 > \beta_0$, we have*

$$\mathbb{P}\left(c_1 \left(\frac{t}{\beta}\right)^{1/3} \leq \operatorname{diam}(R_t) \leq c_2 \left(\frac{t}{\beta}\right)^{1/3} \middle| \mathcal{E}_t \right) \to 1 \tag{1.6}$$

as $t \to \infty$, where the positive constants c_1 and c_2 depend only on β_1.

(ii). For all $d \geq 2$, there exists $\beta_0 = \beta_0(d)$ such that for all $\beta > \beta_1 > \beta_0$,

$$\operatorname{diam}(R_t) \geq c_1 \left(\frac{t}{\beta}\right)^{1/(d+1)}; \text{ and } |R_t| \leq c_2 \left(\frac{t}{\beta}\right)^{d/(d+1)} \tag{1.7}$$

with conditional probability given $\mathcal{E}_t$ tending to 1 as $t \to \infty$, where the positive constants c_1 and c_2 depend only on d and β_1.

As in Theorem 1.1, the result remains valid if $\beta = \beta(t)$ is allowed to depend on t, provided also that $\beta(t)/t \to 0$.

Another variant consists in taking a slightly different Hamiltonian $\tilde{H}$, defined by

$$\tilde{H}(\omega) = \sum_{x \in \partial R_t} L(t, x), \tag{1.8}$$

Thus the penalisation takes into account not only the size of the boundary, but also the amount of time spent on it. Define $d\tilde{\mu} = (\tilde{Z})^{-1} \exp(-\beta \tilde{H}) d\,\mathbb{P}$ on $\mathcal{F}$.

Theorem 1.3. *Theorem 1.1 still holds true with $\tilde{\mu}$ instead on μ.*

Acknowledgements. This work started when AY was a Herschel Smith postdoctoral fellow at the Statistical Laboratory, University of Cambridge. We gratefully acknowledge the financial support of the Herschel Smith fund and EPSRC grant EP/GO55068/1. The first author is grateful for the hospitality of the Theory Group at Microsoft Research, where part of this work was carried out. NB would like to thank Omer Angel, Ori Gurel–Gurevitch, Yuval Peres and Ofer Zeitouni for useful conversations, and Tom Begley for the pictures. We are very grateful to two anonymous referees for their comments which improved the presentation and pointed out some mistakes in an earlier version of the paper.

Updates. Since the first version was posted to the arXiv in 2013, there has been some progress on questions inspired by this paper. For instance, the series of works by Asselah and Schapira [3, 4] which discuss the large deviation principle for the boundary of the range of a simple random walk in dimensions $d \geq 3$. In a different direction, a series of two articles by Biskup and Procaccia [5, 6] study a direct analogue of (1.2) in the two-dimensional case, except that the boundary of the range is understood to mean the edge boundary (whereas we consider here the vertex boundary) and the weight of the edges is also allowed to be random. Then by letting $t \to \infty$ and then $\beta \to \infty$ they are able to obtain a limit theorem for the shape of the range, which is nonrandom but depends on the law of the weights. In the particular of deterministic (nonrandom) edge weights, this limit shape is simply the unit square. This is the analogue of our conjecture here that the limit shape is a diamond when $\beta \to \infty$. (The difference between square and diamond come from the difference between vertex and edge boundary.)

1.3 Main ideas in the proof; organisation of the paper

Since the results mentioned have many technical aspects, and involve careful computations, let us provide a sketch of the main ideas involved. Recall that we are trying to bound the radius and volume of the trace of the random walk under penalisation of having a large local time at its boundary. A very rough heuristic for the size of the diameter is as follows. The probability of staying in a box S of diameter L is approximately $\exp(-\text{const.} \times t/L^2)$, and one can expect the size of the boundary to then be approximately L^{d-1}, so the corresponding energy of such a configuration is of order $\exp(-\beta L^{d-1})$. Balancing energy and entropy gives us $\beta L^{d-1} = t/L^2$ and so $L = (t/\beta)^{d+1}$, which is indeed the conjectured order of the diameter in all dimension $d \geq 2$ (see Theorem 1.1).

The main issue of translating this rough heuristic into a rigorous argument is that the random walk could stay in a box of size L while having a much bigger boundary than L^{d-1}: this will be the case if the boundary is in some sense rough or fractal, which is *a priori* the case at least in small dimensions (recall that in dimension $d = 2$, the dimension of the outer boundary of Brownian motion is 4/3). This raises serious questions about the heuristic argument above: could the probability of staying in a box of size L *and* have a smooth boundary be substantially smaller than $\exp(-\text{const.} \times t/L^2)$? Fortunately we answer by the negative. Correspondingly, our main task is to prove a lower bound on the partition function Z (see Proposition 3.1), which establishes one scenario of probability roughly $\exp(-\text{const.} \times t/L^2)$ where the boundary of the range is of size approximately L^{d-1}.

In order to do this, we first have to guess the profile of local times $(\pi(x))_{x\in S}$, in the box S of size L, achieved by a random walk conditioned to have a small boundary. The trickiest part is to guess the behaviour of this profile close (at micro- and mesoscopic distance) from the boundary of the box. With hindsight we define a specific profile which should be (almost) the optimal one. We then have to compute the cost of achieving this profile, and show that the boundary has the desired size $O(L^{d-1})$ with this profile.

This leads us to a change of measure argument, as done in [9], and we must estimate the Radon–Nikodym derivative under the tilted measure. The main term turns out to be $\exp(\int_0^t \frac{\Delta f}{f}(X_s)ds)$, where $f(x)$ is the square root of the local time profile π which we seek to impose (see Lemma 3.4). If the local times of X are well approximated by the profile π then it is relatively easy to conclude (using careful second-order Taylor expansions, see Lemma 3.5) that this Radon–Nikodym derivative is indeed of order $\exp(-\text{const.} \times t/L^2)$, as desired. Hence what is needed is a precise control of the large deviations of local time at points under the tilted measure. This is achieved by a careful analysis done in Lemma 2.6, and our main use of it is summarised in Corollary 2.8. Roughly speaking, to obtain good large deviation control on the local time at a point x which might be close to the boundary of S, it suffices to show that there is a positive chance to hit the point x, every $1/\pi(x)$ units of time. This is achieved through a quantitative analysis of the tilted measure, using electrical network theory, and is the main purpose of Section 2.

Finally, in Section 3.4, we apply the bound on the partition function bound to control the radius and volume of the penalised random walk trace. This is done mainly using discrete isoperimetric inequalities.

Let us note that our methods work even for the *constant* β regime (not just $\beta(t) \to \infty$). In order to achieve this, it was necessary to correctly pick an accurate local time profile π for the lower bound on the partition function $Z(t, \beta)$. The naive choice in this case was not good enough for constant β. We discuss the required properties of this profile in the beginning of Section 2, adjacent to the definition of π. (For example, the logarithmic terms, and their powers, appearing in the definition of π, are essential for the analysis to work.)

2 Quantitative estimates for lower bound on Z

2.1 Change of measure

A main technical part of this paper consists in deriving good lower bounds on the partition function Z, which holds in all dimensions $d \geq 2$. In order to do this, we introduce a change of measure which is key to our analysis. This is a technique which was already used in Bolthausen's article [9] as well as in the article by Gärtner and den Hollander [19] on intermittency of parabolic Anderson model. However, the precise change of measure which needs to be performed here is significantly more delicate both in its choice and its analysis. The latter in particular will require a host of tools from the quantitative theory of Markov chains: we will need very precise information about how the random walk behaves at microscopic and mesoscopic distances away from the boundary ∂S.

Let $\pi \in \mathcal{M}_1(\mathbb{Z}^d)$ be a probability measure on $\mathbb{Z}^d$, and define a law $\mathbb{Q}$ on $(\Omega, \mathcal{F})$ as the Markov chain on $\mathbb{Z}^d$ having the transition rates $Q(x, y) = \sqrt{\pi(y)/\pi(x)}$ for $x \sim y$ and $\pi(x) > 0$ or, equivalently, infinitesimal generator defined by

$$Qf(x) = \sum_{y \sim x} \sqrt{\frac{\pi(y)}{\pi(x)}}[f(y) - f(x)]. \tag{2.1}$$

It is immediate (but essential) that π is a reversible equilibrium measure under $\mathbb{Q}$. We will also let $\mathbb{Q}_x$ denote the law of this Markov chain started from a given vertex $x \in \mathbb{Z}^d$. The choice of π will be crucial to our proof. Let L be nearest integer above $(t/\beta)^{1/(d+1)}$, i.e., $L = \lceil (t/\beta)^{1/(d+1)} \rceil$. Let $S = [-L, L]^d \cap \mathbb{Z}^d$ be the cube of side length $2L$. Our choice of π is delicate and is determined by the following requirements.

- π must be chosen so that the walk never leaves the cube S, and should spend most of its time in the "bulk" of the cube S.
- π must be chosen so that by time t, points on the boundary ∂S are visited, but typically only a finite (Poisson-like) number of times. At a finite but large distance from the boundary, the mean number of visits should still be finite but large.

- π must be a "reasonably smooth" function near the boundary, so that achieving the profile π is not too unlikely (we are aiming for probability of order $\exp(-ct/L^2)$, which is roughly the probability of staying in a cube of size L for time t).

These three conditions would ensure that the boundary of the range is not much bigger than the boundary of the cube S, while the smoothness condition ensures that π is not too unlikely. Recall in particular that L was chosen so that the entropic cost, $\exp(-ct/L^2)$, balances the energy cost $\exp(-c\beta L^{d-1})$.

The specific choice of π is as follows. For $0 \le r \le L$, let

$$S_r = \{z \in S \ : \ \mathrm{dist}(z, \partial S) = r\} = \{z \in S \ : \ \|z\|_\infty = L - r\},$$

where $\mathrm{dist}(\cdot,\cdot)$ refers to the graph distance on $\mathbb{Z}^d$ and for a point $z = (z_1, \dots, z_d) \in \mathbb{Z}^d$, $\|z\|_\infty = \max_{1\le i\le d} |z_i|$. Then, for $x \in S_r$, set

$$\pi(x) = \mu_r := \begin{cases} C\frac{r}{L^{d+1}(\log(r+2))^2} & \text{if } r \le L/2 \\ \left(\sqrt{\mu_{L/2}} + C\frac{(r-L/2)}{L^{(d+2)/2}}\right)^2 & \text{if } r \ge L/2. \end{cases} \tag{2.2}$$

Let $\pi(z) = 0$ for $z \notin S$, and the constant C is chosen so that $\sum_z \pi(z) = 1$. It can then be checked that C is uniformly bounded away from 0 and infinity and converges to a constant as L tends to infinity.

We comment briefly on the choice of π. In view of large deviation theory and the Donsker–Varadhan principle, the most natural choice *a priori* is to take π to be the square of the first Dirichlet eigenfunction on S, normalised to have unit mass. This is for instance what is used in Bolthausen's work [9] with some additional tweaking near the boundary of the shape (see the definition of $\tilde{\psi}$ on p.893 of [9]). However this turns out to be "too flat" near the boundary, making the second requirement untrue.

Our choice means that the growth of π is much steeper near the boundary. The slightly sublinear growth of π near the boundary, in $r/(\log r)^2$, is in fact the crucial feature of this choice: the linear factor r guarantees that points at a large distance from the boundary have a large mean number of visits, while the correcting factor in $1/(\log r)^2$ ensures that π is smooth enough that achieving a profile π has a probability of the right order of magnitude.

Orientation. At the technical level, we recall that our argument is organised as follows. Roughly speaking, we wish to obtain large deviation bounds on the local time accumulated at a point $y \in S$ under the tilted measure $\mathbb{Q}$ (Lemma 2.6). The key for doing so will be to show that y is hit sufficiently frequently, and in particular to obtain exponential tails on the hitting time of y (Proposition 2.2, using electrical network theory). Once Lemma 2.6 is proved, we use the concentration of local time to estimate the Radon-Nikodym derivative of $\mathbb{P}$ with respect to $\mathbb{Q}$ (Lemma 3.5) and hence estimate the partition function $Z(t, \beta)$ (Proposition 3.1).

2.2 Crude estimate on mixing time

Our first goal is to prove a crude bound on the mixing time of the Markov chain defined by $\mathbb{Q}_x$, which is needed at various points in our argument. We do this by estimating the spectral gap of the Markov chain, using the method of *canonical paths* of Diaconis and Saloff-Coste. We use the standard canonical paths on $\mathbb{Z}^d$: that is, for $x, y \in S$, we choose the path γ_{xy} defined as follows. We first try to match the first coordinate of x and y, then the second coordinate, and so on until the last coordinate. Each time, the change in coordinate is monotone. As an example if $d = 2$ and $x = (x_1; x_2)$ and $y = (y_1; y_2)$, let $z = (y_1; x_2)$. Then $\gamma_{x,y}$ is the union of two straight segments, going horizontally from x to z and then vertically from z to y. We call $|\gamma$ the length (number of edges) of a path γ. If $e = (x, y)$ is an edge, let $q(e) = \pi(x)Q(x, y)$ be the equilibrium flow through e.

Lemma 2.1. *Let E denote the set of edges within S.*

$$B = \max_{e \in E} \left\{ \frac{1}{q(e)} \sum_{x,y: e \in \gamma_{xy}} |\gamma_{xy}| \pi(x) \pi(y) \right\}.$$

Then $B \leq C_{2.1} L^2$ for some constant $C_{2.1} > 0$.

Proof. Fix an edge e and suppose $\mathrm{dist}(e, \partial S) = r$. Say that a point x is below e if $\mathrm{dist}(x, \partial S) \leq r$, and otherwise say that x is above e. Note that if $e \in \gamma_{xy}$, x and y cannot be both above e. Indeed, if $m_i = \min\{x_i, y_i\}$ and $M_i = \max\{x_i, y_i\}$ then $\gamma_{xy} \subset \prod_i [m_i, M_i] \subset S$, and x, y are two corners of this hypercube. So any point on γ_{xy} must be further from ∂S than one of x or y.

Therefore at least one of x or y is below e, say x. In this case $\pi(x)/q(e) \leq O(1)$. Moreover it is elementary to check that for any fixed $y \in S$, there are at most $O(L)$ choices of x such that $e \in \gamma_{xy}$ (this is because going from x to y the path γ_{xy} passes through each direction at most once). Therefore

$$\begin{aligned} \frac{1}{q(e)} \sum_{x,y: e \in \gamma_{xy}} |\gamma_{xy}| \pi(x) \pi(y) &\leq CL \sum_y \pi(y) \sum_{x: e \in \gamma_{xy}} \frac{\pi(x)}{q(e)} \\ &\leq CL^2 \sum_y \pi(y) = CL^2, \end{aligned}$$

as desired. □

By Theorem 3.2.1 in [29], it follows that if γ is the spectral gap of the Markov chain, then $\gamma \geq 1/(C_{2.1} L^2)$. (In fact, that result holds for discrete time chains but it is straightforward to adapt the proof to the continuous time case). Now, it is well known that estimates on the spectral gaps yield estimates on the heat kernel. More precisely,

$$|\mathbb{Q}_x(X_t = y) - \pi(y)| \leq \sqrt{\pi(y)/\pi(x)} e^{-\gamma t}.$$

(See, *e.g.*, the proof of Corollary 2.1.5 of [29]). Let

$$t_{\mathrm{mix}} = \inf \left\{ t \geq 0 \quad : \quad \text{for all } x, y \in S : \left| \frac{\mathbb{Q}_x(X_t = y)}{\pi(y)} - 1 \right| \leq 1/2 \right\}$$

From Lemma 2.1 we deduce that for all $x, y \in S$,

$$|\mathbb{Q}_x(X_t = y) - \pi(y)| \leq \sqrt{L} e^{-t/(C_{2.1} L^2)}.$$

Since $\pi(y) \geq c/L^{d+1}$ for all $y \in S$, it follows that taking $t \geq CL^2 \log L$ with some sufficiently large constant C,

$$|\mathbb{Q}_x(X_t = y) - \pi(y)| \leq \tfrac{1}{2} \pi(y).$$

Thus we have proved:

$$t_{\mathrm{mix}} \leq C_{2.3} L^2 \log L \tag{2.3}$$

2.3 Flows and hitting estimates

In this section we start deriving a key estimate used in the proof, which gives exponential decay of the tail for the hitting time of an arbitrary point y in S (Proposition 2.2 below). Recall that the main use of this result is to derive concentration of local time (Lemma 2.6) which in turns gives us estimates on the Radon-Nikodym derivative of $\mathbb{P}$ with respect to $\mathbb{Q}$, and hence on the partition function $Z(t, \beta)$.

Throughout we will use the notation $T_y := \inf\{t \geq 0 \ : \ X_t = y\}$ for the first hitting time of a vertex y.

Proposition 2.2. *Let (X_t) be a Markov chain with transition rates given by Q. Then uniformly over all $x, y \in S$, for some constants $c_{2.2}, C_{2.2}$ depending only on the dimension d,*

$$\mathbb{Q}_x[T_y > t] \leq \exp\left(-c_{2.2} t \pi(y)\right) \tag{2.4}$$

for all $t \geq C_{2.2}/\pi(y)$, if $d \geq 3$. For $d = 2$, we get

$$\mathbb{Q}_x[T_y > t] \leq \exp\left(-c_{2.2} \frac{t}{\kappa \log(r+2)}\right) \tag{2.5}$$

for all $t \geq C_{2.2} \kappa \log(r+2)$, where $\kappa = 1/\pi(y) \vee L^2 \log L$ and $r = \mathrm{dist}(y, \partial S)$.

Remark 2.3. *Note that when $d = 2$, it is always the case that $(1/\kappa) = \pi(y) \wedge 1/(L^2(\log L))$ satisfies*

$$\frac{1}{\kappa} \geq c \frac{\pi(y)}{\log(r+2)}$$

(consider the cases $\mathrm{dist}(y, \partial S) \leq L/2$ and $\mathrm{dist}(y, \partial S) \geq L/2$ to see this). Hence $c_{2.2}$ can be chosen so that for all $y \in S$ (with $r = \mathrm{dist}(y, \partial S)$), it holds that

$$\mathbb{Q}_x[T_y > t] \leq \exp\left(-c_{2.2} \frac{t\pi(y)}{(\log(r+2))^2}\right) \tag{2.6}$$

We start with the main technical lemma which bounds the local time accumulated at a vertex y until hitting another vertex x. We introduce a box B_1 of side-length $L/100$ at macroscopic distance (of order L) away from ∂S; for now we will take $B_1 = [-L/200, L/200]^d$ but later we will allow B_1 to be centered at a different point such as $(\lfloor L/2 \rfloor, \ldots, \lfloor L/2 \rfloor)$. We take B_2 a box of side length $L/50$ and B_3 a box of side-length $L/10$ both concentric to B_1.

Lemma 2.4. *Assume that $x \in B_1$ and $y \notin B_3$. There is a constant $C_{2.4} = C_{2.4}(d) > 0$ depending only on the dimension d such that*

$$\mathbb{E}_y[L(T_x, y)] \leq \begin{cases} C_{2.4} & \text{if } d \geq 3 \\ C_{2.4} \log(r+2) & \text{if } d = 2, \end{cases}$$

where $r = \mathrm{dist}(y, \partial S)$

Proof. Let $C_y \subset S$ be the cone formed by y and B_1. Let $\Sigma_m = \{z \in C_y \ : \ \mathrm{dist}(z, y) = m\}$. Let $M = \inf\{m : \Sigma_m \cap B_2 \neq \emptyset\}$, note that $M = O(L)$ uniformly in $y \notin B_3$. Let $\tilde{C}_y = \cup_{m=1}^{M} \Sigma_m$. Then let $\tilde{C}_x$ be the cone formed by x and Σ_M (see Figure 2), and let $C = \tilde{C}_x \cup \tilde{C}_y$.

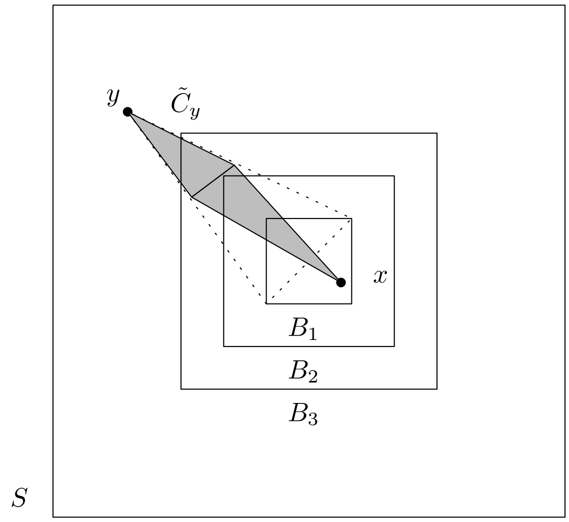

Figure 2: The shaded area is C, the union of the two cones $\tilde{C}_y$ and $\tilde{C}_x$ in the proof of Lemma 2.4

We will bound from below the probability that the Markov chain started from y hits x before returning to y. Recall that our Markov chain with law $\mathbb{Q}$ is reversible. Therefore it is equivalent to a discrete time random walk on a network on S where the weight, or conductance, of the edge $e = (z, w)$ is given by $\pi(z)Q(z, w) = \sqrt{\pi(z)\pi(w)}$. It is equivalent in the sense that both processes visit the same points in the same order, though possibly at different times. Hence it suffices to bound from above the effective resistance $\mathcal{R}_{\text{eff}}(y \to x)$ in this network.

We set all the weights on edges with at distance greater than 1 from C to be 0, then we have only reduced the conductance of all edges. Rayleigh's monotonicity principle (see [25, Chapter 2.4]) tells us that the effective resistance only increases. So it suffices to bound from above the effective resistance between y and x in this modified network.

The approach we use is that of [25, Chapter 2]. Let U be a random variable uniformly distributed (continuously) on the base of the cone, Σ_M, which is at distance of order L from y and from x. Let R be the union of the Euclidean segments $[y, U]$ and $[U, x]$. Let Γ be any monotone path that stays at distance at most $\sqrt{d} = \text{diam}([0, 1]^d)$ from R, until reaching x (how it is chosen exactly subject to the above constraint does not matter). It is easily calculated that for an edge e at distance k from y in $\tilde{C}_y$, the probability that $e \in \Gamma$ is at most $O(k^{-(d-1)})$. Likewise, for an edge e at distance k from x in $\tilde{C}_x$, the probability that e is traversed by Γ is at most $O(k^{-(d-1)})$. Also, the number of edges in $\tilde{C}_y$ (resp. $\tilde{C}_x$) at distance k from y (resp. x) is at most $O(k^{d-1})$. Finally, because Γ is chosen to be monotone, it traverses any edge at most once. So the function $\theta(e) = \mathbb{P}[e \in \Gamma] - \mathbb{P}[\hat{e} \in \Gamma]$ defines a unit flow in C, from y to x where $\hat{e}$ is the directed edge e in the reverse direction (again, see [25, Chapters 2.4 & 2.5]). Thompson's principle ([25, Chapter 2.4]) then implies that the energy of this flow bounds the effective resistance. Noting that $\pi(z) \geq c\pi(y)$ for all $z \in C$ we deduce

$$\begin{aligned}
\mathcal{R}_{\text{eff}}(y \to x) &\leq \sum_e \text{res}(e)\theta(e)^2 \\
&\leq \sum_{k=1}^{M} \frac{1}{c\pi(y)} \cdot O(k^{-2(d-1)}) \cdot O(k^{d-1}) + \sum_{k=1}^{\text{diam}(\tilde{C}_x)} \frac{1}{c\pi(y)} \cdot O(k^{-2(d-1)}) \cdot O(k^{d-1}) \\
&\leq \begin{cases} \frac{1}{c\pi(y)} \cdot O(\log L) & \text{if } d = 2 \\ \frac{1}{c\pi(y)} & \text{if } d \geq 3. \end{cases}
\end{aligned}$$

Thus, letting $w_y = \sum_{x \sim y} \sqrt{\pi(x)\pi(y)}$,

$$\mathbb{Q}_y[T_x < T_y] = \frac{1}{w_y \, \mathcal{R}_{\text{eff}}(y \to x)} \geq \begin{cases} c(\log L)^{-1} & \text{if } d = 2 \\ c & \text{if } d \geq 3. \end{cases} \tag{2.7}$$

In fact, when $d = 2$ we can get a better bound by improving on the estimation of $\text{res}(e)$ used above. Consider first the edges $e \in \tilde{C}_y$, and assume $y \in S_{r_0}$. It is obvious that for $r \geq r_0 + 1$, $|S_r \cap \tilde{C}_y| = O(r - r_0)$. Also, for each edge with at least one end in S_r, the probability that $e \in \Gamma$ is at most $O(r - r_0)^{-1}$. Hence, if $r_0 \leq L/2$,

$$\begin{aligned}
\sum_{e \in \tilde{C}_y} \text{res}(e)\theta(e)^2 &\leq c \sum_{r=r_0+1}^{L} \frac{1}{\mu_r} \frac{1}{r - r_0} \\
&\leq cL^3 \sum_{r=r_0+1}^{L/2} \frac{(\log(r+2))^2}{r(r - r_0)} + c \sum_{r=L/2}^{L} \frac{1}{\mu_{L/2}(L/10)} \\
&\leq cL^3 \frac{(\log(r_0+2))^3}{r_0+2} + cL^2(\log L)^2 = \frac{c}{\mu_{r_0}} \log(r_0+2) + cL^2(\log L)^2 \\
&\leq \frac{c}{\mu_{r_0}} \log(r_0+2).
\end{aligned}$$

In the second line above we have used that

$$\sum_{r=r_0+1}^{\infty} \frac{(\log(r+2))^2}{r(r-r_0)} \leq \frac{4(\log(r_0+2))^2}{r_0+1} \cdot \sum_{r=r_0+1}^{2(r_0+1)} \tfrac{1}{r-r_0} + \sum_{r>2(r_0+1)} \frac{2(\log(r+2))^2}{r^2}.$$

It is immediate that this conclusion also holds if $r_0 \geq L/2$. As for the edges in $\tilde{C}_x$, note that

$$\sum_{e \in \tilde{C}_x} \text{res}(e)\theta(e)^2 \leq c \sum_{k=1}^{L/100} \frac{1}{\mu_{99L/100}} \frac{1}{k} \leq cL^2 \log L$$

and note that this, too, is less or equal to c/μ_{r_0}. Therefore, we deduce that

$$\mathcal{R}_{\text{eff}}(y \to x) \leq (C/\mu_{r_0}) \log(r_0+2).$$

Consequently,

$$\mathbb{Q}_y[T_x < T_y] = \frac{1}{w_y\, \mathcal{R}_{\text{eff}}(y \to x)} \geq \frac{c}{\log(r_0+2)} \tag{2.8}$$

for all $d \geq 2$.

The result follows easily by the strong Markov property and the fact that at each subsequent visit to y, the accumulated local time is an exponential random variable with rate bounded away from 0 and thus has bounded mean. □

Let $\tau_{xy} = \inf\{t \geq 0 : \mathbb{E}_x[L(t,y)] \geq 1\}$, and let $\tau = \max_{x \in B_1} \tau_{xy}$. We immediately deduce from the above:

Lemma 2.5. *Uniformly over $x \in B_1$ and $y \notin B_3$, we have*

$$\mathbb{Q}_x[T_y < \tau] \geq \begin{cases} \frac{c}{\log(r+2)} & \text{if } d = 2 \\ c & \text{if } d \geq 3, \end{cases}$$

where $r = \operatorname{dist}(y, \partial S)$.

Proof. It suffices to prove this with τ replaced by τ_{xy} since $\tau \geq \tau_{xy}$. Now, observe that

$$\mathbb{Q}_x[T_y < \tau_{xy}] = \mathbb{Q}_x[L(\tau_{xy}, y) > 0] = \frac{\mathbb{E}_x[L(\tau_{xy}, y)]}{\mathbb{E}_x[L(\tau_{xy}, y) \mid L(\tau_{xy}, y) > 0]}.$$

Now, by definition, $\mathbb{E}_x[L(\tau_{xy}, y)] = 1$. On the other hand, by the strong Markov property,

$$\begin{aligned} \mathbb{E}_x[L(\tau_{xy}, y) \mid L(\tau_{xy}, y) > 0] &\leq \mathbb{E}_y[L(\tau_{xy}, y)] \leq \mathbb{E}_y[L(T_x, y)] + \mathbb{E}_x[L(\tau_{xy}, y)] \\ &\leq \begin{cases} C\log(r+2) + 1 & \text{if } d = 2 \\ C + 1, & \text{if } d \geq 3. \end{cases} \end{aligned}$$

where $C = C_{2.4}$ is the constant from Lemma 2.4. Thus, $\mathbb{Q}_x[T_y < \tau] \geq 1/(C+1)$ if $d \geq 3$, and $\mathbb{Q}_x[T_y < \tau] \geq 1/(\log(r+2)+1)$ for $d = 2$, as desired. □

Proof of Proposition 2.2. Suppose first that $y \notin B_3$, and let x be arbitrary in S. By (2.3), we know that $t_{\text{mix}} \leq C_{2.3} L^2 \log L$. It follows that, uniformly over $x \in S$, if $t = C_{2.3} L^2 \log L$,

$$\mathbb{Q}_x[X_t \in B_1] \geq c.$$

Define a sequence of times $t_1, t_2, \ldots$ by setting $t_n = n(C_{2.3} L^2 \log L + \tau)$. Then uniformly over $x \in S$ and $y \notin B_3$, we obtain by Lemma 2.5 and the Markov property at time t,

$$\mathbb{Q}_x[T_y > t_1] \leq 1 - h,$$

where $h = c$ if $d \geq 3$ and $h = c/\log(r+2)$ if $d = 2$. Hence, since this estimate is uniform in $x \in S$, we deduce by applying the Markov property at times $t_1, \ldots, t_n$,

$$\mathbb{Q}_x[T_y > t_n] \leq (1-h)^n \leq \exp(-nh).$$

Observe now that for $d \geq 3$, $\tau \leq c/\pi(y)$ for some c large enough. Indeed, $\pi(y) \leq C/L^d$ so if $t = c/\pi(y)$ with c sufficiently large, then $t \geq 2\,t_{\text{mix}}$ (see (2.3)). Hence we have that

$$\mathbb{E}_x[L(t,y)] \geq \int_{t/2}^{t} \mathbb{Q}_x[X_s = y]ds \geq \tfrac{t}{2} \cdot \tfrac{\pi(y)}{2} \geq 1,$$

and hence it follows that for all $x \in B_1$ (and indeed all $x \in S$), $\tau_{xy} \leq t$. Taking the maximum over $x \in B_1$, we obtain as desired $\tau \leq c/\pi(y)$. Observe further that, still in the case $d \geq 3$, we have that $L^2 \log L \leq C/\pi(y)$ hence $t_1 \leq C/\pi(y)$ as well.

On the other hand, if $d = 2$, then the same argument gives $\tau \leq C(L^2 \log L + 1/\pi(y))$ for some $C > 0$ large enough, so that we have $t_1 \leq C\kappa$, where $\kappa = 1/\pi(y) + L^2 \log L$. Hence for $n = \lfloor c\pi(y)t \rfloor$ ($d \geq 3$) or $n = \lfloor ct/\kappa \rfloor$ ($d = 2$),

$$\mathbb{Q}_x(T_y > t) \leq \mathbb{Q}_x(T_y > t_n) \leq \begin{cases} \exp(-ht\pi(y)) & \text{if } d \geq 3 \\ \exp(-ht/\kappa) & \text{if } d = 2, \end{cases}$$

as soon as $t \geq C/\pi(y)$ (for $d \geq 3$) or $t \geq C/\kappa$ ($d = 2$) so that $n \geq 1$.

This immediately implies the result of Proposition 2.2 if $y \notin B_3$. But the restriction $y \notin B_3$ is not essential. Indeed if $y \in B_3$, we can always consider a disjoint cube $\tilde{B}_3$, also of side length $L/100$, and at macroscopic distance (of order L) away from ∂S, for instance the one centered at $(\lfloor L/2 \rfloor, \ldots, \lfloor L/2 \rfloor)$. Throughout this box it will also be the case that $\pi(x) \geq c/L^d$ and so the exact same calculations apply, yielding a similar conclusion for all $y \notin \tilde{B}_3$. Since a given y is either in $S \setminus B_3$ or in $S \setminus \tilde{B}_3$ (as B_3 and $\tilde{B}_3$ are disjoint), Proposition 2.2 follows. □

2.4 Tail estimates for local time

We now turn to Lemma 2.6, which proves concentration of the local time at an arbitrary point $y \in S$, for which the key input is the exponential tails derived in Proposition 2.2. We will then state a corollary summarising our main use of Lemma 2.6.

Lemma 2.6. *There exist constants $C_{2.6}, c_{2.6} > 0$ depending only on d, such that the following holds.*

(i). Assume $d \geq 3$. For any $\delta > 0$, uniformly in $x, y \in S$,

$$\mathbb{Q}_x\big(L(t,y) \geq (1+\delta)\pi(y)t\big) \leq C_{2.6} \exp\big(-c_{2.6}(\delta \wedge \delta^2)\pi(y)t\big). \tag{2.9}$$

(ii). Assume $d = 2$. For any $\delta > 0$, uniformly in $x, y \in S$,

$$\mathbb{Q}_x\big(L(t,y) \geq (1+\delta)\pi(y)t\big) \leq C_{2.6} \exp\Big(-c_{2.6}(\delta \wedge \delta^2)\frac{\pi(y)t}{(\log(r+2))^2}\Big) \tag{2.10}$$

where $r = \mathrm{dist}(y, \partial S)$.

Proof. Fix $y \in S$. In this proof it is convenient to define a time θ by putting

$$\theta = \begin{cases} \frac{c_{2.2}}{\pi(y)}(\log(r+2))^2 & \text{if } d = 2 \\ \frac{c_{2.2}}{\pi(y)} & \text{if } d \geq 3. \end{cases} \tag{2.11}$$

By Proposition 2.2 (and Remark 2.3), y is hit with positive probability every θ units of time.

Let $q_y = \sum_x Q(y,x)$ be the total jump rate from y under $\mathbb{Q}$. Note that q_y is of constant order for L sufficiently large. Fix $\varepsilon > 0$ (in a way which will depend on δ and will be specified below), and let $n = \lceil \pi(y) q_y t(1+\varepsilon)\rceil$.

It will be useful to define

$$T = \inf\{t \geq 0 \ : \ X_t = y\} \qquad T^+ = \inf\left\{t \geq T_{S\setminus\{y\}} \ : \ X_t = y\right\} - T_{S\setminus\{y\}},$$

which are the hitting and return time to y, and also the successive return times to y: $T_0 = T$, and for $k > 0$,

$$\tilde{T}_k = \inf\{t \geq T_{k-1} \ : \ X_t \neq y\} \qquad T_k = \inf\left\{t \geq \tilde{T}_k \ : \ X_t = y\right\}.$$

Note that T_n is the sum of the independent increments

$$T_n = T_0 + \sum_{j=1}^n T_j - T_{j-1} = T_0 + \sum_{j=1}^n T_j - \tilde{T}_j + \sum_{j=1}^n \tilde{T}_j - T_{j-1}. \tag{2.12}$$

For each j, the increments $T_j - \tilde{T}_j$ have the same law. Also, the second sum

$$\sum_{j=1}^n \tilde{T}_j - T_{j-1}$$

is just $L(\tilde{T}_n, y) = L(T_n, y)$, each increment having the law of an independent exponential random variable of rate q_y.

Step I. First, we bound $\mathbb{Q}_x(T_n \leq t)$ by bounding the first sum $\sum_{j=1}^n T_j - \tilde{T}_j$. Note that $T_j - \tilde{T}_j$ has the distribution of T^+ under $\mathbb{Q}_y$. Thus, by Proposition 2.2 and Remark 2.3 (for the $d = 2$ case), $T_j - \tilde{T}_j$ has an exponential tail $\mathbb{Q}_x(T_j - \tilde{T}_j > t) = \mathbb{Q}_y(T^+ > t) \leq e^{-\theta^{-1}t}$ for $t \geq C_{2.2}\theta$, and

$$\begin{aligned}
\mathbb{E}_x[(T_j - \tilde{T}_j)^2] = \mathbb{E}_y[(T^+)^2] &= \int_0^\infty 2t\,\mathbb{P}_y[T^+ > t]dt \\
&\leq \int_0^{C_{2.2}\theta} 2tdt + \int_{C_{2.2}\theta}^\infty 2te^{-t/\theta}dt \\
&\leq A := (2 + (C_{2.2})^2)\cdot\theta^2.
\end{aligned}$$

Also, it is well known that $\mathbb{E}_y[T^+] = \frac{1}{q_y\pi(y)}$, so

$$\sum_{j=1}^n \mathbb{E}_x[T_j - \tilde{T}_j] = n\,\mathbb{E}_y[T^+] \geq t(1+\varepsilon)$$

by our choice of n. Using the inequalities $e^{-\xi} \leq 1 - \xi + \xi^2$, valid for for $\xi > 0$, and $1 + \xi \leq e^\xi$, valid for any $\xi \in \mathbb{R}$, we deduce that for any $\alpha > 0$,

$$\mathbb{E}_x[e^{-\alpha T_n}] \leq \left(\mathbb{E}_y[e^{-\alpha T^+}]\right)^n \leq \left(1 - \alpha\,\mathbb{E}_y[T^+] + \alpha^2 A\right)^n \leq \exp\left(-\alpha n\,\mathbb{E}_y[T^+] + n\alpha^2 A\right). \tag{2.13}$$

Since $\mathbb{Q}_x(T_n \leq t) \leq \mathbb{Q}_x\left(e^{-\alpha T_n} \geq e^{-\alpha t}\right)$, we have

$$\mathbb{Q}_x(T_n \leq t) \leq \exp\left(\alpha^2 nA + \alpha(t - n\,\mathbb{E}_y[T^+])\right),$$

which we may optimise over $\alpha > 0$. We find that the right hand side is minimised for $\alpha = \frac{n\,\mathbb{E}_y[T^+]-t}{2nA}$ (note that $\alpha \geq t\varepsilon/(2nA) > 0$). Substituting, this implies

$$\begin{aligned}\mathbb{Q}_x(T_n \leq t) &\leq \exp\Big(-\frac{(n\,\mathbb{E}_y[T^+]-t)^2}{4nA}\Big) \leq \exp\Big(-\frac{\varepsilon^2 t^2}{4nA}\Big)\\ &\leq \exp\Big(-\frac{\varepsilon^2}{4(2+(C_{2.2})^2)}\cdot\frac{t^2}{n\theta^2}\Big).\end{aligned} \tag{2.14}$$

Step II. Next, we bound $\mathbb{Q}_x(T_n > t, L(t,y) \geq \pi(y)t(1+\delta))$ by bounding the second sum $\sum_{j=1}^n \tilde{T}_j - T_{j-1}$ in (2.12).

Note that

$$\sum_{j=1}^n(\tilde{T}_j - T_{j-1}) = L(\tilde{T}_n, y) = L(T_n, y)$$

has the distribution of $\sum_{j=1}^n E_j$ where $(E_j)_j$ are i.i.d. exponential random variables of rate q_y. Standard concentration bounds on sums of i.i.d. exponential random variables show that for any $\eta > 0$,

$$\mathbb{P}\Big(\sum_{k=1}^n E_k \geq n\tfrac{1+\eta}{q_y}\Big) \leq \exp\big(-n(\eta - \log(1+\eta))\big).$$

Thus, if $\varepsilon > 0$ satisfies $\pi(y)q_y t(1+\delta) \geq n(1+\varepsilon)$ then

$$\mathbb{Q}_x(L(T_n,y) \geq \pi(y)t(1+\delta)) \leq \mathbb{Q}_x(L(T_n,y) \geq n\tfrac{1+\varepsilon}{q_y}) \leq \exp\big(-n(\varepsilon - \log(1+\varepsilon))\big). \tag{2.15}$$

Finally, we have that the event $\{L(t,y) \geq \pi(y)t(1+\delta)\}$ implies that either $T_n \leq t$ or $L(T_n,y) \geq L(t,y) \geq \pi(y)t(1+\delta) \geq n(1+\varepsilon)/q_y$, still assuming that ε is chosen so that $\pi(y)q_y t(1+\delta) \geq n(1+\varepsilon)$. This is the case as soon as

$$\varepsilon \leq \frac{\delta - 1/(\pi(y)q_y t)}{1 + 1/(\pi(y)q_y t)}.$$

By adapting the constant $C_{2.2}$, and since $q_y \to 1$ as L becomes large, we may assume without loss of generality that $\frac{1}{\pi(y)q_y t} < \delta/2$, and also note that $\frac{1}{\pi(y)q_y t} \leq 2/\beta_0$. Hence we can choose

$$\varepsilon = \tfrac{\delta/2}{1+1/\beta_0} \asymp \delta$$

Combining (2.14) and (2.15) we arrive at the conclusion of the lemma, since

$$\begin{aligned}\mathbb{Q}_x\big(L(t,y) \geq \pi(y)t(1+\delta)\big) &\leq \mathbb{Q}_x\big(T_n \leq t\big) + \mathbb{Q}_x\big(L(T_n,y) \geq \pi(y)t(1+\delta)\big)\\ &\leq \exp\Big(-c\tfrac{\varepsilon^2}{1+\varepsilon}(\log(r+2))^{-2}\pi(y)t\Big) + \exp\Big(-c(\varepsilon\wedge\varepsilon^2)\pi(y)t\Big)\\ &\leq 2\exp\Big(-c(\varepsilon^2\wedge\varepsilon)\cdot(\log(r+2))^{-2}\cdot\pi(y)t\Big),\end{aligned} \tag{2.16}$$

where the $(\log(r+2))^{-2}$ term can be removed when $d \geq 3$ and $c > 0$ is a constant depending on $C_{2.2}, c_{2.2}$, and β_0. (We have used (2.15) to get (2.16) and the fact that $(\varepsilon - \log(1+\varepsilon)) \asymp \varepsilon^2 \wedge \varepsilon$ as well as $n \preceq (1+\varepsilon)\pi(y)t$.) The lemma now follows since $\varepsilon \asymp \delta$. □

Remark 2.7. *A similar statement to Lemma 2.6 holds with the upper bound replaced by a lower bound: namely, there exist constant $c_{2.7}, C_{2.7}$ depending only on the dimension and β_0, such that for $d \geq 3$, for any $\delta \in (0,1)$:*

$$\mathbb{Q}_x\big(L(t,y) \leq (1-\delta)\pi(y)t\big) \leq C_{2.7} \exp\big(-c_{2.7}\delta^2\pi(y)t\big); \tag{2.17}$$

and for $d = 2$, uniformly in $x, y \in S$,

$$\mathbb{Q}_x\big(L(t,y) \leq (1-\delta)\pi(y)t\big) \leq C_{2.7} \exp\big(-c_{2.7}\delta^2 \frac{\pi(y)t}{(\log(r+2))^2}\big). \tag{2.18}$$

The proof is essentially similar with a few additional complications because we can no longer use the simple bound $e^{-\xi} \leq 1 - \xi + \xi^2$ which was valid for all $\xi \geq 0$, but when $\xi \leq 0$ is only valid for $-1 \leq \xi \leq 0$ (see (2.13)). However, in order to not overload the paper with technical details, and since this isn't needed for the proof of Theorem 1.1 we have chosen not to include the proof.

Corollary 2.8. *There exist a constant $c_{2.8} > 0$ depending only on d, such that the following holds. For any integer $k > 0$ and any $x, y \in S$,*

$$\mathbb{Q}_x\Big(L(t,y) \geq 2^k\pi(y)t\Big) \leq 2\exp\Big(-c_{2.8}2^k \cdot \beta_0 r(\log(r+2))^{-4}\Big),$$

where $r = \mathrm{dist}(y, \partial S)$.

Proof. This just follows from taking $1 + \delta = 2^k$ in Lemma 2.6, where we also use the fact that $t = \beta L^{d+1}$ so that $\pi(y)t \geq c\beta_0 r(\log(r+2))^{-2}$. □

3 Proof of Theorem 1.1

The goal of this section is to obtain the following lower bound on the partition function.

Proposition 3.1. *Let $\beta_0 > 0$ be fixed and let $\beta > \beta_0$. Then*

$$Z(t,\beta) \geq \exp\left(-\gamma t^{1-2/(d+1)}\beta^{2/(d+1)}\right)$$

where γ is a constant depending only on β_0 and $d \geq 2$.

3.1 Good event

For any $0 < r \leq L$ recall the definition of S_r:

$$S_r = \{z \in S \ : \ \mathrm{dist}(z, \partial S) = r\} = \{z \in S \ : \ \|z\|_\infty = L - r\}.$$

For $z \in S_r$ let $\langle z \rangle = \#\{1 \leq j \leq d \ : \ |z_j| = \|z\|_\infty\}$ (which is between 1 and d). Define

$$D_r = \{z \in S_r \ : \ \langle z \rangle > 1\}.$$

In two dimensions the vertices of D_r are exactly the four corners of the square S_r, while in three dimensions these are the edges of the cube defined by S_r. More generally the vertices of D_r are those which are in the intersections of faces of the hypercube defined by S_r.

For any $k \geq 1$ define the (random) subset

$$\mathcal{X}_k = \{x \in S \ : \ \frac{L(t,x)}{t\pi(x)} \in [2^k, 2^{k+1})\},$$

and consider the set of vertices

$$S_{r,k} = S_r \cap \mathcal{X}_k \qquad D_{r,k} = D_r \cap \mathcal{X}_k.$$

Define the "good" events:

$$\mathcal{S}_{r,k} = \{|S_{r,k}| \leq |S_r| \exp(-c_{3.2} 2^k r(\log(r+2))^{-4})\}$$

where $c_{3.2} = \frac{1}{2} c_{2.8}$. Define:

$$\mathcal{S}_r = \bigcap_{k \geq k_{3.2}} \mathcal{S}_{r,k} \qquad \mathcal{S} = \bigcap_{r=1}^{L} \mathcal{S}_r$$

where $k_{3.2}$ will be chosen below, large enough. Likewise, define

$$\mathcal{D}_{r,k} = \{|D_{r,k}| \leq |D_r| \exp(-c_{3.2} 2^k r(\log(r+2))^{-4})\}$$

where $c_{3.2} = \frac{1}{2} c_{2.8}$, and

$$\mathcal{D}_r = \bigcap_{k \geq k_{3.2}} \mathcal{D}_{r,k} \qquad \mathcal{D} = \bigcap_{r=1}^{L} \mathcal{D}_r$$

as above.

Fixing some γ large enough (which will be chosen later) we define the event

$$\mathcal{B} := \{|\partial R_t| \leq \gamma L^{d-1}\},$$

Finally, define the event

$$\mathcal{G} = \mathcal{B} \cap \mathcal{S} \cap \mathcal{D}. \tag{3.1}$$

We will now proceed to show that the probability of the good event $\mathbb{Q}_x(\mathcal{G})$ is bounded below uniformly in L. We will allow the starting point to be any fixed arbitrary $x \in S$ (although we only require these results with $x = 0$).

Lemma 3.2. *Fix $\varepsilon > 0$ and $\beta_0 > 0$. We can choose $k_{3.2} = k_{3.2}(\beta_0, \varepsilon)$ such that for any $\beta \geq \beta_0$. and for all t sufficiently large, we have $\mathbb{Q}_x(\mathcal{S}) \geq 1 - \varepsilon$ and $\mathbb{Q}_x(\mathcal{D}) \geq 1 - \varepsilon$ for all $x \in S$.*

Proof. We only show the proof for $\mathcal{S}$, as the proof for $\mathcal{D}$ is very similar.

By Corollary 2.8, taking expectation under $\mathbb{Q}_x$,

$$\begin{aligned}\mathbb{E}_x |S_{k,r}| &\leq |S_r| \max_{y \in S_r} \mathbb{Q}_x(y \in \mathcal{X}_k) \leq |S_r| \max_{y \in S_r} \mathbb{Q}_x(L(t,y) \geq 2^k \pi(y) t) \\ &\leq |S_r| \exp\big(-2c_{3.2} 2^k r(\log(r+2))^{-4}\big).\end{aligned}$$

Applying a union bound and Markov's inequality, we deduce that

$$\mathbb{Q}_x(\mathcal{S}^c) \leq \sum_{k \geq k_{3.2}} \sum_{r \geq 1} \exp\big(-c_{3.2} 2^k r(\log(r+2))^{-4}\big)$$

and so can be made arbitrarily small by choosing $k_{3.2}$ large enough (depending only on β_0), as desired. ∎

Now, we estimate $\mathbb{E}_x |\partial R_t|$ under $\mathbb{Q}_x$.

Lemma 3.3. *Let $\beta_0 > 0$. There exists $C_{3.3} > 0$ (depending only on β_0) such that for all $\beta \geq \beta_0$, under $\mathbb{Q}_x$ we have $\mathbb{E}_x |\partial R_t| \leq C_{3.3} L^{d-1}$.*

Proof. Since the maximal degree is $2d$ we obtain that $|\partial R_t| \leq 2d|S \setminus R_t| + CL^{d-1}$, where the second term represents all vertices on ∂S. Note that if $y \in S_r$ then $\pi(y)t \geq c_{2.2}\beta r(\log(r+2))^{-2}$ for some constant $c > 0$. Using Proposition 2.2 (and Remark 2.3),

$$\mathbb{E}_x |S \setminus R_t| = \sum_{y \in S} \mathbb{Q}_x[y \notin R_t] = \sum_{y \in S} \mathbb{Q}_x[T_y > t] \leq \sum_r |S_r| \exp\left(-c_{2.2}\beta_0 r(\log(r+2))^{-4}\right) \leq CL^{d-1},$$

as desired. ◻

We deduce from Lemma 3.3 and Markov's inequality that

$$\mathbb{Q}_x\left(|\partial R_t| \geq 2C_{3.3}L^{d-1}\right) \leq 1/2.$$

In particular, together with Lemma 3.2, if we take $\gamma \geq 2C_{3.3}$ (so altogether γ is chosen large enough in a way which depends only on β_0), we obtain for L sufficiently large

$$\mathbb{Q}_x(\mathcal{G}) \geq 1/4. \tag{3.2}$$

3.2 Radon-Nikodym derivative estimates

The following lemma is well known but very useful, see e.g. [28], IV, (22.8). We include it for completeness.

Lemma 3.4. *Let $f(z) = \sqrt{\pi(z)}$. Let $\Delta f(x) = \sum_{y \sim x} f(y) - f(x)$ be the discrete Laplacian. Then*

$$\frac{d\,\mathbb{P}_x}{d\mathbb{Q}_x}|_{\mathcal{F}_t} = \frac{f(X_0)}{f(X_t)} \exp\left(\int_0^t \frac{\Delta f}{f}(X_s)ds\right).$$

Proof. This follows easily from a discrete Feynman–Kac representation (see e.g. Lemma 11 in [19]). An alternative elementary proof is as follows. Suppose the successive states visited by ω up to time t are $x_0, \ldots, x_n$, with the path staying a time $\tau_0, \ldots, \tau_n$ at respectively at these locations. (Hence $\tau_1 + \ldots + \tau_n = t$.) If $x \in \mathbb{Z}^d$, then the total rate at which the particle would jump out of x under $\mathbb{Q}$ is given by $q(x) = f(x)^{-1} \sum_{y \sim x} f(y)$. Then letting $d(x) = 2d$ be the total rate of leaving x under $\mathbb{P}$,

$$\begin{aligned}
\frac{d\,\mathbb{P}_x}{d\mathbb{Q}_x}(\omega) &= \frac{e^{-d(x_0)\tau_0} \ldots e^{-d(x_n)\tau_n}}{Q(x_0,x_1)e^{-q(x_0)\tau_0} \ldots Q(x_{n-1},x_n)e^{-q(x_n)\tau_n}} \\
&= \frac{f(x_0)}{f(x_n)} \prod_{i=0}^n \exp\left((q-d)(x_i)\tau_i\right) = \frac{f(x_0)}{f(x_n)} \prod_{x \in \mathbb{Z}^d} \exp\left((q-d)(x)L(t,x)\right) \\
&= \frac{f(x_0)}{f(x_n)} \exp\left(\sum_{x \in \mathbb{Z}^d} \frac{\Delta f(x)}{f(x)} L(t,x)\right).
\end{aligned}$$

The result follows immediately. ◻

Lemma 3.5. *Recall the events $\mathcal{S}, \mathcal{D}$ defined above (3.1). On the event $\mathcal{S} \cap \mathcal{D}$ we have*

$$\int_0^t \frac{\Delta f}{f}(X_s)ds \geq -\gamma t L^{-2},$$

where $\gamma > 0$ is some constant (depending only on the dimension d and on β_0).

Proof. To ease the presentation, write $\mu(r) = \mu_r$, and consider μ as a function on real positive numbers. We want to estimate $\sum_x \frac{\Delta f(x)}{f(x)} L(t,x)$ from below.

Recall that for $x \in S_r$ we have

$$f(x) = \sqrt{\mu(r)} = \begin{cases} CL^{-(d+1)/2} \frac{\sqrt{r}}{\log(r+2)} & \text{if } r \leq L/2 \\ \sqrt{\mu(L/2)} + C\frac{r-L/2}{L^{(d+2)/2}} & \text{if } r > L/2 \end{cases}$$

where C is such that $\sum_x f(x)^2 = 1$. A careful second order Taylor expansion provides the following estimate for all $x \in S_r, y \in S_{r+1}$ with $r+1 \leq L/2$:

$$\tfrac{f(y)}{f(x)} - 1 = \tfrac{1}{2r} - \tfrac{1}{(r+2)\log(r+2)} + O(r^{-2}).$$

Now, for $x \in S_r \setminus D_r$ with $r+1 \leq L/2$, there are two neighbours of x one in S_{r+1} and one in S_{r-1}, and all other neighbours are in S_r. Hence, for some $\xi \in [r-1, r]$,

$$\begin{aligned} \tfrac{\Delta f(x)}{f(x)} &= \tfrac{1}{2r} - \tfrac{1}{(r+2)\log(r+2)} - \tfrac{1}{2(r-1)} + \tfrac{1}{(r+1)\log(r+1)} + O(r^{-2}) \\ &= -\tfrac{1}{2\xi^2} + \tfrac{\log(\xi+2)+1}{(\xi+2)^2(\log(\xi+2))^2} + O(r^{-2}) \geq -cr^{-2}, \end{aligned}$$

for some constant $c > 0$. For $r > L/2$ and $r = L/2$, we have similarly to the above that if $x \in S_r \setminus D_r$, then $\Delta f(x) \geq -cr^{-2}$.

On the other hand, if $x \in D_r$ for $r < L/2$, then $2d - \langle x \rangle$ neighbours of x are in S_r, and $\langle x \rangle$ neighbours are in S_{r-1}. So,

$$\tfrac{\Delta f(x)}{f(x)} = -\tfrac{\langle x \rangle}{2(r-1)} + \tfrac{\langle x \rangle}{(r+1)\log(r+1)} + O(r^{-2}) \geq -d \cdot cr^{-1}.$$

Also, if $x \in D_r$ for $r > L/2$ then,

$$\tfrac{\Delta f(x)}{f(x)} = \langle x \rangle \cdot \frac{\sqrt{\mu(r-1)} - \sqrt{\mu(r)}}{\sqrt{\mu(r)}} = -\langle x \rangle \cdot \frac{C}{L^{(d+2)/2}} \geq -cr^{-1}.$$

Finally, in the case $r = L/2$, the same computations as above leave us with the estimate $\Delta f(x)/f(x) \geq -d \cdot cL^{-1} \geq -cr^{-1}$, for all $x \in S_{L/2}$. Altogether, whatever the value of r, if $x \in D_r$,

$$\tfrac{\Delta f(x)}{f(x)} \geq -cr^{-1}.$$

Now let us estimate the contribution coming from points in $S_r \setminus D_r$. Denote

$$\varphi(r) = \begin{cases} r(\log(r+2))^{-2} & \text{if } r \leq L/2, \\ r & \text{if } r > L/2. \end{cases}$$

If $x \in S_{r,k} \setminus D_r$ we have that $\frac{\Delta f(x)}{f(x)} \geq -cr^{-2}$ and also $L(t,x) \leq 2^{k+1}\pi(x)t \leq c2^{k+1}\beta\varphi(r)$. So, on the event $\mathcal{S}_{r,k}$,

$$\sum_{x \in S_{r,k} \setminus D_r} \tfrac{\Delta f(x)}{f(x)} L(t,x) \geq -c|S_{r,k} \setminus D_r| \cdot 2^{k+1}\beta\varphi(r)r^{-2}$$
$$\geq -c|S_r| \cdot 2^{k+1}\beta\varphi(r)r^{-2} \cdot \exp\big(-c_{3.2}2^k r(\log(r+2))^{-4}\big).$$

Summing over k gives that on the event $\mathcal{S}$,

$$\sum_{x \in S_r \setminus D_r} \tfrac{\Delta f(x)}{f(x)} L(t,x) \geq -cL^{d-1} \cdot \beta\varphi(r)r^{-2} \cdot \Big(k_{3.2}2^{k_{3.2}+1} + \sum_{k \geq k_{3.2}} 2^{k+1}e^{-c_{3.2}2^k r(\log(r+2))^{-4}}\Big) \geq -ctL^{-2}\varphi(r)r^{-2},$$

where the final constant $c > 0$ depends on $k_{3.2}$. Likewise, on the good event $\mathcal{D}$, the contribution from $x \in D_r$ is

$$\sum_{x \in D_r} \tfrac{\Delta f(x)}{f(x)} L(t,x) \geq -c|D_r| \cdot \beta\varphi(r)r^{-1} \cdot \Big(k_{3.2}2^{k_{3.2}+1} + \sum_{k \geq k_{3.2}} 2^{k+1}e^{-c_{3.2}2^{k/2} r(\log(r+2))^{-4}}\Big)$$
$$\geq -cL^{d-2} \cdot \beta\varphi(r)r^{-1} \geq -ctL^{-3}\varphi(r)r^{-1} \geq -ctL^{-2}\varphi(r)r^{-2}.$$

Consequently, summing these two contributions and summing over r we get,

$$\sum_{x \in S} \tfrac{\Delta f(x)}{f(x)} L(t,x) \geq -ctL^{-2}$$

because

$$\sum_{r=1}^{L} \varphi(r)r^{-2} \leq \sum_{1 \leq r \leq L/2} \tfrac{1}{r(\log(r+2))^2} + \sum_{L/2 < r \leq L} \tfrac{1}{r} \leq C. \tag{3.3}$$

This concludes the proof of Lemma 3.5. Note that it is in (3.3) that we see the importance of the logarithmic correction terms in the choice of the local time profile $\pi(x)$ in (2.2). □

We now turn to the proof of the lower bound on the partition function.

Proof of Proposition 3.1. Let $d \geq 2$, and let $x = 0$ be the starting point of the walk. Using the definition of $\mathcal{G}$, Lemma 3.5, and (3.2), and the fact that $\pi(x) \leq \pi(0)$ for any $x \in S$, we obtain:

$$Z(t,\beta) = \mathbb{E}_0[\exp(-\beta|\partial R_t|)] \geq \mathbb{E}_0[1_{\mathcal{G}} \exp(-\beta|\partial R_t|)]$$
$$\geq \exp(-\beta\gamma L^{d-1})\, \mathbb{P}_0(\mathcal{G}) = \exp(-\beta\gamma L^{d-1})\mathbb{Q}_0(1_{\mathcal{G}} \frac{d\,\mathbb{P}_0}{d\mathbb{Q}_0})$$
$$\geq \tfrac{1}{4} \exp(-\gamma(\beta L^{d-1} + tL^{-2})).$$

Recall that our choice of $L = (t/\beta)^{1/(d+1)}$ guarantees that both terms βL^{d-1} and tL^{-2} in the exponential are of the same order of magnitude, namely $t^{1-2/(d+1)}\beta^{2/(d+1)}$. (This is saying that the entropic and energetic costs balance each other out). This finishes the proof of Proposition 3.1 for a sufficiently large γ (depending only on β_0 and the dimension d). □

3.3 Discrete isoperimetry

We now state and prove a modified isoperimetric inequality which deals with the outer boundary of a set. We first need some definitions. For a set $G \subset \mathbb{Z}^d$, let $\mathrm{Ext}(G)$ be the unique unbounded connected component of $\mathbb{Z}^d \setminus G$. Let the outer vertex boundary $\partial^* G$ be defined by

$$\partial^* G = \{x \in G \ : \ \exists y \in \mathrm{Ext}(G) \ , \ x \sim y\}.$$

The outer edge boundary, denoted by $\partial_e^* G$, consists of those edges $e = (x, y)$ with $x \in G$ and $y \in \mathrm{Ext}(G)$.

Lemma 3.6. *Let $A \subset \mathbb{Z}^d$.*

(i). Assume $d = 2$, and let R be the smallest rectangle in $\mathbb{Z}^2$ containing A. Then,

$$|\partial^* R| \leq 3|\partial^* A|.$$

(ii). For any $d \geq 2$,

$$|\partial^* A| \geq \frac{2d}{2d-1} |A|^{\frac{d-1}{d}}$$

Proof. We start with the proof for $d = 2$. For any connected set A, we have that $|\partial^* A| \leq |\partial_e^* A| \leq 3|\partial^* A|$ since adjacent to any vertex in $\partial^* S$ there are at most 3 edges in $\partial_e^* S$ (and at least one edge must connect that vertex to the set itself). Thus, it suffices to prove that $|\partial_e^* R| \leq |\partial_e^* A|$.

Let $(x, x + e) \in \partial_e^* A$, for some $e \in \{\pm e_i\}$, where e_i are the standard unit vectors of $\mathbb{Z}^2$, where $x \in A$ and $x + e \notin A$. Note that since $xinA$, we also have $x \in R$. Thus, there exists a (necessarily unique) $k \geq 0$ such that $x + ke \in R$ and $x + ze \notin R$ for all $z > k$. Thus, $(x + ke, x + (k+1)e) \in \partial_e^* R$. Define a map $\phi : \partial_e^* A \to \partial_e^* R$ by $\phi((x, x + e)) = (x + ke, x + (k + 1)e)$ for this k.

We claim that $\phi : \partial_e^* A \to \partial_e^* R$ is onto. This follows since if $(y, y + e) \in \partial_e^* R$, then considering the line $L = \{y - ke \ : \ k \geq 0\}$, it must be that $L \cap A \neq \emptyset$, since otherwise either A would not be connected or R would not be the smallest rectangle containing A. (This relies on the assumption that $d = 2$.) Thus, there must exist some $k \geq 0$ such that $y - ke \in A$ and $y - ze \notin A$ for any $z < k$. Thus, the edge $(y - ke, y - (k-1)e)$ is in $\partial_e^* A$, and it is immediate that $\phi(y - ke, y - (k-1)e) = (y, y + e)$.

This proves that there is a map from $\partial_e^* A$ onto $\partial_e^* R$. So $|\partial^* R| \leq |\partial_e^* R| \leq |\partial_e^* A| \leq 3|\partial^* A|$.

For the general case $d \geq 3$ we use the discrete Loomis–Whitney inequality (Theorem 2 in [24]), which states that if A_i is the projection of A onto $\mathbb{Z}^{d-1}$ along the ith coordinate then

$$|A|^{d-1} \leq \prod_{i=1}^{d} |A_i|. \tag{3.4}$$

For each vertex in $z \in A_i$ consider the line L going through z and which is parallel to the ith coordinate axis. It intersects A in at least one place (assume for simplicity and without loss of generality that A does not intersect any hyperplane where one of the coordinates is 0). The first and last such intersections with A necessarily correspond to two edges in $\partial_e^* A$, since the rest of the line lies in $\mathbb{Z}^d \setminus A$ and is unbounded. Thus to each vertex in A_i one can associate an edge $e \in \partial_e^* A$.

This edge is necessarily unique, since its direction indicates the coordinate used for the projection. We deduce, using the arithmetic-geometric inequality,

$$\prod_{i=1}^{d}|A_i| \le \left(\frac{1}{d}\sum_{i=1}^{d}|A_i|\right)^d \le \left(\frac{1}{2d}|\partial_e^* A|\right)^d \le \left(\frac{2d-1}{2d}|\partial^* A|\right)^d,$$

(we have used that to any vertex in $\partial^* A$ there are at most $2d-1$ corresponding edges in $\partial_e^* A$, so $|\partial_e^* A| \le (2d-1)|\partial^* A|$). Combining with (3.4) this gives the desired result. □

3.4 Proof of condensation

We will prove the following more precise statement of Theorem 1.1.

Theorem 3.7. *Let $d \ge 2$. Fix $\beta_0 > 0$ and let $\beta > \beta_0$. Let γ be as in Proposition 3.1. Then,*

$$\mu_t\Big[\operatorname{diam}(R_t) \ge \frac{1}{\sqrt{\gamma}}\left(\tfrac{t}{\beta}\right)^{1/(d+1)}\Big] \ge 1 - C\exp\left(-\gamma t^{1-2/(d+1)}\beta^{2/(d+1)}\right), \tag{3.5}$$

and if $d = 2$ then

$$\mu_t\Big[\operatorname{diam}(R_t) \le 6\gamma\left(\tfrac{t}{\beta}\right)^{1/3}\Big] \ge 1 - C\exp\left(-\gamma t^{1/3}\beta^{2/3}\right).$$

Moreover, for all $d \ge 2$,

$$\mu_t\Big[|R_t| \le (2\gamma)^{d/(d-1)}\left(\tfrac{t}{\beta}\right)^{d/(d+1)}\Big] \ge 1 - C\exp\left(-\gamma t^{1-2/(d+1)}\beta^{2/(d+1)}\right).$$

Proof. Recall that by Proposition 3.1

$$Z(t,\beta) \ge \exp\left(-\gamma t^{1-2/(d+1)}\beta^{2/(d+1)}\right). \tag{3.6}$$

We start with the lower bound on the diameter. We require the following standard estimate. Let $R_t^{\square}$ denote the smallest d-dimensional box containing R_t. For $1 \le i \le d$, let J_t^i denote the length of the projection of $R_t^{\square}$ (or equivalently R_t) onto the ith coordinate axis.

Lemma 3.8. *We have*

$$\mathbb{P}_0[J_t^i \le n] \le n\exp\left(-t\frac{\pi^2}{4n^2}\right)$$

Proof. Under $\mathbb{P}_0$, the coordinates $X_t^1, \dots, X_t^d$ are independent rate 1 simple random walks on $\mathbb{Z}$. We just focus on the first coordinate, $X_t = X_t^1$, and compute $\mathbb{P}_x(T > t)$ where $x \in \{1, \dots, J\}$ and $T = \inf\{t \ge 0 : X_t \notin [1, J]\}$. Let $\mathcal{L}$ denote the generator of (rate 1) simple random walk on $\mathbb{Z}$, and let $\phi(x) = e^{i\pi x/J}$. It is trivial to check that

$$\mathcal{L}\phi(x) = -\lambda\phi(x)$$

for all $x \in \mathbb{Z}$, where $\lambda = 1 - \cos(\pi/J)$. Thus if we let $\psi(t,x) = e^{\lambda t}\sin(\pi x/J)$ we have

$$\frac{\partial}{\partial t}\psi + \mathcal{L}\psi = 0$$

and hence $M_t := e^{\lambda t}\sin(\pi X_t/J)$ is a martingale. Consequently, applying the optional stopping time theorem at the time $t \wedge T$ (which is bounded), and the inequality $\sin(u) \geq (2/\pi)u$ valid for $0 \leq u \leq \pi/2$, yields

$$\begin{aligned}\sin(\pi x/J) &= \mathbb{E}_x(e^{\lambda t}\sin(\pi X_{t\wedge T}/J)) \\ &\geq e^{\lambda t}\frac{2}{J}\mathbb{P}_x(T > t).\end{aligned}$$

Therefore,

$$\mathbb{P}_x(T > t) \leq J e^{-\lambda t}.$$

Now, $\lambda = 1 - \cos(\pi/J) \geq \pi^2/4J^2$ for J large enough, and the result follows. □

We now combine Lemma 3.8 with the previously obtained upper bound to obtain a lower bound on the diameter of R_t under μ_t. Let $J_t^1, \dots, J_t^d$ be the side-lengths of $R_t^\square$. Let $N = \min\{J_t^1, \dots, J_t^d\}$. We will prove the stronger statement that $N \geq c(t/\beta)^{1/(d+1)}$ with high probability. This will also be needed in the upper bound in dimension 2.

By Lemma 3.8 and Proposition 3.1, for an integer $n > 0$,

$$\begin{aligned}\mu_t[J_t^1 \leq n] &\leq Z(t,\beta)^{-1} n C \exp\left(-\frac{\pi^2 t}{4n^2}\right) \\ &\leq Cn \exp\left(\gamma t^{1-2/(d+1)}\beta^{2/(d+1)} - \frac{\pi^2}{4n^2}t\right).\end{aligned}$$

Thus, if $n = (1/\sqrt{\gamma})(t/\beta)^{1/(d+1)}$, we get that

$$\mu_t[J_t^1 \leq n] \leq \exp\left(-\gamma t^{1-2/(d+1)}\beta^{2/(d+1)}\right).$$

Of course, we get the same bound replacing J_t^1 by J_t^i. Therefore,

$$\mu_t[N \leq n] \leq d \exp\left(-\gamma t^{1-2/(d+1)}\beta^{2/(d+1)}\right). \tag{3.7}$$

In particular, it holds that with high μ_t-probability

$$\mathrm{diam}(R_t) \geq \frac{1}{\sqrt{\gamma}}(t/\beta)^{1/(d+1)}.$$

We now turn to the upper bound on the diameter in dimension $d = 2$. We make the following observation. In dimension $d = 2$, if we know that the diameter of a shape G is $\geq M$ for some large M then it automatically follows by Lemma 3.6 that $|\partial G| \geq cM$. This ensures that the energy associated to this particular shape is at least $c\beta M$. This is enough for proving the theorem in the $d = 2$ case. [On the other hand, in dimension 3 and higher, such a simple relationship is no longer true: if $\mathrm{diam}(G) \geq M$ then we can only infer that $|\partial G| \geq cM$, translating into an energy cost of $c\beta M$. This is far less than what we need, since we believe the relevant energy contributions are of order βM^{d-1}. The issue is that a shape could have a big diameter in one direction and be very "thin" along other directions.]

Recall that $R_t^\square$ is a $J_t^1 \times J_t^2$ rectangle. Lemma 3.6 tells us that $|\partial R_t| \geq |\partial^* R_t| \geq \frac{1}{3}|\partial^* R_t^\square| = \frac{2}{3} \cdot (J_t^1 + J_t^2)$. Thus,

$$\begin{aligned} \mu_t[J_t^i > m] &\leq Z(t,\beta)^{-1} \sum_{k=m+1}^{\infty} \exp(-\beta \tfrac{2}{3} k) \\ &\leq C \exp(-\beta \tfrac{2}{3} m + \gamma t^{1/3} \beta^{2/3}). \end{aligned}$$

If $m = 3\gamma(t/\beta)^{1/3}$ this probability is at most of order $\exp(-\gamma t^{1/3}\beta^{2/3})$. A union bound over $i = 1, 2$ give that in particular, $\mathrm{diam}(R_t) \leq 2m$ with high probability, which concludes the proof of the first part of Theorem 3.7.

We turn to the second part of the proof which yields an upper bound on the volume of R_t in all dimensions $d \geq 2$. For this we note that by Lemma 3.6, if $|R_t| \geq m$ then $|\partial R_t| \geq |\partial^* R_t| \geq m^{(d-1)/d}$, and so almost surely on this event, $\exp(-\beta H(\omega)) \leq \exp(-\beta m^{(d-1)/d})$. Consequently,

$$\begin{aligned} \mu_t[|R_t| \geq m] &\leq Z(t,\beta)^{-1} \exp(-\beta m^{(d-1)/d}) \cdot \mathbb{P}[|R_t| \geq m] \\ &\leq \exp\left(-\beta m^{(d-1)/d} + \gamma t^{1-2/(d+1)} \beta^{2/(d+1)}\right). \end{aligned}$$

If $m^{(d-1)/d} = 2\gamma(\frac{t}{\beta})^{(1-2/(d+1))}$, or equivalently, $m = (2\gamma)^{d/(d-1)}(\frac{t}{\beta})^{d/(d+1)}$, this probability is at most $\exp(-\gamma t^{1-2/(d+1)}\beta^{2/(d+1)})$. This completes the upper bound on the volume in all dimensions. □

3.5 Proof of Theorems 1.3 and 1.2

We explain how to adapt the arguments of the proof of Theorem 1.1 to give the proof of Theorem 1.2. Let $K > 0$ be large enough and let $S' = \cup_{r > K} S_r$. Let $\mathcal{B}' = \{\forall x \in S' : L(t,x) \geq \beta\}$. Let $\mathcal{G}' = \mathcal{G}'_t = \mathcal{B}' \cap \mathcal{S} \cap \mathcal{D}$. Then the same arguments as in (3.2) show that $\mathbb{Q}(\mathcal{G}') \geq 1/4$, provided that K is a sufficiently large constant. The only difference with (3.2) is that it no longer suffices to bound the expected number of vertices that were not visited by time t as in Lemma 3.3, which followed directly from Proposition 2.2. Instead, we need to show that the local time at every vertex in S' is greater than β with probability greater than $1/2$ say. However this is a direct consequence of the lower bound large deviations discussed in Remark 2.7.

We deduce that

$$\mathbb{P}(\mathcal{G}') \geq \exp(-\gamma t^{1-2/(d+1)} \beta^{2/(d+1)})$$

for some large enough constant γ depending only on β_0 and d. Assume that $\mathcal{G}'_t$ holds. In the next t units of time, we make sure that the each of the remaining $O(KL^{d-1})$ vertices of $S \setminus S'$ are visited at least β units of time each, as follows. For each $1 \leq k \leq K$, we visit each vertex in S_k in clockwise order, starting from $(k, 0, \ldots, 0)$. At each new vertex, the walk remains at least β and at most 2β units of time. When the walk has visited each vertex of S_k, it moves on to S_{k+1}. The total amount of time spent doing so is at most $2\beta K L^{d-1} \leq 2Kt/L^2$, which is much less than the t units of time in which we want to achieve this, since by assumption $\beta = o(t)$. In the remaining amount of time, the walk is free to do what it wants, provided it stays in S.

If all these conditions are fulfilled, it is clear that $R_{2t} = S$ and that each vertex has a local time greater than β, so $\mathcal{E}_{2t}$ holds. The probability of visiting every vertex in this prescribed order immediately after t is at least $\exp(-c\beta K L^{d-1})$ for some $c < \infty$. The probability of remaining in

S after that (for a time necessarily shorter than t) is easily seen to be at least $\exp(-ct/L^2)$ and hence at least $\exp(-ct^{1-2/(d+1)}\beta^{2/(d+1)})$. All in all, we deduce

$$\mathbb{P}(\mathcal{E}_{2t}) \geq \exp(-\gamma t^{1-2/(d+1)}\beta^{2/(d+1)}), \tag{3.8}$$

and thus (changing t into $t/2$) the same inequality holds with the left hand side replaced by $\mathbb{P}(\mathcal{E}_t)$. This argument also shows that if $\tilde{Z}(t,\beta)$ is the partition function corresponding to the Hamiltonian $\tilde{H} = \sum_{x\in\partial R_t} L(t,x)$ in (1.8), then

$$\tilde{Z}(t,\beta) \geq \exp(-\gamma t^{1-2/(d+1)}\beta^{2/(d+1)}). \tag{3.9}$$

(In fact, this could also be deduced from Corollary 2.8.)

Now, we claim that for any finite set G of vertices,

$$\mathbb{P}(R_t = G, \mathcal{E}_t) \leq \exp(-\beta|\partial G|). \tag{3.10}$$

For each $x \in \partial G$, let y be a neighbour of x such that $y \notin G$. Consider the event $J_{xy}(t)$ that by time t there has never been a jump from x to y. On $\mathcal{E}_t$, x is visited at least β units of time. While at x, the rate of jumping to y is of course 1. Let E_{xy} be independent exponential random variables with rate 1, which represents the amount of time a particle would have to wait before jumping to y. Thus $J_{xy}(t) \cap \mathcal{E}_t \subset \{E_{xy} > \beta\}$. Hence

$$\mathbb{P}(R_t = G, \mathcal{E}_t) \leq \mathbb{P}(\cap_{x\in\partial G} J_{xy}(t) \cap \mathcal{E}_t) \leq \mathbb{P}(\cap_{x\in\partial G} E_{xy} > \beta) \leq e^{-\beta|\partial G|}$$

by independence of the random variables E_{xy}. Thus (3.10) is established.

Putting together (3.8) and (3.10) (resp. (3.9) and the definition of $\tilde{\mu}$), the proof of Theorem 1.2 (resp. Theorem 1.3) proceeds essentially as in Theorem 3.7. More precisely, let $J_t^1, \ldots, J_t^d$ be the dimensions of R_t in each coordinate. The lower bound in (3.8) implies exactly as in (3.7) that

$$\mathbb{P}(\min(J_t^1, \ldots, J_t^d) \geq n | \mathcal{E}_t) \to 1$$

as $t \to \infty$, where $n = (1/\sqrt{\gamma})(t/\beta)^{1/(d+1)}$. In particular, conditioned on $\mathcal{E}_t$, with high probability we have $\operatorname{diam}(R_t) \geq n$.

For the upper-bound on $\operatorname{diam}(R_t)$ in the case $d = 2$, or the upper bound on $|R_t|$ in the general case $d \geq 2$, we proceed as follows. We focus on the bound on $|R_t|$ in the general case $d \geq 2$, which requires a few more ideas. For each edge e, consider the unit area plaquette $p(e)$, orthogonal to e and such that the centre of $p(e)$ coincides with the midpoint of the edge e.

Definition 3.9. By a self-avoiding surface, we mean a connected union of plaquettes with disjoint $(d-1)$-dimensional interior.

When $d = 2$, this is essentially equivalent to a self-avoiding walk. Let $\mathcal{S}_n$ denote the set of self-avoiding surfaces with n plaquettes, and contained in a ball of radius n about the origin. Let $c_n = |\mathcal{S}_n|$ and let

$$\alpha = \alpha(d) = \limsup_{n\to\infty} c_n^{1/n} \qquad \beta_0 = \log\alpha. \tag{3.11}$$

Note that when $d = 2$, the limsup is a limit and is (essentially by definition) equal to the connective constant of $\mathbb{Z}^2$. It is easy to check that $1 \leq \alpha \leq (2d)^{2d} < \infty$ in general, which is all we will use.

To each finite $G \subset \mathbb{Z}^d$ we can associate a finite self-avoiding surface, where the plaquettes are obtained by considering each of the edges $e = (x, y)$, with $x \in G$ and $y \in \mathrm{Ext}(G)$. Let $\mathcal{S}_{j_1,\dots,j_d}$ denote the set of surfaces where the diameter in each direction $1, \dots, d$, does not exceed $j_1, \dots, j_d$ respectively. Let Σ be the (random) self-avoiding surface associated with R_t. For a given self-avoiding surface $\sigma \in \mathcal{S}_j$, we have by the same argument as in (3.10) (since each plaquette corresponds to an edge (x, y) such that the corresponding exponential random variable E_{xy} sattisfies $E_{xy} > \beta$, and these events are independent even for edges which share vertices)

$$\mathbb{P}(\Sigma = \sigma, \mathcal{E}_t) \le \exp(-\beta j), \tag{3.12}$$

Let $\beta_1 > \beta_0 = \log \alpha$ and assume that $\beta > \beta_1$. Let $\beta_1' = (\beta_0 + \beta_1)/2$. Note that for n large enough, we have $|\mathcal{S}_n| \le \exp(\beta_1' n)$.

Therefore, by (3.12),

$$\begin{aligned}\mathbb{P}[\Sigma \in \mathcal{S}_n | \mathcal{E}_t] &\le \mathbb{P}(\mathcal{E}_t)^{-1} \sum_{j=n}^{\infty} e^{\beta_1' j} e^{-\beta j} \\ &\le C \exp\{\gamma t^{1-2/(d+1)} \beta^{2/(d+1)} - (\beta - \beta_1')n\}\end{aligned}$$

where $C = \sum_{j \ge 0} \exp(-(\beta - \beta_1')j) \le \sum_j \exp(-j(\beta_1 - \beta_1')/2) < \infty$ since $\beta_1 > \beta_1'$. Let

$$n = \left\lceil \frac{\gamma t^{1-2/(d+1)} \beta^{2/(d+1)}}{2(\beta - \beta_1')} \right\rceil \le C\gamma (t/\beta)^{\frac{d-1}{d+1}}$$

where C depends only on β_1. Then we deduce

$$\mathbb{P}[\Sigma \in \mathcal{S}_n | \mathcal{E}_t] \to 0,$$

Hence $|\partial^* R_t^\square| \le n$ with high conditional probability given $\mathcal{E}_t$, and thus (by Lemma 3.6)

$$|R_t| \le [(2d-1)n]^{d/(d-1)} \le C\gamma (t/\beta)^{\frac{d}{d+1}}$$

with high conditional probability, as desired.

Remark 3.10. *It is interesting to note that the lower bound on* $\mathrm{diam}(R_t)$ *is valid for all* $\beta > 0$ *(i.e., does not assume* $\beta > \beta_0$*).*

4 Open problems and conjectures

We now discuss several open problems. The most basic ones is to ask whether the constants c_1 and c_2 appearing in Theorem 1.1 really need to be different from one another, and if indeed $t^{1/(d+1)}$ is the right order of magnitude in all dimensions $d \ge 2$.

Problem 1. *Does there exist a constant c such that* $\mathrm{diam}(R_t)/t^{1/(d+1)}$ *converges to c in μ_t-probability as $t \to \infty$?*

Another immediate question is to obtain the limiting shape of the random walk. Suppose $d = 2$ for simplicity.

Problem 2. *Find a closed bounded set $S = S(\beta) \subset \mathbb{R}^2$ such that*

$$\inf_{z\in\mathbb{R}^2} d_{\mathrm{Haus}}\left(\frac{R_t}{\mathrm{diam}(R_t)}; z + S\right) \to 0$$

in probability, where d_{Haus} stands for Hausdorff distance.

The reason for looking for sets of the form $z + S$ to approximate the normalised range of the walk, is that we believe that with this formulation S can be chosen to be nonrandom: that is, a certain random translation of the deterministic shape S gives a good approximation of the normalised range.

Once the existence of S is established one may ask numerous questions about its geometry. For instance, does it have any flat facet? We conjecture in fact that as $\beta \to \infty$, $S(\beta)$ converges in the Hausdorff sense to a diamond of unit diameter. Indeed it is easy to see that the diamond is the minimiser of the isoperimetric problem for the vertex-boundary: $\min_{|S|=k} |\partial S|$ is attained for a diamond $S = \{x, y : |x| + |y| \le n\}$, whenever $k = 2n(n + 1)$.

It is interesting to compare this situation for the case of a random walk conditioned on $\{L_t(x) \ge \beta, \forall x \in R_t\}$. There again, we believe a similar shape theorem holds, but the limiting shape of $S(\beta)$ as $\beta \to \infty$ should instead be a square with unit diameter. This is because, (3.10) in fact shows that

$$\mathbb{P}[R_t = G, \mathcal{E}_t] \le \exp(-\beta|\partial_e G|)$$

where $\partial_e G$ denotes the edge boundary of a graph G. Thus, when $\beta \to \infty$, it is reasonable to guess that $S(\beta)$ should minimise its edge boundary, rather than its vertex boundary, and hence be a square rather than a diamond.

Typical fluctuations of the shape (in the case where $\beta > \beta_0$ say) raise intriguing questions. In the percolation construction of the Wulff crystal when β is large but finite, these fluctuations are known with considerable precision. For instance, the so-called maximal local roughness, which measures the maximal distance from a point on the boundary of the shape to the polygonal hull of that shape, is of order $n^{1/3}$ if the diameter of the Wulff shape is of order n: see [30], [1]; more recently an extremely precise result in this direction was recently established by Hammond [21]. These exponents are identical to those arising in the Kardar–Parisi–Zhang (KPZ) universality class. While the Wulff crystal is not believed to belong to this universality class, there is nevertheless an analogy between the two situations, a fact which guides the intuition in the approach of [21].

Problem 3. *What are the maximum local roughness and facet length of the shape R_t when $t \to \infty$?*

We speculate that in two dimensions, the exponent for the fluctuations take the value in agreement with the Wulff crystal (and hence also with KPZ): this suggests that the maximum facet length is of order $t^{2/9}$ and the maximum local roughness of order $t^{1/9}$.

Finally, we end with a question on the random walk conditioned on the event $\mathcal{E}_t$.

Problem 4. *Suppose $0 < \beta < \beta_0$. What can be said about the random walk conditioned on $\mathcal{E}_t$?*

It would be interesting to know whether for some small values of β the behavior can be markedly different to the case where $\beta > \beta_0$. This appears to be a rather delicate question.

References

[1] K. Alexander. Cube-root boundary fluctuations for droplets in random cluster models. *Comm. Math. Phys.*, 224(3):733–781, 2001.

[2] K. Alexander, J.T. Chayes and L. Chayes. The Wulff construction and asymptotics of the finite cluster distribution for two-dimensional Bernoulli percolation. *Comm. Math. Phys.* 131, 1-50 (1990).

[3] A. Asselah and B. Schapira. Boundary of the range I: Typical behaviour. arXiv:1507.01031

[4] A. Asselah and B. Schapira. Boundary of the range II: Lower tails. arXiv:1601.03957

[5] M. Biskup and E. Procaccia. Eigenvalue vs perimeter in a shape theorem for self-interacting random walks. arXiv:1603.03817

[6] M. Biskup and E. B. Proccacia. Shapes of drums with lowest base frequency under non-isotropic perimeter constraints. arXiv:1603.03871

[7] T. Bodineau. The Wulff construction in three and more dimensions. *Commun. Math. Phys.*, 207(1):197–229, 1999.

[8] T. Bodineau, D. Ioffe, and Y. Velenik. Rigorous probabilistic analysis of equilibrium crystal shapes. *J. Math. Phys.,* 41(3):1033–1098, 2000.

[9] E. Bolthausen. Localization of a two-dimensional random walk with an attractive path interaction. *Ann. Probab.*, 22, 875–918 (1994).

[10] R. Cerf. The Wulff crystal in Ising and Percolation models. *Lecture Notes in Mathematics* 1878, Springer. (Ecole d'été de Probabilités de Saint-Flour 2004.)

[11] R. Cerf and Á. Pisztora. On the Wulff crystal in the Ising model. *Ann. Probab.,* 28(3):947–1017, 2000.

[12] P. Diaconis and L. Saloff-Coste. Comparison techniques for random walks on finite groups. *Ann. Probab.*, 21, 2131–2156 (1993).

[13] A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications.* Springer, Applications of Mathematics, Stochastic Modelling and Applied Probability, Second Edition.

[14] R.L. Dobrushin and O. Hryniv. Fluctuations of shapes of large areas under paths of random walks. *Probab. Theor. Rel. Fields.* 102 (No. 3) (1995) 313–330.

[15] R. L. Dobrushin, R. Kotecký, S. B. Shlosman. Wulff crystal: a global shape from local interaction. AMS, Translations Of Mathematical Monographs 104, Providence, Rhode Island, 1992.

[16] Donsker, M. D. and Varadhan, S. R. S. (1975). Asymptotic evaluation of certain Wiener integrals for large time. Proceedings of International Conference of Function Space Integration, Oxford, (1974), 15–33.

[17] Donsker, M. D. and Varadhan, S. R. S. (1975). Asymptotic evaluation of certain Markov process expectations for large time. I- IV *Comm. Pure Appl. Math.* 28 (1–47), (279–301). 29, (389–461). 36, (183–212).

[18] Donsker, M. D. and Varadhan, S. R. S., (1979). On the number of distinct points visited by a random walk. *Comm. Pure Appl. Math.* 32 721–747.

[19] Gärtner, J. and den Hollander, F. (1999). Correlation structure of intermittency in the parabolic Anderson model. *Probab. Theory Relat. Fields* 114, 1–54.

[20] Grimmett, G. *Percolation.* Grundlehren der mathematischen Wissenschaften, vol 321, Springer, 1999 (second edition).

[21] A. Hammond. Phase separation in random cluster models I: uniform upper bounds on local deviation. *Comm. Math. Phys.*, to appear.

[22] F. den Hollander. *Large deviations.* Fields institute monographs, American Mathematical Society.

[23] D. Ioffe and R. H. Schonmann. Dobrushin-Kotecký-Shlosman theorem up to the critical temperature. *Commun. Math. Phys.*, 199(1):117–167, 1998

[24] L. H. Loomis and H. Whitney. An inequality related to the isoperimetric inequality. Bull. Amer. Math. Soc. 55 (1949), 961-962.

[25] R. Lyons and Y. Peres. *Probability on trees and networks.* Book available on R. Lyons' website.

[26] Minlos, R. A. F., and Sinai, J. G. (1967). The phenomenon of phase separation at low temperatures in some lattice models of a gas. I. Sbornik: Mathematics, 2(3), 335–395.

[27] C.-E. Pfister. Large deviations and phase separation in the two-dimensional Ising model. *Helv. Phys. Acta* 64, n.7, 953–1054 (1991).

[28] L.C.G. Rogers and D. Williams, 2000. Diffusions, Markov processes and martingales: Volume 2, Itô calculus (Vol. 2). Cambridge university press.

[29] Saloff-Coste, L. *Lectures on finite Markov chains.* Lectures on probability theory and statistics (Saint-Flour, 1996), Lecture Notes in Math., vol. 1665, Springer, Berlin, 1997, pp. 301–413. MR1490046 (99b:60119).

[30] H. Uzun and K. Alexander. Lower bounds for boundary roughness for droplets in Bernoulli percolation. *Probab. Theory Related Fields*, 127(1):62–88, 2003.

[31] Wulff, G. Zur Frage der Geschwindigkeit des Wachstums und der Auflösung der Kristallflächen. *Zeitschrift fur Krystallographie und Mineralogie*, 34 (1901), 5/6, 449–530.