ANNALES
DE L'INSTITUT
HENRI
POINCARÉ

PROBABILITÉS
ET STATISTIQUES

Condensation of a self-attracting random walk

Nathanaël Berestycki[a,b,1] and Ariel Yadin[c]

[a]*Statistical Laboratory, Centre for Mathematical Sciences, University of Cambridge, Cambridge CB3 0WB, United Kingdom.*
E-mail: nberestycki@gmail.com
[b]*Universität Wien, Oskar Morgenstern Platz 1, Wien 1090, Austria*
[c]*Department of Mathematics, Ben Gurion University of the Negev, Beersheba, Israel. E-mail: ariel.yadin@gmail.com*

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 $(t/\beta)^{1/(d+1)}$. 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.

Résumé. Nous introduisons une mesure de Gibbs sur les chemins de longueur t dans le réseau Euclidien de dimension d, telle qu'un chemin donné est penalisé par un facteur proportionnel à la taille de sa frontière et l'inverse d'une température $\beta > 0$. Nous montrons qu'en dimension $d = 2$, la marche aléatoire se condense dans un ensemble de diamètre $(t/\beta)^{1/3}$ à une constante multiplicative près. En dimensions $d \geq 3$, nous montrons que la marche occupe un volume inférieur à $(t/\beta)^{d/(d+1)}$ et son diamètre est au moins $(t/\beta)^{1/(d+1)}$. Des résultats similaires sont obtenus pour une marche aléatoire conditionnée à avoir un temps local supérieur à β en chaque point visité, pourvu que β soit supérieur à une constante explicite qui en deux dimensions est égale au logarithme de la constante de connectivité.

MSC: 60K35; 60J27; 60F10

Keywords: Gibbs measure; Condensation; Self-attractive random walk; Wulff crystal; Large deviations; Donsker–Varadhan principle

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 where each edge has rate one started from x, and 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 $\mu = \mu_t$ 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

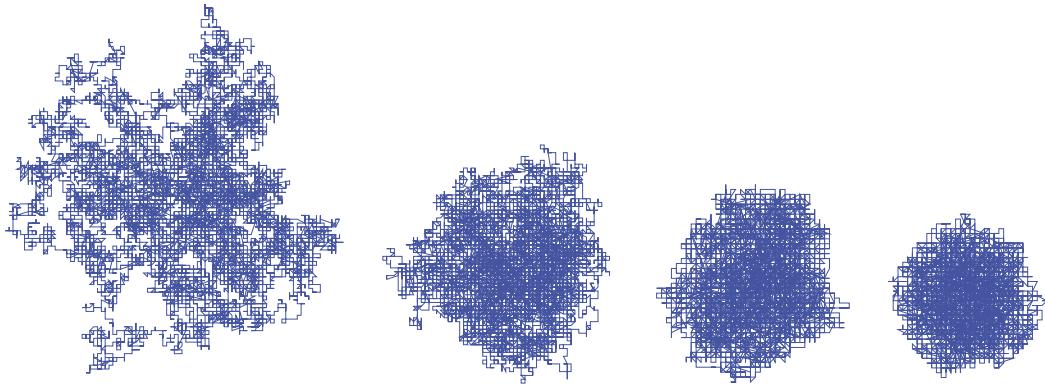$$H(\omega) = |\partial R_t|, \tag{1.1}$$

Fig. 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 $\beta = 2$.

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

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

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 (see Figure 1 for a simulation). Interpreting the random walk $(X_0, \ldots, X_t)$ as a chain of monomers, the Gibbs measure μ describes the law of a diluted polymer in a poor solvent.

We will be interested in describing the geometry of R_t. With a constraint on its maximal volume and a penalty in case of a large boundary, this problem is closely related to the question of *phase separation* in statistical mechanics. This is a classical topic with a long and distinguished history, for which we mention only a few major milestones. Traditionally, microscopic models for this phenomenon have been based on the framework of either percolation or the Ising model. Either way, a key goal is to prove a shape theorem for the cluster. Such a result can then be viewed as a microscopic justification for the Wulff construction ([33]), which is a method to determine the equilibrium shape of crystals based on surface energy minimisation. 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 slightly earlier in a celebrated work of Dobrushin, Kotecký and Shlosman [17] (for which preliminary announcements can be found in [16,25], as noted by an anonymous referee). This result was derived again by Pfister [29], see also the papers by Ioffe and Schonmann [24] extending the results of [17] 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 [28] 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.

As mentioned above, until now the question of phase separation has been studied rigorously mostly 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 [15] that a limiting Wulff shape arises.

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

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

While we are currently unable to derive a shape theorem, our main result suggests that the limit shape, if it exists, has diameter of order $t^{1/(d+1)}$.

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 have for all* β_0 *and for all* $\beta > \beta_0$,

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

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 $\beta(t)/t \to 0$ as $t \to \infty$.

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 3.8 for a precise statement. In dimensions $d \geq 3$, we conjecture that $\operatorname{diam}(R_t)$ 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 [18,19]. 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. 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. A related major difficulty is that two local time profiles can be macroscopically close in the ℓ^1 sense, say, even though the sizes of their boundaries are of widely different orders of magnitude.

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:

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

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 (3.19)), 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 a behaviour similar to Theorem 1.1 may be expected. 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, typical points are 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 (3.20) 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*

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

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 $\beta > \beta_1 > \beta_0$,*

$$\mathbb{P}\left(\mathrm{diam}(R_t) \geq c_1 \left(\frac{t}{\beta} \right)^{1/(d+1)} ; \ \text{and } |R_t| \leq c_2 \left(\frac{t}{\beta} \right)^{d/(d+1)} \,\middle|\, \mathcal{E}_t \right) \to 1 \tag{1.7}$$

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 $\beta(t)/t \to 0$ as $t \to \infty$.

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 $d\tilde{\mu}_t = (\tilde{Z})^{-1} \exp(-\beta \tilde{H}) \, d\mathbb{P}$ on $\mathcal{F}$.

Theorem 1.3. *Theorem 1.1 still holds true with $\tilde{\mu}_t$ instead of μ_t.*

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] discusses a 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 case of deterministic (nonrandom) edge weights, this limit shape is simply the unit square. This is the analogue of our conjecture here (see Section 4) that the limit shape is a diamond when $\beta \to \infty$. The difference between their square and our diamond comes from the difference between edge and vertex boundary in the formulation of the problem in [5,6].

1.3. *Main ideas in the proof*; *organisation of the paper*

Since the proof of Theorem 1.1 has many technical aspects, and involves plenty of careful computations, let us provide a sketch of the main ideas involved. Recall that we are trying to estimate the radius and volume of the trace of the random walk penalised by its 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 $\exp(-\text{const.} \times t/L^2)$, since the random walk has a probability of order 1 to escape S every L^2 units of time. In this case, 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 $L = (t/\beta)^{1/(d+1)}$, which is indeed the conjectured order of the diameter in all dimensions $d \geq 2$ (see Theorem 1.1).

The main issue in translating this rough heuristic to a rigorous argument is that the random walk could stay in a box of size L while having a boundary much larger 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}. From this point of view, the most delicate case appears to be $d = 2$; yet surprisingly this is where our results are also the most precise.

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 diameter 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) to the boundary of the box. We define a specific profile which with hindsight 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. This is one of the most technical parts of the paper, and is particularly delicate in the case $d = 2$ (reflecting the above mentioned difficulty).

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 (essentially the normalised squared principal eigenfunction) was not good enough for constant β because of how it behaves near the boundary. We discuss the required properties of this profile in the beginning of Section 2, adjacent to the definition of π. (For example, the polylogarithmic terms appearing in the definition of π are essential for the analysis to work, but would be absent in the naive choice of π.)

2. Quantitative estimates for lower bound on Z

2.1. *Change of measure*

As mentioned above, a main technical part of this paper consists in deriving good lower bounds on the partition function Z, which hold 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 relatively standard technique in large deviations (see, e.g., Bolthausen's article [9] as well as

the work of Gärtner and den Hollander [21] on intermittency of parabolic Anderson model; see [12] and [13] for a general introduction to large deviations theory). However, the precise change of measure which needs to be performed here is much more delicate than usual. The analysis of the titled walk in particular will require a host of tools from the quantitative theory of Markov chains: we will need very precise information about how the tilted 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

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

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 least integer greater than $(t/\beta)^{1/(d+1)}$, 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 extremely 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, so $\pi(x) \asymp 1/L^d$ near the centre.
- 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, so $\pi(x) \asymp 1/t$ near the boundary, i.e., $\pi(x) \asymp 1/L^{d+1}$. 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 $\exp(-c\beta L^{d-1})$.

In view of the above requirements it might be natural to take $\pi(x) \asymp \text{dist}(x, \partial S)/L^{d+1}$, i.e., increases linearly with the distance to the boundary of S. While this clearly fulfils the first and second point, it turns out that the Dirichlet energy of $\sqrt{\pi}$ (which ends up governing how likely it is to achieve π) is too high by a logarithmic factor. Instead, the specific choice of π is as follows. For $0 \leq r \leq L$, let

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

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

$$\pi(x) = C_{(2.2)} \mu_r \quad \text{where } \mu_r := \begin{cases} \frac{r+1}{L^{d+1}(\log(r+2))^2} & \text{if } r \leq L/2, \\ \big(\sqrt{\mu_{L/2}} + \frac{(r-L/2)}{L^{(d+2)/2}} \big)^2 & \text{if } r \geq L/2. \end{cases} \tag{2.2}$$

Let $\pi(z) = 0$ for $z \notin S$, and the constant $C_{(2.2)}$ is chosen so that $\sum_z \pi(z) = 1$. It can then be checked that $C_{(2.2)}$ 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 [14]. We use the standard canonical paths on $\mathbb{Z}^d$: that is, for $x, y \in S$, we define the path $\gamma_{x,y}$ 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.*

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

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 $e \in \gamma_{x,y}$, 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 $\gamma_{x,y} \subset \prod_i [m_i, M_i] \subset S$, and x, y are two corners of this hypercube. So any point on $\gamma_{x,y}$ 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 the number of pairs of points $x, y \in S$ such that $e \in \gamma_{x,y}$ is at most $O(L^{d+1})$. Indeed, suppose that the two endpoints of e differ only in the ith coordinate with $1 \leq i \leq d$. Then the coordinates $1, \ldots, i$ of x can be chosen arbitrarily among $O(L)$ possibilities (while the remaining coordinates are fixed and imposed by those of either endpoint of e). Conversely, the coordinates $i, \ldots, d$ can be chosen arbitrarily among $O(L)$ choices for y, and the remaining coordinates are fixed and imposed by those of either endpoint of e. Consequently, the total number of choices for x and y such that $e \in \gamma_{x,y}$ is at most $O(L^i) \times O(L^{d-i+1}) = O(L^{d+1})$.

Therefore, using the facts that $\pi(y) \leq O(1/L^d)$ and $|\gamma_{x,y}| = O(L)$,

$$\frac{1}{q(e)} \sum_{x,y:e \in \gamma_{x,y}} |\gamma_{x,y}| \pi(x)\pi(y) \leq CL \frac{1}{L^d} \#\{x, y \in S : e \in \gamma_{x,y}\} \leq CL^2$$

as desired. $\square$

By Theorem 3.2.1 in [31], it follows that if gap is the spectral gap of the Markov chain, then $\mathsf{gap} \geq 1/(C_{2.1}L^2)$. (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,

$$\left| \mathbb{Q}_x(X_t = y) - \pi(y) \right| \leq \sqrt{\pi(y)/\pi(x)} e^{-\mathsf{gap}\, t}.$$

(See, *e.g.*, the proof of Corollary 2.1.5 of [31]). Let

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

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

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

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,

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

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 turn 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. *Uniformly over all $x, y \in S$, for some positive 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 $\kappa = \pi(y)^{-1} \vee L^2 \log L$ 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 $y \in S$,

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

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

We start the proof of Proposition 2.2 with a 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.
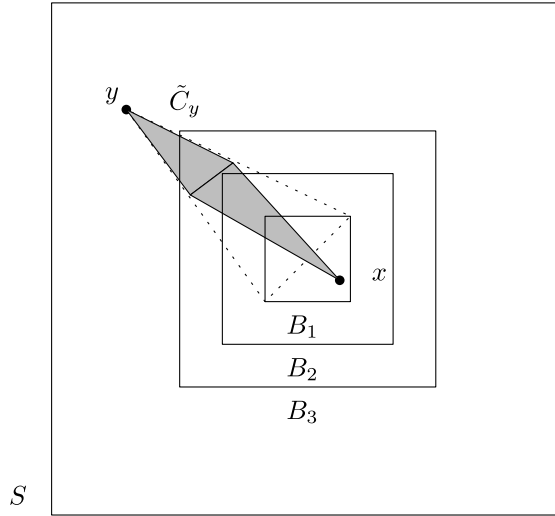
Fig. 2. The shaded area is C, the union of the two cones $\tilde{C}_y$ and $\tilde{C}_x$ in the proof of Lemma 2.4.

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

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

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

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

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 with respect to π. 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 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 [27, 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 [27, Chapter 2] (see, e.g., (2.17)). Let U be a random variable uniformly distributed on the base of the cone, Σ_M. Let R be the union of the Euclidean segments $[y, U]$ and $[U, x]$. Given U, let Γ be some choice of a monotone path in S that stays as close as possible from the two segments forming R, starting at y and ending at x. By monotone we mean that each coordinate changes monotonically along Γ. (Γ is thus a random monotone path connecting x and y through Σ_M, which stays at distance at most $\sqrt{d}$ from R; the exact way of choosing Γ given U does not matter.) Because Γ is chosen to be monotone, it traverses any edge at most once. So the function $\theta(e) = \mathbb{P}[e \in \Gamma] - \mathbb{P}[\hat{e} \in \Gamma]$ (where $\hat{e}$ is the directed edge e in the reverse direction) defines a unit flow in C, from y to x. Indeed θ can be viewed as the expectation of a random variable which itself defines a unit flow almost surely. (Again, see [27, Chapters 2.4 & 2.5], and especially (2.17).) Moreover, 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 $O(k^{d-1})$.

Thompson's principle ([27, 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

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

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. $\qquad\square$

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

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

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

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

$$\mathbb{E}_x\big[L(\tau_{xy}, y) \mid L(\tau_{xy}, y) > 0\big] \leq \mathbb{E}_y\big[L(\tau_{xy}, y)\big] \leq \mathbb{E}_y\big[L(T_x, y)\big] + \mathbb{E}_x\big[L(\tau_{xy}, y)\big]$$

$$\leq \begin{cases} C \log(r+2) + 1 & \text{if } d = 2, \\ C + 1, & \text{if } d \geq 3, \end{cases}$$

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 $\mathbb{Q}_x[T_y < \tau] \geq 1/(C \log(r+2) + 1)$ for $d = 2$, as desired. $\qquad\square$

We are now able to complete the proof of Proposition 2.2.

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

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

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 $d \geq 3$. Uniformly in $x, y \in S$, for any $\delta > 2/(\pi(y)q_y t)$,*

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

(ii) *Assume $d = 2$. Uniformly in $x, y \in S$, for any $\delta > 2/(\pi(y)q_y t)$,*

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

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

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

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

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

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

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

$$\mathbb{E}_x\big[(T_j - \tilde{T}_j)^2\big] = \mathbb{E}_y\big[(T^+)^2\big] = \int_0^{\infty} 2t\, \mathbb{P}_y\big[T^+ > t\big]\, dt$$

$$\leq \int_0^{C_{2.2}\theta} 2t\, dt + \int_{C_{2.2}\theta}^{\infty} 2t e^{-t/\theta}\, dt$$

$$\leq A := \big((C_{2.2})^2 + 2\big) \cdot \theta^2.$$

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

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

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 $\alpha > 0$,

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

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

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

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

$$\mathbb{Q}_x(T_n \leq t) \leq \exp\left(-\frac{(n\, \mathbb{E}_y[T^+] - t)^2}{4nA}\right) \leq \exp\left(-\frac{\varepsilon^2 t^2}{4nA}\right)$$

$$\leq \exp\left(-\frac{\varepsilon^2}{4(2 + (C_{2.2})^2)} \cdot \frac{t^2}{n\theta^2}\right). \tag{2.14}$$

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 $\eta > 0$,

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

Thus, if $\varepsilon > 0$ satisfies

$$\pi(y)q_y t(1 + \delta) \geq n(1 + \varepsilon) \tag{2.15}$$

then

$$\mathbb{Q}_x\big(L(T_n, y) \geq \pi(y)t(1+\delta)\big) \leq \mathbb{Q}_x\left(L(T_n, y) \geq n\frac{1+\varepsilon}{q_y}\right) \leq \exp\big(-n\big(\varepsilon - \log(1+\varepsilon)\big)\big). \tag{2.16}$$

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 satisfies the constraint (2.15). Since $n \leq \pi(y)q_y t(1 + \varepsilon) + 1$ and $\delta \geq 2/(\pi(y)q_y t)$ by assumption on δ in the theorem, this is the case as soon as

$$\big(2\varepsilon + \varepsilon^2\big) \leq \frac{\delta/2}{1 + 1/(\pi(y)q_y t)}.$$

Note that $\pi(y)q_y t \geq \beta_0 C_{(2.2)}/2$ as $q_y \to 1$ uniformly and $t \geq \beta L^{d+1}$. Hence we can choose

$$\varepsilon \asymp \min(\delta, \sqrt{\delta})$$

so that (2.15) is satisfied, where the implied constants depend on β_0 and d only. Combining (2.14) and (2.16) we arrive at the conclusion of the lemma, since

$$\mathbb{Q}_x\big(L(t, y) \geq \pi(y)t(1+\delta)\big) \leq \mathbb{Q}_x(T_n \leq t) + \mathbb{Q}_x\big(L(T_n, y) \geq \pi(y)t(1+\delta)\big)$$

$$\leq \exp\left(-c\frac{\varepsilon^2}{1+\varepsilon}\big(\log(r+2)\big)^{-2}\pi(y)t\right) + \exp\big(-c\varepsilon^2\pi(y)t\big)$$

$$\leq 2\exp\big(-c\big(\varepsilon^2 \wedge \varepsilon\big) \cdot \big(\log(r+2)\big)^{-2} \cdot \pi(y)t\big), \tag{2.17}$$

where the $(\log(r + 2))^{-2}$ term can be removed when $d \geq 3$ and $c > 0$ is a constant depending on $C_{(2.2)}$, $C_{2.2}$, $c_{2.2}$, β_0 and d. (We have used (2.16) to get (2.17) and the fact that $(\varepsilon - \log(1 + \varepsilon)) \asymp \varepsilon^2 \wedge \varepsilon$ as well as $n \leq (1 + \varepsilon)\pi(y)t$.) The lemma now follows since $\varepsilon \asymp \min(\delta, \sqrt{\delta})$, so that $\varepsilon^2 \wedge \varepsilon \asymp \min(\sqrt{\delta}, \delta, \delta^2) \asymp \min(\sqrt{\delta}, \delta^2)$. $\square$

Remark 2.7. A similar statement to Lemma 2.6 holds with the upper bound replaced by a lower bound: 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. *For all $\beta_0 > 0$ there exists a constant $c_{2.8} > 0$ depending only on d and β_0, such that the following holds. For any integer $k > 0$ and any $x, y \in S$,*

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

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 facts that $q_y \to 1$ uniformly over S, and that $t = \beta L^{d+1}$ so that $\pi(y)t \geq C_{(2.2)}\beta_0 r(\log(r + 2))^{-2}$. $\square$

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:

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

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

$$D_r = \left\{z \in S_r : \langle z \rangle > 1\right\}. \tag{3.1}$$

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

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

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:

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

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

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

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 $c_{(3.2)}$ large enough (which will be chosen later) we define the event

$$\mathcal{B} := \left\{|\partial R_t| \leq c_{(3.2)} L^{d-1}\right\}, \tag{3.2}$$

and finally, we define the good event $\mathcal{G}$ as follows:

$$\mathcal{G} = \mathcal{B} \cap \mathcal{S} \cap \mathcal{D}. \tag{3.3}$$

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 $\mathbb{Q}_x$,

$$\mathbb{E}_x |S_{k,r}| \leq |S_r| \max_{y \in S_r} \mathbb{Q}_x(y \in \mathcal{X}_k) \leq |S_r| \max_{y \in S_r} \mathbb{Q}_x \big(L(t, y) \geq 2^k \pi(y) t \big)$$

$$\leq |S_r| \exp\big(-2c_{3.2} 2^{k/2} r \big(\log(r+2) \big)^{-4} \big).$$

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

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

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

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 $|\partial R_t| \leq 2d |S \setminus R_t| + C L^{d-1}$, 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),

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

as desired. $\qquad\square$

We deduce from Lemma 3.3 and Markov's inequality that

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

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

$$\mathbb{Q}_x(\mathcal{G}) \geq 1/4. \tag{3.4}$$

3.2. *Radon–Nikodym derivative estimates*

The following lemma is well known but very useful, see e.g. [30], 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*

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

Proof. This follows easily from a discrete Feynman–Kac representation (see e.g. Lemma 11 in [21]). 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 $\tau_1 + \cdots + \tau_n = t$.) 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 $\mathbb{P}$,

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

The result follows immediately. $\qquad\square$

Lemma 3.5. *Recall the events* $\mathcal{S}, \mathcal{D}$ *defined above* (3.3). *On the event* $\mathcal{S} \cap \mathcal{D}$ *we have*

$$
\int_0^t \frac{\Delta f}{f}(X_s)\, ds \geq -c_{3.5} t L^{-2},
$$

where $c_{3.5} > 0$ *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. The terms $x \in S_r \setminus D_r$, $r \neq L/2$ are the "main terms" and all the other terms ($r = L/2$ or $x \in D_r$) are a kind of error which we need to estimate.

Step 1: contribution of main terms. We will show that

$$
\sum_{r \neq L/2} \sum_{x \in S_r \setminus D_r} \frac{\Delta f(x)}{f(x)} L(t, x) \geq -c \cdot \frac{t}{L^2}. \tag{3.5}
$$

Note that $\Delta f(x)/f(x)$ does not change if we mutiply f by a nonzero constant. Hence for this calculation we may take $C_{(2.2)} = 1$ in the definition of $\pi(x)$. Thus we have for $x \in S_r$

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

A second order Taylor expansion provides the following estimate for all $x \in S_r$, $y \in S_{r+1}$ with $r + 1 \leq L/2$:

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

Now, for $x \in S_r \setminus D_r$ with $r + 1 \leq L/2$, all neighbours are in S_r (and hence do no contribute to the Laplacian) except for one in S_{r+1} and one in S_{r-1}. Hence, for some $\xi \in [r - 1, r]$,

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

for some constant $c > 0$. For $r > L/2$, we have that if $x \in S_r \setminus D_r$, then $\Delta f(x) = 0$ since f is affine in this range. Hence let us estimate the contribution to the Radon–Nikodym derivative (3.5) 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 $L(t,x) \leq 2^{k+1}\pi(x)t \leq c2^{k+1}(t/L^{d+1})\varphi(r)$. So, on the event $\mathcal{S}_{r,k}$,

$$\sum_{x \in S_{r,k} \setminus D_r} \frac{\Delta f(x)}{f(x)} L(t,x) \geq -c|S_{r,k} \setminus D_r| \cdot 2^{k+1} \frac{t}{L^{d+1}} \varphi(r) r^{-2}$$

$$\geq -c|S_r| \cdot 2^{k+1} \frac{t}{L^{d+1}} \varphi(r) r^{-2} \cdot \exp\left(-c_{3.2} 2^{k/2} r \left(\log(r+2)\right)^{-4}\right).$$

Summing over k and since $|S_r| \leq L^{d-1}$, we obtain that on the event $\mathcal{S}$,

$$\sum_{x \in S_r \setminus D_r} \frac{\Delta f(x)}{f(x)} L(t,x) \geq -tL^{-2} \cdot \varphi(r) r^{-2} \cdot \left(2^{k_{3.2}+1} + \sum_{k \geq k_{3.2}} 2^{k+1} e^{-c_{3.2} 2^{k/2} r (\log(r+2))^{-4}}\right)$$

$$\geq -ctL^{-2}\varphi(r) r^{-2},$$

where the final constant $c > 0$ depends on $k_{3.2}$. Hence, summing over $r \neq L/2$, the contribution of the main terms to the Radon–Nikodym derivative is

$$\sum_{r \neq L/2} \sum_{x \in S_r \setminus D_r} \frac{\Delta f(x)}{f(x)} L(t,x) \geq -ctL^{-2}$$

as desired in (3.5), because

$$\sum_{r=1}^{L} \varphi(r) r^{-2} \leq \sum_{1 \leq r \leq L/2} \frac{1}{r(\log(r+2))^2} + \sum_{L/2 < r \leq L} \frac{1}{r} \leq C. \tag{3.6}$$

Step 2: contribution of $\mathcal{D}_r, r < L/2$. 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} (here recall that $\langle x \rangle = \#\{1 \leq j \leq d : |x_j| = \|x\|_\infty\}$ is the number of coordinates which achieve the sup norm of x, as defined in (3.1)). So,

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

Hence noting that $|D_r| = O(L^{d-2})$, on the event $\mathcal{D}$ the contribution to (3.5) coming from $D_r, r < L/2$, is:

$$\sum_{x \in D_r} \frac{\Delta f(x)}{f(x)} L(t,x) \geq -c|D_r| \cdot \frac{t}{L^{d+1}} \varphi(r) r^{-1} \cdot \left(2^{k_{3.2}+1} + \sum_{k \geq k_{3.2}} 2^{k+1} e^{-c_{3.2} 2^{k/2} r (\log(r+2))^{-4}}\right)$$

$$\geq -cL^{d-2} \cdot \frac{t}{L^{d+1}} \varphi(r) r^{-1} \geq -ctL^{-3}\varphi(r) r^{-1}$$

and thus summing over $r < L/2$ we get

$$\sum_{r < L/2} \sum_{x \in D_r} \frac{\Delta f(x)}{f(x)} L(t,x) \geq -c\frac{t}{L^2(\log L)^2}. \tag{3.7}$$

Step 3: contribution of $\mathcal{D}_r$ with $r \geq L/2$. If $x \in D_r$ for $r \geq L/2$ then

$$\frac{\Delta f(x)}{f(x)} = -\langle x \rangle \cdot \frac{1}{L^{(d+2)/2} f(x)} \geq -c \frac{1}{L^{d/2+1} f(x)}.$$

Thus for $r \geq L/2$ and $k \geq 1$ we have, on $\mathcal{D}_{r,k}$:

$$\sum_{x \in D_{r,k}} \frac{\Delta f(x)}{f(x)} L(t,x) \geq -\frac{1/L^{d/2+1}}{f(x)} 2^{k+1} t f(x)^2 \times |D_r| e^{-c_{3.2} 2^{k/2} r (\log(r+2))^{-4}}$$

$$\geq -\frac{t}{L^{d/2+1}} L^{d-2} \left(\sqrt{\mu_{L/2}} + \frac{r - L/2}{L^{d/2+1}} \right) \times e^{-c_{3.2} 2^{k/2} r (\log(r+2))^{-4}}$$

so that summing over k, on $\mathcal{D}$, reasoning as above,

$$\sum_{x \in D_r} \frac{\Delta f(x)}{f(x)} L(t,x) \geq -c \frac{t}{L^{d/2+1}} L^{d-2} \left(\frac{1}{L^{d/2} \log L} + \frac{r - L/2}{L^{d/2+1}} \right)$$

and then summing over $r \geq L/2$:

$$\sum_{r \geq L/2} \sum_{x \in D_r} \frac{\Delta f(x)}{f(x)} L(t,x) \geq -c \left(\frac{t}{L^2 \log L} + \frac{t}{L^2} \right) \geq -c \frac{t}{L^2}. \tag{3.8}$$

Step 4. The contribution coming from $r = L/2$ is estimated in a similar way: for any $x \in S_{L/2}$,

$$\frac{\Delta f(x)}{f(x)} \geq -c \frac{\log L}{L}$$

and hence for the same reason as above, on the good event $\mathcal{S}$,

$$\sum_{x \in S_{L/2}} \frac{\Delta f(x)}{f(x)} L(t,x) \geq -c L^{d-1} \frac{\log L}{L} \times t \frac{1}{L^d (\log L)^2}$$

$$\geq -c \frac{t}{L^2 \log L}. \tag{3.9}$$

Conclusion

Combining the results of all four steps above ((3.5), (3.7), (3.8), and (3.9)), we deduce

$$\sum_{x \in S} \frac{\Delta f(x)}{f(x)} L(t,x) \geq -c \frac{t}{L^2}.$$

This concludes the proof of Lemma 3.5. $\qquad\square$

Remark 3.6. Note that it is in (3.6) that we see the importance of the logarithmic correction terms in the choice of the local time profile $\pi(x)$ in (2.2).

With this lemma it is now easy to conclude 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 (3.4), and the fact that $\pi(x) \leq \pi(0)$ for any $x \in S$, we obtain:

$$Z(t, \beta) = \mathbb{E}_0 \big[\exp \big(-\beta |\partial R_t| \big) \big] \geq \mathbb{E}_0 \big[1_{\mathcal{G}} \exp \big(-\beta |\partial R_t| \big) \big]$$

$$\geq \exp\!\left(-\beta c_{(3.2)} L^{d-1}\right) \mathbb{P}_0(\mathcal{G}) = \exp\!\left(-\beta c_{(3.2)} L^{d-1}\right) \mathbb{Q}_0\!\left(1_{\mathcal{G}}\, \frac{d\,\mathbb{P}_0}{d\mathbb{Q}_0}\right)$$

$$\geq \frac{1}{4} \exp\!\left(-\left(\beta c_{(3.2)} L^{d-1} + c_{3.5} t L^{-2}\right)\right).$$

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

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

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

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 3.7. *Let $A \subset \mathbb{Z}^d$ be a finite, connected set with $|A| \geq 2$.*

(i) *Assume $d = 2$, and let R be the smallest rectangle in $\mathbb{Z}^2$ containing A (i.e., R is the intersection of all rectangles containing A). Then,*

$$\left|\partial^* R\right| \leq 3\left|\partial^* A\right|. \tag{3.10}$$

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

$$\left|\partial^* A\right| \geq \frac{2d}{2d-1} |A|^{\frac{d-1}{d}}. \tag{3.11}$$

Proof. For any connected set A such that $2 \leq |A| < \infty$, we have that

$$\left|\partial^* A\right| \leq \left|\partial_e^* A\right| \leq (2d-1)\left|\partial^* A\right|. \tag{3.12}$$

Indeed, for the first inequality simply note that the map which associates to an edge $e \in \partial_e^* A$ the endpoint of e which belongs to A is a map from $\partial_e^* A$ to $\partial^* A$ which is clearly onto. This proves the first inequality. Moreover, any $x \in \partial^* A$ has at most $2d - 1$ pre-images in this map (since for any $x \in \partial^* A$ then there are at most $2d - 1$ edges in $\partial_e^* A$ connected to x, and at least one other edge must connect x to the rest of A, as A is connected). This proves the second inequality and thus (3.12).

Consider the case $d = 2$. We claim that $|\partial_e^* R| \leq |\partial_e^* A|$ (then we will see that (3.10) follows directly from (3.12)). 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 $x \in A$, we also have $x \in R$. Thus, there exists a (necessarily unique) $k \geq 0$ such that $y = x + ke \in R$ and $x + ze \notin R$ for all $z > k$. Thus, $(x + ke, x + (k+1)e) \in \partial_e^* R$. Hence we can define a map $\phi : \partial_e^* A \to \partial_e^* R$ by setting:

$$\phi\big((x, x + e)\big) = (y, y + e).$$

In words, we start from $(x, x + e)$ and travel in the direction e until we leave R. This defines an edge in the outer edge boundary of R.

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 $L \cap A \neq \varnothing$, 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 $\partial_e^* R$, and hence $|\partial_e^* R| \leq |\partial_e^* A|$. Therefore, by (3.12), $|\partial^* R| \leq |\partial_e^* R| \leq |\partial_e^* A| \leq 3|\partial^* A|$, which proves (3.10).

For the general case $d \geq 3$ we use the discrete Loomis–Whitney inequality (Theorem 2 in [26]), which states that if A_i is the projection of A onto $\mathbb{Z}^{d-1}$ along the ith coordinate then

$$|A|^{d-1} \leq \prod_{i=1}^d |A_i|. \tag{3.13}$$

For each $1 \leq i \leq d$ and 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 vertex (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 two edges in $\partial_e^* A$. Note that for two distinct vertices z and w the corresponding edges will be pairwise distinct. Hence $|A_i| \leq |\partial_e^* A|/2$ for each $1 \leq i \leq d$. We deduce, using the arithmetic geometric inequality and (3.12),

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

Combining with (3.13) this gives the desired result. $\qquad\square$

3.4. *Proof of condensation*

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

Theorem 3.8. *Let $d \geq 2$. Fix $\beta_0 > 0$ and let $\beta > \beta_0$. Let γ be as in Proposition 3.1. Then,*

$$\mu_t\left[\operatorname{diam}(R_t) \geq \frac{1}{\sqrt{2\gamma}} \left(\frac{t}{\beta} \right)^{1/(d+1)} \right] \geq 1 - C \exp\left(-\gamma t^{1-2/(d+1)} \beta^{2/(d+1)} \right), \tag{3.14}$$

and if $d = 2$ then

$$\mu_t\left[\operatorname{diam}(R_t) \leq 6\gamma \left(\frac{t}{\beta} \right)^{1/3} \right] \geq 1 - C \exp\left(-\gamma t^{1/3} \beta^{2/3} \right).$$

Moreover, for all $d \geq 2$,

$$\mu_t\left[|R_t| \leq (2\gamma)^{d/(d-1)} \left(\frac{t}{\beta} \right)^{d/(d+1)} \right] \geq 1 - C \exp\left(-\gamma t^{1-2/(d+1)} \beta^{2/(d+1)} \right).$$

Proof. Recall that by Proposition 3.1

$$Z(t, \beta) \geq \exp\left(-\gamma t^{1-2/(d+1)} \beta^{2/(d+1)} \right). \tag{3.15}$$

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 $1 \leq i \leq d$, let J_t^i denote the length of the projection of $R_t^\square$ (or equivalently R_t) onto the ith coordinate axis.

Lemma 3.9. *We have*

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

Proof. Under $\mathbb{P}_0$, the coordinates $X_t^1, \ldots, X_t^d$ are independent continuous time (with rate 2) 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 $x \in \{1, \ldots, J\}$ and $T = \inf\{t \geq 0 : X_t \notin [1, J-1]\}$. 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 $\lambda = 2(1 - \cos(\pi/J))$. 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

$$\sin(\pi x/J) = \mathbb{E}_x\big(e^{\lambda t} \sin(\pi X_{t \wedge T}/J)\big)$$

$$\geq e^{\lambda t} \frac{2}{J} \mathbb{P}_x(T > t).$$

Therefore,

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

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

We now deduce from Lemma 3.9 a lower bound on the diameter of R_t under μ_t. Let $J_t^1, \ldots, J_t^d$ be the side-lengths of $R_t^\square$. Let $N = \min\{J_t^1, \ldots, J_t^d\}$. We will prove the stronger statement that $N \geq c(t/\beta)^{1/(d+1)}$ with high probability. By Lemma 3.9 and Proposition 3.1, for an integer $n > 0$,

$$\mu_t\big[J_t^1 \leq n\big] \leq Z(t, \beta)^{-1} nC \exp\left(-\frac{\pi^2 t}{4dn^2}\right)$$

$$\leq Cn \exp\left(\gamma t^{1-2/(d+1)} \beta^{2/(d+1)} - \frac{\pi^2}{2n^2}t\right).$$

Thus, if $n = (1/\sqrt{2\gamma})(t/\beta)^{1/(d+1)}$ and since $\pi^2 \geq 2$, we get that

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

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

$$\mu_t[N \leq n] \leq O(t/\beta)^{d+1} \exp\big(-\gamma t^{1-2/(d+1)} \beta^{2/(d+1)}\big). \tag{3.16}$$

In particular, it holds that with high μ_t-probability

$$\mathrm{diam}(R_t) \geq \frac{1}{\sqrt{2\gamma}}(t/\beta)^{1/(d+1)}.$$

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 we will see that it automatically follows (by Lemma 3.7) 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.]

More precisely, recall that $R_t^\square$ is a $J_t^1 \times J_t^2$ rectangle. Lemma 3.7 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,

$$\mu_t\big[J_t^i > m\big] \leq Z(t,\beta)^{-1} \sum_{k=m+1}^{\infty} \exp\Big(-\beta \frac{2}{3}k\Big)$$

$$\leq C \exp\Big(-\beta \frac{2}{3}m + \gamma t^{1/3}\beta^{2/3}\Big).$$

If $m = 3\gamma(t/\beta)^{1/3}$ this probability is at most $C \exp(-\gamma t^{1/3}\beta^{2/3})$. 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 3.8.

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 3.7, if $|R_t| \geq m$ then $|\partial R_t| \geq |\partial^* R_t| \geq (2d/(2d-1))m^{(d-1)/d} \geq m^{(d-1)/d}$, and so almost surely on this event, $\exp(-\beta H(\omega)) \leq \exp(-\beta m^{(d-1)/d})$. Consequently,

$$\mu_t\big[|R_t| \geq m\big] \leq Z(t,\beta)^{-1} \exp\big(-\beta m^{(d-1)/d}\big) \cdot \mathbb{P}\big[|R_t| \geq m\big]$$

$$\leq \exp\big(-\beta m^{(d-1)/d} + \gamma t^{1-2/(d+1)}\beta^{2/(d+1)}\big).$$

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 $\exp(-\gamma \times t^{1-2/(d+1)}\beta^{2/(d+1)})$. This completes the upper bound on the volume in all dimensions and thus the proof of the theorem. $\square$

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 $S' = \bigcup_{r>K} S_r$. 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 (3.4) show that $\mathbb{Q}(\mathcal{G}_t') \geq 1/4$, provided that K is a sufficiently large constant. The only difference with (3.4) 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

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

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

$$\mathbb{P}\big(\mathcal{E}_{2t}\big) \geq \exp\big(-\gamma t^{1-2/(d+1)}\beta^{2/(d+1)}\big), \tag{3.17}$$

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

$$\tilde{Z}(t, \beta) \geq \exp\left(-\gamma t^{1-2/(d+1)} \beta^{2/(d+1)}\right). \tag{3.18}$$

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

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

$$\mathbb{P}(R_t = G, \mathcal{E}_t) \leq \exp\left(-\beta|\partial G|\right). \tag{3.19}$$

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

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

by independence of the random variables E_{xy}. Thus (3.19) is established.

Putting together (3.17) and (3.19) (resp. (3.18) and the definition of $\tilde{\mu}$), the proof of Theorem 1.2 (resp. Theorem 1.3) proceeds essentially as in Theorem 3.8. More precisely, let $J_t^1, \ldots, J_t^d$ be the dimensions of R_t in each coordinate. The lower bound in (3.17) implies exactly as in (3.16) that

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

as $t \to \infty$, where $n = (1/\sqrt{2\gamma})(t/\beta)^{1/(d+1)}$. In particular, conditioned on $\mathcal{E}_t$, with high probability we have $\text{diam}(R_t) \geq n$.

For the upper-bound on $\text{diam}(R_t)$ 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 3.10. 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

$$\alpha = \alpha(d) = \limsup_{n \to \infty} c_n^{1/n}, \qquad \beta_0 = \log \alpha. \tag{3.20}$$

Note that when $d = 2$, the limsup is a limit and is (essentially by definition) equal to the connective constant of $\mathbb{Z}^2$ [22]. 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 $y \in \text{Ext}(G)$. Let $\mathcal{S}_{j_1,\ldots,j_d}$ denote the set of surfaces where the diameter in each direction $1, \ldots, d$, does not exceed $j_1, \ldots, j_d$ 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 (3.19) (since each plaquette corresponds to an edge (x, y) such that the corresponding exponential random variable E_{xy} satisfies $E_{xy} > \beta$, and these events are independent even for edges which share vertices)

$$\mathbb{P}(\Sigma = \sigma, \mathcal{E}_t) \leq \exp(-\beta j). \tag{3.21}$$

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 $|\mathcal{S}_n| \leq \exp(\beta_1' n)$.

Therefore, by (3.21),

$$\mathbb{P}[\Sigma \in \mathcal{S}_n \mid \mathcal{E}_t] \le \mathbb{P}(\mathcal{E}_t)^{-1} \sum_{j=n}^{\infty} e^{\beta_1' j} e^{-\beta j}$$

$$\le C \exp\{\gamma t^{1-2/(d+1)} \beta^{2/(d+1)} - (\beta - \beta_1')n\},$$

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

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

Hence $|\partial^* R_t^{\square}| \le n$ with high conditional probability given $\mathcal{E}_t$, and thus (by Lemma 3.7)

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

with high conditional probability, as desired.

Remark 3.11. 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 finish the paper with a brief discussion of some open problems raised by our results.

Limit shape theorem

The most basic question 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 $d \ge 2$. We make the following more precise conjecture:

Conjecture. *There exists a nonrandom closed, bounded and convex set $S = S(\beta) \subset \mathbb{R}^d$ such that*

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

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

An equivalent way of stating the conjecture is that there exists a deterministic S (compact and convex) such that if we translate the range R_t to have a centre of mass at the origin, then the resulting set is close to S with high probability in the Hausdorff sense. This is similar to the situation in [9].

Once the existence of S is established one may ask numerous questions about its geometry. For instance, does it have any (macroscopic) flat facet?

Studying the extreme cases $\beta \to \infty$ and $\beta \to 0$ should also be interesting. Further to the above open problem, we conjecture that as $\beta \to \infty$, $S(\beta)/\mathrm{diam}(S(\beta))$ converges in the Hausdorff sense to a diamond of unit diameter. This is because 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 $k = 2n(n+1)$. Since this conjecture was first made, a very closely related result has been proved by Biskup and Procaccia [5,6]. At the other extreme, as $\beta \to 0$ it is natural to believe

that the lattice effects become less and less relevant, so that the limit shape becomes rotationally invariant. Thus we conjecture that $S(\beta)/\mathrm{diam}(S(\beta))$ converges as $\beta \to 0$ in the Hausdorff sense to a ball of unit diameter. This seems intuitively related to the result of Duminil-Copin on the limit of the Wulff crystal for percolation as $p \to p_c$ on the triangular lattice ([20]).

We make similar conjectures for the case of a random walk conditioned on $\{L_t(x) \geq \beta, \forall x \in R_t\}$. However, in the case $\beta \to \infty$ we believe that the limit should be a square with unit diameter instead of a diamond. This is because by (3.19)

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

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. As we do not yet know whether the behaviour described in Theorem 1.2 persists for $\beta \leq \beta_0$, we do not make any conjecture for the case $\beta \to 0$.

Fluctuations

The question of the roughness of the boundary of the shape is of considerable interest. In the case of two-dimensional percolation, these fluctuations are known with considerable precision. For instance (see [32] and [1]), the *maximal local roughness*, which measures the maximal distance from a point on the boundary of the shape to the polygonal hull of that shape (and hence the size of inward deviations), is of order (diameter)$^{1/3}$. More recently, Hammond [23] established an extremely precise result in this direction which gives a sharp logarithmic power-law correction (stated in the greater generality of the q-state Potts model with $q \geq 1$). This exponent and related ones are common to a large class of two-dimensional interfaces, including the KPZ (Kardar–Parisi–Zhang) universality class. We conjecture that this is the case here as well; and since the diameter itself is of order $t^{1/3}$, this leads us to the following:

Conjecture. *For any $\beta > 0$, with high μ_t-probability, the maximum local roughness of R_t is of order $t^{1/9}$, up to logarithmic corrections.*

Acknowledgements

References

[1] K. Alexander. Cube-root boundary fluctuations for droplets in random cluster models. *Comm. Math. Phys.* **224** (3) (2001) 733–781. MR1871907

[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** (1990) 1–50. MR1062747

[3] A. Asselah and B. Schapira. Moderate deviations for the range of a transient random walk: Path concentration. *Ann. Sci. Éc. Norm. Supér. (4)* **50** (2017) 755–786. MR3665554

[4] A. Asselah and B. Schapira. Boundary of the range of transient random walk. *Probab. Theory Related Fields* **168** (2017) 691–719. MR3663629

[5] M. Biskup and E. Procaccia. Eigenvalue vs perimeter in a shape theorem for self-interacting random walks. *Ann. Appl. Probab.* **28** (2018) 340–377. MR3770879

[6] M. Biskup and E. B. Proccacia. Shapes of drums with lowest base frequency under non-isotropic perimeter constraints. *Trans. Amer. Math. Soc.* To appear, 2018. Available at arXiv:1603.03871.

[7] T. Bodineau. The Wulff construction in three and more dimensions. *Comm. Math. Phys.* **207** (1) (1999) 197–229. MR1724851

[8] T. Bodineau, D. Ioffe and Y. Velenik. Rigorous probabilistic analysis of equilibrium crystal shapes. *J. Math. Phys.* **41** (3) (2000) 1033–1098. MR1757951

[9] E. Bolthausen. Localization of a two-dimensional random walk with an attractive path interaction. *Ann. Probab.* **22** (1994) 875–918. MR1288136

[10] R. Cerf. *The Wulff Crystal in Ising and Percolation Models*. *Lecture Notes in Mathematics* **1878**. Springer, Berlin, 2004. MR2241754

[11] R. Cerf and Á. Pisztora. On the Wulff crystal in the Ising model. *Ann. Probab.* **28** (3) (2000) 947–1017. MR1797302

[12] A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications*, 2nd edition. *Applications of Mathematics* **38**. Springer, New York, 1998. MR1619036

[13] F. den Hollander. *Large Deviations*. *Fields Institute Monographs*. Amer. Math. Soc., Providence, RI, 2008. MR1739680

[14] P. Diaconis and L. Saloff-Coste. Comparison techniques for random walks on finite groups. *Ann. Probab.* **21** (1993) 2131–2156. MR1245303

[15] R. L. Dobrushin and O. Hryniv. Fluctuations of shapes of large areas under paths of random walks. *Probab. Theory Related Fields* **102** (3) (1995) 313–330. MR1402652

[16] R. L. Dobrushin, R. Kotecký and S. B. Shlosman. Equilibrium crystal shapes – A microscopic proof of the Wulff construction. In *Proc. of the XXIVth Karpacz Winter School Stochastic Methods in Mathematics and Physics* (*Karpacz, 1988*) 221–229. R. Gielerak and W. Karwowski (Eds). World Scientific, Singapore, 1989. MR1124695

[17] R. L. Dobrushin, R. Kotecký and S. B. Shlosman. *Wulff Crystal: A Global Shape from Local Interaction*. *Translations of Mathematical Monographs* **104**. Amer. Math. Soc., Providence, RI, 1992. MR1181197

[18] M. D. Donsker and S. R. S. Varadhan. Asymptotic evaluation of certain Markov process expectations for large time. I–IV. *Comm. Pure Appl. Math.* **28** (1975) 1–47; ibid. **28** 279–301, **29** 389–461, **36** 183–212. MR0386024

[19] M. D. Donsker and S. R. S. Varadhan. On the number of distinct points visited by a random walk. *Comm. Pure Appl. Math.* **32** (1979) 721–747. MR0539157

[20] H. Duminil-Copin. Limit of the Wulff Crystal when approaching criticality for site percolation on the triangular lattice. *Electron. Commun. Probab.* **18**, Article ID 93 (2013). MR3151749

[21] J. Gärtner and F. den Hollander. Correlation structure of intermittency in the parabolic Anderson model. *Probab. Theory Related Fields* **114** (1999) 1–54. MR1697138

[22] G. Grimmett. *Percolation*, 2nd edition. *Grundlehren der Mathematischen Wissenschaften* **321**. Springer, Berlin, 1999. MR1707339

[23] A. Hammond. Phase separation in random cluster models I: Uniform upper bounds on local deviation. *Comm. Math. Phys.* **310** (2) (2012) 455–509. MR2890306

[24] D. Ioffe and R. H. Schonmann. Dobrushin–Kotecký–Shlosman theorem up to the critical temperature. *Comm. Math. Phys.* **199** (1) (1998) 117–167. MR1660207

[25] R. Kotecký. Statistical mechanics of interfaces and equilibrium crystal shapes. In *Proc. of the IXth International Congress of Mathematical Physics* (*Swansea, 1988*) 148–163. B. Simon, A. Truman and I. M. Davies (Eds). Adam Hilger, Bristol, 1989. MR1033760

[26] L. H. Loomis and H. Whitney. An inequality related to the isoperimetric inequality. *Bull. Amer. Math. Soc.* **55** (1949) 961–962. MR0031538

[27] R. Lyons and Y. Peres. *Probability on Trees and Networks*, **42**. Cambridge University Press, New York, 2016. MR3616205

[28] R. A. F. Minlos and J. G. Sinai. The phenomenon of phase separation at low temperatures in some lattice models of a gas. I. *Sb. Math.* **2** (3) (1967) 335–395. MR0213133

[29] C.-E. Pfister. Large deviations and phase separation in the two-dimensional Ising model. *Helv. Phys. Acta* **64** (7) (1991) 953–1054. MR1149430

[30] L. C. G. Rogers and D. Williams. *Diffusions, Markov Processes and Martingales: Volume 2. Itô Calculus*. Cambridge University Press, Cambridge, 2000. MR1780932

[31] L. Saloff-Coste. Lectures on finite Markov chains. In *Lectures on Probability Theory and Statistics (Saint-Flour, 1996)* 301–413. *Lectures on Probability Theory and Statistics* **1665**. Springer, Berlin, 1997. MR1490046

[32] H. Uzun and K. Alexander. Lower bounds for boundary roughness for droplets in Bernoulli percolation. *Probab. Theory Related Fields* **127** (1) (2003) 62–88. MR2006231

[33] Wulff and G. Zur. Frage der Geschwindigkeit des Wachstums und der Auflösung der Kristallflächen. *Z. Krystallogr. Mineral.* **34** (1901) 449–530.