Moderate deviations for the range of a transient random walk: path concentration

Amine Asselah* & Bruno Schapira†

Abstract

We study downward deviations of the boundary of the range of a transient walk on the Euclidean lattice. We describe the optimal strategy adopted by the walk in order to shrink the boundary of its range. The technics we develop apply equally well to the range, and provide pathwise statements for the *Swiss cheese* picture of Bolthausen, van den Berg and den Hollander [BBH01].

Keywords and phrases. Large deviations; capacity; range of a random walk; boundary of the range.
MSC 2010 *subject classifications.* Primary 60F10; 60G50.

1 Introduction

In this paper we study downward deviations of the boundary of the range of a simple random walk $(S_n,\ n \in \mathbb{N})$ on $\mathbb{Z}^d$, with $d \geq 3$. The range at time n, denoted $\mathcal{R}_n$, is the set of visited sites $\{S_0, \dots, S_n\}$, and its boundary, denoted $\partial\mathcal{R}_n$, is the set of sites of $\mathcal{R}_n$ with at least one neighbor outside $\mathcal{R}_n$. Our previous study [AS15] focused on the typical behavior of the boundary of the range, whereas this work is devoted to downward deviations and applications to a hydrophobic polymer model. The zest of the paper is about describing the optimal strategy adopted in order to shrink the boundary of the range, and our approach shed some light on the shape of the walk realizing such a deviation. In [AS15], we emphasized the ways in which, for a transient walk, the range and its boundary share a similar nature. Thus, even though the boundary of the range is our primary interest, we mention at the outset that the technics we develop apply equally well to the range. Since this last issue has been the focus of many celebrated works, let us describe first the state of the art there.

Deviations of the range. A pioneering large deviation study of Donsker and Varadhan [DV75] establishes asymptotics for downward deviations of the volume of the Wiener sausage $t \mapsto W^a(t)$, that is the Lebesgue measure of an a-neighborhood of the standard Brownian motion. The main result of [DV75] establishes, in any dimension and for any $\beta > 0$, the following asymptotics

$$\lim_{t\to\infty} t^{-\frac{d}{d+2}} \log \mathbb{E}[\exp\big(-\beta W^a(t)\big)] = \frac{d+2}{2}\ \beta(\frac{2\lambda_D}{d\beta})^{\frac{d}{d+2}}, \tag{1.1}$$

where λ_D is the first eigenvalue of the Laplacian with Dirichlet condition on the boundary of a sphere of volume one. The asymptotics (1.1), obtained in the random walk setting in

*Aix-Marseille Université & Université Paris-Est Créteil E-mail: amine.asselah@u-pec.fr
†Aix-Marseille Université, CNRS, Centrale Marseille, I2M, UMR 7373, 13453 Marseille, France. E-mail: bruno.schapira@univ-amu.fr

[DV79], correspond to downward deviation of the volume of the range $\{|\mathcal{R}_n| \le f(n)\}$ where $|\mathcal{R}_n|$ denotes the volume of $\mathcal{R}_n$ and $f(n)$ is of order $n^{\frac{d}{d+2}}$. They suggest that during time n a random walk is localized in a ball of radius $(n/\beta)^{\frac{1}{d+2}}$ filled without holes. Bolthausen [B90] and Sznitman [S90], with different technics, extended the result of [DV75] to cover downward deviations corresponding to $f(n) = n^{1-\delta}$ for any $\delta > 0$. A consequence of their analysis is that for $0 < \gamma \le 2$

$$\lim_{t\to\infty} t^{-\frac{d+\gamma-2}{d+\gamma}} \log \mathbb{E}[\exp\big(-\beta\, t^{-\frac{2-\gamma}{d+\gamma}} W^a(t)\big)] = -\frac{d+2}{2}\beta(\frac{2\lambda_D}{d\beta})^{\frac{d}{d+2}}. \tag{1.2}$$

Then, three deep studies dealt with the trajectory conditioned on realizing a large deviation by Sznitman [S91], Bolthausen [B94] and Povel [Pov99]. The case $\gamma = 0$ in (1.2) was recognized as *critical* by Bolthausen [B90], and indeed a different behavior was later proved to hold [BBH01]. The series of papers on downward deviations culminated in a paper of Bolthausen, van den Berg and den Hollander [BBH01] which covers the *critical regime* $\{|\mathcal{R}_n| - \mathbb{E}[|\mathcal{R}_n|] \le \varepsilon n\}$. The latter contribution offers a precise Large Deviation Principle, but no pathwise statement characterizing the most likely scenario. The present paper is a step towards filling this gap and providing answers to their motto *How a Wiener sausage turns into a Swiss cheese*? Let us quote their mathematical results. In dimension $d \ge 3$, $\mathbb{E}[W^a(t)]$ grows linearly and the limit of $\frac{1}{t}\mathbb{E}[W^a(t)]$ is denoted κ_a (the Newtonian capacity of a ball of radius a). It is proved in [BBH01] that for any $0 < \varepsilon < 1$

$$\lim_{t\to\infty} \frac{1}{t^{\frac{d-2}{d}}} \log \mathbb{P}\big[W^a(t) - \mathbb{E}[W^a(t)] \le -\varepsilon\kappa_a t\big] = -I_a(\varepsilon), \tag{1.3}$$

where

$$I_a(\varepsilon) = \frac{1}{2\kappa_a^{2/d}} \inf\{\|\nabla f\|_2 :\ f \in H^1(\mathbb{R}^d), \|f\|_2 = 1,\ \int_{\mathbb{R}^d} \big(1 - \exp(-f^2(x))\big)dx\ \le 1-\varepsilon\}. \tag{1.4}$$

A similar result for simple random walks is obtained in Phetpradap's thesis [Phet12]: κ_a becomes the non-return probability say κ_d, and the factor $1/2\kappa_a^{2/d}$ in (1.4) becomes $1/2d\kappa_d^{2/d}$. When $d = 3$ or $d = 4$, the minimizers of (1.4) are strictly positive on $\mathbb{R}^d$, and decrease in the radial component. This is interpreted as saying that Wiener sausage *"looks like a Swiss cheese" with random holes whose sizes are of order 1 and whose density varies on scale $t^{1/d}$*. On the other hand, when $d \ge 5$, and when the parameter ε in (1.3) is small, there is no minimizer for the variational problem (1.4), suggesting that *the optimal strategy is time-inhomogeneous.*

Boundary of the range. The boundary of the range, in spite of not receiving much attention, enters naturally into the modelling of *hydrophobic polymers*. Indeed, a *polymer* is a succession of monomers centered at the positions of the walk (and thus covering $\mathcal{R}_n$), the complement of the range is occupied by the aqueous solvent, and being *hydrophobic* means that the monomers try to hide from it. A natural model is then the following polymer measure depending on two parameters: its length n, and its inverse temperature β.

$$d\widetilde{\mathbb{Q}}_n^\beta = \frac{1}{\widetilde{Z}_n(\beta)} \exp(-\beta|\partial\mathcal{R}_n|)\, d\mathbb{P}_n,$$

where $\mathbb{P}_n$ denotes the law of the simple random walk up to time n and $\widetilde{Z}_n(\beta)$, the partition function, is a normalizing factor. Biology suggests that as one tunes β, for a fixed polymer length, a phase transition appears. The recent results of Berestycki and Yadin [BY13] treat an asymptotic regime of length going to infinity, and suggest that for any positive β a long enough polymer, that is under $\widetilde{\mathbb{Q}}_n^\beta$, is localized in a ball of radius ρ_n with ρ_n^{d+1} of order n. Thus, to capture the insight from Biology, we rather scale β with $n^{2/d}$, when n is taken to infinity. We therefore consider

$$d\mathbb{Q}_n^\beta = \frac{1}{Z_n(\beta)} \exp\Big(- \frac{\beta}{n^{2/d}} \big(|\partial \mathcal{R}_n| - \mathbb{E}[|\partial \mathcal{R}_n|]\big)\Big)\, d\mathbb{P}_n. \tag{1.5}$$

The centering of $|\partial\mathcal{R}_n|$ is a matter of taste, but the scaling of β by $n^{2/d}$ is crucial, and corresponds to *a critical regime* for the boundary of the range reminiscent of (1.2) for $\gamma = 0$. Indeed, understanding the polymer measure is linked with analyzing the scenarii responsible for shrinking the boundary of the range on the scale of its mean. However, before tackling deviations, let us recall some typical behavior of the boundary of the range. Okada [Ok15] has proved a law of large numbers in dimension $d \geq 3$, and when dimension is two, he proved that

$$\frac{\pi^2}{2} \leq \lim_{n\to\infty} \frac{\mathbb{E}[|\partial\mathcal{R}_n|]}{n/(\log n)^2} \leq 2\pi^2.$$

Note that Benjamini, Kozma, Yadin and Yehudayoff [BKYY10] in their study of the entropy of the range have obtained the correct order of magnitude for $\mathbb{E}[|\partial\mathcal{R}_n|]$ in $d = 2$, and have linked the entropy of the range to the size of its boundary.

In addition, a central limit theorem for the boundary of the range was proved in [AS15] in dimension $d \geq 4$. When dimension is three, the variance is expected to grow like $n \log n$, and only an upper bound of the right order is known [AS15]. We henceforth focus on the ways in which a random walk reduces the boundary of its range.

Capacity of the range. A key object used to probe the shape of the random walk is the *capacity of its range*. We first define it, and then state our result. For $\Lambda \subset \mathbb{Z}^d$, let H_Λ^+ be the time needed by the walk to return to Λ. The capacity of Λ, denoted $\text{cap}(\Lambda)$, is

$$\text{cap}(\Lambda) = \sum_{x\in\Lambda} \mathbb{P}_x\big[H_\Lambda^+ = \infty\big]. \tag{1.6}$$

Let us recall one of its basic property. There exists a positive constant c_{cap}, such that for all finite subset $\Lambda \subset \mathbb{Z}^d$

$$c_{\text{cap}}\, |\Lambda|^{1-\frac{2}{d}} \leq \text{cap}(\Lambda) \leq |\Lambda|. \tag{1.7}$$

The upper bound follows by definition and the lower bound is well known (see the proof of Proposition 2.5.1 in [L13]). In a weak sense, the capacity of a set characterizes the shape of a set: the closer it is to a ball, the smaller is its capacity. In view of (1.7), this is captured by the index $\mathcal{I}_d$, defined for finite subsets Λ in $\mathbb{Z}^d$, by

$$\mathcal{I}_d(\Lambda) := \frac{\text{cap}(\Lambda)}{|\Lambda|^{1-\frac{2}{d}}}. \tag{1.8}$$

It is known that the index $\mathcal{I}_d$ of a ball is bounded by some constant, independently of the radius of the ball (see (5.1) below). On the other hand, the capacity of the range $\mathcal{R}_n$ has been studied in [ASS16, JO69, L13, RS12], and it is known that its mean is of order $\sqrt{n}$ in dimension three, of order $n/\log n$ in dimension four, and grows linearly in dimension five and larger. Thus typically, for a transient walk, $\mathcal{I}_d(\mathcal{R}_n)$ goes to infinity with n.

Results. We have found that the strategy for shrinking the (size of the) boundary of the range $|\partial\mathcal{R}_n|$ below its mean by εn is different in dimension three and in dimensions five and larger. In dimension three (see (1.9) below), the walk has to spend a positive fraction of its time in a set with $\mathcal{I}_d$-index of order 1 and volume of order n/ε, whereas in dimensions five and larger (see (1.11) below), it has to spend a fraction of order ε of its time in a set of $\mathcal{I}_d$-index of order 1 and volume of order n.

In addition, we prove that spending a positive fraction (resp. a fraction ε) of the time in a ball of radius $(n/\varepsilon)^{1/d}$ (resp. $n^{1/d}$), leads to reducing the boundary of the range by a factor ε in dimension three (resp. four and larger), and this gives us the lower bounds for the large deviations in (1.10), (1.12) and (1.14), see also Section 4 for more details.

For a subset Λ of $\mathbb{Z}^d$, we denote by $\ell_n(\Lambda)$ the time spent by the walk inside Λ up to time n. For an integrable random variable X, we also denote by $\overline{X}$ the centered variable $X - \mathbb{E}[X]$.

Theorem 1.1. *Assume that $d = 3$. There exist $\alpha \in (0,1)$ and $C > 0$, such that for all $\varepsilon \in (0, \nu_3/2)$,*

$$\lim_{n\to\infty} \mathbb{P}\Big[\exists \Lambda \subset \mathbb{Z}^3 \,:\, \ell_n(\Lambda) \geq \alpha n,\, \mathcal{I}_d(\Lambda) \leq C,\, \frac{n}{C\varepsilon} \leq |\Lambda| \leq C\frac{n}{\varepsilon} \;\Big|\; \overline{|\partial\mathcal{R}_n|} \leq -\varepsilon n\Big] = 1. \tag{1.9}$$

Moreover, there exist positive constants $\underline{\kappa}_3$ and $\bar{\kappa}_3$, such that for all $\varepsilon \in (0, \nu_3/2)$, and n large enough

$$\exp\big(-\bar{\kappa}_3 \cdot \varepsilon^{\frac{2}{3}} n^{\frac{1}{3}}\big) \;\leq\; \mathbb{P}\Big[|\partial\mathcal{R}_n| - \mathbb{E}[|\partial\mathcal{R}_n|] \leq -\varepsilon n\Big] \;\leq\; \exp\big(-\underline{\kappa}_3 \cdot \varepsilon^{\frac{2}{3}} n^{\frac{1}{3}}\big). \tag{1.10}$$

Theorem 1.2. *Assume that $d \geq 5$. There exist $\alpha \in (0,1)$ and $C > 0$, such that for all $\varepsilon \in (0, \nu_d/2)$,*

$$\lim_{n\to\infty} \mathbb{P}\Big[\exists \Lambda \subset \mathbb{Z}^d \,:\, \ell_n(\Lambda) \geq \alpha\varepsilon n,\, \mathcal{I}_d(\Lambda) \leq C,\, \frac{\varepsilon n}{C} \leq |\Lambda| \leq C\varepsilon n \;\Big|\; \overline{|\partial\mathcal{R}_n|} \leq -\varepsilon n\Big] = 1. \tag{1.11}$$

Moreover, there exist positive constants $\underline{\kappa}_d$ and $\bar{\kappa}_d$, such that for all $\varepsilon \in (0, \nu_d/2)$, and n large enough

$$\exp\big(-\bar{\kappa}_d \cdot (\varepsilon n)^{1-\frac{2}{d}}\big) \;\leq\; \mathbb{P}\Big[|\partial\mathcal{R}_n| - \mathbb{E}[|\partial\mathcal{R}_n|] \leq -\varepsilon n\Big] \;\leq\; \exp\big(-\underline{\kappa}_d \cdot (\varepsilon n)^{1-\frac{2}{d}}\big). \tag{1.12}$$

In dimension 4, our result is slightly less precise: our upper bound for the large deviations is larger than the lower bound by a factor $|\log\varepsilon|^{1/2}$, and we only describe the localization phenomenon for ε away from 0 and ν_4.

Theorem 1.3. *Assume that $d = 4$. For any $\varepsilon \in (0, \nu_4/2)$, there exist $\alpha = \alpha(\varepsilon) \in (0,1)$, and $C = C(\varepsilon) > 0$, such that*

$$\lim_{n\to\infty} \mathbb{P}\Big[\exists \Lambda \subset \mathbb{Z}^4 \,:\, \ell_n(\Lambda) \geq \alpha n,\, \mathcal{I}_d(\Lambda) \leq C,\, \frac{n}{C} \leq |\Lambda| \leq Cn \;\Big|\; \overline{|\partial\mathcal{R}_n|} \leq -\varepsilon n\Big] = 1. \tag{1.13}$$

Moreover, there exist positive constants $\underline{\kappa}_4$ and $\bar{\kappa}_4$, such that for all $\varepsilon \in (0, \nu_4/2)$, and n large enough

$$\exp\left(-\bar{\kappa}_4 \cdot (\varepsilon n)^{\frac{1}{2}}\right) \;\leq\; \mathbb{P}\left[|\partial\mathcal{R}_n| - \mathbb{E}[|\partial\mathcal{R}_n|] \leq -\varepsilon n\right] \;\leq\; \exp\left(-\underline{\kappa}_4 \cdot \frac{(\varepsilon n)^{1/2}}{|\log\varepsilon|^{1/2}}\right). \tag{1.14}$$

Remark 1.4. Our theorems play with two parameters n and ε, and consider the regime where n is large and ε is small. They suggest that conditioned on having a small boundary of the range, in dimension three a (large) fraction of the walk is localized in a ball of volume n/ε, whereas in dimension five or larger a length εn of the trajectory is localized in a ball of volume of order εn. This is the time-inhomogeneous nature of the trajectory that was mentioned earlier.

Remark 1.5. We state the results for $|\partial\mathcal{R}_n|$, and they stand as well for $|\mathcal{R}_n|$. In this case, (1.9), (1.11), and (1.13) are new, and complement the results of van den Berg, Bolthausen, and den Hollander [BBH01]. Our proof is based on a recursive slicing of a strand of random walk, as we explain in Section 2, whereas the proof of [BBH01] consists of many steps: (i) first a compactification obtained by wrapping the trajectory on a Torus of side $An^{1/d}$, with A large, (ii) fixing the position of the walk at times multiple of $n^{2/d}$, the so-called *skeleton*, and using concentration to showing that the range and *averaged range conditioned on the skeleton* are close and finally (iii) representing the averaged range conditioned on the skeleton as a continuous functional of the pair empirical measure, and invoking Donsker Varadhan Large Deviation Theory. Note that step (ii) uses that the range is a union of ranges over periods of length $n^{2/d}$, and thus a Lipshitz function of the collection of smaller ranges. The boundary of the range is neither a union of boundaries, nor a monotonous function of time, and we had to follow a rougher but more robust method (see below for a sketch of our approach).

Remark 1.6. It is noted in [BBH01] that there is an anomaly in the scaling of the rate function for ε close to zero, which does not connect with the central limit theorem. We quote from [BBH01] "*The anomaly for $d \geq 3$ is somewhat surprising. It suggests that the central limit behavior is controlled by the local fluctuation [...], while the moderate and large deviations are controlled by the global fluctuations.*" This is exactly what our decomposition shows: the central limit theorem obtained in [AS15], and the downward deviations obtained here correspond to two distinct parts of the slicing down of a random trajectory: the *self-similar independent parts* contribute to the fluctuations, whereas *the mutual-intersection parts* contribute to downward deviations. We will come back on this later.

The capacity of the range plays a central role in our results. Our main technical contribution, interesting on its own, is the following estimate which generalizes an inequality (1.8) of [AC07], as well as Proposition 1.5 of [A08]. We bound the probability of multiple visits to non-overlapping spheres, all of the same radius, and our bound involves the capacity of the union of the spheres. Let us first introduce handy notation. For $x \in \mathbb{Z}^d$ and $r > 0$, we denote with $B(x,r)$ the Euclidean ball of radius r and center x, and for any subset $\mathcal{C} \subset \mathbb{Z}^d$, we let

$$B(\mathcal{C},r) := \bigcup_{x\in\mathcal{C}} B(x,r).$$

Additionally, for $r > 0$, $\mathcal{A}(r)$ is the collection of finite sets of centers defining non-overlapping spheres of radius $2r$

$$\mathcal{A}(r) := \big\{\mathcal{C} \subset \mathbb{Z}^d : \ |\mathcal{C}| < \infty \ \text{ and } \ \|x-y\| \geq 4r \text{ for all } x \neq y \in \mathcal{C}\big\}. \tag{1.15}$$

Proposition 1.7. *Assume that $d \geq 3$. There exist positive constants C and κ, such that for any $t > 0$, $r \geq 1$, $\mathcal{C} \in \mathcal{A}(r)$, and $n \geq 1$, one has*

$$\mathbb{P}\Big[\ell_n(B(x,r)) \geq t \ \text{ for all } x \in \mathcal{C}\Big] \leq C\,(|\mathcal{C}|\,n)^{|\mathcal{C}|}\,\exp\left(-\kappa\cdot\frac{t|\mathcal{C}|}{|B(\mathcal{C},r)|^{\frac{2}{d}}}\cdot\mathcal{I}_d(B(\mathcal{C},r))\right). \tag{1.16}$$

This result is useful when the combinatorial term in (1.16) is innocuous, and using (1.7) we see that this holds when for a constant δ small enough

$$|B(\mathcal{C},r)|^{2/d}\log n \ \leq \ \delta t. \tag{1.17}$$

Note also that on the right hand side of (1.16), $t|\mathcal{C}|$ is a lower bound for the total time spent in $B(\mathcal{C},r)$, under the event of the left hand side. If this total time is comparable with the volume of $B(\mathcal{C}, r)$, then the estimate discriminates sets of comparable volume with different indices $\mathcal{I}_d$. This is exactly the kind of situation we will encounter here. To be more precise we will use (1.16) in cases where both $t|\mathcal{C}|$ and $|B(\mathcal{C}, r)|$ are of order n. In these cases, (1.17) holds when $r^d \geq C n^{2/d} \log n$, for some large enough constant C, and (1.16) becomes

$$\mathbb{P}\Big[\ell_n(B(x,r)) \geq t \ \text{ for all } x \in \mathcal{C}\Big] \ \leq \ C \exp\big(- \kappa' \cdot \mathcal{I}_d(B(\mathcal{C}, r)) \cdot n^{1-2/d}\big), \tag{1.18}$$

with κ' some other positive constant. Now in our results (1.9), (1.11) and (1.13), the set Λ will be a certain union of balls as here, and thus (1.18) explains why its $\mathcal{I}_d$-index should be bounded.

Sketch of our approach (in $d = 3$). Our approach treats separately the upper and lower bounds of the downward deviation in (1.10). The upper bound relies on the inequality (3.1) which roughly bounds the centered boundary of the range by self-similar terms, whose deviations are costly, minus an increasing process $\xi_n(T)$ (defined in (3.12)) measuring the *folding of the trajectory* over small strands of length T. Since this term is the key term driving the moderate deviations, let us mention that it can be written in terms of Green's function up to time T, denoted G_T and defined later (and the $\sim$ means that the *correct* definition of $\xi_n(T)$ requires more notation),

$$\xi_n(T) \sim \frac{1}{T} \sum_{k=1}^{n} \sum_{z \in \mathcal{R}_k} G_T(z - S_k).$$

What we show first is that a decrease in $\partial\mathcal{R}_n$ most likely translates into an increase of $\xi_n(T)$, for an appropriate scale T . Then, we show that the event that $\xi_n(T)$ large means that *many* balls are visited *often* by the walk: see (2.19) and Lemma 2.4 for a precise statement. The probability of this latter event is controlled by the capacity of the collection of balls as in Proposition 1.7. We believe that a similar structure shows up in the capacity of the range in dimension five or more, which has been the focus of recent studies (see [ASS16] and references therein). The lower bound has a different flavor, and relies on a covering result saying that when the walk is localized in a ball of volume n/ε, it likely visits an ε-fraction of all fixed *large* subsets of the ball (see Proposition 4.1).

Application to the polymer measure. Recall that

$$Z_n(\beta) = \mathbb{E}\left[\exp\Big(- \frac{\beta}{n^{2/d}}\big(|\partial\mathcal{R}_n| - \mathbb{E}[|\partial\mathcal{R}_n|]\big)\Big)\right].$$

Note first that by Jensen's inequality, one has $Z_n(\beta) \geq 1$, for all $\beta \geq 0$. Thus one can define for $\beta \geq 0$,

$$F^+(\beta) = \limsup_{n\to\infty} \frac{1}{n^{1-\frac{2}{d}}} \log Z_n(\beta),$$

and

$$F^-(\beta) = \liminf_{n\to\infty} \frac{1}{n^{1-\frac{2}{d}}} \log Z_n(\beta).$$

Denote by $\mathcal{A}(r, v)$ (with $v > 0$) the subset of $\mathcal{A}(r)$, whose elements $\mathcal{C}$ satisfy $|B(\mathcal{C}, r)| \leq v$.

Theorem 1.8. *Assume that $d \geq 3$. The following statements hold.*

1. *The functions F^+ and F^- are non decreasing in β. As a consequence there exist $0 \leq \beta_d^+ \leq \beta_d^- \leq +\infty$, such that*

$$F^{\pm}(\beta) = 0 \quad \textit{for } \beta < \beta_d^{\pm} \quad \textit{and} \quad F^{\pm}(\beta) > 0 \quad \textit{for } \beta > \beta_d^{\pm}.$$

2. In fact β_d^+ and β_d^- are positive and finite.

Let $(r_n,\ n \in \mathbb{N})$ be any fixed sequence of reals satisfying $n^{2/d}(\log n)^2 \le r_n^d \le \frac{n}{\log n}$, for all $n \ge 1$. Then,

3. For any $\beta < \beta_d^+$, $\alpha \in (0,1)$ and $A > 0$,

$$\lim_{n\to\infty} \mathbb{Q}_n^\beta \left[\exists \mathcal{C} \in \mathcal{A}(r_n, An) \ :\ \ell_n(B(x,r_n)) \ge \alpha \frac{n}{|\mathcal{C}|} \text{ for all } x \in \mathcal{C} \right] = 0.$$

Moreover, for any $\varepsilon \in (0, \nu_d)$,

$$\lim_{n\to\infty} \mathbb{Q}_n^\beta \left[|\partial \mathcal{R}_n| - \mathbb{E}[|\partial \mathcal{R}_n|] \le -\varepsilon n \right] = 0.$$

4. For any $\beta > \beta_d^-$, there exists $\varepsilon(\beta) \in (0, \nu_d)$, such that

$$\lim_{n\to\infty} \mathbb{Q}_n^\beta \left[|\partial \mathcal{R}_n| - \mathbb{E}[|\partial \mathcal{R}_n|] \le -\varepsilon(\beta) n \right] = 1.$$

Moreover, when $d = 3$, there exist $\alpha \in (0,1)$ and $A > 0$, such that

$$\lim_{n\to\infty} \mathbb{Q}_n^\beta \left[\exists \mathcal{C} \in \mathcal{A}(r_n, \frac{An}{\varepsilon(\beta)}) \ :\ \mathcal{I}_d(B(\mathcal{C}, r_n)) \le A \text{ and } \ell_n(B(x,r)) \ge \alpha \frac{n}{|\mathcal{C}|} \text{ for all } x \in \mathcal{C} \right] = 1. \tag{1.19}$$

When $d \ge 4$, there exist $\alpha \in (0,1)$ and $A > 0$, such that

$$\lim_{n\to\infty} \mathbb{Q}_n^\beta \left[\exists \mathcal{C} \in \mathcal{A}(r_n, An),\ \mathcal{I}_d(B(\mathcal{C}, r_n)) \le A \text{ and } \ell_n(B(x,r_n)) \ge \alpha\varepsilon(\beta) \frac{n}{|\mathcal{C}|} \text{ for all } x \in \mathcal{C} \right] = 1. \tag{1.20}$$

5. In the previous part, one can choose $\varepsilon(\beta)$ such that $\varepsilon(\beta) \to \nu_d$, as $\beta \to +\infty$.

The rest of the paper is organized as follows. Section 2 sets the preliminaries, where Lemma 2.4 is our key technical tool, interesting on its own. Section 3 presents the upper bounds for our large deviation estimates, and Section 4 presents the corresponding lower bounds. A large part of Section 4 deals with the three dimensional case and proves Proposition 4.1 which establishes lower bound on the probability of covering fraction of domains. Section 5 introduces capacities, and contains the proof of Proposition 1.7, as well as establishing statements (1.9), (1.11) and (1.13). Finally, our application to hydrophobic polymer, that is Theorem 1.8, is explained in Section 6.

2 Preliminaries

2.1 Notation

For any integers $n \le m$ we write

$$\mathcal{R}(n,m) = \{S_n, \dots, S_m\}.$$

We denote by $\|\cdot\|$ the Euclidean norm on $\mathbb{R}^d$, and by

$$B(z,r) = \left\{ y \in \mathbb{Z}^d \ :\ \|z - y\| \le r \right\},$$

the (discrete) ball of radius r centered at z. For $z \in \mathbb{Z}^d$ and Λ a subset of $\mathbb{Z}^d$, we write

$$d(z, \Lambda) = \inf\{\|z - y\| \ : \ y \in \Lambda\},$$

for the distance between z and Λ, and let $|\Lambda|$ be the size (which we also call the volume) of Λ. We also define

$$\partial\Lambda = \{z \in \Lambda \ : \ d(z, \Lambda^c) = 1\} \quad \text{and} \quad \Lambda^+ = \Lambda \cup \left\{z \in \mathbb{Z}^d \ : \ d(z, \Lambda) = 1\right\}.$$

Furthermore,

$$H_\Lambda = \inf\{k \geq 0 \ : \ S_k \in \Lambda\},$$

is the hitting time of Λ, that we abbreviate in H_z when $\Lambda = \{z\}$. and we recall that for $n \geq 0$,

$$\ell_n(\Lambda) = \ |\{k \leq n \ : \ S_k \in \Lambda\}|,$$

is the number of steps spent in Λ before time n. We next consider Green's function,

$$G(z) = \sum_{k \geq 0} \mathbb{P}[S_k = z] = \mathbb{P}[H_z < \infty] \times G(0) \qquad \text{for } z \in \mathbb{Z}^d,$$

and its restricted version for any positive integer T

$$G_T(z) = \sum_{k=0}^{T} \mathbb{P}[S_k = z].$$

2.2 Range and its boundary.

A powerful idea, going back at least to Le Gall [LG86], is to cut a trajectory into two pieces, and read the total range as a union of two ranges minus the mutual intersection of two *independent* strands. From a set-theoretical point of view, the volumes satisfy an exclusion-inclusion formula:

$$|\Lambda_1 \cup \Lambda_2| = |\Lambda_1| + |\Lambda_2| - |\Lambda_1 \cap \Lambda_2| \qquad \text{for all } \Lambda_1,\ \Lambda_2 \subset \mathbb{Z}^d. \tag{2.1}$$

The key *probabilistic* point is that if $\Lambda_1 = \{S_0, \dots, S_n\}$ and $\Lambda_2 = \{S_n, \dots, S_{n+m}\}$, then by translating sets by S_n, we have

$$|\Lambda_1 \cap \Lambda_2| = \left|\{S_0 - S_n, S_1 - S_n, \dots, 0\} \cap \{0, \dots, S_{n+m} - S_n\}\right|, \tag{2.2}$$

and the two sets on the right hand side are independent. It is then possible, using the symmetry of the walk, to compute the expectation:

$$\mathbb{E}[|\Lambda_1 \cap \Lambda_2|] = \sum_{z \in \mathbb{Z}^d} \mathbb{P}[H_z \leq n]\, \mathbb{P}[H_z \leq m]. \tag{2.3}$$

In view of (2.1) and (2.2), it is no surprise that Bolthausen, van den Berg and den Hollander studied at the same time downward deviations of the Wiener sausage [BBH01] and upward deviation of the intersection of two Wiener sausages in [BBH04]. The second paper is independently motivated by an older paper of Khanin, Mazel, Shlosman, and Sinai [KMSS94] studying bounds for the intersection of two independent ranges in an infinite time-horizon, a problem which is still open.

Now, the boundary of the range does not quite satisfy the exclusion-inclusion equality (2.1). However, if $z \in \partial\Lambda_1 \backslash \Lambda_2^+$, then the neighbors of z are not in Λ_2, and at least one of them is not in Λ_1. This means that $z \in \partial(\Lambda_1 \cup \Lambda_2)$, and therefore

$$\big(\partial\Lambda_1 \backslash \Lambda_2^+\big) \cup \big(\partial\Lambda_2 \backslash \Lambda_1^+\big) \subset \partial\big(\Lambda_1 \cup \Lambda_2\big). \tag{2.4}$$

When taking the volume, this reads

$$\begin{aligned} |\partial\big(\Lambda_1 \cup \Lambda_2\big)| \geq & |\partial\Lambda_1| + |\partial\Lambda_2| - \Big(|\partial\Lambda_1 \cap \Lambda_2^+| + |\Lambda_1^+ \cap \partial\Lambda_2|\Big) \\ \geq & |\partial\Lambda_1| + |\partial\Lambda_2| - 2|\Lambda_1^+ \cap \Lambda_2^+|. \end{aligned} \tag{2.5}$$

On the other hand, if $z \notin \partial\Lambda_1$, but $z \in \Lambda_1$, then $\{z\}^+ \subset \Lambda_1$, and z is not in the boundary of $\Lambda_1 \cup \Lambda_2$. In other words,

$$|\partial\big(\Lambda_1 \cup \Lambda_2\big)| \leq |\partial\Lambda_1| + |\partial\Lambda_2| - |\Lambda_1 \cap \partial\Lambda_2|. \tag{2.6}$$

Thus, (2.5) and (2.6) make up for the equality (2.1).

By exploiting (2.5), it has been shown in [AS15] that there exist positive dimension-dependent constants ν_d and C_d, such that

$$\big|\mathbb{E}[|\partial\mathcal{R}_n|] - \nu_d n\big| \leq C_d\, \psi_d(n), \tag{2.7}$$

for all $n \geq 1$, where

$$\psi_3(n) = \sqrt{n}, \qquad \psi_4(n) = \log n, \qquad \psi_d(n) = 1 \ \text{ for } d \geq 5. \tag{2.8}$$

2.3 On Green's function

First, we recall the asymptotics of Green's function (see [LL10, Theorem 4.3.1])

$$G(z) = \mathcal{O}\big(\frac{1}{1 + \|z\|^{d-2}}\big). \tag{2.9}$$

For Green's function restricted to the first T steps, the following holds.

Lemma 2.1. *Assume that $d \geq 3$. There exist positive constants c and C, such that for any $T > 0$ and $z \in \mathbb{Z}^d$,*

$$G_T(z) \ \leq \ C\, \frac{T}{1 + \|z\|^d}\, \exp\left(-c\, \frac{\|z\|^2}{T}\right).$$

Proof. One can assume that $\|z\| \geq \sqrt{T}$, as otherwise the result follows from (2.9). One result of [HSC93] ensures that there exist constants c and C, such that

$$\mathbb{P}(S_n = z) \leq C\, \frac{1}{n^{d/2}}\, \exp(-c\|z\|^2/n) \quad \text{for all } z \text{ and } n \geq 1. \tag{2.10}$$

This implies the following bound.

$$\begin{aligned} G_T(z) = \sum_{k=1}^{T} \mathbb{P}[S_k = z] \leq & \ C \sum_{k=1}^{T} k^{-d/2} \exp(-c\|z\|^2/k) \\ \leq & \ C\, \|z\|^{-d} \sum_{k=1}^{T} \exp(-\frac{c}{2}\|z\|^2/k) \leq \ C\, \frac{T}{\|z\|^d}\, \exp(-\frac{c}{2}\|z\|^2/T). \end{aligned}$$

□

For $x \in \mathbb{Z}^d$ and $r \in \mathbb{N}$, we consider the discrete cube centered on x and of side r:

$$Q(x,r) = \big(x +]-r,r]^d\big) \cap \mathbb{Z}^d.$$

The elements of a partition of $\mathbb{Z}^d$ obtained from translates of $Q(0,r)$ is denoted $\mathcal{P}_r$ with

$$\mathcal{P}_r = \{Q(x,r) : \ x \in 2r\mathbb{Z}^d\}. \tag{2.11}$$

For $T > 0$, we denote by $\mathcal{P}_r(T)$ the elements of $\mathcal{P}_r$ whose intersection with $B(0,T)$ is not empty. We will need the following covering result.

Lemma 2.2. *Assume that $d \geq 3$. There exists a constant $C > 0$, such that for any $r \geq 1$, $\gamma \in (0,1)$, and any collection of subsets $\{\Lambda_Q,\ Q \in \mathcal{P}_r\}$ satisfying*

$$\Lambda_Q \subset Q \quad \text{and} \quad |\Lambda_Q| \leq \gamma\, r^d \qquad \text{for all } Q \in \mathcal{P}_r, \tag{2.12}$$

one has for all $T \geq r$,

$$\sum_{Q \in \mathcal{P}_r} \sum_{z \in \Lambda_Q} G_T(z) \ \leq \ C\,\big(\gamma^{2/d} r^2 + \gamma T\big). \tag{2.13}$$

Note that since the elements of $\mathcal{P}_r$ form a partition of the space (in particular they are disjoints), the left hand side of (2.13) is also equal to the expectation of the time spent before T in the union of all the Λ_Q, with $Q \in \mathcal{P}_r$. In particular, when $\Lambda_Q = \Lambda \cap Q$ for some set Λ, then the left hand side of (2.13) is simply the expected value of the number of steps spent in Λ before time T.

Proof. We denote by C a constant whose value might change from line to line. We decompose the sum in the left hand side of (2.13) into $\Sigma_I + \Sigma_{II}$, with

$$\Sigma_I := \sum_{Q \in \mathcal{P}_r(\sqrt{T})} \sum_{z \in \Lambda_Q} G_T(z) \quad \text{and} \quad \Sigma_{II} := \sum_{Q \in \mathcal{P}_r(T) \setminus \mathcal{P}_r(\sqrt{T})} \sum_{z \in \Lambda_Q} G_T(z).$$

For Σ_I we use (2.9) and (2.12). This gives

$$\begin{aligned}
\Sigma_I &\leq \sum_{z \in \Lambda_{Q(0,r)}} G_T(z) + \sum_{Q \in \mathcal{P}_r(\sqrt{T}) \setminus \{Q(0,r)\}} \sum_{z \in \Lambda_Q} G_T(z) \\
&\leq C \left(|\Lambda_{Q(0,r)}|^{2/d} + \sum_{Q \in \mathcal{P}_r(\sqrt{T}) \setminus \{Q(0,r)\}} \frac{\gamma r^d}{d(0,Q)^{d-2}} \right) \\
&\leq C \left(\gamma^{2/d} r^2 + \gamma r^d \sum_{k=1}^{[\sqrt{T}/r]+1} \frac{k^{d-1}}{(kr)^{d-2}} \right) \\
&\leq C\big(\gamma^{2/d} r^2 + \gamma T\big),
\end{aligned}$$

where at the second line we used a well known bound for the first sum (see for instance the proof of Proposition 2.5.1 in [L13]). We deal now with Σ_{II} and use this time Lemma 2.1 instead of (2.9). This gives

$$\begin{aligned}
\Sigma_{II} &\le C\,\gamma r^d T \sum_{Q\in\mathcal{P}_r(T)\setminus\mathcal{P}_r(\sqrt{T})} \frac{\exp\big(-cd(0,Q)^2/T\big)}{d(0,Q)^d} \\
&\le C\,\gamma r^d T \sum_{k=[\sqrt{T}/r]}^{[T/r]+1} k^{d-1}\frac{\exp(-ck^2r^2/T)}{(kr)^d} \\
&\le C\,\gamma T \sum_{k\ge[\sqrt{T}/r]} \frac{1}{k}\exp\left(-c\frac{k^2r^2}{T}\right) \\
&\le C\,\gamma T.
\end{aligned}$$

□

2.4 Rolling a Ball

The main result of this section requires further notation. For V a subset of $\mathbb{Z}^d$, define

$$\mathcal{K}_n(V,t) := \{k\in\{1,\dots,n\}:\ \ell_k(S_k+V)>t\}. \tag{2.14}$$

Thus, as the set V rolls along the random walk trajectory, $\mathcal{K}_n(V,t)$ records the times before n when the number of visits in the moving window V exceeds t. We use here only the case where V is a ball, and our first concern is to express $\{|\mathcal{K}_n(V,t)|>L\}$ in terms of occupation of non-overlapping balls.

Lemma 2.3. *For any $v\in\mathbb{Z}^d$, $r\ge1$, $t>0$, $L>0$ and $n\ge1$, there is a random subset $\mathcal{C}\in\mathcal{A}(r)$, which is measurable with respect to $\sigma(S_1,\dots,S_n)$, such that*

$$\{|\mathcal{K}_n(B(v,r),t)|>L\}\ \subset\ \{\ell_n(B(x+v,r)\ge t\ \text{ for all } x\in\mathcal{C}\}\cap\{\ell_n\big(B(\mathcal{C},4r)\big)\ge L\}. \tag{2.15}$$

Proof. Assume $\{|\mathcal{K}_n(B(v,r),t)|>L\}$ and define

$$k_1 := \inf\{k\ge0\ :\ \ell_k(S_k+B(v,r))>t\}.$$

Since $L>0$, we have $k_1\le n$, and we set $\mathcal{C}_1=\{S_{k_1}\}$. If all the elements S_k with $k\in\mathcal{K}_n(B(v,r),t)$ are in $B(\mathcal{C}_1,4r)$, set $\mathcal{C}=\mathcal{C}_1$. Otherwise, proceed by induction and assume that for $i>1$, $\mathcal{C}_{i-1}\in\mathcal{A}(r)$ and k_{i-1} satisfy

$$\ell_{k_{i-1}}\Big(B(x+v,r)\Big)>t \qquad \text{for all } x\in\mathcal{C}_{i-1}. \tag{2.16}$$

If $\ell_n(B(\mathcal{C}_{i-1},4r))\ge L$, then $\mathcal{C}=\mathcal{C}_{i-1}$. Otherwise, define

$$k_i := \inf\{k>k_{i-1}:\ S_k\notin B(\mathcal{C}_{i-1},4r) \text{ and } \ell_k(S_k+B(v,r))>t\}.$$

Note that $k_i\le n$, and define $\mathcal{C}_i=\mathcal{C}_{i-1}\cup\{S_{k_i}\}$. Since $k_i\ge k_{i-1}+1$, the construction stops in a finite number of steps and yields a finite set $\mathcal{C}\in\mathcal{A}(r)$. □

We introduce now some new notation. For any positive integers n and m, and any positive r, t and L, define

$$\mathcal{G}_n(r,t,m) := \bigcup_{\mathcal{C}\in\mathcal{A}(r)\,:\,|\mathcal{C}|=m} \Big\{\ell_n\Big(B(x,r)\Big)>t\ \text{ for all } x\in\mathcal{C}\Big\}, \tag{2.17}$$

and

$$\mathcal{H}_n(r,L,m) := \bigcup_{\mathcal{C}\in\mathcal{A}(r)\,:\,|\mathcal{C}|=m} \Big\{\ell_n\Big(B(\mathcal{C},4r)\Big) \ge L\Big\}. \tag{2.18}$$

Note that $\mathcal{G}_n(r,t,m) \subset \mathcal{G}_n(r,t,m-1)$ whereas $\mathcal{H}_n(r,L,m-1) \subset \mathcal{H}_n(r,L,m)$. Therefore (2.15) implies that for *any* positive integer m

$$\{|\mathcal{K}_n(B(v,r),t)| > L\} \;\subset\; \mathcal{G}_n(r,t,m) \cup \mathcal{H}_n(r,L,m). \tag{2.19}$$

We recall next a useful inequality proved in [AC07, Lemma 1.2]. There exist positive constants κ_0 and C, such that for all $t > 0$ and all (nonempty) subsets $\Lambda \subset \mathbb{Z}^d$,

$$\mathbb{P}\left[\ell_\infty(\Lambda) \ge t\right] \;\le\; C\,\exp\left(-\kappa_0 \cdot \frac{t}{|\Lambda|^{2/d}}\right). \tag{2.20}$$

We are now ready to estimate the probability of the event $\{|\mathcal{K}_n(B(v,r),t)| > L\}$.

Lemma 2.4. *Assume that $d \ge 3$. There exist positive constants κ and C, such that for any $r \ge 1$, $L \ge 1$, $t \ge 1$, and $n \ge 2$, satisfying*

$$\left(1 + \frac{L}{t}\right)^{2/d} \log n \;\le\; \kappa\, \frac{t}{r^2}, \tag{2.21}$$

we have for any $v \in \mathbb{Z}^d$,

$$\mathbb{P}\big[|\mathcal{K}_n(B(v,r),t)| \ge L\big] \;\le\; C\,\exp\left\{-\kappa\,\frac{t}{r^2}\left(1+\frac{L}{t}\right)^{1-\frac{2}{d}}\right\}. \tag{2.22}$$

Moreover, for all $K > 0$, there exists $\delta \in (0,1)$, such that

$$\mathbb{P}\Big[|\mathcal{K}_n(B(v,r),t)| \ge L,\, \mathcal{G}_n^c(r,t,[\delta L/t])\Big] \;\le\; C\,\exp\left\{-K\,\frac{t}{r^2}\left(1+\frac{L}{t}\right)^{1-\frac{2}{d}}\right\}, \tag{2.23}$$

with the convention that $\mathcal{G}_n^c(r,t,0)$ is the empty set.

Proof. We start with the proof of (2.22). First note that if $0 < L < t$, then on the event $\{|\mathcal{K}_n(B(v,r),t)| \ge L\}$, there is $k \le n$ such that $\ell_n(B(S_k+v,r)) > t$, and then (2.20) gives the result. Thus we can assume that $L \ge t$, and in fact also that $L \le n$ (as otherwise there is nothing to prove). Similarly by taking κ small enough, one can assume that $r < n$. Choose now $m^* = 1 + [L/t]$ and note that if κ is small enough, then (2.21) implies that

$$(m^*)^{2/d} \log(4n) \;\le\; \frac{\kappa_0}{18d}\,\frac{t}{r^2},$$

with κ_0 as in (2.20). Now, inclusion (2.19) gives

$$\mathbb{P}\big[|\mathcal{K}_n(B(v,r),t)| > L\big] \le \mathbb{P}\big[\mathcal{H}_n(r,L,m^*)\big] + \mathbb{P}\big[\mathcal{G}_n(r,t,m^*)\big]. \tag{2.24}$$

By (2.20), (noting that $B(x,r)$ can be visited before time n only if $\|x\| < 2n$, and that its cardinality is smaller than $(3r)^d$),

$$\begin{aligned}\mathbb{P}[\mathcal{G}_n(r,t,m^*)] &\le \sum_{\mathcal{C}\in\mathcal{A}(r)\,:\,|\mathcal{C}|=m^*} \mathbb{P}\left[\ell_n\Big(B(\mathcal{C},r)\Big) \ge m^* t\right] \\ &\le C(4n)^{dm^*} \exp\left(-\frac{\kappa_0}{9}\,(m^*)^{1-\frac{2}{d}}\,\frac{t}{r^2}\right) \\ &\le C \exp\left(-\frac{\kappa_0}{18}\,(m^*)^{1-\frac{2}{d}}\,\frac{t}{r^2}\right).\end{aligned} \tag{2.25}$$

Likewise, if (2.21) holds with κ small enough, one has

$$(m^*)^{1+2/d}\log(10n) \;\le\; \frac{\kappa_0}{18\cdot 16d}\,\frac{L}{r^2},$$

and as in (2.25) this leads, for any $m \le m^*$, to

$$\begin{aligned}
\mathbb{P}[\mathcal{H}_n(r,L,m)] \;&\le\; \sum_{\mathcal{C}\in\mathcal{A}(r)\,:\,|\mathcal{C}|=m} \mathbb{P}\Big[\ell_n\Big(B(\mathcal{C},4r)\Big)\ge L\Big] \\
&\le\; C\,(10n)^{dm}\,\exp\left(-\frac{\kappa_0}{9\cdot 16}\,\frac{L}{m^{2/d}r^2}\right) \\
&\le\; C\,\exp\left(-\frac{\kappa_0}{18\cdot 16}\,\frac{L}{m^{2/d}r^2}\right) \\
&\le\; C\,\exp\left(-\frac{\kappa_0}{36\cdot 16}\cdot\frac{1+L/t}{m^{2/d}}\cdot\frac{t}{r^2}\right),
\end{aligned} \tag{2.26}$$

where at the last line we used that $L/t \ge (1+L/t)/2$ (which holds since $L/t\ge 1$). Then (2.22) follows from (2.24), (2.25) and (2.26).

To prove (2.23), note that (2.19) implies that

$$\{|\mathcal{K}_n(B(v,r),t)|\ge L\}\cap\mathcal{G}_n^c(r,t,[\delta L/t]) \;\subset\; \mathcal{H}_n(r,L,[\delta L/t]).$$

Then take δ small enough, so that $\kappa_0/(36\cdot 16\delta^{2/d}) > K$, and use (2.26) with $m=[\delta L/t]$. □

3 Upper Bounds

3.1 Slicing of a trajectory and first estimates

The main idea is that by slicing a trajectory into small pieces, the boundary of the range divides into a sum of boundaries of these pieces minus mutual intersections. The main result here is the following lower bound on $|\partial\mathcal{R}_n|$.

Proposition 3.1. *For any positive integers T and n with $T\le n$, we have*

$$\overline{|\partial\mathcal{R}_n|} \;\ge\; \mathcal{S}_n(T)+\mathcal{M}_n(T)-\frac{4d}{T}\sum_{k=1}^{n}\sum_{z\in\mathcal{R}_k^{++}} G_T(z-S_k)+\mathcal{E}_n(T), \tag{3.1}$$

where $\mathcal{S}_n(T)$ behaves like a sum of n/T independent centered terms bounded by T, and $\mathcal{M}_n(T)$ behaves like a martingale with increments bounded by T. In particular there exists a positive constant c, such that for any positive ε, any $T=T(n)$ going to infinity with n, and n large enough, we have

$$\mathbb{P}\big[\mathcal{S}_n(T)+\mathcal{M}_n(T)\le -\varepsilon n\big]\le T\exp\big(-c\frac{\varepsilon^2 n}{T}\big). \tag{3.2}$$

Moreover, $\mathcal{E}_n(T)=\mathcal{O}(T+n\psi_d(T)/T)$, with $\psi_d(\cdot)$ as in (2.8).

Proof. Recall that for any subsets Λ_1 and Λ_2 in $\mathbb{Z}^d$, one has

$$|\partial(\Lambda_1\cup\Lambda_2)|\ge|\partial\Lambda_1|+|\partial\Lambda_2|-2|\Lambda_1^+\cap\Lambda_2^+|.$$

Now, if we consider N subsets $\{\Lambda_1,\dots,\Lambda_N\}$, by induction one obtains

$$|\partial\big(\Lambda_1\cup\dots\cup\Lambda_N\big)|\ge\sum_{i=1}^{N}|\partial\Lambda_i|-2\sum_{i=2}^{N}|\Lambda_i^+\cap\big(\cup_{j<i}\Lambda_j^+\big)|. \tag{3.3}$$

Applying this to the boundary of the range, this gives for any positive integers $T \leq n$, any $i \in \{-1, \dots, T-2\}$, and with the notation $K_n(T) = \lfloor n/T \rfloor - 2$,

$$|\partial \mathcal{R}_n| \geq \left(\sum_{j=0}^{K_n(T)} |\partial \mathcal{R}(i + jT + 1, i + (j+1)T)| \right) - X_{K_n(T)}(i, T) - \mathcal{O}(T), \tag{3.4}$$

with for $k \leq K_n(T)$

$$X_k(i, T) = 2 \sum_{j=1}^{k} |\mathcal{R}_{i+jT}^+ \cap \mathcal{R}^+(i + jT + 1, i + (j+1)T)|. \tag{3.5}$$

Moreover, the elements of the sum in (3.4) are i.i.d. and distributed like $|\partial \mathcal{R}_{T-1}|$. For simplicity, denote the j-th term of this sum by $U_{i,j}(T)$. Now, using (2.7) leads to

$$\overline{|\partial \mathcal{R}_n|} \geq \left(\sum_{j=0}^{K_n(T)} \overline{U}_{i,j}(T) \right) - X_{K_n(T)}(i, T) - \mathcal{O}\left(T + \frac{n}{T} \psi_d(T) \right). \tag{3.6}$$

Denote by $\mathcal{F}_k(i, T)$ the σ-field generated by $\{S_0, \dots, S_{i+kT}\}$ for $1 \leq k \leq K_n(T)$. Then, for $k \leq K_n(T)$, define $M_k(i, T)$ by

$$X_k(i, T) = M_k(i, T) + 2 \sum_{j=1}^{k} \mathbb{E}\left[|\mathcal{R}_{i+jT}^+ \cap \mathcal{R}^+(i + jT + 1, i + (j+1)T)| \,\Big|\, \mathcal{F}_j(i, T) \right]. \tag{3.7}$$

Note that for each (i, T), $(M_k(i, T),\ k \leq K_n(T))$ is an $(\mathcal{F}_k(i, T),\ k \leq K_n(T))$-martingale. On the other hand, using that

$$|\Lambda_1^+ \cap \Lambda_2^+| \leq 2d\, |\Lambda_1^{++} \cap \Lambda_2| \qquad \text{for all nonempty } \Lambda_1, \Lambda_2 \subset \mathbb{Z}^d,$$

with $\Lambda_1^{++} = (\Lambda_1^+)^+$, we get

$$\begin{aligned} \mathbb{E}\left[|\mathcal{R}_{i+jT}^+ \cap \mathcal{R}^+(i + jT + 1, i + (j+1)T)| \,\Big|\, \mathcal{F}_{i+jT} \right] &\leq 2d \sum_{z \in \mathcal{R}_{i+jT}^{++}} \mathbb{P}_{S_{i+jT}}\left[1 \leq H_z \leq T \right] \\ &\leq 2d \sum_{z \in \mathcal{R}_{i+jT}^{++}} G_T(z - S_{i+jT}). \end{aligned} \tag{3.8}$$

Now, define

$$\mathcal{S}_n(T) = \frac{1}{T} \sum_{i=-1}^{T-2} \sum_{j=0}^{K_n(T)} \overline{U}_{i,j}(T) \quad \text{and} \quad \mathcal{M}_n(T) = -\frac{1}{T} \sum_{i=-1}^{T-2} M_n(i, T),$$

so that using (3.6), (3.7), and (3.8), we establish (3.1). Now assume that $T = T(n)$ satisfies $T(n) \to \infty$, as $n \to \infty$. Then, using that $|U_{i,j}(T)| \leq T$, and Azuma's inequality, we see that for any i, T, and n large enough

$$\mathbb{P}\left[\sum_{j=0}^{K_n(T)} \overline{U}_{i,j}(T) \leq -\frac{\varepsilon n}{2} \right] \leq \exp\left(-c \frac{\varepsilon^2 n}{T} \right), \tag{3.9}$$

for some constant $c > 0$. Therefore a union bound gives

$$\mathbb{P}\left[\frac{1}{T}\sum_{i=-1}^{T-2}\sum_{j=0}^{K_n(T)}\overline{U}_{i,j}(T) \le -\frac{\varepsilon n}{2}\right] \le T\exp\left(-c\frac{\varepsilon^2 n}{T}\right). \tag{3.10}$$

Likewise, one obtains

$$\mathbb{P}\left[M_n(T) \ge \frac{\varepsilon n}{2}\right] \le T\exp\left(-c\,\frac{\varepsilon^2 n}{T}\right). \tag{3.11}$$

Inequality (3.2) follows. □

It remains now to evaluate the probability that the remaining sum in (3.1) be larger than εn, which we will do in the next subsections. But let us already introduce the notation

$$\xi_n(T) = \frac{1}{T}\sum_{k=1}^{n}\sum_{z\in\mathcal{R}_k^{++}} G_T(z - S_k). \tag{3.12}$$

3.2 Case of dimension $d = 3$

The desired estimate on $\xi_n(T)$ will be derived from the following lemma.

Lemma 3.2. *There exist positive constants η, c and C, such that for any $\varepsilon \in (0, \nu_3)$, $n \ge 2$, and any $r \le T$, satisfying*

$$C\,\varepsilon^{-5/3}\,n^{2/3}\,\log n \ \le\ r^3 \ \le\ c\,(\varepsilon^{1/2}T^{3/2} \wedge \varepsilon^{3/2}n), \tag{3.13}$$

the two following statements hold. First

$$\{\xi_n(T) \ge \varepsilon n\} \ \subseteq\ \bigcup_{j=0}^{J}\bigcup_{x\in\mathbb{Z}^3}\left\{\left|\mathcal{K}_n\left(B(x,r), \eta\varepsilon 2^j\, r^3\right)\right| \ge \frac{n}{2^{\frac{3}{2}j}}\right\}, \tag{3.14}$$

with J the smallest integer such that $2^J \ge 1/(\eta\varepsilon)$. Furthermore, there is κ_3 such that for any $0 \le j \le J$,

$$\mathbb{P}\left[\bigcup_{x\in\mathbb{Z}^3}\left\{\left|\mathcal{K}_n\left(B(x,r), \eta\varepsilon 2^j\, r^3\right)\right| \ge \frac{n}{2^{\frac{3}{2}j}}\right\}\right] \le C\,\exp\left(-\kappa_3\cdot 2^{j/6}\,\varepsilon^{2/3}n^{1/3}\right). \tag{3.15}$$

Remark 3.3. The upper bound in (1.10) follows from (3.1), (3.2), and Lemma 3.2. Indeed one can just take $T = \varepsilon^{4/3}n^{2/3}$, and then any r satisfying (3.13), for instance $r = c^{1/3}\varepsilon^{5/6}n^{1/3}$, with c thereof.

Proof. For $j \le J$, let

$$t_j = 8d^2\,\eta\,\varepsilon 2^j r^3 \quad\text{and}\quad L_j = n2^{-3j/2},$$

with η to be determined later. We first observe that, once η is fixed, if $r^3 \le c\varepsilon^{3/2}n$, with c small enough, then $L_j/t_j \ge 1$, for all $j \le J$. Moreover, it is not difficult to see that if in addition the lower bound in (3.13) holds with C large enough, then (2.21) is satisfied for $t = t_j$ and $L = L_j$, for all $j \ge 0$. Thus, with these choices of C and c, (3.15) follows from (2.22) in Lemma 2.4 and a union bound. Now, for (3.14), we have to show that when η is small enough,

$$\bigcap_{0\le j\le J}\bigcap_{x\in\mathbb{Z}^3}\left\{\left|\mathcal{K}_n\left(B(x,r), \eta\varepsilon 2^j r^3\right)\right| < L_j\right\} \subset \{\xi_n(T) < \varepsilon n\}.$$

To see this, we rather work with cubes $Q \in \mathcal{P}_{r'}$, with $r' = r/\sqrt{d} - 2$, and one assume that $r' \geq 1$, since this is a consequence of (3.13) when C is large enough. The reason to introduce r' is that we want

$$\Lambda \subset [-r', r']^d \quad \Longrightarrow \quad \Lambda^{++} \subset B(0, r), \tag{3.16}$$

which holds for this choice of r'. Now for $j \in \{1, \dots, J\}$, define

$$\mathcal{K}_n^*(Q, j) := \left\{ k \in \{1, \dots, n\} \ : \ t_{j-1} < |\mathcal{R}_k^{++} \cap (S_k + Q)| \leq t_j \right\},$$

and

$$\mathcal{K}_n^*(Q, 0) := \left\{ k \in \{1, \dots, n\} \ : \ |\mathcal{R}_k^{++} \cap (S_k + Q)| \leq t_0 \right\}.$$

Note that

$$\ell_k(Q^{++}) \ \geq \ \frac{|\mathcal{R}_k^{++} \cap Q|}{4d^2},$$

and combined with (3.16), this shows that

$$\mathcal{K}_n^*(Q, j) \ \subset \ \mathcal{K}_n\left(B(x, r), \eta\varepsilon 2^j r^3\right) \qquad \text{for all } j \geq 0.$$

Note now that by definition of J, one has $t_J \geq 8r^3 \geq |Q|$, for $Q \in \mathcal{P}_{r'}$ and therefore

$$\{1, \dots, n\} \ \subset \ \bigcup_{j=0}^{J} \mathcal{K}_n^*(Q, j). \tag{3.17}$$

Then, let

$$\Xi_n(r, \eta, \varepsilon) := \bigcap_{0 \leq j \leq J} \bigcap_{Q \in \mathcal{P}_r} \left\{ |\mathcal{K}_n^*(Q, \ j)| < L_j \right\}.$$

For $j \in \{0, \dots, J\}$ and $Q \in \mathcal{P}_{r'}$, let $k^* \in \mathcal{K}_n^*(Q, j)$ be the index which maximizes the sum $\sum_{z \in (-S_k + \mathcal{R}_k^{++}) \cap Q} G_T(z)$, and define

$$\Lambda_Q(j) := (-S_{k^*} + \mathcal{R}_{k^*}^{++}) \cap Q.$$

Note that by definition of $\mathcal{K}_n^*(Q, j)$, one has $|\Lambda_Q(j)| \leq \gamma_j r^3$, with $\gamma_j = 8d^2 \eta \varepsilon 2^j$. Hence, by using Lemma 2.2 and (3.17), we see that on the event $\Xi_n(r, \eta, \varepsilon)$, for some constant $A > 0$, using the trivial bound $|\mathcal{K}_n^*(Q, 0)| \leq n$.

$$\begin{aligned}
\xi_n(T) &\leq \frac{1}{T} \sum_{j=0}^{J} \sum_{Q \in \mathcal{P}_{r'}} \sum_{k \in \mathcal{K}_n^*(Q, j)} \sum_{z \in (-S_k + \mathcal{R}_k^{++}) \cap Q} G_T(z) \\
&\leq \frac{1}{T} \sum_{j=0}^{J} \sum_{Q \in \mathcal{P}_{r'}} |\mathcal{K}_n^*(Q, j)| \sum_{z \in \Lambda_Q(j)} G_T(z) \\
&\leq \frac{A}{T} \left\{ n(\gamma_0^{2/3} r^2 + \gamma_0 T) + \sum_{j=1}^{J} L_j \left(\gamma_j^{2/3} r^2 + \gamma_j T \right) \right\}.
\end{aligned} \tag{3.18}$$

Now, if $r^3 \leq c\varepsilon^{1/2} T^{3/2}$, with c small enough, then $\gamma_j^{2/3} r^2 \leq \gamma_j T$, for all $j \geq 0$. Therefore (3.18) shows that if η is small enough, then

$$\Xi_n(r, \eta, \varepsilon) \ \subset \ \{\xi_n(T) < \varepsilon n\}. \tag{3.19}$$

Together with (3.19) this concludes the proof of the lemma. □

3.3 Case of dimensions $d \geq 5$

The following is an analogue of Lemma 3.2:

Lemma 3.4. *Assume that $d \geq 5$. There are positive constants η, κ_d, c and C, such that for any $\varepsilon \in (0, \nu_d)$, $n \geq 2$, and any $r \leq T$, satisfying*

$$C\varepsilon^{-\frac{d^2-4}{3d}} n^{2/d} \log n \;\leq\; r^d \;\leq\; c\left(\varepsilon^{\frac{(d-2)^2}{6}} T^{d/2} \wedge \varepsilon n\right), \tag{3.20}$$

the two following statements hold. First

$$\{\xi_n(T) \geq \varepsilon n\} \;\subseteq\; \bigcup_{j=0}^{J} \bigcup_{x \in \mathbb{Z}^d} \left\{|\mathcal{K}_n\big(B(x,r), \eta 2^{-j} r^d\big)| \geq \varepsilon 2^{\frac{2,5j}{d-2}} n\right\}, \tag{3.21}$$

with J the smallest integer such that $2^{2,5J/(d-2)} \geq 1/\varepsilon$. Furthermore, for any $0 \leq j \leq J$,

$$\mathbb{P}\left[\bigcup_{x \in \mathbb{Z}^d} \left\{|\mathcal{K}_n\big(B(x,r), \eta 2^{-j} r^d\big)| \geq \varepsilon 2^{\frac{2,5j}{d-2}} n\right\}\right] \leq C \exp\left(-\kappa_d \cdot 2^{\frac{j}{2d}} (\varepsilon n)^{1-\frac{2}{d}}\right). \tag{3.22}$$

Proof. Since the proof is entirely similar to the case of dimension 3, we will not reproduce it here. One just has to use this time $t_j = (8d^2)\eta 2^{-j} r^d$ and $L_j = \varepsilon 2^{2,5j/(d-2)} n$, for $j \leq J$. □

Note that here, for proving the upper bound in (1.12), one can choose $T = \varepsilon^{1+2/d} n^{2/d}$.

3.4 Case of dimension $d = 4$

In this case we obtain a weaker statement.

Lemma 3.5. *Assume that $d = 4$. There are positive constants η, κ_4, c and C, such that for any $\varepsilon \in (0, \nu_4)$, $n \geq 2$, and any $r \leq T$, satisfying*

$$C\,\varepsilon^{-3/2} |\log \varepsilon|^{3/2} n^{1/2} \log n \;\leq\; r^4 \;\leq\; c\,\frac{\varepsilon}{|\log \varepsilon|}(T^2 \wedge n), \tag{3.23}$$

the two following statements hold. First

$$\{\xi_n(T) \geq \varepsilon n\} \;\subseteq\; \bigcup_{j=0}^{J} \bigcup_{x \in \mathbb{Z}^4} \left\{|\mathcal{K}_n\big(B(x,r), \eta 2^{-j} r^4\big)| \geq \frac{2^j}{J} \varepsilon n\right\}, \tag{3.24}$$

with J the smallest integer such that $2^J \geq J/\varepsilon$. Furthermore, for any $0 \leq j \leq J$,

$$\mathbb{P}\left[\bigcup_{x \in \mathbb{Z}^4} \left\{|\mathcal{K}_n\big(B(x,r), \eta 2^{-j} r^4\big)| \geq \frac{2^j}{J} \varepsilon n\right\}\right] \leq C \exp\left(-\kappa_4 \cdot \frac{(\varepsilon n)^{1/2}}{|\log \varepsilon|^{1/2}}\right). \tag{3.25}$$

Proof. Since the proof is entirely similar to the previous cases, we leave the details to the reader. Just for (3.25) one can notice that J is of order $|\log \varepsilon|$ by definition. □

4 Lower bounds

4.1 Lower bound in dimension $d \geq 4$

This is the easiest case. We impose that the walk stays a time $\alpha\varepsilon n$ in a ball of volume $\varepsilon n/2$, with $\alpha = 2/\nu_d$. During this short time, the boundary of the range is necessarily smaller than $\varepsilon n/2$, and the rest of the trajectory cannot make up for this loss. To be more precise, first write

$$|\partial\mathcal{R}_n| \leq |\partial\mathcal{R}_{\alpha\varepsilon n}| + |\partial\mathcal{R}(\alpha\varepsilon n, n)|,$$

which gives after centering and using (2.7),

$$\overline{|\partial\mathcal{R}_n|} \leq \overline{|\partial\mathcal{R}_{\alpha\varepsilon n}|} + \overline{|\partial\mathcal{R}(\alpha\varepsilon n, n)|} + \mathcal{O}(\log n).$$

Let ρ_n be the maximal radius such that $|B(0,\rho_n)| \leq \varepsilon n/2$. On the event $\{\mathcal{R}_{\alpha\varepsilon n} \subset B(0,\rho_n)\}$, one has

$$|\partial\mathcal{R}_{\alpha\varepsilon n}| \leq |B(0,\rho_n)| \leq \frac{\varepsilon n}{2},$$

and therefore also (remembering that $\alpha = 2/\nu_d$)

$$\overline{|\partial\mathcal{R}_n|} \leq -2\varepsilon n + \frac{\varepsilon n}{2} + \overline{|\partial\mathcal{R}(\alpha\varepsilon n, n)|} + \mathcal{O}(\log n).$$

Hence, for n large enough

$$\mathbb{P}\big[\overline{|\partial\mathcal{R}_n|} \leq -\varepsilon n\big] \geq \mathbb{P}\big[\mathcal{R}_{\alpha\varepsilon n} \subset B(0,\rho_n)\big] - \mathbb{P}\Big[\overline{|\partial\mathcal{R}(\alpha\varepsilon n, n)|} \geq \frac{1}{4}\varepsilon n\Big]. \tag{4.1}$$

Then one can use Okada's results [Ok15] which show that upper large deviations have exponentially small probability. More precisely, his results imply in particular that for any $\varepsilon > 0$, there exists a constant $c = c(\varepsilon) > 0$, such that

$$\mathbb{P}\left[\overline{|\partial\mathcal{R}_n|} \geq \varepsilon n\right] \leq \exp(-cn), \tag{4.2}$$

for n large enough. On the other hand it is well known that there exists a constant $\kappa > 0$, such that for any $r \geq 1$ and $n \geq 1$,

$$\mathbb{P}\big[\mathcal{R}_n \subset B(0,r)\big] \geq \exp\left(-\kappa \cdot \frac{n}{r^2}\right). \tag{4.3}$$

Combining (4.1), (4.2) and (4.3) gives the lower bounds in (1.12) and (1.14).

4.2 Lower bound in dimension $d = 3$

This case is more delicate than the previous one. We distinguish two regimes: when ε is small, say $\varepsilon \leq \varepsilon_0$, for some small enough constant ε_0, and when $\varepsilon \in (\varepsilon_0, \nu_3/2)$. The latter case can be handled using the same argument as in higher dimensions, and we omit it, and rather concentrate on the former.

We will see that to shrink the boundary of the range, the random walk needs to localize a time n in a ball of radius ρ_n with ρ_n^3 of order n/ε. The heuristics behind this picture relies on the following key relation already mentioned in Subsection 2.2: for any integers n and m

$$|\partial\mathcal{R}(0, n+m)| \leq |\partial\mathcal{R}(0,n)| + |\partial\mathcal{R}(n, n+m)| - |\mathcal{R}(0,n) \cap \partial\mathcal{R}(n, n+m)|. \tag{4.4}$$

Without constraint, the intersection of two strands $|\mathcal{R}(0,n) \cap \partial\mathcal{R}(n, 2n)|$ is typically of order $\sqrt{n}$ in dimension three, and does not influence the (linear) growth of the boundary of the range.

However, when the walk is localized in a ball of volume n/ε, it likely visits an ε-fraction of all fixed *large* volume – see Proposition 4.1 below for a precise statement – so that the former intersection is typically of order εn when we choose $m = n$, and realize that $|\partial\mathcal{R}(n, 2n)|$ is typically of order n. The fact that this scenario leads to the correct cost follows from the bound (4.3), which was already instrumental in higher dimensions.

Let us give now some details. The proof is based on the following covering result.

Proposition 4.1. *There are positive constants c, C, and ε_0, such that for all $\varepsilon \in (0, \varepsilon_0)$, and n large enough, we have*

$$\mathbb{P}\Big[|\mathcal{R}_n \cap \Lambda| > \varepsilon|\Lambda|\Big] \geq \exp(-\varepsilon^{2/3}n^{1/3}) \qquad \text{for all } \Lambda \subset B\Big(0, c(\frac{n}{\varepsilon})^{1/3}\Big) \text{ with } |\Lambda| \geq \frac{C}{\varepsilon^3}. \tag{4.5}$$

Before proving this result, let us deduce the desired lower bound.

Proof of the lower bound in (1.10)**:** We center the variables of (4.4) using (2.7) to obtain

$$\overline{|\partial\mathcal{R}(0,2n)|} \leq \overline{|\partial\mathcal{R}(0,n)|} + \overline{|\partial\mathcal{R}(n,2n)|} - |\mathcal{R}(0,n) \cap \partial\mathcal{R}(n,2n)| + \mathcal{O}(\sqrt{n}).$$

By invariance of time-inversion and using the Markov property, we see that the intersection term in this inequality is equal in law to the intersection of two independent ranges. Hence Okada's large deviations estimate (4.2) yields, for n large enough,

$$\begin{aligned}
\mathbb{P}[\overline{|\partial\mathcal{R}_{2n}|} \leq -2\varepsilon n] &\geq \mathbb{P}[|\mathcal{R}_n \cap \partial\widetilde{\mathcal{R}}_n| \geq 4\varepsilon n] - 2\mathbb{P}[\overline{|\partial\mathcal{R}_n|} \geq \varepsilon n/2] \\
&\geq \mathbb{P}[|\mathcal{R}_n \cap \partial\widetilde{\mathcal{R}}_n| \geq 4\varepsilon n] - 2\exp(-cn),
\end{aligned}$$

with $\widetilde{\mathcal{R}}_n$ an independent copy of $\mathcal{R}_n$, and $c = c(\varepsilon) > 0$ a constant. We are therefore left with showing the existence of $\kappa > 0$ such that for all ε small enough and n large enough

$$\mathbb{P}\big[|\mathcal{R}_n \cap \partial\widetilde{\mathcal{R}}_n| \geq \varepsilon n\big] \geq \exp(-\kappa\,\varepsilon^{2/3}n^{1/3}). \tag{4.6}$$

To this end, we first claim that localizing a random walk a time n inside $B(0, \rho_n)$ with ρ_n^3 of order n/ε still produces a boundary of the range of order n. Indeed,

$$\mathbb{P}\big[|\partial\widetilde{\mathcal{R}}_n| \geq \nu_3 n/2,\ \widetilde{\mathcal{R}}_n \subset B(0,\rho_n)\big] \geq \mathbb{P}\big[\widetilde{\mathcal{R}}_n \subset B(0,\rho_n)\big] - \mathbb{P}\big[|\partial\widetilde{\mathcal{R}}_n| \leq \nu_3 n/2\big].$$

Then using (4.3), we get

$$\mathbb{P}\big[\widetilde{\mathcal{R}}_n \subset B(0,\rho_n)\big] \geq \exp(-\kappa\,\varepsilon^{2/3}n^{1/3}), \tag{4.7}$$

for some constant $\kappa > 0$ (here we take $\rho_n = c(n/\varepsilon)^{1/3}$, with c as in Proposition 4.1) and from our upper bound, we have for some constant $\kappa' > 0$,

$$\mathbb{P}\Big[|\partial\widetilde{\mathcal{R}}_n| \leq \nu_3 n/2\Big] \leq \exp(-\kappa'\,n^{1/3}).$$

Thus, if we take ε small enough, we see that for some possibly larger constant $\kappa > 0$

$$\mathbb{P}\Big[|\partial\widetilde{\mathcal{R}}_n| \geq \nu_3 n/2,\ \widetilde{\mathcal{R}}_n \subset B(0,\rho_n)\Big] \geq \exp(-\kappa\,\varepsilon^{2/3}n^{1/3}). \tag{4.8}$$

Then by using the independence between $\mathcal{R}_n$ and $\widetilde{\mathcal{R}}_n$, Proposition 4.1, and (4.8) we have

$$\begin{aligned}
\mathbb{P}\Big[|\mathcal{R}_n \cap \partial\widetilde{\mathcal{R}}_n| \geq \varepsilon n\Big] &\geq \sum_{V \subset B(0,\rho_n),\, |V| \geq \nu_3 n/2} \mathbb{P}\big[|\mathcal{R}_n \cap V| \geq \varepsilon n\big] \times \mathbb{P}[\partial\widetilde{\mathcal{R}}_n = V] \\
&\geq \exp(-\varepsilon^{2/3}n^{1/3}) \times \mathbb{P}\big[|\partial\widetilde{\mathcal{R}}_n| \geq \nu_3 n/2,\ \widetilde{\mathcal{R}}_n \subset B(0,\rho_n)\big] \\
&\geq \exp(-(1+\kappa)\,\varepsilon^{2/3}n^{1/3}).
\end{aligned}$$

This proves (4.6). □

We are left with proving Proposition 4.1.

Proof of Proposition 4.1. Define ρ_n and K_n by $\rho_n = c(n/\varepsilon)^{1/3}$, and $K_n = K\varepsilon\rho_n$, with c and K some constants to be chosen later. Observe that a walk covers a fraction ε of a set $\Lambda \subset B(0,\rho_n)$, if it makes K_n excursions between $B(0,2\rho_n)$ and $\partial B(0,5\rho_n)$ before time n. Indeed, each excursion has a chance of order $1/\rho_n$ to visit any given site of $B(0,\rho_n)$ (since ρ_n is the typical distance between such site and $\partial B(0,2\rho_n)$), so that K_n *independent* excursions have a chance of order $K\varepsilon$ to cover any given site. If K is large enough, one deduces well that K_n excursions cover at least a fraction ε of Λ. Now, a Gambler's ruin estimate shows that the probability to hit the ball $B(0,2\rho_n)$ before exiting $B(0,8\rho_n)$ starting from $\partial B(0,5\rho_n)$ is bounded away from 0, so this happens K_n times at a cost $\exp(-K_n)$ which is of the right order.

Finally note that the length of a typical excursion is of order ρ_n^2, so that with high probability K_n of them have to occur before time n, provided c is small enough, and this concludes the heuristics of the proof.

Let us proceed now to the details. We first define the excursions in a standard way as follows. Set $\sigma_0 = 0$, and then for $i \geq 0$,

$$\tau_i := \inf\{t \geq \sigma_i \ : \ S_t \notin B(0,5\rho_n)\},$$

and

$$\sigma_{i+1} := \inf\{t \geq \tau_i \ : \ S_t \in B(0,2\rho_n)\}.$$

Also define τ^* to be the exit time from $B(0,8\rho_n)$.

The number of excursions (after τ_0) from $\partial B(0,2\rho_n)$ to $\partial B(0,5\rho_n)$ is defined to be

$$\mathcal{N} := \sup\ \{k \geq 0 \ : \ \sigma_k < \infty\}.$$

Recall that $K_n = K\varepsilon\rho_n$, for K to be fixed later. Let $\mathcal{G}_\mathcal{N}$ be the sigma-field generated by $\mathcal{N}$, and the starting and end points of the excursions.

$$\mathcal{G}_\mathcal{N} := \sigma\left(\mathcal{N},\, X_{\sigma_i}, X_{\tau_i},\ i \leq \mathcal{N}\right).$$

Define $\Lambda_1 = \Lambda$, and for $i \geq 1$, define

$$\mathcal{R}^{(i)} = \{X_{\sigma_i}, \dots, X_{\tau_i}\} \quad \text{and} \quad \Lambda_{i+1} = \Lambda \setminus (\cup_{j\leq i}\mathcal{R}^{(j)} \cap \Lambda).$$

Then, set

$$X_i := |\mathcal{R}^{(i)} \cap \Lambda_i|\ \mathbb{1}_{\{\sigma_i<\infty\}}. \tag{4.9}$$

Notice that conditionally on $\mathcal{G}_\mathcal{N}$, the variables $\{X_i,\ i \leq \mathcal{N}\}$ are independent of the event $\{\sigma_{K_n} \leq \tau^*\}$. Therefore,

$$\mathbb{P}\left[\sum_{i=1}^{K_n} X_i > \varepsilon|\Lambda|,\, \sigma_{K_n} \leq \tau^*\right] = \mathbb{E}\left[\mathbb{1}_{\{\mathcal{N}\geq K_n\}}\mathbb{P}\Big[\sum_{i=1}^{K_n} X_i > \varepsilon|\Lambda| \ \Big|\ \mathcal{G}_\mathcal{N}\Big]\mathbb{P}\Big[\sigma_{K_n} \leq \tau^* \mid \mathcal{G}_\mathcal{N}\Big]\right]. \tag{4.10}$$

Let $\mathcal{H}_i$ the sigma-field generated by the walk up to the stopping time σ_i. Define

$$M_n = \sum_{i=1}^{\mathcal{N}\wedge K_n} (X_i - \mathbb{E}[X_i \mid \mathcal{H}_i,\, \mathcal{G}_\mathcal{N}]), \tag{4.11}$$

and note that

$$\mathbb{E}[M_n \mid \mathcal{G}_\mathcal{N}] = 0 \quad \text{and} \quad \mathbb{E}[M_n^2 \mid \mathcal{G}_\mathcal{N}] \le 2 \sum_{i=1}^{\mathcal{N}\wedge K_n} \mathbb{E}[X_i^2 \mid \mathcal{G}_\mathcal{N}]. \tag{4.12}$$

Now, for any $i \le \mathcal{N}$, we have

$$\mathbb{E}[X_i \mid \mathcal{H}_i, \mathcal{G}_\mathcal{N}] = \sum_{x\in\Lambda_i} \mathbb{P}[x \in \mathcal{R}^{(i)} \mid \mathcal{G}_\mathcal{N}] = \sum_{x\in\Lambda_i} \mathbb{P}\left[x \in \mathcal{R}^{(i)} \mid X_{\sigma_i}, X_{\tau_i}\right], \tag{4.13}$$

since (conditionally on $\mathcal{G}_\mathcal{N}$), the law of X_i depends on $\mathcal{H}_i$ only through Λ_i. Moreover, as a consequence of the Harnack principle there exists a constant $c_0 > 0$, such that for any $x \in B(0, \rho_n)$,

$$\mathbb{P}[x \in \mathcal{R}^{(i)} \mid X_{\sigma_i}, X_{\tau_i}] \ge c_0\, \mathbb{P}[x \in \mathcal{R}^{(i)} \mid X_{\sigma_i}]. \tag{4.14}$$

Indeed, for any $y \in \partial B(0, 5\rho_n)$, the Markov property yields

$$\mathbb{P}[x \in \mathcal{R}^{(i)} \mid X_{\sigma_i}, X_{\tau_i} = y] = \mathbb{P}[x \in \mathcal{R}^{(i)} \mid X_{\sigma_i}] \times \mathbb{P}_x\left[X_{H_{\partial B(0,5\rho_n)}} = y\right],$$

and now the Harnack principle, see [LL10, Theorem 6.3.9], shows that there exists $c_0 > 0$, independent of x, y and X_{σ_i}, such that almost surely,

$$\mathbb{P}_x[X_{H_{\partial B(0,5\rho_n)}} = y] \ge c_0\, \mathbb{P}[X_{\tau_i} = y \mid X_{\sigma_i}],$$

which proves (4.14). Then for any $z \in \partial B(0, 2\rho_n)$, by using standard estimates on the Green's function (see [LL10, Theorem 4.3.1]), we can write

$$\begin{aligned}
\mathbb{P}[x \in \mathcal{R}^{(i)} \mid X_{\sigma_i} = z] &\ge \mathbb{P}_z[H_x < \infty] - \sup_{y\in\partial B(0,5\rho_n)} \mathbb{P}_y[H_x < \infty] \\
&\ge \frac{c_{\mathrm{Gr}}}{\|x - z\|} - \sup_{y\in\partial B(0,5\rho_n)} \frac{c_{\mathrm{Gr}}}{\|y - x\|} - \mathcal{O}\left(\frac{1}{\rho_n^2}\right) \\
&\ge \frac{c_{\mathrm{Gr}}}{3\rho_n} - \frac{c_{\mathrm{Gr}}}{4\rho_n} - \mathcal{O}\left(\frac{1}{\rho_n^2}\right) \\
&\ge \frac{c_{\mathrm{Gr}}}{20\rho_n},
\end{aligned}$$

for some constant $c_{\mathrm{Gr}} > 0$, and n large enough. Combining this with (4.13) and (4.14), we obtain that for any $i \le \mathcal{N}$, almost surely

$$\mathbb{E}[X_i \mid \mathcal{H}_i, \mathcal{G}_\mathcal{N}] \ge \frac{c_1}{\rho_n} |\Lambda_i|, \tag{4.15}$$

for some constant $c_1 > 0$. Now, choose $K := 4/c_1$, and use the previous inequality. This shows that on the event $\{\mathcal{N} \ge K_n\}$,

$$\sum_{i=1}^{\mathcal{N}} \mathbb{E}[X_i \mid \mathcal{H}_i, \mathcal{G}_\mathcal{N}] \ge 4\varepsilon |\Lambda_{K_n}|.$$

Since $|\Lambda_{K_n}| = |\Lambda| - \sum_{i=1}^{K_n-1} X_i$, we have for any $\varepsilon \le 1/2$, on the event $\{\mathcal{N} \ge K_n\}$,

$$\begin{aligned}
\mathbb{P}\left[\sum_{i=1}^{K_n} X_i \le \varepsilon|\Lambda| \;\Big|\; \mathcal{G}_{\mathcal{N}}\right] &= \mathbb{P}\left[\sum_{i=1}^{K_n} X_i \le \varepsilon|\Lambda|,\, |\Lambda_{K_n}| \ge |\Lambda|/2 \;\Big|\; \mathcal{G}_{\mathcal{N}}\right] \\
&\le \mathbb{P}[|M_n| \ge \varepsilon|\Lambda| \mid \mathcal{G}_{\mathcal{N}}] \\
&\le \frac{\mathbb{E}\left[M_n^2 \mid \mathcal{G}_{\mathcal{N}}\right]}{\varepsilon^2|\Lambda|^2} \\
&\le \frac{2}{\varepsilon^2|\Lambda|^2}\sum_{i=1}^{K_n}\mathbb{E}\left[X_i^2 \mid \mathcal{G}_{\mathcal{N}}\right],
\end{aligned} \tag{4.16}$$

using (4.12) at the last line. Moreover, for any $i \le \mathcal{N}$, using again the Harnack inequality at the third line, and (2.9) at the last line,

$$\begin{aligned}
\mathbb{E}\left[X_i^2 \mid \mathcal{H}_i,\, \mathcal{G}_{\mathcal{N}}\right] &= \sum_{(z,z')\in\Lambda_i\times\Lambda_i}\mathbb{P}\left[z\in\mathcal{R}^{(i)},\, z'\in\mathcal{R}^{(i)} \mid X_{\sigma_i},\, X_{\tau_i}\right] \\
&\le 2\sum_{(z,z')\in\Lambda_i\times\Lambda_i}\mathbb{P}\left[z\in\mathcal{R}^{(i)},\, z'\in\mathcal{R}^{(i)},\, H_z<H_{z'} \mid X_{\sigma_i},\, X_{\tau_i}\right] \\
&\le C_0\sum_{(z,z')\in\Lambda_i\times\Lambda_i}\mathbb{P}\left[z\in\mathcal{R}^{(i)},\, z'\in\mathcal{R}^{(i)},\, H_z<H_{z'} \mid X_{\sigma_i}\right] \\
&\le C_0\sum_{(z,z')\in\Lambda_i\times\Lambda_i}\mathbb{P}_{X_{\sigma_i}}\left[H_z<H_{z'}<\infty\right] \le C_0\sum_{z,z'\in\Lambda_i}\frac{1}{\rho_n\left(\|z-z'\|+1\right)},
\end{aligned} \tag{4.17}$$

for some constant $C_0 > 0$. Now it is not difficult to see that for any set Λ, and any $y \in \mathbb{Z}^d$, one has $\sum_{z\in\Lambda} 1/(\|z-y\|+1) = \mathcal{O}(|\Lambda|^{2/3})$, with an implicit constant which is uniform in Λ and y, since the worst case is easily seen to be reached when Λ is a ball and y is the center of this ball. Therefore (4.17) gives

$$\mathbb{E}\left[X_i^2 \mid \mathcal{H}_i,\, \mathcal{G}_{\mathcal{N}}\right] \le \frac{C_0'}{\rho_n}|\Lambda|^{5/3},$$

for some constant $C_0' > 0$. Together with (4.16) if we assume that $|\Lambda| \ge C/\varepsilon^3$, with $C := (4KC_0')^3$, one deduces that, on the event $\{\mathcal{N} \ge K_n\}$,

$$\mathbb{P}\left[\sum_{i=1}^{K_n} X_i \le \varepsilon|\Lambda| \;\Big|\; \mathcal{G}_{\mathcal{N}}\right] \le \frac{2KC_0'}{\varepsilon|\Lambda|^{1/3}} \le \frac{1}{2}.$$

Coming back to (4.10) we get

$$\begin{aligned}
\mathbb{P}\left[\sum_{i=1}^{K_n} X_i > \varepsilon|\Lambda|,\, \sigma_{K_n}\le\tau^*\right] &\ge \frac{1}{2}\mathbb{P}\Big[\sigma_{K_n}\le\tau^*\Big] \\
&\ge \frac{1}{2}\exp(-\kappa\cdot Kc^{1/3}\cdot\varepsilon^{2/3}n^{1/3}),
\end{aligned} \tag{4.18}$$

for some constant $\kappa > 0$, since for any $x \in \partial B(0, 5\rho_n)$ the probability starting from x to hit the ball $B(0, 2\rho_n)$ before τ^* is bounded away from 0 (see [L13, Proposition 1.5.10]). Now, observe that

$$\begin{aligned}
\mathbb{P}[|\mathcal{R}_n\cap\Lambda| \ge \varepsilon|\Lambda|] &\ge \mathbb{P}\left[\sum_{i=1}^{K_n} X_i > \varepsilon|\Lambda|,\, \sigma_{K_n}<\tau^*,\, \tau_{K_n}\le\tau^*\wedge n\right] \\
&\ge \mathbb{P}\left[\sum_{i=1}^{K_n} X_i > \varepsilon|\Lambda|,\, \sigma_{K_n}\le\tau^*\right] - \mathbb{P}[n<\tau_{K_n}\le\tau^*].
\end{aligned} \tag{4.19}$$

We now bound the last term in the right hand side of (4.19). To this end we let

$$\xi_i = (\sigma_{i+1} \wedge \tau^* - \tau_i)1\!\!1_{\{\tau_i<\infty\}} \quad \text{and} \quad \tilde{\xi}_i = (\tau_i - \sigma_i)1\!\!1_{\{\sigma_i<\infty\}} \quad \text{for } i \geq 0,$$

and observe that for any integer $k \geq 1$, on the event $\{\tau_k \leq \tau^*\}$, we have

$$\tau_k = \tilde{\xi}_k + \sum_{i=0}^{k-1}(\xi_i + \tilde{\xi}_i). \tag{4.20}$$

Then we claim that the random variables (ξ_i/ρ_n^2) and $(\tilde{\xi}_i/\rho_n^2)$ are (up to a multiplicative constant) dominated by i.i.d. geometric random variables. To see this, one can use that $(\|S_n\|^2 - n)_{n\geq 0}$ is a martingale. It implies, using also the optional stopping time theorem, that for any $x \in B(0, 8\rho_n)$,

$$(8\rho_n + 1)^2 \geq \mathbb{E}_x\left[\|S_{\tau^*\wedge 200\rho_n^2}\|\right] \geq \mathbb{E}_x\left[\tau^* \wedge 200\rho_n^2\right] \geq 200\rho_n^2\, \mathbb{P}_x[\tau^* \geq 200\rho_n^2].$$

Hence for any $x \in B(0, 8\rho_n)$,

$$\mathbb{P}_x[\tau^* < 200\rho_n^2] \geq 1/2.$$

We deduce that there exist (G_i) and $(\tilde{G}_i)$, i.i.d. geometric random variables with parameter 1/2, such that

$$\xi_i/(200\rho_n^2) \leq G_i \quad \text{and} \quad \tilde{\xi}_i/(200\rho_n^2) \leq \tilde{G}_i \quad \text{for all } i \geq 0.$$

Now assume that $c < 1/(2000K)$, so that $1/(200Kc) \geq 10$. Then with (4.20), and using Markov's exponential inequality, this gives

$$\begin{aligned}
\mathbb{P}[n < \tau_{K_n} \leq \tau^*] &\leq \mathbb{P}\left[\sum_{i=0}^{K_n}(G_i + \tilde{G}_i) \geq \frac{n}{200\rho_n^2}\right] \\
&\leq \mathbb{P}\left[\frac{1}{K_n}\sum_{i=0}^{K_n}(G_i + \tilde{G}_i) \geq \frac{1}{200Kc}\right] \\
&\leq \exp(-\kappa \cdot c^{-2/3} \cdot \varepsilon^{2/3} n^{1/3}),
\end{aligned}$$

for some constant $\kappa > 0$. Finally by taking c small enough and using (4.18) and (4.19) this proves the desired lower bound. □

5 Bound on the capacity

In this Section, we prove Proposition 1.7, and then use it for proving (1.9), (1.11) and (1.13).

We first recall some important property of the capacity of a ball (see (2.16) in [L13] for a stronger statement): there is a positive constant C_{cap}, such that for all $x \in \mathbb{Z}^d$ and $r > 0$

$$\text{cap}(B(x,r)) \leq C_{\text{cap}}\, |B(x,r)|^{1-2/d}. \tag{5.1}$$

Proof of Proposition 1.7. First, consider the case when $\mathcal{I}_d(B(\mathcal{C}, r))$ is bounded above by some constant $\overline{C}$. Then by using (2.20) (see [AC07, Lemma 1.2]), we obtain

$$\begin{aligned}
\mathbb{P}\big[\ell_n(B(x,r)) \geq t \;\text{ for all } x \in \mathcal{C}\big] &\leq \mathbb{P}\big[\ell_n(B(\mathcal{C},r)) \geq t|\mathcal{C}|\big] \\
&\leq C \exp\Big(-\kappa_0 \cdot \frac{|\mathcal{C}|t}{|B(\mathcal{C},r)|^{2/d}}\Big),
\end{aligned} \tag{5.2}$$

which proves the desired inequality (1.16), with $\kappa = \kappa_0/\overline{C}$. Henceforth, we assume that

$$\frac{\mathrm{cap}(B(\mathcal{C},r))}{|\mathcal{C}|^{1-2/d}\cdot r^{d-2}} > 2C_{\mathrm{cap}}. \tag{5.3}$$

We need now a few intermediate lemmas. We first show that spending a time t in a ball $B(x,r)$ implies making order t/r^2 excursions from $\partial B(x,r)$ to $\partial B(x,2r)$, with high probability. So as in the proof of Proposition 4.1, set $\sigma_0(x,r)=0$, and for $j\geq 0$,

$$\tau_j(x,r) := \inf\{t\geq \sigma_j(x,r)\ :\ S_t\in\partial B(x,2r)\},$$

and

$$\sigma_{j+1}(x,r) := \inf\{t\geq \tau_j(x,r)\ :\ S_t\in B(x,r)\}.$$

Then for $n\geq 1$, let

$$N_n(x,r) := \sup\{j\ :\ \sigma_j(x,r)\leq n\}.$$

Lemma 5.1. *There exists a constant $c>0$, such that for any $r\geq 1$, $\mathcal{C}\in\mathcal{A}(r)$, and $n\geq 1$,*

$$\mathbb{P}\left[\ell_n(B(x,r))\geq t \text{ and } N_n(x,r)\leq ct/r^2 \text{ for all } x\in\mathcal{C}\right]\leq \exp\big(-c\,|\mathcal{C}|\,\frac{t}{r^2}\big).$$

Proof. Exactly as in the proof of Proposition 4.1, one can see that there exist i.i.d. geometric random variables $\{G_j(x),\ j\in\mathbb{N}, x\in\mathcal{C}\}$, with parameter 1/2, such that for any $j\in\mathbb{N}$, and $x\in\mathcal{C}$,

$$(\tau_j(x,r)-\sigma_j(x,r))\mathbb{1}_{\{\sigma_j(x,r)<\infty\}}\leq 100r^2G_j(x).$$

Moreover, for any $x\in\mathcal{C}$,

$$\ell_n(B(x,r))\leq \sum_{j=0}^{N_n(x,r)}(\tau_j(x,r)-\sigma_j(x,r)).$$

Indeed, we include the 0-th order term $\tau_0(x,r)$ to cover the case where the walk starts in $B(x,2r)$. Therefore, by standard large deviation estimates, for any $\delta\in(0,1/400)$, there is $\gamma>0$ (independent of n,t and r), such that

$$\begin{aligned}\mathbb{P}\left[\ell_n(B(x,r))\geq t,\ N_n(x,r)\leq \delta\frac{t}{r^2}\text{ for all } x\in\mathcal{C}\right] &\leq \mathbb{P}\left[\sum_{j=0}^{\delta t/r^2}G_j(x)\geq \frac{t}{100r^2}\text{ for all } x\in\mathcal{C}\right]\\ &\leq \prod_{x\in\mathcal{C}}\mathbb{P}\left[\sum_{j=0}^{\delta t/r^2}G_j(x)\geq\frac{1}{200\delta}\frac{2\delta t}{r^2}\right]\\ &\leq \exp(-\gamma|\mathcal{C}|\frac{t}{r^2}).\end{aligned}$$

The result follows as we take $c=\min(\delta,\gamma)$. □

The next result relates the probability to never return to $B(\mathcal{C},r)$ starting from a boundary point of $B(x,2r)$ to the probability of the same event starting from a uniformly chosen site on $\partial B(x,r)$.

Lemma 5.2. *With the notation and hypothesis of Proposition 1.7, there exists $\theta>0$ (independent of r and $\mathcal{C}\in\mathcal{A}(r)$), such that for any $x\in\mathcal{C}$ and $z\in\partial B(x,2r)$,*

$$\mathbb{P}_z[H_{B(\mathcal{C},r)}=+\infty]\geq\theta\cdot\frac{1}{r^{d-2}}\sum_{y\in\partial B(x,r)}\mathbb{P}_y[H^+_{B(\mathcal{C},r)}=+\infty].$$

Proof. We first argue that there exists $\theta_1 > 0$, such that for any $z' \in \partial B(x, 2r)$

$$\mathbb{P}_z[H_{B(\mathcal{C},r)} = +\infty] \;\geq\; \theta_1 \cdot \mathbb{P}_{z'}[H_{B(\mathcal{C},r)} = +\infty]. \tag{5.4}$$

To this end, let τ be the exit time from the annulus $B(x, 5r/2)\backslash B(x, 3r/2)$. By using the optional stopping time theorem, we obtain

$$\mathbb{P}_z[H_{B(\mathcal{C},r)} = +\infty] = \sum_v \mathbb{P}_z[S_\tau = v]\,\mathbb{P}_v[H_{B(\mathcal{C},r)} = +\infty].$$

Then, Harnack's inequality (see [LL10, Theorem 6.3.9]) shows that there exists a constant θ_1, such that for all $z' \in \partial B(x, 2r)$ and v,

$$\mathbb{P}_z[S_\tau = v] \;\geq\; \theta_1 \cdot \mathbb{P}_{z'}[S_\tau = v].$$

Inequality (5.4) follows. Now using Proposition 1.5.10 in [L13], we see that there exists $\theta_2 > 0$, such that

$$\mathbb{P}_y[H_{\partial B(x,2r)} < H^+_{\partial B(x,r)}] \;\leq\; \theta_2\, r^{-1}, \tag{5.5}$$

for all $y \in \partial B(x, r)$. Next, using Lemma 6.3.7 and Proposition 6.4.4 in [LL10], we deduce that there exists $\theta_3 > 0$ satisfying for all $y \in \partial B(x, r)$ and $w \in \partial B(x, 2r)$,

$$\mathbb{P}_y[S_{\tau'} = w] \;\leq\; \theta_3 \cdot \frac{1}{r^d} \qquad \text{with } \tau' = H^+_{\partial B(x,2r) \cup B(x,r)}. \tag{5.6}$$

Then by using that the size of the boundary of a ball of radius r or $2r$ is of order r^{d-1}, we deduce that

$$\frac{1}{|\partial B(x, 2r)|} \sum_{w \in \partial B(x,2r)} \mathbb{P}_w[H_{B(\mathcal{C},r)} = +\infty] \;\geq\; \theta_4 \cdot \frac{1}{r^{d-2}} \sum_{y \in \partial B(x,r)} \mathbb{P}_y[H^+_{B(\mathcal{C},r)} = +\infty].$$

Combining this with (5.4) gives the result. □

Lemma 5.3. *For $x \in \mathcal{C}$, define*

$$q_x = \exp\left(- \min_{z \in \partial B(x,2r)} \mathbb{P}_z[H_{B(\mathcal{C},r)} = +\infty]\right).$$

Then, for any $x_0 \in \mathcal{C}$, any $y \in \partial B(x_0, 2r)$, and any integers n and $(n_x, x \in \mathcal{C})$,

$$\mathbb{P}_y\left[N_n(x, r) = n_x \;\; \text{for all } x \in \mathcal{C}\right] \;\leq\; \frac{q_{x_0}}{\min_{x \in \mathcal{C}} q_x} \prod_{x \in \mathcal{C}} q_x^{n_x}. \tag{5.7}$$

As a consequence, we have

$$\mathbb{P}_y\left[N_n(x, r) \geq n_x \;\; \text{for all } x \in \mathcal{C}\right] \;\leq\; n^{|\mathcal{C}|}\, \frac{q_{x_0}}{\min_x q_x} \prod_{x \in \mathcal{C}} q_x^{n_x}. \tag{5.8}$$

Proof. We only need to prove (5.7), since (5.8) follows immediately using that in any ball $B(x, r)$, there can be at most n excursions before time n. So we prove (5.7) by induction on

$$N := \sum_{x \in \mathcal{C}} n_x.$$

First assume that $N=0$. Then we just bound the left-hand side in (5.7) by 1, and observe that the right-hand side is equal to $q_{x_0}/\min_x q_x$, which is well always larger than or equal to one. Next assume that $N=1$. In this case the left-hand side in (5.7) is bounded above by

$$\mathbb{P}_y[H_{B(\mathcal{C},r)} \le n] \le 1 - \mathbb{P}_y[H_{B(\mathcal{C},r)} = +\infty] \le \exp(-\mathbb{P}_y[H_{B(\mathcal{C},r)} = +\infty]) \le q_{x_0},$$

which proves the case $N=1$. We now prove the induction step, and assume for this that $N \ge 2$. We write $P_{n,y}(n_x,\ x\in\mathcal{C})$ for the term on the left-hand side of (5.7), and define

$$\tau = \inf\{t \ge H_{B(\mathcal{C},r)} \ :\ S_t \in \partial B(\mathcal{C},2r)\}.$$

Now, by the Markov property

$$P_{n,y}(n_x,\ x\in\mathcal{C}) = \sum_{x'\in\mathcal{C}} \sum_{k=1}^{n} \sum_{z\in\partial B(x',2r)} \mathbb{P}_y\left[\tau = k,\ S_\tau = z\right]\ P_{n-k,z}(n_x - 1\!\!1_{\{x=x'\}},\ x\in\mathcal{C}),$$

where the first sum is over all centers x', such that $n_{x'} \ge 1$. By using the induction hypothesis, this gives

$$P_{n,y}(n_x,\ x\in\mathcal{C}) \le \frac{\prod_{x\in\mathcal{C}} q_x^{n_x}}{\min_{x\in\mathcal{C}} q_x} \times \mathbb{P}_y[H_{B(\mathcal{C},r)} \le n],$$

and we conclude as in the case $N=1$. $\square$

We are now in position to give the proof of Proposition 1.7.

Proof of Proposition 1.7. Let k_0 be the integral part of $\mathrm{cap}(B(\mathcal{C},r))/(2C_{\mathrm{cap}}r^{d-2})$, with C_{cap} as in (5.1). Note that by (5.3), one has $k_0 \ge 1$, and therefore

$$\frac{1}{4C_{\mathrm{cap}}} \cdot \frac{\mathrm{cap}(B(\mathcal{C},r))}{r^{d-2}} \ \le\ k_0 \ \le\ \frac{1}{2C_{\mathrm{cap}}} \cdot \frac{\mathrm{cap}(B(\mathcal{C},r))}{r^{d-2}}. \tag{5.9}$$

Now, for $x\in\mathcal{C}$ define

$$\mathrm{cap}_x(B(\mathcal{C},r)) := \sum_{y\in\partial B(x,r)} \mathbb{P}_y[H^+_{B(\mathcal{C},r)} = +\infty].$$

Using (5.1), one can see that for any $x\in\mathcal{C}$,

$$\mathrm{cap}_x(B(\mathcal{C},r)) \le \mathrm{cap}(B(x,r)) \le C_{\mathrm{cap}}\ r^{d-2}.$$

Moreover, by definition,

$$\mathrm{cap}(B(\mathcal{C},r)) = \sum_{x\in\mathcal{C}} \mathrm{cap}_x(B(\mathcal{C},r)).$$

Therefore for any subset $I\subset\mathcal{C}$ with cardinality $m-k_0$, one has

$$\sum_{x\in I} \mathrm{cap}_x(B(\mathcal{C},r)) \ge \mathrm{cap}(B(\mathcal{C},r)) - k_0 C_{\mathrm{cap}}\ r^{d-2} \ge\ \frac{\mathrm{cap}(B(\mathcal{C},r))}{2}. \tag{5.10}$$

Next let

$$E = \{\ell_n(B(x,r)) \ge t \text{ for all } x\in\mathcal{C}\},$$

and

$$F = \left\{N_n(x,r) \ge ct/r^2 \text{ for at least } |\mathcal{C}| - k_0 \text{ elements } x\in\mathcal{C}\right\},$$

with c as in Lemma 5.1. By using Lemma 5.1 we get for some positive constant C,

$$\mathbb{P}\left[E \cap F^c\right] \ \le \ C \cdot \binom{|\mathcal{C}|}{k_0} \cdot \exp\left(-c \cdot k_0 \cdot \frac{t}{r^2}\right). \tag{5.11}$$

On the other hand Lemma 5.2, (5.8) and (5.10) show that for some positive constant C, with θ as in Lemma 5.2,

$$\mathbb{P}[F] \ \le \ C \cdot \binom{|\mathcal{C}|}{k_0} \cdot n^{|\mathcal{C}|-k_0} \exp\left(-\theta c \cdot \frac{\mathrm{cap}(B(\mathcal{C},r))}{2} \cdot \frac{t}{r^d}\right). \tag{5.12}$$

Then, (5.9), (5.11) and (5.12) prove the proposition. □

Proof of (1.9)**,** (1.11) **and** (1.13)**.** We only give the proof of (1.9) which corresponds to the case of dimension $d = 3$. The other cases are similar and we leave the details to the reader.

Define

$$E_n(\varepsilon) = \{|\partial \mathcal{R}_n| - \mathbb{E}[|\partial \mathcal{R}_n|] \le -\varepsilon n\}, \tag{5.13}$$

and

$$F_n(\alpha, A, \varepsilon) = \{\exists \Lambda \subset \mathbb{Z}^3 \ : \ \ell_n(\Lambda) \ge \alpha n, \, \mathcal{I}_d(\Lambda) \le A, \, \frac{n}{A\varepsilon} \le |\Lambda| \le \frac{An}{\varepsilon}\}. \tag{5.14}$$

We have to prove the existence of $\alpha \in (0,1)$ and $A > 0$, such that for any $\varepsilon \in (0, \nu_3/2)$,

$$\mathbb{P}[F_n(\alpha, A, \varepsilon)^c \mid E_n(\varepsilon)] \to 0, \tag{5.15}$$

as $n \to \infty$. Our strategy for this is to define five sets $E_{n,1}, \dots, E_{n,5}$ with $E_{n,1} = E_n(\varepsilon)$, $E_{n,5} = F_n(\alpha, A, \varepsilon)$, and

$$\mathbb{P}[E_{n,j} \cap E_{n,j+1}^c] = o\big(p_n\big) \qquad \text{for } j = 1, \dots, 4, \tag{5.16}$$

with $p_n = \mathbb{P}[E_{n,1}] = \mathbb{P}[E_n(\varepsilon)]$. This will well prove (5.15), since

$$E_{n,1} \cap E_{n,5}^c \ \subset \ \bigcup_{j=1}^{4} E_{n,j} \cap E_{n,j+1}^c.$$

Set $T = n^{5/9}$. Since $T = o(n^{2/3})$, Proposition 3.1 and the lower bound in (1.10) show that

$$\mathbb{P}\left[E_{n,1} \cap E_{n,2}^c\right] = o(p_n) \qquad \text{with} \quad E_{n,2} = \left\{\xi_n(T) \ge \frac{\varepsilon n}{100}\right\}. \tag{5.17}$$

Now take η as in Lemma 3.2 and define

$$r := n^{1/4} \quad \text{and} \quad t := \frac{\eta \varepsilon}{100} r^3, \tag{5.18}$$

so that (3.13) is satisfied for n large enough. Let j_0 be the smallest integer such that $2^{j_0/6} \ge 25\bar{\kappa}_3/\kappa_3$, with $\bar{\kappa}_3$ and κ_3 as in (1.10) and (3.15) respectively. Then (3.14) and (3.15) show that

$$\mathbb{P}\left[E_{n,2} \cap E_{n,3}^c\right] = o(p_n) \qquad \text{with} \quad E_{n,3} = \bigcup_{x \in \mathbb{Z}^d} \{|\mathcal{K}_n(B(x,r),t)| \ge \frac{n}{2^{3j/2}}\}. \tag{5.19}$$

Observe next that with our choice (5.18) and $L = n/2^{3j/2}$,

$$\frac{t}{r^2}\left(1 + \frac{L}{t}\right)^{1/3} \ge \frac{t^{2/3}L^{1/3}}{r^2} = \eta^{2/3} 2^{-j_0/2} 100^{-2/3} \cdot \varepsilon^{2/3} n^{1/3}.$$

Therefore a union bound and (2.23) show that, with the notation of Lemma 2.4,

$$\mathbb{P}\left[E_{n,3}\cap E_{n,4}^c\right]=o(p_n)\qquad\text{with}\quad E_{n,4}=\mathcal{G}_n(r,t,[\delta L/t]),\tag{5.20}$$

where δ is the constant associated to $K=2\cdot 100^{2/3}\bar{\kappa}_3\eta^{-2/3}2^{j_0/2}$. Finally let $\alpha=(\delta/2)\cdot 2^{-3j_0/2}$, so that $t[\delta L/t]\geq \delta L/2=\alpha n$, for n large enough. Note that if Λ is the union of $m_\delta=[\delta L/t]$ disjoint balls of radius r, then at least for n large enough, we have

$$m_\delta(r/\sqrt{3})^3\ \leq\ |\Lambda|\ \leq\ m_\delta(3r)^3\ \leq\ \frac{10^4\alpha}{\eta}\,\frac{n}{\varepsilon}.$$

Therefore, if we take

$$A=\max\left(\frac{2^{j_0+1}\sqrt{3}\,\bar{\kappa}_3}{\delta^{1/3}\eta^{2/3}\kappa},\ \frac{10^4\alpha}{\eta}\right),$$

with κ as in Proposition 1.7, then this proposition and a union bound show that $\mathbb{P}(E_{n,4}\cap E_{n,5}^c)=o(p_n)$, using also that $L/t=\mathcal{O}(n^{1/4})$, to remove all the combinatorial terms, and (2.20) for the lower bound on the volume of Λ. Combining this with (5.17), (5.19) and (5.20) gives (5.16), and concludes the proof of (1.9). □

6 Proof of Theorem 1.8

We start with the proof of the first statement. Define the positive part of a real x as $x_+=\max(x,0)$ and $x_-=x_+-x=max(-x,0)$. Define now

$$Z_n^-(\beta)=\mathbb{E}\left[\exp\left(-\frac{\beta}{n^{2/d}}\left(|\partial\mathcal{R}_n|-\mathbb{E}[|\partial\mathcal{R}_n|]\right)_+\right)\right],$$

and

$$Z_n^+(\beta)=\mathbb{E}\left[\exp\left(\frac{\beta}{n^{2/d}}\left(|\partial\mathcal{R}_n|-\mathbb{E}[|\partial\mathcal{R}_n|]\right)_-\right)\right].$$

Since $e^{-x}=e^{x_-}+e^{-x_+}-1$, we have the following simple relations (using also Jensen's inequality)

$$Z_n(\beta)=Z_n^+(\beta)+Z_n^-(\beta)-1,\qquad 0<Z_n^-(\beta)\leq 1,\quad\text{and}\qquad 1\leq Z_n(\beta)\leq Z_n^+(\beta).$$

Thus

$$F^+(\beta)=\limsup\frac{1}{n^{1-2/d}}\,\log Z_n^+(\beta)\quad\text{and}\quad F^-(\beta)=\liminf\frac{1}{n^{1-2/d}}\,\log Z_n^+(\beta).$$

The result follows since Z_n^+ is nondecreasing in β by construction.

For the second statement, note first that

$$Z_n(\beta)\geq\exp\left(\frac{\beta\nu_d}{4}\,n^{1-2/d}\right)\,\mathbb{P}\left[|\partial\mathcal{R}_n|-\mathbb{E}[|\partial\mathcal{R}_n|]\leq-\frac{\nu_d}{4}n\right].$$

Then the fact that β_d^- is finite follows from the lower bounds in (1.10), (1.12) and (1.14). On the other hand, using this time the upper bounds in the latter inequalities, one get

$$\begin{aligned}Z_n(\beta)\ &\leq\ e^{\beta\varepsilon n^{1-2/d}}+\sum_{k=1}^{|\log_2\varepsilon|}\exp(\beta 2^{-k+1}n^{1-2/d})\,\mathbb{P}\left[|\partial\mathcal{R}_n|-\mathbb{E}[|\partial\mathcal{R}_n|]\leq-2^{-k}n\right]\\ &\leq\ e^{\beta\varepsilon n^{1-2/d}}+\sum_{k=1}^{|\log_2\varepsilon|}\exp\left((\beta 2^{-k+1}-\underline{\kappa}_d\frac{2^{-2k/3}}{\sqrt{k\log 2}})\cdot n^{1-2/d}\right).\end{aligned}$$

So we see that for β small enough, one has for all $\varepsilon > 0$,

$$Z_n(\beta) \le e^{\varepsilon n^{1-2/d}} + C|\log \varepsilon| \exp(-c\varepsilon^{3/4} n^{1-2/d}),$$

for some positive constants c and C. The fact that β_d^+ is positive follows.

Let us prove the third statement now. So let $\beta < \beta_d^+$, $\alpha \in (0,1)$, and $A > 0$ be given. Since $r_n \ge n^{2/d}(\log n)^2$, Proposition 1.7 shows (see the discussion after its statement in the introduction, in particular (1.18)) that there exists a positive constant $c = c(\alpha, A)$, such that

$$\mathbb{P}\Big[\exists \mathcal{C} \in \mathcal{A}(r_n, An) \,:\, \ell_n(B(x, r_n)) \ge \alpha \frac{n}{|\mathcal{C}|} \text{ for all } x \in \mathcal{C}\Big] \le \exp(-cn^{1-2/d}),$$

at least for n large enough. Then, we have

$$\mathbb{Q}_n^\beta \left[\exists \mathcal{C} \in \mathcal{A}(r_n, An) : \ell_n(B(x, r_n)) \ge \alpha \frac{n}{|\mathcal{C}|} \, \forall x \in \mathcal{C}\right] \le \frac{e^{-\frac{c}{2} n^{1-\frac{2}{d}}}}{Z_n(\beta)} + \mathbb{Q}_n^\beta \left[\overline{|\partial \mathcal{R}_n|} \le -\frac{c}{2\beta} n\right]. \quad (6.1)$$

The first term on the right-hand side goes to 0, since $Z_n(\beta) \ge 1$. For the second term, define $\beta' = (\beta + \beta_d^+)/2$, and note that for any $\varepsilon > 0$,

$$\begin{aligned} \mathbb{Q}_n^\beta \left[\overline{|\partial \mathcal{R}_n|} \le -\varepsilon n\right] \le & \frac{Z_n(\beta')}{Z_n(\beta)} \cdot \mathbb{Q}_n^{\beta'} [\mathbb{1}_{\{\overline{|\partial \mathcal{R}_n|} \le -\varepsilon n\}} \exp\big(-\frac{(\beta - \beta')}{n^{2/d}} \overline{|\partial \mathcal{R}_n|}\big)] \\ \le & \frac{Z_n(\beta')}{Z_n(\beta)} \cdot \exp\Big(-(\beta' - \beta)\varepsilon n^{1-2/d}\Big). \end{aligned} \quad (6.2)$$

Then by using that $\beta' < \beta_d^+$, we deduce that this last term goes to 0 as n tends to infinity.

We prove now the fourth statement. Fix $\beta > \beta_d^-$. By definition, there exists $c > 0$, such that for n large enough,

$$Z_n(\beta) \ge \exp(cn^{1-2/d}). \quad (6.3)$$

It follows that as $n \to \infty$, with $\varepsilon = c/(2\beta)$, we have

$$\mathbb{Q}_n^\beta [|\partial \mathcal{R}_n| - \mathbb{E}[|\partial \mathcal{R}_n|] \ge -\varepsilon n] \to 0. \quad (6.4)$$

Now, the arguments used for the proof of (1.9) in the previous section show that there exist α and A, such that

$$\begin{aligned} \mathbb{P}\Big[&\overline{|\partial \mathcal{R}_n|} \le -\varepsilon n, \ \min_{x \in \mathcal{C}} \ell_n(B(x, r_n)) < \frac{\alpha \varepsilon n}{|\mathcal{C}|} \text{ for all } \mathcal{C} \in \mathcal{A}(r_n, An) \text{ satisfying } \mathcal{I}_d(B(\mathcal{C}, r_n)) \le A\Big] \\ &\le \exp(-2\beta n^{1-\frac{2}{d}}), \end{aligned}$$

for n large enough (in the proof of the previous section, take for instance $T = T(n) = (n/\sqrt{\log n})^{2/d}$, so that $r_n^d = o(T^{d/2})$). Using again that $Z_n(\beta) \ge 1$, this implies that

$$\mathbb{Q}_n^\beta\Big[\min_{x \in \mathcal{C}} \ell_n(B(x, r_n)) < \frac{\alpha \varepsilon n}{|\mathcal{C}|} \text{ for all } \mathcal{C} \in \mathcal{A}(r_n, An) \text{ satisfying } \mathcal{I}_d(B(\mathcal{C}, r_n)) \le A\Big] \to 0,$$

as $n \to \infty$, and this proves (1.19). The proof of (1.20) is similar and left to the reader.

It remains now to prove the last statement. In fact as the proof above shows, it suffices to see that for any fixed $\chi \in (0,1)$, (6.3) holds true with $c = c(\beta) = (1-\chi)\beta\nu_d$, for all β large enough (since then the proof of the previous statement works as well with $\varepsilon = (1-\chi)c/\beta$). For this, notice that (4.3) shows that there exists a constant $C > 0$, such that for n large enough,

$$\mathbb{P}[|\partial \mathcal{R}_n| - \mathbb{E}[|\partial \mathcal{R}_n|] \le -(1-\chi/2)\nu_d n] \ge \exp(-Cn^{1-2/d}), \quad (6.5)$$

as if the walk spends the first n steps in a ball of volume $(\chi/4)\nu_d n$, then using also (2.7) we deduce that $\overline{|\partial\mathcal{R}_n|}$ is smaller than $(\chi/2-1)\nu_d n$, for n large enough. Therefore, for β large enough

$$Z_n(\beta) \;\geq\; \exp\Big((\beta(1-\chi/2)\nu_d - C)n^{1-2/d}\Big) \;\geq\; \exp\Big(\beta(1-\chi)\nu_d n^{1-2/d}\Big),$$

which concludes the proof. □

Acknowledgements. A.A. received support of the A*MIDEX grant (ANR-11-IDEX-0001-02) funded by the French Government "Investissements d'Avenir" program.

References

[A08] **A. Asselah**, *On large intersection and self-intersection local times in dimension five or more*, arXiv:0801.3918.

[AC07] **A. Asselah and F. Castell**, *A note on random walk in random scenery*, Ann. Inst. H. Poincaré Probab. Statist. 43 (2007), 163–173.

[AS15] **A. Asselah and B. Schapira**, *Boundary of the range of transient random walk*, arXiv:1507.01031.

[ASS16] **A. Asselah, B. Schapira, P. Sousi**, *Capacity of the range of random walk on* $\mathbb{Z}^d$, in preparation.

[BKYY10] **I. Benjamini, G. Kozma, A. Yadin, A. Yehudayoff**, *Entropy of random walk range*, Ann. Inst. Henri Poincaré Probab. Stat. 46 (2010), 1080–1092.

[BY13] **N. Berestycki and A. Yadin**, *Condensation of random walks and the Wulff crystal*, arXiv:1305.0139.

[BBH01] **M. van den Berg, E. Bolthausen, F. den Hollander**, *Moderate deviations for the volume of the Wiener sausage*, Ann. of Math. (2) 153 (2001), 355–406.

[BBH04] **M. van den Berg, E. Bolthausen, F. den Hollander**, *On the volume of the intersection of two Wiener sausages.* Ann. of Math. (2) 159 (2004), 741–782.

[B90] **E. Bolthausen**, *On the volume of the Wiener sausage*, Ann. Probab. 18, (1990), 1576–1582.

[B94] **E. Bolthausen**, *Localization of a two-dimensional random walk with an attractive path interaction*, Ann. Probab. 22, (1994), 875–918.

[DV75] **M. D. Donsker, S.R.S. Varadhan**, *Asymptotics for the Wiener sausage*, Comm. Pure Appl. Math. 28, (1975), 525–565.

[DV79] **M. D. Donsker, S. R. S. Varadhan**, *On the number of distinct sites visited by a random walk*, Comm. Pure Appl. Math. 32, (1979), 721–747.

[JO69] **N. C. Jain, S. Orey** *On the range of random walk*, Israel J. Math. 6, (1969), 373–380.

[KMSS94] **K. M. Khanin, A. E. Mazel, S. B. Shlosman, Y. G. Sinai**, *Loop condensation effects in the behavior of random walks*, The Dynkin Festschrift, 167–184, Progr. Probab., 34, Birkhuser Boston, Boston, MA, (1994).

[HSC93] **W. Hebisch, L. Saloff-Coste**, *Gaussian estimates for Markov chains and random walks on groups*, Ann. Probab. 21, (1993), 673–709.

[L13] **G. F. Lawler**, Intersections of random walks, Reprint of the 1996 edition. Modern Birkhauser Classics. Birkhauser/Springer, New York, (2013), iv+223 pp.

[LL10] **G. F. Lawler and V. Limic**, *Random walk: a modern introduction.* Cambridge Studies in Advanced Mathematics, 123. Cambridge University Press, Cambridge, (2010). xii+364 pp.

[LG86] **J.-F. Le Gall**, *Propriétés d'intersection des marches aléatoires. I. Convergence vers le temps local d'intersection*, Comm. Math. Phys. 104 (1986), 471–507.

[Ok15] **I. Okada** *The inner boundary of random walk range*, to appear in J. Math. Soc. Japan.

[Phet12] **P. Phetpradap** *Ph.D. thesis*, University of Bath, (2012).

[Pov99] **T. Povel** *Confinement of Brownian motion among Poissonian obstacles in $\mathbb{R}^d$, $d \geq 3$*, Probab. Theory Related Fields 114, (1999), 177–205.

[RS12] **B. Rath, A. Sapozhnikov**, *Connectivity properties of random interlacement and intersection of random walks*, ALEA Lat. Am. J. Probab. Math. Stat. 9 (2012), 67–83.

[S90] **A.-S. Sznitman**, *Long time asymptotics for the shrinking Wiener sausage* Comm. Pure App. Math. 43, (1990), 809–820.

[S91] **A.-S. Sznitman**, *On the confinement property of two-dimensional Brownian motion among poissonian obstacles*, Comm. Pure App. Math. 44, (1991), 1137–1170.