Random Walk, Local times, and subsets maximizing capacity

Amine Asselah [*] Bruno Schapira[†]

Abstract

We prove that in any finite set, in $\mathbb{Z}^d$ with $d \geq 3$, there is a subset whose capacity and volume are both of the same order as the capacity of the initial set. This implies some optimal estimates on probabilities of various *folding* events of a random walk. For instance, knowing that a random walk folds into an *atypical high density* region, we show that this random region is most likely ball-like.

Keywords and phrases. Random Walk; Local times; Capacity; Range.
MSC 2010 *subject classifications.* Primary 60F05, 60G50.

1 Introduction

We consider a random walk $\{S_n\}_{n\in\mathbb{N}}$ on $\mathbb{Z}^d$, with $d \geq 3$. For $n \in \mathbb{N} \cup \{\infty\}$, and site $z \in \mathbb{Z}^d$, the local time is defined as

$$\ell_n(z) := \sum_{k\in[0,n)} \mathbf{1}\{S_k = z\} \quad \text{and for} \quad \Lambda \subset \mathbb{Z}^d,\ \ell_n(\Lambda) := \sum_{z\in\Lambda} \ell_n(z). \tag{1.1}$$

One of the main result of this note relates the tail behavior of the occupation time of Λ to the harmonic capacity of $\Lambda \subset \mathbb{Z}^d$ whose definition is as follows:

$$\text{cap}(\Lambda) = \sum_{x\in\Lambda} \mathbb{P}_x\big(H_\Lambda^+ = \infty\big), \tag{1.2}$$

where $\mathbb{P}_x$ denotes the law of the walk starting from x, and H_Λ^+ is the first return time to Λ. Alternatively it can also be viewed as a probability to hit Λ by a walk starting from infinity, conveniently renormalized:

$$\text{cap}(\Lambda) = \lim_{\|z\|\to\infty} \frac{1}{G(z)} \mathbb{P}_z\big(H_\Lambda^+ < \infty\big), \tag{1.3}$$

where G is Green's function:

$$G(z) := \sum_{n\geq 0} \mathbb{P}(S_n = z).$$

[*]LAMA, Univ Paris Est Creteil, Univ Gustave Eiffel, UPEM, CNRS, F-94010, Créteil, France; amine.asselah@u-pec.fr

[†]Aix-Marseille Université, CNRS, Centrale Marseille, I2M, UMR 7373, 13453 Marseille, France; bruno.schapira@univ-amu.fr

A third equivalent way to define the capacity is given by the variational formula

$$\frac{1}{\text{cap}(\Lambda)} = \inf\Big\{\sum_{x\in\Lambda}\sum_{y\in\Lambda} G(x-y)\mu(x)\mu(y): \ \mu \text{ probability on } \Lambda\Big\}. \tag{1.4}$$

The infimum is reached for the *harmonic (or equilibrium) measure* $\mu(x) = \mathbb{P}_x(H_\Lambda^+ = \infty)/\text{cap}(\Lambda)$. The many writings of the harmonic capacity, help obtain useful bounds. There exists a constant $a_d > 0$ (which depends on dimension d), such that

$$a_d \cdot |\Lambda|^{1-2/d} \le \text{cap}(\Lambda) \le |\Lambda|. \tag{1.5}$$

The upper bound comes from (1.2), and the lower bound comes from taking μ in (1.4) to be uniform, and rearranging the sum using that Green's function is bounded by a symmetric radially decreasing function, and

$$G(z) = \frac{c_G}{\|z\|^{d-2}} + \mathcal{O}\big(\frac{1}{\|z\|^d}\big), \tag{1.6}$$

with c_G some explicit positive constant (see Theorem 4.3.1 of [LL10]). In the continuous setting (1.4) allows to define the (usual) electrostatic capacity on $\mathbb{R}^d$ with $G(x) = \|x\|^{2-d}$. The Poincaré-Faber-Szegö's inequality [Polya] states that for some $\alpha_d > 0$, $\text{cap}(\Lambda) \ge \alpha_d \cdot |\Lambda|^{1-2/d}$, for any compact set Λ with smooth boundary, with equality reached only when Λ is a Euclidean ball.

The capacity appeared as a central object in many remarkable studies, and we would like to highlight some of them. In the forties, Kakutani [K44] discovers that a compact set of $\mathbb{R}^d$, is hit by Brownian motion with positive probability, if and only if it has positive electrostatic capacity. Fifty years later, Benjamini, Pemantle and Peres [BPP95] make quantitative Kakutani's discovery, and establish sharp upper and lower bounds between a related capacity (Martin's one) and the hitting probability in the context of Markov Chains. In the late eightees, Kesten [Kes90] bounds the growth rate of diffusion limited aggregation (DLA), a celebrated model of discrete random growth on $\mathbb{Z}^d$ where sites in the boundary of the cluster are chosen according to the harmonic measure (of the boundary of the cluster). For doing so Kesten introduces a martingale whose compensator is the sum of inverses of capacities of the growing cluster. This in itself is inspiring: understanding the growth of the capacity of the cluster plays a key role in understanding the reinforcement phenomenon behind the ramified tree-like shape of DLA (see also [LT19] for a related model). Finally, ten years ago, Sznitman [S10] introduced a model called random interlacements which is a homogeneous Poisson point process on $\mathbb{Z}^d$ such that the number of trajectories (the points of the process) hitting a given compact K is a Poisson random variable with mean $u \cdot \text{cap}(K)$, and whose hitting sites (of K) distribution is according to the harmonic measure (of K). The model of random interlacements proves (or is conjectured) to be adapted to the study of many phenomena where a random walk realizes atypical high densities: (i) either by reducing its range, and in a certain regime this is the *Swiss Cheese* problem (see [BBH]), (ii) or by disconnecting say the sphere $B_n := \{z \in \mathbb{Z}^d : \|z\| < n\}$, from the complement of B_{2n}, see [S17, NS20], and many more sophisticated events which we gather under the name of *folding*. One very basic example of folding is captured by the following event. For $r \ge 1$, a set $\mathcal{C}$ of sites at distance at least $4r$ one from each others, and for a density $\rho > 0$, we consider

$$\mathcal{F}(r,\rho,\mathcal{C}) := \{\forall x \in \mathcal{C}, \quad \ell_\infty(B_r + x) > \rho|B_r|\}. \tag{1.7}$$

In many folding problems, one central issue is to characterize the size and the shape of the folding region $\mathcal{C}$.

Our key result concerning such folding events requires some notation. Let $\mathcal{X}_r$ be the collection of finite subsets of $\mathbb{Z}^d$, whose points are at distance at least $4r$ one from each others, and if $\mathcal{C} \in \mathcal{X}_r$, let $B_r(\mathcal{C}) = \bigcup_{x\in\mathcal{C}}(B_r + x)$.

Theorem 1.1. *Assume $d \geq 3$. There exist positive constants β and κ, such that for any $r \geq 1$, $\rho > 0$ and $\mathcal{C} \in \mathcal{X}_r$, satisfying*

$$\rho r^{d-2} > \beta \log |\mathcal{C}|, \tag{1.8}$$

one has

$$\mathbb{P}\big(\forall x \in \mathcal{C},\ \ell_\infty(B_r + x) > \rho r^d\big) \leq \exp\big(-\kappa \cdot \rho \cdot cap(B_r(\mathcal{C}))\big). \tag{1.9}$$

Note that when $r = 1$, $B_r(x)$ reduces to $\{x\}$, and $B_r(\mathcal{C})$ reduces to $\mathcal{C}$.

Remark 1.2. We obtained a similar inequality in Proposition 1.7 of [AS17], but for local times taken at some finite time n, and with an additional combinatorial factor of order $(n|\mathcal{C}|)^{|\mathcal{C}|}$ in the right hand side of (1.9). The extension to infinite time horizon comes from a slight modification of our previous proof, but the removal of the combinatorial factor requires an original ingredient, given by Theorem 1.4 below.

Remark 1.3. Sznitman obtained results with a similar flavor as Theorem 1.1 in the context of the Gaussian Free Field (GFF) and in the model of Random Interlacements, respectively in [S15, Corollaty 4.4] and [S17, Theorem 4.2] (see also [LS15] for related results). This is consistent with the facts that on one hand the random walk is a limiting case of Random Interlacements, as the intensity parameter goes to zero; yet in our case it does not seem possible to deduce our results from those of [S17] or [LS15]. On the other hand, it is well known that the GFF and the field of local times of a random walk are strongly connected via the celebrated Dynkin's isomorphism (see e.g. [E94] or [S12]), and actually many results concerning the percolative properties of the field of local times for Random Interlacements find their exact analogues in the setting of the level-sets of the GFF, as illustrated in the recent works [CN20a, CN20b, RS13, S15, S17].

Theorem 1.1 relies on the following observation, which to the best of our knowledge is new.

Theorem 1.4. *Assume $d \geq 3$. There exists $\alpha > 0$, such that for any $r \geq 1$ and $\mathcal{C} \in \mathcal{X}_r$, there is a subset $U \subseteq \mathcal{C}$, satisfying*

$$(i) \quad cap(B_r(U)) \geq \alpha \cdot r^{d-2}|U| \quad \text{and} \quad (ii) \quad r^{d-2}|U| \geq \alpha \cdot cap(B_r(\mathcal{C})). \tag{1.10}$$

To appreciate the usefulness of Theorem 1.4, we note that in many settings, the folding event concerns a walk of length n and is of the form $\cup_{\mathcal{C}} \mathcal{F}(r, \rho, \mathcal{C})$ where the union is over all $\mathcal{C} \subseteq [-n, n]^d$, with only a lower bound on their volume, say $|\mathcal{C}| \geq L$. In other words, the folding region is often unknown. Then Theorem 1.1 and a naive union bound gives

$$\mathbb{P}(\cup_{\mathcal{C}} \mathcal{F}(r, \rho, \mathcal{C})) \leq (2n)^{d \cdot L} \cdot \exp(-\kappa\rho \cdot a_d \cdot r^{d-2} L^{1-2/d}),$$

using also that $\text{cap}(\mathcal{B}_r(\mathcal{C})) \geq a_d r^{d-2}|\mathcal{C}|^{1-2/d}$, by (1.5), which is useful only when (ignoring constants)

$$L^{2/d} \cdot \log(n) \leq \rho \cdot r^{d-2}. \tag{1.11}$$

Now Theorem 1.4 allows to go beyond this condition (1.11), and gives

$$\mathbb{P}(\cup_{\mathcal{C}} \mathcal{F}(r, \rho, \mathcal{C})) \leq \exp(-\kappa\rho \cdot a_d \cdot r^{d-2} L^{1-2/d}),$$

under the weaker assumption (ignoring constants):

$$\log(n) \leq \rho \cdot r^{d-2}.$$

The latter can be used to characterize the shape of a localization region for a random walk (which corresponds to the level-sets for the local times in the particular case $r = 1$, see below), which we now describe in details. First, we introduce more notation. To obtain a neat partition of $\mathbb{Z}^d$ we switch to cubes, rather than balls. Define for $r \geq 1$,

$$Q_r := [-r/2, r/2)^d \cap \mathbb{Z}^d, \quad \text{and for} \quad x \in \mathbb{Z}^d, \quad Q_r(x) = x + Q_r.$$

Define further for $\rho > 0$ and $n \geq 1$,

$$\mathcal{C}_n(r,\rho) := \{x \in r\mathbb{Z}^d \,:\, \ell_n(Q_r(x)) \geq \rho r^d\}, \quad \text{and} \quad \mathcal{V}_n(r,\rho) := \bigcup_{x \in \mathcal{C}_n(r,\rho)} Q_r(x). \tag{1.12}$$

We can now state our third result.

Theorem 1.5. *Assume $d \geq 3$. There are positive constants $\underline{\kappa}$, $\overline{\kappa}$, and C, such that for any n, r and L positive integers and $\rho > 0$, satisfying*

$$\rho r^{d-2} \geq C \cdot \log(n), \quad \textit{and} \quad n \geq C\rho r^d L, \tag{1.13}$$

one has

$$\exp\big(-\underline{\kappa} \cdot \rho \cdot r^{d-2} \cdot L^{1-2/d}\big) \leq \mathbb{P}\big(|\mathcal{C}_n(r,\rho)| > L\big) \leq \exp\big(-\overline{\kappa} \cdot \rho \cdot r^{d-2} \cdot L^{1-2/d}\big). \tag{1.14}$$

In addition there exists $A > 0$, such that

$$\lim_{n\to\infty} \inf_{(r,\rho,L)} \mathbb{P}\big(\mathit{cap}(\mathcal{V}_n(r,\rho)) \leq A \cdot |\mathcal{V}_n(r,\rho)|^{1-2/d} \mid |\mathcal{C}_n(r,\rho)| > L\big) = 1, \tag{1.15}$$

where the infimum is taken over all triples (r, ρ, L) satisfying (1.13).

Remark 1.6. The result is interesting on its own right for $r = 1$, in which case it concerns the so-called *level-sets* of the local times, that is the sets of the form

$$\mathcal{L}_n(\rho) := \{z \in \mathbb{Z}^d \,:\, \ell_n(z) > \rho\}.$$

Specializing Theorem 1.5 to these sets gives that for $\rho \geq C \cdot \log(n)$ and $n \geq C\rho \cdot L$ (i.e. (1.13))

$$\exp(-\underline{\kappa} \cdot \rho \cdot L^{1-2/d}) \leq \mathbb{P}(|\mathcal{L}_n(\rho)| > L) \leq \exp(-\overline{\kappa} \cdot \rho \cdot L^{1-2/d}).$$

Furthermore, asymptotically as n goes to infinity, conditionally on being non-empty, the shape of $\mathcal{L}_n(\rho)$ is ball-like in the following sense. There is $A > 0$ such that for ρ_n, L_n satisfying $\rho_n \geq C \cdot \log(n)$ and $n \geq C\rho_n \cdot L_n$

$$\lim_{n\to\infty} \mathbb{P}\big(\mathrm{cap}(\mathcal{L}_n(\rho_n)) \leq A \cdot |\mathcal{L}_n(\rho_n)|^{1-2/d} \mid |\mathcal{L}_n(\rho_n)| > L_n\big) = 1.$$

Remark 1.7. Theorem 1.5 is also interesting in its full generality for the question of the moderate deviations for the volume and the capacity of the range. Indeed, in our previous works [AS17, AS19a], we found that conditionally on shrinking the range, the walk folds in a ball-like (random) domain in a distinct manner according to whether $d = 3$ or $d \geq 5$. By ball-like we mean that as the walk length n increases, the capacity of the region of high-density, say $\mathcal{V}_n$, remains of the order of $|\mathcal{V}_n|^{1-2/d}$ that is of the order of the capacity of a ball of volume $|\mathcal{V}_n|$. However, the characterization of the deviation scenarii in [AS17, AS19a] was limited to a subregime of the moderate deviations. With the help of Theorem 1.5, we can now complete the picture, and get that the folded region is typically ball-like.

Remark 1.8. Note that the condition (1.13) is (nearly) optimal. Indeed, concerning the first part, observe that when the walk hits a ball of radius r, it stays typically a time of order r^2 in it, so that the density of occupation is at least $1/r^{d-2}$. Moreover, concerning the second part, which is only needed for the lower bound in (1.14), note that one needs at least $n \geq \rho r^d L$, for the set $\mathcal{C}_n(r,\rho)$ to be non-empty, and the role of the large constant C is simply to allow some room for proving the lower bound. Actually we also obtain a similar result as Theorem 1.5 for a closely related set, where instead of recording the time spent in small cubes, we count the number of visited sites, see Proposition 4.1 for details.

The paper is organized as follows. Section 2 contains our main technical novelty: the proof of Theorem 1.4. In Section 3, we prove Theorem 1.1. Finally, in Section 4 we prove Theorem 1.5. The proof is divided into a short upper bound, and a technical lower bound in Section 4.2 where we actually state Proposition 4.1 which deals with the (slightly more difficult) problem of covering a certain partition of space, rather than with local times.

2 Proof of Theorem 1.4

We start with the proof in the case $r = 1$, which we think is instructive and more transparent.

Case $r = 1$. We need to show that in any finite set $\Lambda \subseteq \mathbb{Z}^d$, there exists a subset U, whose capacity and cardinality are both of the order of the capacity of Λ. The proof is an instance of the probabilistic method. Indeed, we build a random set $\mathcal{U}$ which satisfies the desired constraints with positive probability.

First, choose a family of i.i.d. trajectories $(\gamma^x, x \in \mathbb{Z}^d)$ with the same law as the walk $S = \{S_n\}_{n\geq 0}$ starting from the origin, and denote their joint law by $\mathbb{P}$. The hitting time of Λ by a (random) path $\gamma : \mathbb{N} \to \mathbb{Z}^d$ is denoted by $H_\Lambda(\gamma)$, the return time to Λ by $H_\Lambda^+(\gamma)$, and set $\gamma_x = \gamma^x + x$. Now, the random set $\mathcal{U}$ is

$$\mathcal{U} := \{x \in \Lambda : \ H_\Lambda^+(\gamma_x) = \infty\}.$$

Note that the volume of $\mathcal{U}$ is a sum of independent Bernoulli random variables, and thus

$$\mathbb{E}[|\mathcal{U}|] = \sum_{x\in\Lambda} \mathbb{P}\big(H_\Lambda^+(\gamma_x) = \infty\big) = \mathrm{cap}(\Lambda), \quad \text{and} \quad \mathrm{var}(|\mathcal{U}|) \leq \mathrm{cap}(\Lambda).$$

Thus $|\mathcal{U}|$ is concentrated around its mean and by Chebychev's inequality

$$\mathbb{P}\big(|\mathcal{U}| < \frac{1}{2}\mathbb{E}[|\mathcal{U}|]\big) \leq \frac{4}{\mathrm{cap}(\Lambda)}, \quad \text{and} \quad \mathbb{P}\big(|\mathcal{U}| > 2\mathbb{E}[|\mathcal{U}|]\big) \leq \frac{1}{\mathrm{cap}(\Lambda)}. \tag{2.1}$$

We can assume $\mathrm{cap}(\Lambda) > 16$, (as for sets with bounded capacity one can always choose α small enough) so that (2.1) reads

$$\mathbb{P}\big(2\mathrm{cap}(\Lambda) \geq |\mathcal{U}| \geq \frac{1}{2}\mathrm{cap}(\Lambda)\big) \geq \frac{2}{3}. \tag{2.2}$$

Now, we show that $\mathrm{cap}(\mathcal{U})$ is of order its volume. By (1.4), if we choose for μ the uniform measure on $\mathcal{U}$, then

$$\frac{\mathrm{cap}(\mathcal{U})}{|\mathcal{U}|} \geq \Big(\frac{1}{|\mathcal{U}|}\sum_{x,y\in\mathcal{U}} G(x-y)\Big)^{-1}. \tag{2.3}$$

Let us compute the expression on the right hand side of (2.3).

$$\sum_{x,y\in\mathcal{U}} G(x-y) = \sum_{x\in\Lambda}\sum_{y\in\Lambda} \mathbb{1}\{H_\Lambda^+(\gamma_x)=\infty,\ H_\Lambda^+(\gamma_y)=\infty\}\cdot G(x-y)$$
$$= G(0)\cdot|\mathcal{U}| + \sum_{x\in\Lambda}\sum_{y\in\Lambda\setminus\{x\}} \mathbb{1}\{H_\Lambda^+(\gamma_x)=\infty,\ H_\Lambda^+(\gamma_y)=\infty\}\cdot G(x-y).$$

Note that if $x\neq y$, then γ_x and γ_y are independent. Therefore,

$$\mathbb{E}\left[\sum_{x,y\in\mathcal{U}} G(x-y)\right] \le G(0)\cdot\mathbb{E}[|\mathcal{U}|] + \sum_{x,y\in\Lambda}\mathbb{P}\big(H_\Lambda^+(\gamma_x)=\infty\big)G(x-y)\mathbb{P}\big(H_\Lambda^+(\gamma_y)=\infty\big).$$

By a last passage decomposition (see Proposition 4.6.4 in [LL10]), for $x\in\Lambda$,

$$1=\mathbb{P}\big(H_\Lambda(\gamma_x)<\infty\big)=\sum_{y\in\Lambda} G(x-y)\mathbb{P}\big(H_\Lambda^+(\gamma_y)=\infty\big).$$

Thus,

$$\mathbb{E}\left[\sum_{x,y\in\mathcal{U}} G(x-y)\right] \le \big(G(0)+1\big)\cdot \mathrm{cap}(\Lambda), \quad \text{and} \quad \mathbb{P}\left(\sum_{x,y\in\mathcal{U}} G(x-y)\le 4(G(0)+1)\cdot\mathrm{cap}(\Lambda)\right)\ge\frac{3}{4}.$$

Together with (2.2), we obtain

$$\mathbb{P}\left(2\mathrm{cap}(\Lambda)\ge|\mathcal{U}|\ge\frac{1}{2}\mathrm{cap}(\Lambda),\ \sum_{x,y\in\mathcal{U}} G(x-y)\le 4\big(G(0)+1\big)\cdot\mathrm{cap}(\Lambda)\right)\ge\frac{5}{12}. \tag{2.4}$$

By (2.3) and (2.4), we deduce that for some $\alpha>0$,

$$\mathbb{P}\left(2\mathrm{cap}(\Lambda)\ge|\mathcal{U}|\ge\frac{1}{2}\mathrm{cap}(\Lambda),\ \mathrm{cap}(\mathcal{U})\ge\alpha\cdot\mathrm{cap}(\Lambda)\right)\ge\frac{5}{12}. \tag{2.5}$$

Thus, we conclude that (1.10) holds for a random set $\mathcal{U}$, when $r=1$.

We now prove the general case by refining the previous argument.

General case $r\ge 1$. The proof follows the same steps after we choose an appropriate random subset of the set of centers $\mathcal{C}$. For simplicity, for $r>0$, we set $\Lambda_r=B_r(\mathcal{C})$, and V_r to be the complement of $B_{2r}(\mathcal{C})$. We need now the hitting time of Λ_r after exiting $B_{2r}(\mathcal{C})$. For a trajectory γ, define

$$H_{\Lambda_r}^r(\gamma)=\inf\{k>H_{V_r}(\gamma):\gamma(k)\in\Lambda_r\}.$$

Then choose a family of i.i.d. trajectories $(\gamma^x, x\in\mathbb{Z}^d)$ with the same law as S, denote the joint law by $\mathbb{P}$, and set $\gamma_x=\gamma^x+x$. Our random set reads now

$$\mathcal{U}:=\{x\in\mathcal{C}:\ H_{\Lambda_r}^r(\gamma_x)=\infty\}.$$

Thus, each center $x\in\mathcal{C}$ is kept in $\mathcal{U}$ if a random walk launched from x escapes Λ_r after exiting $B_{2r}(\mathcal{C})$. The reason to force first to exit $B_{2r}(\mathcal{C})$ stems from the following lemma taken from [AS17].

Lemma 2.1. *There is $\theta>1$, such that for any $r\ge1$, $\mathcal{C}\in\mathcal{X}_r$, and $x\in\mathcal{C}$,*

$$\theta\mathbb{P}\big(H_{\Lambda_r}^r(\gamma_x)=\infty\big)\ge\frac{1}{r^{d-2}}\sum_{y\in\partial B_r(x)}\mathbb{P}\big(H_{\Lambda_r}^+(y+S)=\infty\big)\ge\frac{1}{\theta}\mathbb{P}\big(H_{\Lambda_r}^r(\gamma_x)=\infty\big). \tag{2.6}$$

Now, note that $|B_r(\mathcal{U})|/|B_r|$ is a sum of $|\mathcal{C}|$ independent Bernoulli random variables, and therefore

$$\mathrm{var}(|B_r(\mathcal{U})|) \le |B_r| \cdot \mathbb{E}[|B_r(\mathcal{U})|].$$

Furthermore, thanks to Lemma 2.1, there are positive constants c_1 and c_2, such that

$$c_1 r^2 \cdot \mathrm{cap}(B_r(\mathcal{C})) \le \mathbb{E}[|B_r(\mathcal{U})|] = |B_r| \cdot \sum_{x\in\mathcal{C}} \mathbb{P}\big(H^r_{\Lambda_r}(\gamma_x) = \infty\big) \le c_2 r^2 \cdot \mathrm{cap}(B_r(\mathcal{C})).$$

It follows that for a positive constant c_3,

$$\mathbb{P}\left(\frac{1}{2}\mathbb{E}[|B_r(\mathcal{U})|] \le |B_r(\mathcal{U})| \le 2\mathbb{E}[|B_r(\mathcal{U})|]\right) \ge 1 - c_3 \frac{r^{d-2}}{\mathrm{cap}(B_r(\mathcal{C}))}.$$

We can assume that $\mathrm{cap}(B_r(\mathcal{C})) \ge 4c_3 r^{d-2}$ (as otherwise we conclude by taking $\alpha < 1/(4c_3)$), in which case it follows that

$$\mathbb{P}\left(\frac{1}{2}\mathbb{E}[|B_r(\mathcal{U})|] \le |B_r(\mathcal{U})| \le 2\mathbb{E}[|B_r(\mathcal{U})|]\right) \ge \frac{3}{4}.$$

Thus, with probability larger than 3/4, the random set $\mathcal{U}$ satisfies (ii) of (1.10). Let us check now (i). By (1.4), we obtain a lower bound on $\mathrm{cap}(B_r(\mathcal{U}))$ as we choose a measure on $B_r(\mathcal{U})$. Taking the uniform measure on the *boundary of* $B_r(\mathcal{U})$ gives

$$\frac{1}{(|\partial B_r| \cdot |\mathcal{U}|)^2} \sum_{x,x'\in\mathcal{U}} \sum_{y\in\partial B_r(x)} \sum_{y'\in\partial B_r(x')} G(y-y') \ge \frac{1}{\mathrm{cap}(B_r(\mathcal{U}))}. \tag{2.7}$$

We need to show that the left hand side of (2.7) is smaller than $1/(r^{d-2}|\mathcal{U}|)$. First, we treat the case $x' = x$. Note that by Green's function asymptotic (1.6), there is $c > 0$, such that

$$\forall y \in \partial B_r(x) \qquad \sum_{y'\in\partial B_r(x)} G(y-y') \le c \cdot r.$$

Thus, as we further sum over $y \in \partial B_r(x)$, and $x \in \mathcal{U}$, we obtain

$$\sum_{x\in\mathcal{U}} \sum_{y\in\partial B_r(x)} \sum_{y'\in\partial B_r(x)} G(y-y') \le c \cdot r \cdot r^{d-1} \cdot |\mathcal{U}|. \tag{2.8}$$

Now, to deal with the terms with $x' \neq x$, we take expectation first, and we bound $G(y-y')$ by $c_4 \cdot G(x-x')$ uniformly in $y \in \partial B_r(x)$ and $y' \in \partial B_r(x')$. Therefore,

$$\begin{aligned}
\mathbb{E}\left[\sum_{x\neq x'\in\mathcal{U}} \sum_{y\in\partial B_r(x)} \sum_{y'\in\partial B_r(x')} G(y-y')\right] &\le c_4 |\partial B_r|^2 \cdot \mathbb{E}\big[\sum_{x\neq x'\in\mathcal{U}} G(x-x')\big] \\
&\le c_4 |\partial B_r|^2 \sum_{x\neq x'\in\mathcal{C}} \mathbb{P}\big(H^r_{\Lambda_r}(\gamma_x)=\infty\big) G(x-x') \mathbb{P}\big(H^r_{\Lambda_r}(\gamma_{x'})=\infty\big) \\
&\le c_5 (r^{d-1})^2 \mathbb{E}[|\mathcal{U}|] \sup_{x\in\mathcal{C}} \sum_{x'\neq x} G(x-x') \mathbb{P}\big(H^r_{\Lambda_r}(\gamma_{x'})=\infty\big).
\end{aligned}$$

By using (2.6) of Lemma 2.1, and a last passage decomposition we have for a constant $c_6 > 0$, and any $x \in \mathcal{C}$,

$$\begin{aligned}
1 = \mathbb{P}\big(H_{\Lambda_r}(\gamma_x) < \infty\big) &\ge \sum_{\substack{x'\in\mathcal{C}\\ x'\neq x}} \sum_{y\in\partial B_r(x')} G(x-y) \mathbb{P}\big(H^+_{\Lambda_r}(\gamma_y)=\infty\big) \\
&\ge c_6 r^{d-2} \sum_{\substack{x'\in\mathcal{C}\\ x'\neq x}} G(x-x') \mathbb{P}\big(H^r_{\Lambda_r}(\gamma_{x'})=\infty\big).
\end{aligned}$$

This implies that for a constant $c_7 > 0$,

$$\mathbb{E}\left[\sum_{x\neq x'\in\mathcal{U}}\sum_{y\in\partial B_r(x)}\sum_{y'\in\partial B_r(x')} G(y-y')\right] \le c_7 r^d \cdot \mathbb{E}[|\mathcal{U}|].$$

Chebychev's inequality now allows us to conclude as in the proof of the case $r = 1$.

3 Proof of Theorem 1.1

The result extends a previous one proved in [AC07] stating that there are positive constants C and κ_0 such that for any finite $\Lambda \subset \mathbb{Z}^d$ with $d \ge 3$, and any $T > 0$,

$$\mathbb{P}\big(\ell_\infty(\Lambda) > T\big) \le C\exp(-\kappa_0\frac{T}{|\Lambda|^{2/d}}), \quad \text{where} \quad \ell_\infty(\Lambda) = \sum_{x\in\Lambda}\ell_\infty(x). \tag{3.1}$$

Note that this result already proves our desired estimate for small values of $|\mathcal{C}|$, i.e. one can now suppose that $|\mathcal{C}|$ is large enough.

The starting point of the proof is Proposition 1.7 of [AS17]. We showed that for any $n \ge 1$, any positive r and ρ, and any $\mathcal{C} \in \mathcal{X}_r$,

$$\mathbb{P}\big(\forall x\in\mathcal{C},\ \ell_n(B_r(x)) > \rho r^d\big) \le n^{|\mathcal{C}|}\cdot|\mathcal{C}|!\ \exp\big(-\kappa\cdot\rho\cdot\text{cap}(B_r(\mathcal{C}))\big), \tag{3.2}$$

for some constant $\kappa > 0$.

Note that if we set $\Lambda = B_r(\mathcal{C})$, the event estimated in (3.2) corresponds to a *uniform* covering of Λ, whereas (3.1) only considers the total time spent in Λ. The proof of (3.1) uses that $\ell_\infty(\Lambda)$ is bounded by a geometric variable of mean of order $|\Lambda|^{2/d}$, whereas the proof of (3.2) is by way of induction and transforms visits of $B_r(\mathcal{C})$ into excursions between $B_r(x)$ and $B_{2r}(x)$ for $x \in \mathcal{C}$. Recall that an excursion is a period of time that the walk needs to go from the inside of $B_r(x)$ to the outside of $B_{2r}(x)$. The number of such excursions is denoted $\mathcal{N}_x$, and our proof of (3.2) in [AS17] establishes that for any $x_0 \in \mathcal{C}$, any $y \in \partial B_r(x_0)$, and any sequence of positive integers $(n_x)_{x\in\mathcal{C}}$,

$$\mathbb{P}_y\big(\forall x\in\mathcal{C},\, \mathcal{N}_x = n_x\big) \le \frac{q_{x_0}}{\min_{z\in\mathcal{C}} q_z}\prod_{x\in\mathcal{C}} q_x^{n_x}, \quad \text{with} \quad \log(q_x) = -\min_{z\in\partial B_{2r}(x)}\mathbb{P}_z\big(H_\Lambda = \infty\big). \tag{3.3}$$

Note that $\exp(-1) \le q_x \le 1$, and since we wish to sum (3.3) over all $k \ge n_x$ and all $x \in \mathcal{C}$, we wish to restrict the product over $x \in \mathcal{C}$ in the right hand side of (3.3) to terms where q_x are uniformly bounded away from one. For this purpose we introduce the notation, with $\Lambda = B_r(\mathcal{C})$,

$$\text{cap}_x(\Lambda) = \sum_{z\in\partial B_r(x)}\mathbb{P}_z(H_\Lambda^+ = \infty), \quad \text{and} \quad \mathcal{C}^* = \{x\in\mathcal{C}:\ \text{cap}_x(\Lambda) > \frac{1}{2}\cdot\frac{\text{cap}(\Lambda)}{|\mathcal{C}|}\}. \tag{3.4}$$

Note that

$$\sum_{x\in\mathcal{C}\setminus\mathcal{C}^*}\text{cap}_x(\Lambda) \le \frac{1}{2}\text{cap}(\Lambda), \qquad \text{so that} \qquad \sum_{x\in\mathcal{C}^*}\text{cap}_x(\Lambda) > \frac{1}{2}\text{cap}(\Lambda). \tag{3.5}$$

Now, recall that by Lemma 5.2 of [AS17], there is $\theta > 0$ such that

$$\min_{z\in\partial B_{2r}(x)}\mathbb{P}_z(H_\Lambda = \infty) \ge \frac{\theta}{r^{d-2}}\text{cap}_x(\Lambda).$$

Thus, for any $x \in \mathcal{C}^*$, we have

$$q_x \leq \exp\big(-\frac{\theta}{r^{d-2}}\mathrm{cap}_x(\Lambda)\big) \leq \exp\big(-\frac{\theta\cdot \mathrm{cap}(\Lambda)}{2r^{d-2}\cdot |\mathcal{C}|}\big). \tag{3.6}$$

Using next (1.5), we deduce that $\mathrm{cap}(\Lambda) \geq cr^{d-2}|\mathcal{C}|^{1-2/d}$, for some constant $c > 0$ (that might change from line to line), and therefore

$$\prod_{x\in\mathcal{C}^*} \frac{1}{1-q_x} \leq (c|\mathcal{C}|^{2/d})^{|\mathcal{C}|} \leq c|\mathcal{C}|!. \tag{3.7}$$

Thus, by summing (3.3) over all $k \geq n_x$, and all $x \in \mathcal{C}^*$, we get

$$\mathbb{P}_y\big(\forall x \in \mathcal{C}, \mathcal{N}_x \geq n_x\big) \leq e \prod_{x\in\mathcal{C}^*} \frac{1}{1-q_x} \prod_{x\in\mathcal{C}^*} q_x^{n_x} \leq c|\mathcal{C}|! \cdot \prod_{x\in\mathcal{C}^*} q_x^{n_x}. \tag{3.8}$$

Now, if we set $n_x = \alpha\rho r^{d-2}$, we obtain using (3.5), (3.8) and the first inequality of (3.6),

$$\mathbb{P}_y\big(\forall x \in \mathcal{C}, \mathcal{N}_x \geq \alpha\rho r^{d-2}\big) \leq c|\mathcal{C}|! \cdot \exp\big(-\kappa\cdot\rho\cdot \mathrm{cap}(B_r(\mathcal{C}))\big).$$

From this point, we can follow the end of the proof of Proposition 1.7 in [AS17], and deduce that

$$\mathbb{P}\big(\forall x \in \mathcal{C},\ \ell_\infty(B_r(x)) > \rho r^d\big) \leq c|\mathcal{C}|! \cdot \exp\big(-\kappa\cdot\rho\cdot \mathrm{cap}(B_r(\mathcal{C})\big). \tag{3.9}$$

Then Theorem 1.4 gives the existence of a subset $U \subseteq \mathcal{C}$, with $\mathrm{cap}(B_r(U))$ of the same order as both $|U|r^{d-2}$ and $\mathrm{cap}(B_r(\mathcal{C}))$. Moreover, for a given integer k, the number of possible choices for U with cardinality k is of the order of $|\mathcal{C}|^k$, which altogether allows to remove the combinatorial factor in (3.9), using the hypothesis (1.8). This concludes the proof of the theorem.

4 Application to Folding, and Theorem 1.5

The proof of Theorem 1.5 is divided in two parts. In the first part (see Subsection 4.1 below), we show that for some positive constants $\widetilde{\kappa}$ and A_0, for any $A > A_0$, any $n \geq 1$, and any (r, ρ, L) satisfying (1.13),

$$\mathbb{P}\left(|\mathcal{C}_n(r,\rho)| > L,\, \mathrm{cap}(\mathcal{V}_n(r,\rho)) > A|\mathcal{V}_n(r,\rho)|^{1-2/d}\right) \leq \exp(-\widetilde{\kappa}A\cdot\rho\cdot r^{d-2}L^{1-2/d}). \tag{4.1}$$

In the second part (see Subsection 4.2 below) we prove the lower bound in (1.14). Note that altogether this gives (1.15) as well, and thus proves Theorem 1.5.

4.1 The upper bound: proof of (4.1)

We introduce the notation $Q_r(U)$ for $\cup_{x\in U} Q_r(x)$, and we use Theorem 1.4, with the condition (1.13) and then (1.9) as follows.

$$\begin{aligned}
\mathbb{P}\big(|\mathcal{C}_n(r,\rho)| > L,\ \mathrm{cap}(\mathcal{V}_n(r,\rho)) \geq A\cdot|\mathcal{V}_n(r,\rho)|^{1-\frac{2}{d}}\big) &\leq \sum_{k>L}\mathbb{P}\big(|\mathcal{C}_n(r,\rho)|=k, \mathrm{cap}(\mathcal{V}_n(r,\rho)) \geq A\cdot r^{d-2}k^{1-2/d}\big)\\
&\leq \sum_{L<k\leq n}\mathbb{P}\big(\exists\, \mathcal{U}\subset[-n,n]^d:\ k\geq|\mathcal{U}|>\alpha A k^{1-2/d},\ \mathrm{cap}(Q_r(\mathcal{U}))\geq \alpha r^{d-2}|\mathcal{U}|\big)\\
&\leq \sum_{L<k\leq n}\ \sum_{\alpha A k^{1-2/d}<i\leq k}\mathbb{P}\big(\exists\, \mathcal{U}\subset[-n,n]^d:\ |\mathcal{U}|=i,\ \mathrm{cap}(Q_r(\mathcal{U}))\geq\alpha r^{d-2}i\big)\\
&\leq \sum_{L<k\leq n}\ \sum_{\alpha A k^{1-2/d}<i\leq k} c(2n)^{d\cdot i}\cdot i!\cdot\exp\big(-\kappa\alpha\rho r^{d-2}\cdot i\big)\\
&\leq\ c\sum_{k>L}\exp\big(-\widetilde{\kappa}A\cdot\rho r^{d-2}k^{1-2/d}\big)\leq\ c\exp\big(-2\widetilde{\kappa}A\cdot\rho r^{d-2}L^{1-2/d}\big).
\end{aligned}$$

The combinatorial factor $(2n)^{d\cdot i}\cdot i!$ was swallowed after using the condition that $r^{d-2}\cdot\rho > C\log(n)$, and choosing A large enough.

4.2 Lower bound

In this subsection, we establish a result which slightly differs from the lower bound in (1.14), and deals with covering rather than occupation. For this purpose we introduce for $n\geq 1$, the range of the walk $\mathcal{R}_n := \{S_0,\dots,S_n\}$, and for any $r\geq 1$ and $\rho\in[0,1]$,

$$\widetilde{\mathcal{C}}_n(r,\rho) := \{z\in r\mathbb{Z}^d:\ |\mathcal{R}_n\cap Q_r(z)|\geq\rho|Q_r|\}.$$

Our result is as follows.

Proposition 4.1. *There exist positive constants c and C, such that for any $n\geq 1$, $r>0$, $\rho\in(0,1/2)$, and $L\geq 1$, satisfying*

$$\rho r^{d-2}\geq 1,\quad \textit{and}\quad n\geq C\rho r^d L,$$

one has

$$\mathbb{P}\big(|\widetilde{\mathcal{C}}_n(r,\rho)|\geq L\big)\geq c\,\exp(-c\,\rho r^{d-2}L^{1-\frac{2}{d}}).$$

Thus this result is exactly the same as the lower bound in (1.14), but for this new set $\widetilde{\mathcal{C}}_n(r,\rho)$. Since the proof for $\mathcal{C}_n(r,\rho)$ is easier, we will only provide the proof for the set $\widetilde{\mathcal{C}}_n(r,\rho)$. Note also that since $\widetilde{\mathcal{C}}_n(r,\rho)\subseteq\mathcal{C}_n(r,\rho)$, the upper bound in (1.14), as well as (1.15), also hold for the set $\widetilde{\mathcal{C}}_n(r,\rho)$.

Remark 4.2. The hypothesis $\rho<1/2$ in Proposition 4.1 could be replaced by $\rho<1-\eta$, for any fixed constant $\eta>0$, and the constants c and C would then depend on η. However, when ρ gets close to 1, we fall in another regime, and for instance when $\rho=1$ an extra $\log r$ factor is needed in the exponential (and in the time needed to achieve the covering).

Proof. The scenario we choose to produce the desired event is to localize the walk long enough in a ball so that its occupation density is ρ. It is convenient to transform localization into a statement about excursions. We call *excursion* from a set U to a set V, the part of the random walk path starting from a point in U up to its hitting time of V which is maximal (for the inclusion of paths). To produce L boxes of side r, and filled to a density above ρ, we expect to localize the walk in a ball of radius R for a time T with

$$R = \lfloor L^{1/d} r \rfloor, \quad \text{and} \quad T = \lfloor C_1 \rho R^d \rfloor \quad \text{(and } C_1 \text{ large enough).}$$

Let now $\mathcal{N}_R$ be the number of excursions from ∂Q_{2R} to ∂Q_{4R} before exiting Q_{8R}, and consider

$$A := \{\mathcal{N}_R \geq N\}, \quad \text{with} \quad N = \lfloor C_2 \rho R^{d-2} \rfloor. \tag{4.2}$$

Now, for any C_2 (that is any number of excursions N), one makes the event A typical by choosing C_1 large (that is a large localization time T), and assume $n \geq T$. On event A, we define $X = (X_1, \dots, X_N)$ and $Y = (Y_1, \dots, Y_N)$, as respectively the starting and ending points of the N first excursions from ∂Q_{2R} to ∂Q_{4R}. Then given some fixed $\mathbf{x} = (x_1, \dots, x_N) \in \partial Q_{2R}^N$ we let $\mathbb{P}_{\mathbf{x}}$ be the law of N excursions starting from $\{x_i,\ i \leq N\}$, up to ∂Q_{4R}. We still denote by Y the set of ending points of the N excursions under $\mathbb{P}_{\mathbf{x}}$. We let M be the cardinality of the set $r\mathbb{Z}^d \cap Q_{R-r}$, and number its elements in some arbitrary order, say $v_1, \dots, v_M$. We define

$$B := \left\{ \begin{array}{c} \text{The } N \text{ (first) excursions visit at least half of the boxes} \\ \{Q_r(v_i)\}_{i \leq M}, \text{ a fraction at least } \rho \text{ of their sites} \end{array} \right\}.$$

Assume for a moment that,

$$\forall \mathbf{x} \in (\partial Q_{2R})^N, \qquad \mathbb{P}_{\mathbf{x}}(B) \geq 1/2. \tag{4.3}$$

With $\sigma = \inf\{n \geq 1 \,:\, S_n \in Q_{2R} \cup Q_{8R}^c\}$, we have

$$\mathbb{P}(A \cap B,\, X = \mathbf{x},\, Y = \mathbf{y}) = \prod_{i=1}^{n-1} \mathbb{P}_{y_i}\big(S(\sigma) = x_{i+1}\big) \cdot \mathbb{P}_{\mathbf{x}}(B, Y = \mathbf{y}). \tag{4.4}$$

Using Harnack's inequality (see [LL10, Theorem 6.3.9]), for some constant $c_H > 0$, for any $x \in \partial Q_{2R}$,

$$\inf_y \mathbb{P}_y\big(S(\sigma) = x\big) \geq c_H \mathbb{P}_{y^*}\big(S(\sigma) = x\big), \quad \text{with} \quad y^* := (4R, 0, \dots, 0). \tag{4.5}$$

Thus, using that there is a positive lower bound (uniform in R) for the probability that a walk starting from y^* hits ∂Q_{2R} before ∂Q_{8R}, we have $c_1 > 0$, such that

$$\begin{aligned} \mathbb{P}(A \cap B) = \sum_{\mathbf{x}} \sum_{\mathbf{y}} \mathbb{P}(A \cap B, X = \mathbf{x}, Y = \mathbf{y}) &\geq \sum_{\mathbf{x}} c_H^N \prod_{i=1}^{n} \mathbb{P}_{y^*}\big(S(\sigma) = x_i\big) \sum_{\mathbf{y}} \mathbb{P}_{\mathbf{x}}(B, Y = \mathbf{y}) \\ \geq c_H^N \inf_{\mathbf{x}} \mathbb{P}_{\mathbf{x}}(B) \sum_{\mathbf{x}} \prod_{i=1}^{n} \mathbb{P}_{y^*}\big(S(\sigma) = x_i\big) &\geq \frac{c_H^N}{2} \prod_{i=1}^{n} \mathbb{P}_{y^*}\big(S(\sigma) \in B_{2R}\big) \geq e^{-c_1 N}. \end{aligned} \tag{4.6}$$

Finally, define

$$C := \{\text{The walk makes at least } N \text{ excursions from } \partial Q_{4R} \text{ to } \partial Q_{2R} \text{ before time } T\}.$$

Using that on the event $A \cap C^c$, the walk spends a time at least T in Q_{8R}, we deduce that for some constant $c > 0$,

$$\mathbb{P}(A \cap C^c) \leq \exp(-c \frac{T}{R^2}). \tag{4.7}$$

Then the proposition readily follows from (4.6) and (4.7), once we choose $C_1 c > 2C_2 c_1$ and use

$$\mathbb{P}\big(A \cap B \cap C\big) \geq \mathbb{P}\big(A \cap B\big) - \mathbb{P}\big(A \cap C^c\big).$$

We now prove (4.3). We fix some $\mathbf{x} \in \partial Q_{2R}^N$, and in the remaining part of the proof, we work under $\mathbb{P}_{\mathbf{x}}$. We denote by $\mathcal{R}^N$ the range produced by the N excursions. We note that it suffices to show that for C_2 large enough, one has for any $i \leq M$,

$$\mathbb{P}_{\mathbf{x}}\big(|\mathcal{R}^N \cap Q_r(v_i)| > \rho |Q_r|\big) \geq 3/4. \tag{4.8}$$

Indeed, letting Z be the number of boxes whose fraction of visited sites exceeds ρ, (4.8) shows that $\mathbb{E}[Z] \geq (3/4)M$, and using also that Z is bounded by M, it implies that $\mathbb{P}(Z \leq M/2) \leq 1/2$, as wanted.

Thus, we are lead to prove (4.8) for $i \leq M$. For a chosen $i \leq M$, we introduce new notation. Let $\mathcal{N}_r$ be the number of excursions which hit $\partial Q_{2r}(v_i)$, and $\mathcal{G}$ be the $\sigma-$field generated by $\mathcal{N}_r$ and the hitting points of $\partial Q_{2r}(v_i)$ by these excursions. Finally we let $\mathcal{X} \subseteq Q_r(v_i)$ be the set of vertices visited by these excursions. Since any vertex in $Q_r(v_i)$ has a probability of order r^{2-d} to be visited by a walk starting from $\partial Q_{2r}(v_i)$, uniformely in its starting point, we have for some constant $c_0 > 0$, almost surely

$$\mathbb{E}_{\mathbf{x}}[|\mathcal{X}| \mid \mathcal{G}] \geq \Big(1 - (1 - \frac{c_0}{r^{d-2}})^{\mathcal{N}_r}\Big) \cdot |Q_r| \geq \Big(1 - \exp\big(-c_0 \frac{\mathcal{N}_r}{r^{d-2}}\big)\Big) \cdot |Q_r|, \tag{4.9}$$

using that $1 - u \leq e^{-u}$, for all $u \geq 0$. We choose a large constant K whose value will be fixed later. Since any excursion has a probability of order at least $(r/R)^{d-2}$ to hit $\partial Q_{2r}(v_i)$, it is possible to choose C_1 (and C_2) so that for this chosen K,

$$\mathbb{P}_{\mathbf{x}}\big(\mathcal{N}_r \geq K\rho r^{d-2}\big) \geq \sqrt{7/8}. \tag{4.10}$$

We distinguish now a high and a low density regimes.

High density. If $1 - \exp(-c_0\, K\rho) \geq \sqrt{7/8}$, then (4.9) and (4.10) imply that

$$\mathbb{E}_{\mathbf{x}}[|\mathcal{X}|] \geq (7/8) \cdot |Q_r|.$$

Using that $\mathcal{X} \subseteq Q_r$, and as a consequence that $|\mathcal{X}|$ is bounded by $|Q_r|$, we obtain with (4.8),

$$\mathbb{P}_{\mathbf{x}}(|\mathcal{X}| < \rho|Q_r|) \leq \mathbb{P}_{\mathbf{x}}(|\mathcal{X}| < |Q_r|/2) \leq 1/4 \qquad (\text{using } \rho < 1/2).$$

Low density. If ρ is such that $1 - \exp(-c_0\, K\rho) < \sqrt{7/8}$, then by (4.9) we may choose K large enough, so that on the event $\{\mathcal{N}_r \geq K\rho r^{d-2}\}$,

$$\mathbb{E}_{\mathbf{x}}[|\mathcal{X}| \mid \mathcal{G}] \geq \sqrt{K}\rho \cdot |Q_r|. \tag{4.11}$$

Our strategy now is to use a second (conditional) moment method, and show that the conditional variance of $\mathcal{X}$ is small. We denote with $\mathcal{X}_1$ the set of pairs of vertices $(y, z) \in \mathcal{X} \times \mathcal{X}$, for which there exists an excursion going through both y and z, and let $\mathcal{X}_2$ be the complement of $\mathcal{X}_1$ in $\mathcal{X} \times \mathcal{X}$ and note that $|\mathcal{X}|^2 = |\mathcal{X}_1| + |\mathcal{X}_2|$. Since for any $y \in Q_r(v_i)$ the mean number of vertices in $Q_r(v_i)$ which are visited by a walk starting from y is of order r^2, one has for some constant $c > 0$,

$$\mathbb{E}_{\mathbf{x}}[|\mathcal{X}_1| \mid \mathcal{G}] \leq cr^2 \cdot \mathbb{E}_{\mathbf{x}}[|\mathcal{X}| \mid \mathcal{G}].$$

Recalling $\rho r^{d-2} \geq 1$ and (4.11), it follows that on the event $\{\mathcal{N}_r \geq K\rho r^{d-2}\}$, for a possibly larger constant c,

$$\mathbb{E}_{\mathbf{x}}[|\mathcal{X}_1| \mid \mathcal{G}] \leq \frac{c}{\sqrt{K}} \cdot \mathbb{E}_{\mathbf{x}}[|\mathcal{X}| \mid \mathcal{G}]^2. \tag{4.12}$$

We can take K such that $\sqrt{K} > 64c$. It remains to bound the conditional mean of $|\mathcal{X}_2|$ knowing $\mathcal{G}$. One can find a constant $K' > 0$, such that

$$\mathbb{P}_{\mathbf{x}}(D) \geq 7/8, \quad \text{for} \quad D := \left\{ K\rho r^{d-2} \leq \mathcal{N}_r \leq K'\rho r^{d-2} \right\}.$$

We denote by $\mathcal{E}_1, \dots, \mathcal{E}_{\mathcal{N}_r}$, the $\mathcal{N}_r$ excursions hitting $Q_r(v_i)$ under $\mathbb{P}_{\mathbf{x}}$. Fix some $y, z \in Q_r(v_i)$, and let

$$\mathcal{I}_y := \{k \leq \mathcal{N}_r \,:\, y \in \mathcal{E}_k\}.$$

By definition, for any $k \leq \mathcal{N}_r$,

$$\mathbb{P}_{\mathbf{x}}(z \in \mathcal{E}_k \mid \mathcal{G},\, k \notin \mathcal{I}_y) \leq \frac{\mathbb{P}_{\mathbf{x}}(z \in \mathcal{E}_k \mid \mathcal{G})}{\mathbb{P}_{\mathbf{x}}(y \notin \mathcal{E}_k \mid \mathcal{G})} \leq \frac{\mathbb{P}_{\mathbf{x}}(z \in \mathcal{E}_k \mid \mathcal{G})}{1 - cr^{2-d}} \leq \mathbb{P}_{\mathbf{x}}(z \in \mathcal{E}_k \mid \mathcal{G}) + \mathcal{O}(r^{2(2-d)}),$$

for some constant $c > 0$. As a consequence, on the event D, and for r large enough,

$$\begin{aligned}
\mathbb{P}_{\mathbf{x}}\left(z \in \bigcup_{k \notin \mathcal{I}_y} \mathcal{E}_k \;\middle|\; \mathcal{G},\, \mathcal{I}_y \right) &= 1 - \prod_{k \notin \mathcal{I}_y} \Big(1 - \mathbb{P}_{\mathbf{x}}(z \in \mathcal{E}_k \mid \mathcal{G},\, \mathcal{I}_y)\Big) \\
&\leq 1 - \prod_{k \notin \mathcal{I}_y} \Big(1 - \mathbb{P}_{\mathbf{x}}(z \in \mathcal{E}_k \mid \mathcal{G}) - \mathcal{O}(r^{2(2-d)})\Big) \\
&\leq 1 - \prod_{k \leq \mathcal{N}_r} \Big(1 - \mathbb{P}_{\mathbf{x}}(z \in \mathcal{E}_k \mid \mathcal{G})\Big) + \mathcal{O}\left(\frac{\mathcal{N}_r}{r^{2(d-2)}}\right) \\
&= \mathbb{P}_{\mathbf{x}}(z \in \mathcal{X} \mid \mathcal{G}) + \mathcal{O}\left(\frac{\mathcal{N}_r}{r^{2(d-2)}}\right),
\end{aligned} \tag{4.13}$$

where at the penultimate line we use that on D, and when r is large enough, the term $\mathcal{O}(\mathcal{N}_r/r^{2(d-2)})$ can be made smaller than 1. Then on the event D, we get from (4.13),

$$\mathbb{P}_{\mathbf{x}}((y,z) \in \mathcal{X}_2 \mid \mathcal{G}) \leq \Big(\mathbb{P}_{\mathbf{x}}(z \in \mathcal{X} \mid \mathcal{G}) + \mathcal{O}(\rho r^{2-d})\Big) \cdot \mathbb{P}_{\mathbf{x}}(y \in \mathcal{X} \mid \mathcal{G}).$$

Summing over $y, z \in Q_r$, we deduce from (4.11), that on the event D,

$$\mathbb{E}_{\mathbf{x}}[|\mathcal{X}_2| \mid \mathcal{G}] \leq \mathbb{E}_{\mathbf{x}}[|\mathcal{X}| \mid \mathcal{G}]^2 (1 + \mathcal{O}(r^{2-d})).$$

Combining this with (4.12), we get for r large enough,

$$\mathrm{var}_{\mathbf{x}}(|\mathcal{X}| \mid \mathcal{G}) = \mathbb{E}_{\mathbf{x}}[|\mathcal{X}_2| \mid \mathcal{G}] + \mathbb{E}_{\mathbf{x}}[|\mathcal{X}_1| \mid \mathcal{G}] - \mathbb{E}_{\mathbf{x}}[|\mathcal{X}| \mid \mathcal{G}]^2 \leq \frac{1}{32} \cdot \mathbb{E}_{\mathbf{x}}[|\mathcal{X}| \mid \mathcal{G}]^2.$$

Together with (4.11), it follows that for r large enough, on the event D,

$$\mathbb{P}_{\mathbf{x}}(|\mathcal{X}| \leq \rho |Q_r| \mid \mathcal{G}) \leq \mathbb{P}_{\mathbf{x}}\left(|\mathcal{X}| \leq \frac{1}{2} \mathbb{E}_{\mathbf{x}}[|\mathcal{X}| \mid \mathcal{G}] \mid \mathcal{G} \right) \leq \frac{4\, \mathrm{var}_{\mathbf{x}}(|\mathcal{X}| \mid \mathcal{G})}{\mathbb{E}_{\mathbf{x}}[|\mathcal{X}| \mid \mathcal{G}]^2} \leq \frac{1}{8}.$$

Finally, using that $\mathbb{P}_{\mathbf{x}}(D) \geq 7/8$, we obtain the desired bound for r large enough,

$$\mathbb{P}_{\mathbf{x}}(|\mathcal{X}| \leq \rho |Q_r|) \leq \mathbb{P}_{\mathbf{x}}(D^c) + \mathbb{P}_{\mathbf{x}}(|\mathcal{X}| \leq \rho |Q_r|,\, D) \leq 1/4,$$

On the other hand, for small values of r, the result is immediate. This concludes the proof of (4.3) and the Proposition. □

Aknowledgements A.Asselah and B.Schapira were supported by public grants overseen by the french National Research Agency, ANR SWiWS (ANR-17-CE40-0032-02) and ANR MALIN (ANR-16-CE93-0003).

References

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

[AS17] Asselah Amine; Schapira Bruno. Moderate deviations for the range of a transient random walk: path concentration. Ann. Sci. Éc. Norm. Supér. (4) 50 (2017), 755–786.

[AS19a] Asselah Amine; Schapira Bruno. Moderate deviations for the range of a transient random walk, part II. arXiv:1909.01925.

[AS19b] Asselah Amine; Schapira Bruno. Deviations for the capacity of the range of a random walk. arXiv:1807.02325.

[BPP95] Benjamini, Itai; Pemantle, Robin; Peres, Yuval. Martin capacity for Markov chains. Ann. Probab. 23 (1995), 1332–1346.

[BBH] M. van den Berg, E. Bolthausen, F. den Hollander. Moderate deviations for the volume of the Wiener sausage. Ann. of Math. (2) 153 (2001), no. 2, 355–406.

[CN20a] Chiarini Alberto; Nitzschner Maximilian. Entropic repulsion for the Gaussian free field conditioned on disconnection by level-sets. To appear in Probab. Theory Related Fields.

[CN20b] Chiarini Alberto; Nitzschner Maximilian. Entropic repulsion for the occupation-time field of random interlacements conditioned on disconnection. To appear in Ann. Probab.

[E94] Eisenbaum, Nathalie. Dynkin's isomorphism theorem and the Ray-Knight theorems. Probab. Theory Related Fields 99 (1994), 321–335.

[K44] Kakutani, Shizuo. Two-dimensional Brownian motion and harmonic functions. Proc. Imp. Acad. Tokyo 20 (1944), 706–714.

[Kes90] Kesten, Harry. Upper bounds for the growth rate of DLA. Phys. A 168 (1990), 529–535.

[LL10] Lawler, Gregory F. and Limic, Vlada. Random walk: a modern introduction. Cambridge Studies in Advanced Mathematics, 123. Cambridge University Press, Cambridge, 2010.

[LS15] Li, Xinyi; Sznitman, Alain-Sol. Large deviations for occupation time profiles of random interlacements. Probab. Theory Related Fields 161 (2015), 309–350.

[LT19] Liddle, George; Turner, Amanda. Scaling limits and fluctuations for random growth under capacity rescaling, arXiv:1907.09321.

[NS20] Nitzschner, Maximilian; Sznitman, Alain-Sol. Solidification of porous interfaces and disconnection. To appear in J. Eur. Math. Soc.

[Polya] G.Pòlya & G.Szegö, Isoperimetric Inequalities in Mathematical Physics, Annals of MathematicsStudies, no. 27, Princeton University Press, Princeton, N. J., 1951.

[RS13] Rodriguez, Pierre-François; Sznitman, Alain-Sol. Phase transition and level-set percolation for the Gaussian free field. Comm. Math. Phys. 320 (2013), 571–601.

[S10] Sznitman, Alain-Sol. Vacant set of random interlacements and percolation, Ann. of Math. (2) 171, (2010), 2039–2087.

[S12] Sznitman, Alain-Sol. Topics in occupation times and Gaussian free fields. Zurich Lectures in Advanced Mathematics. European Mathematical Society (EMS), Zürich, 2012. viii+114 pp.

[S15] Sznitman, Alain-Sol. Disconnection and level-set percolation for the Gaussian free field. J. Math. Soc. Japan 67 (2015), 1801–1843.

[S17] Sznitman, Alain-Sol. Disconnection, random walks, and random interlacements. Probab. Theory Related Fields 167 (2017), 1–44.