Extracting subsets maximizing capacity and folding of random walks

Amine Asselah *   Bruno Schapira†

Abstract

We prove that given any finite set of $\mathbb{Z}^d$, 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. As an application, we obtain estimates on the probability a transient random walk *covers uniformly* a finite set. Finally, we characterize some *folding* events, under optimal hypotheses. For instance, knowing that a random walk folds to produce an *atypically high occupation density* somewhere, we show that the *folding region* is most likely ball-like, asymptotically as the length of the walk goes to infinity.

Keywords and phrases. Random walk; local times; capacity; range.
MSC 2010 *subject classifications.* Primary 60F05, 60G50.

Sous-ensembles de capacité maximale et marches aléatoires

Nous montrons que de toute partie finie de $\mathbb{Z}^d$, en dimension trois et plus, on peut extraire un sous-ensemble dont la capacité et le volume sont du même ordre de grandeur que la capacité de la partie initiale. Cette observation nous permet d'obtenir, sous des hypothèses optimales, des estimations de la probabilité qu'une marche aléatoire *recouvre uniformément* un ensemble fini. Enfin, nous caractérisons certains évènements de repliement de la marche. Par exemple, lorsque l'on sait qu'une marche aléatoire se replie pour produire une densité d'occupation atypiquement grande, alors la région de repliement a typiquement la forme d'une boule, au sens où sa capacité est du même ordre de grandeur que celle d'une boule.

Mots clés. Temps locaux; capacité; marche aléatoire.

1 Introduction

This note deals with *capacity* in the context of a random walk on $\mathbb{Z}^d$, with $d \geq 3$. If $\mathbb{P}_x$ is the law of the random walk starting from x, Λ is a non-empty finite subset of $\mathbb{Z}^d$ and H_Λ^+ is the return time into Λ, then the capacity of Λ is

$$\mathrm{cap}(\Lambda) := \sum_{x\in\Lambda} \mathbb{P}_x\big(H_\Lambda^+ = \infty\big). \tag{1.1}$$

Our main observation is that in any finite subset of $\mathbb{Z}^d$, say made of disjoint balls with common radius r, there exists a subset whose size and capacity are both of order the capacity of the initial

*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

set. To state precisely this result, let us introduce the needed notation. For $x \in \mathbb{Z}^d$, and $r \geq 1$, we define $B_r(x) = \{z \in \mathbb{Z}^d : \|z - x\| < r\}$, with $\| \cdot \|$ the Euclidean norm, and for $\mathcal{C} \subset \mathbb{Z}^d$, we let $B_r(\mathcal{C}) := \cup_{x \in \mathcal{C}} B_r(x)$.

In the whole paper, we deal with space dimension three and higher, and all our results assume this hypothesis.

Theorem 1.1. *There exists $\alpha > 0$, such that for any $r \geq 1$ and any finite $\mathcal{C} \subset \mathbb{Z}^d$, there is a subset $U \subseteq \mathcal{C}$, satisfying*

$$(i) \quad \mathrm{cap}(B_r(U)) \geq \alpha \cdot r^{d-2}|U| \quad \text{and} \quad (ii) \quad r^{d-2}|U| \geq \alpha \cdot \mathrm{cap}(B_r(\mathcal{C})). \tag{1.2}$$

We now present two applications of this result, Theorems 1.2 and 1.4 below. The former deals informally with the event that a random walk *covers uniformly* a fraction ρ of a set, and bounds the probability of such event by exponential minus ρ times the capacity of the set, under some optimal assumptions on ρ and the scale at which we measure the occupation density. The latter, Theorem 1.4, deals with the shape of the folding region for a walk conditioned on squeezing part of its range, and shows that this region is typically ball-like in the sense that its capacity is of smallest possible order, that is with capacity of order its volume to the power $1 - 2/d$, as it is for balls. This has some natural applications in the context of moderate deviations for the volume or the capacity of the range of the walk, as shown in [AS17, AS21a, AS21c].

Let us also mention that Theorem 1.1 has found application in the context of random interlacements [Sz20, Sz21].

To be more precise now, for $\Lambda \subset \mathbb{Z}^d$ made of disjoint balls of radius r, consider the event obtained by asking the random walk to spend a time $\rho \cdot r^d$ in each ball making Λ, for some $\rho > 0$. We have shown in [AS17] how to relate the probability of such covering event with the capacity of Λ. Let $\{S_n\}_{n \in \mathbb{N}}$ denote the discrete-time simple random walk, and $\mathbb{P}$ be its law when starting from the origin. At a time $n \in \mathbb{N} \cup \{\infty\}$, and site $z \in \mathbb{Z}^d$, the local time reads

$$\ell_n(z) := \sum_{k=0}^{n} \mathbf{1}\{S_k = z\} \quad \text{and for } \Lambda \subset \mathbb{Z}^d, \ \ell_n(\Lambda) := \sum_{z \in \Lambda} \ell_n(z). \tag{1.3}$$

Theorem 1.2. *There exist positive constants A and κ, such that for any $r \geq 1$ and $\rho > 0$ satisfying*

$$\rho r^{d-2} > A, \tag{1.4}$$

one has for any finite $\mathcal{C} \subset \mathbb{Z}^d$

$$\mathbb{P}\big(\ell_\infty(B_r(x)) > \rho r^d, \quad \forall x \in \mathcal{C}\big) \leq \exp\big(-\kappa \cdot \rho \cdot \mathrm{cap}(B_r(\mathcal{C}))\big). \tag{1.5}$$

The condition (1.4) improves the condition in Proposition 1.7 of [AS17]: $\rho r^{d-2} > A|\mathcal{C}|^{2/d} \cdot \log(n)$. Eliminating the term $|\mathcal{C}|^{2/d}$ is a serious issue which requires Theorem 1.1. Going to an infinite-time setting is straightforward, and is explained in the proof of Theorem 1.2 in Section 4.2.

Note that (1.4) is optimal since typically a walk spends a time of order r^2 in a ball of radius r, conditionally on visiting it. In the case $r = 1$ (when $B_r(x) = \{x\}$ for all x), we obtain a stronger and more general result. First, the result holds true for any $\rho > 0$ and we show that we can take the constant κ equal to one in (1.5). Furthermore, we can deal with non-uniform covering and general transient walks. We refer to Theorem 4.1 in Section 4 for a precise statement.

Remark 1.3. We note that Sznitman obtained results with a similar flavor as Theorem 1.2 in the context of the Gaussian free field (GFF) and in the model of random interlacements, respectively in [Sz15, Corollary 4.4] and [Sz17, Theorem 4.2] (see also [LSz15] for related results).

Our second application deals with finite times. For $r \geq 1$, $\rho > 0$, $n \geq 1$, and $\mathcal{C} \subset \mathbb{Z}^d$ finite, we consider

$$\mathcal{F}_n(r, \rho, \mathcal{C}) := \{\forall x \in \mathcal{C}, \quad \ell_n(B_r(x)) > \rho r^d\}. \tag{1.6}$$

In many folding problems, one central issue is to characterize the size and the shape of the folding region $\mathcal{C}$, which might be random. More precisely, one may consider folding events of the form $\cup_{\mathcal{C}} \mathcal{F}_n(r, \rho, \mathcal{C})$, where the union is over all $\mathcal{C} \subseteq [-n, n]^d$, with only a lower bound on their volume, say $|\mathcal{C}| \geq L$, when $B_r(\mathcal{C})$ is made of disjoint balls. Then Theorem 1.2 and a naive union bound gives

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

using the lower bound on capacity (2.5). The bound just obtained is useful only when

$$\rho \cdot r^{d-2} \geq C L^{2/d} \cdot \log(n). \tag{1.7}$$

Now Theorem 1.1 allows to go beyond this condition (1.7), and gives

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

under the weaker assumption:

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

The latter is of crucial importance in [AS21c], and can also be used to characterize the shape of a localization region for a random walk, 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 $r \geq 1$, and $x \in \mathbb{Z}^d$,

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

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

$$\mathcal{C}_n(r, \rho) := \{x \in r\mathbb{Z}^d \, : \, \ell_n(Q_r(x)) > \rho r^d\}, \quad \text{and} \quad \mathcal{V}_n(r, \rho) := \bigcup_{x \in \mathcal{C}_n(r, \rho)} Q_r(x). \tag{1.8}$$

We can now state our third result.

Theorem 1.4. *There are positive constants $\underline{\kappa}$, $\overline{\kappa}$, and C, such that for any $n \geq 2$, r and L positive integers and $\rho > 0$, satisfying*

$$\rho r^{d-2} \geq C \cdot \log(n), \quad \textit{and} \quad n \geq C \rho r^d L, \tag{1.9}$$

one has

$$\exp\big(-\underline{\kappa} \cdot \rho \cdot r^{d-2} \cdot L^{1-2/d}\big) \leq \mathbb{P}\big(|\mathcal{C}_n(r, \rho)| > L\big) \leq \exp\big(-\overline{\kappa} \cdot \rho \cdot r^{d-2} \cdot L^{1-2/d}\big). \tag{1.10}$$

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

$$\lim_{n \to \infty} \inf_{(r, \rho, L)} \mathbb{P}\big(\operatorname{cap}(\mathcal{V}_n(r, \rho)) \leq A \cdot |\mathcal{V}_n(r, \rho)|^{1-2/d} \, \big| |\mathcal{C}_n(r, \rho)| > L\big) = 1, \tag{1.11}$$

where the infimum is taken over all triples (r, ρ, L) satisfying (1.9).

Let us stress that conditions (1.9) are optimal in the following sense. Concerning the first inequality, just recall that in time n, the walk typically fills balls with an occupation density of order $r^{2-d}\log n$; and for the second inequality, which is only needed for the lower bound in (1.10), note that one needs at least $n \geq \rho r^d L$, for the set $\{|\mathcal{C}_n(r,\rho)| > L\}$ to be non-empty. Let us also mention here that we obtain a similar result as Theorem 1.4, where instead of recording the time spent in small cubes, we count the number of visited sites, see Proposition 5.1 for details.

Remark 1.5. Note that the result is interesting in its own right even 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 1.4 to these sets gives that for $\rho \geq C \cdot \log(n)$ and $n \geq C\rho \cdot L$,

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

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 $n \geq C\rho_n \cdot L_n$

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

Remark 1.6. For simplicity, we focus here on the case of the discrete time simple random walk, but our results would likely adapt to a more general setting, such as finite range random walks.

Historical account. Let us put our results into perspective. Capacity appears as a central object in many remarkable studies, and we would like to highlight some of them. In the thirties, Wiener introduces his celebrated test, where the electrostatic capacity plays the key role, and is adapted to random walk context by Itô and McKean much later [IK60]. 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. Much later, 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 [LT21] for a related model). Finally, ten years ago, Sznitman [Sz10] introduces 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 set K is a Poisson random variable with mean $u \cdot \operatorname{cap}(K)$, and whose hitting sites distribution on K 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 atypically high densities: (i) either by reducing its range, and in a certain regime this is the *Swiss cheese* problem (see [BBH01]), (ii) or by disconnecting the ball $B_n(0)$ from the complement of $B_{2n}(0)$, and many more sophisticated events, see in particular [Sz17, NSz20, Sz20, Sz21].

Concerning the deviations for local times, a rich literature exists on large deviations for the field of renormalized local times, initiated by Donsker and Varadhan [DV75], or for self-intersection local times, see [Chen09] and references therein. However, it seems that not much is known concerning the deviations of local times of a random walk on a fixed finite set of cardinal at least two.

The paper is organized as follows. In Section 2.2 we recall some basic facts on the capacity. Section 3 contains our main technical novelty: the proof of Theorem 1.1. In Section 4, we prove Theorem 1.2, and introduce a related result Theorem 4.1 of a similar flavor, but dealing with the local times of sites. Finally, in Section 5 we prove Theorem 1.4. The proof is divided into a short upper bound, and a technical lower bound in Section 5.2 where we actually state Proposition 5.1 which deals with the (slightly more difficult) problem of covering a certain partition of the space, rather than with local times.

2 Preliminaries

2.1 Notations

In the rest of the paper, we use c, C as generic constants, changing from place to place, depending only on the dimension d. Similarly, when the constants are numbered, except that they no more change value from line to line.

We need first to define the hitting and return times of the random walk to a non-empty set Λ in $\mathbb{Z}^d$ as respectively

$$H_\Lambda = \inf\{n \geq 0 : \ S_n \in \Lambda\}, \quad \text{and} \quad H_\Lambda^+ = \inf\{n \geq 1 : \ S_n \in \Lambda\},$$

which we shall sometimes also write as $H(\Lambda)$ and $H^+(\Lambda)$ respectively. We recall that Green's function $G(x)$ is the average number of visits to site $x \in \mathbb{Z}^d$ when the random walk starts at 0, and the well known asymptotics (see precise bounds in Theorem 4.3.1 of [LL10])

$$G(z) := \sum_{n\geq 0} \mathbb{P}(S_n = z), \quad \text{and} \quad G(z) \leq \frac{C}{1 + \|z\|^{d-2}}. \tag{2.1}$$

We write $B_r(x) = \{y \in \mathbb{Z}^d : \ \|x - y\| < r\}$ for the Euclidean ball of radius r and center x, B_r stands for $B_r(0)$, and Q_r stands for the cube $Q_r(0)$. For a non-empty set Λ in $\mathbb{Z}^d$, we call $\partial\Lambda$ its outer boundary, i.e. $\partial\Lambda := \{y \in \mathbb{Z}^d \backslash \Lambda : \ \exists x \in \Lambda, \|x - y\| = 1\}$, whereas $\partial_i \Lambda$ stands for the inner boundary, that is $\partial\Lambda^c$, where $\Lambda^c = \mathbb{Z}^d \backslash \Lambda$.

2.2 On capacity

We recall here some alternative definitions of the capacity, and refer to [LL10] for proofs of these standard facts. The first alternative and equivalent definition is in terms of hitting time, rather than escape probabilities (see Proposition 6.5.1 in [LL10]):

$$\operatorname{cap}(\Lambda) = \lim_{\|z\|\to\infty} \frac{1}{G(z)} \cdot \mathbb{P}_z\big(H_\Lambda < \infty\big). \tag{2.2}$$

A third classical way to define the capacity is given by the following variational formula (see for instance Proposition 1.9 of [Sz12] for a proof).

$$\frac{1}{\operatorname{cap}(\Lambda)} = \inf\Big\{ \sum_{x\in\Lambda}\sum_{y\in\Lambda} G(x-y)\mu(x)\mu(y) : \ \mu \text{ probability on } \Lambda \Big\}. \tag{2.3}$$

The infimum is reached for the *equilibrium measure* e_Λ, defined for $x \in \Lambda$ by $e_\Lambda(x) = \mathbb{P}_x(H_\Lambda^+ = \infty)/\operatorname{cap}(\Lambda)$.

As we already mentioned, the capacity of a ball $B_r(x)$ is of order r^{d-2}, and more generally, there exists a constant $c > 0$, such that for any $\Lambda \subset \mathbb{Z}^d$,

$$\mathrm{cap}(\Lambda) \geq c|\Lambda|^{1-2/d}. \tag{2.4}$$

(Note that this follows from (2.1) and taking μ uniform in (2.3).) When applied to a union of disjoint balls, this gives

$$\mathrm{cap}(B_r(\mathcal{C})) \geq c \cdot r^{d-2}|\mathcal{C}|^{1-2/d}. \tag{2.5}$$

This bound cannot be improved. Looking now for an upper bound of the capacity of a union of balls, subadditivity of the capacity gives that it is always bounded (up to constant) by the number of balls times r^{d-2}. However, one can improve this crude bound using (2.2) yielding

$$\mathrm{cap}(B_r(\mathcal{C})) \leq C \cdot r^{d-2} \cdot \mathrm{cap}(\mathcal{C}). \tag{2.6}$$

We shall also need the following lemma. For $r \geq 1$, we denote by $\mathcal{X}_r$ the set of finite $\mathcal{C} \subset \mathbb{Z}^d$, whose points are all at Euclidean distance at least $4r$ from each other.

Lemma 2.1. *There exists a constant $c > 0$, such that for any finite $\mathcal{C} \subset \mathbb{Z}^d$, there exists a subset $\mathcal{C}' \subset \mathcal{C}$, with $\mathcal{C}' \in \mathcal{X}_r$, satisfying*

$$\mathrm{cap}(B_r(\mathcal{C}')) \geq c \cdot \mathrm{cap}(B_r(\mathcal{C})).$$

Proof. We define recursively $\mathcal{C}_n \subset \mathcal{C}$, for $1 \leq n \leq |\mathcal{C}|$ as follows. First pick a point x_1 in $\mathcal{C}$, and set $\mathcal{C}_1 := \{x_1\}$. Then assuming $\mathcal{C}_n$ has been defined for some $n < |\mathcal{C}|$, define $\mathcal{C}_{n+1}$ as the union of $\mathcal{C}_n$ and an arbitrarily chosen point of $\mathcal{C} \setminus (\cup_{x \in \mathcal{C}_n} B_{4r}(x))$, if this set is nonempty. Otherwise, set $\mathcal{C}_{n+1} := \mathcal{C}_n$. Define $\mathcal{C}'$ as the set one eventually obtains. Note that by construction $\mathcal{C}' \in \mathcal{X}_r$.

We express now the hitting time of $B_r(\mathcal{C}')$ and use that $B_r(\mathcal{C}) \subset B_{5r}(\mathcal{C}')$ to obtain for any $z \in \mathbb{Z}^d$,

$$\begin{aligned}
\mathbb{P}_z(H(B_r(\mathcal{C}')) < \infty) =& \mathbb{P}_z(H(B_{5r}(\mathcal{C}')) < \infty) \times \mathbb{P}_z(H(B_r(\mathcal{C}')) < \infty \mid H(B_{5r}(\mathcal{C}')) < \infty) \\
\geq& \mathbb{P}_z(H(B_r(\mathcal{C})) < \infty) \times \mathbb{P}_z(H(B_r(\mathcal{C}')) < \infty \mid H(B_{5r}(\mathcal{C}')) < \infty).
\end{aligned}$$

Since after arriving on the boundary of a ball $B_{5r}(x)$, for $x \in \mathcal{C}'$, the random walk hits $B_r(x)$ with a positive probability, say c independent of r and $\mathcal{C}'$, we obtain by the strong Markov property

$$\mathbb{P}_z(H(B_r(\mathcal{C}')) < \infty \mid H(B_{5r}(\mathcal{C}')) < \infty) \geq c.$$

The proof ends as we recall (2.2), normalize $\mathbb{P}_z(H(B_r(\mathcal{C}')) < \infty)$ by $G(z)$ and let $\|z\|$ tend to infinity. $\square$

3 Proof of Theorem 1.1

The proof is an instance of the probabilistic method: we define an appropriately chosen random subset of $\mathcal{C}$, and show that it satisfies the desired constraints with nonzero probability.

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

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

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

$$\mathbb{P}\big(|\mathcal{U}| < \frac{1}{2}\mathbb{E}\big[|\mathcal{U}|\big]\big) \leq \frac{4}{\mathrm{cap}(\Lambda)}, \quad \text{and} \quad \mathbb{P}\big(|\mathcal{U}| > 2\mathbb{E}\big[|\mathcal{U}|\big]\big) \leq \frac{1}{\mathrm{cap}(\Lambda)}. \tag{3.1}$$

We can assume $\mathrm{cap}(\Lambda) > 16$, (as for sets with bounded capacity one can always choose α small enough) so that (3.1) reads

$$\mathbb{P}\big(2\,\mathrm{cap}(\Lambda) \geq |\mathcal{U}| \geq \frac{1}{2}\,\mathrm{cap}(\Lambda)\big) \geq \frac{2}{3}. \tag{3.2}$$

Now, we show that $\mathrm{cap}(\mathcal{U})$ is typically of order its volume. Assume $|\mathcal{U}| > 0$, and choose for μ the uniform measure on $\mathcal{U}$. By (2.3), we have

$$\frac{\mathrm{cap}(\mathcal{U})}{|\mathcal{U}|} \geq \Big(\frac{1}{|\mathcal{U}|}\sum_{x,y\in\mathcal{U}} G(x-y)\Big)^{-1}. \tag{3.3}$$

Let us compute the expression on the right hand side of (3.3).

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

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

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

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,

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

Together with (3.2), we obtain

$$\mathbb{P}\left(2\,\mathrm{cap}(\Lambda) \geq |\mathcal{U}| \geq \frac{1}{2}\,\mathrm{cap}(\Lambda),\ \sum_{x,y\in\mathcal{U}} G(x-y) \leq 4\big(G(0)+1\big)\cdot\mathrm{cap}(\Lambda)\right) \geq \frac{5}{12}. \tag{3.4}$$

By (3.3) and (3.4), we deduce that for some $\alpha > 0$,

$$\mathbb{P}\left(2\,\mathrm{cap}(\Lambda) \geq |\mathcal{U}| \geq \frac{1}{2}\,\mathrm{cap}(\Lambda),\ \mathrm{cap}(\mathcal{U}) \geq \alpha \cdot \mathrm{cap}(\Lambda)\right) \geq \frac{5}{12}. \tag{3.5}$$

Since the right-hand side is positive, we conclude that (1.2) holds for a random set $\mathcal{U}$, when $r = 1$. We now prove the general case by refining the previous argument.

General case $r \geq 1$. The proof follows the same steps after we choose an appropriate random subset of the set of centers $\mathcal{C}$. First, by Lemma 2.1 one can assume that $\mathcal{C} \in \mathcal{X}_r$ (i.e. that points of $\mathcal{C}$ are all at distance at least $4r$ from each other). For any $r > 0$, we set $\Lambda_r = B_r(\mathcal{C})$, and $V_r = \mathbb{Z}^d \backslash B_{2r}(\mathcal{C})$. We need now the hitting time of Λ_r after exiting $B_{2r}(\mathcal{C})$. For a trajectory γ, define

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

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

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

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 simple Lemma, whose proof is recalled at the end of this section for the reader's convenience.

Lemma 3.1. *There exists a constant $\theta > 1$, such that for any $r \geq 1$, $\mathcal{C} \in \mathcal{X}_r$, and $x \in \mathcal{C}$,*

$$\theta\mathbb{P}\big(H^r_{\Lambda_r}(\gamma_x) = \infty\big) \geq \frac{1}{r^{d-2}} \sum_{y\in\partial_i B_r(x)} \mathbb{P}\big(H^+_{\Lambda_r}(y + S) = \infty\big) \geq \frac{1}{\theta}\mathbb{P}\big(H^r_{\Lambda_r}(\gamma_x) = \infty\big). \tag{3.6}$$

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

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

Furthermore, thanks to Lemma 3.1, there are positive constants c_1 and c_2, such that

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

It follows that for a positive constant c_3,

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

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

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

Thus, with probability larger than or equal to $3/4$, the random set $\mathcal{U}$ satisfies (ii) of (1.2). Let us check now (i). By (2.3), 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 inner boundary of $B_r(\mathcal{U})$ gives

$$\frac{1}{(|\partial_i B_r| \cdot |\mathcal{U}|)^2} \sum_{x,x'\in\mathcal{U}} \sum_{y\in\partial_i B_r(x)} \sum_{y'\in\partial_i B_r(x')} G(y - y') \geq \frac{1}{\mathrm{cap}(B_r(\mathcal{U}))}. \tag{3.7}$$

We need to show that the left hand side of (3.7) is smaller than a constant times $1/(r^{d-2}|\mathcal{U}|)$. First, we treat the case $x' = x$. Note that by Green's function asymptotic (2.1), there is $C > 0$, such that

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

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

$$\sum_{x \in \mathcal{U}} \sum_{y \in \partial_i B_r(x)} \sum_{y' \in \partial_i B_r(x)} G(y - y') \le C \cdot r \cdot r^{d-1} \cdot |\mathcal{U}|. \tag{3.8}$$

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 $y \in \partial_i B_r(x)$ and $y' \in \partial_i B_r(x')$. Therefore,

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

By using (3.6) of Lemma 3.1, and a last passage decomposition we have for a constant $c_6 > 0$, and any $x \in \mathcal{C}$,

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

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

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

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

We end this section with a proof of Lemma 3.1.

Proof of Lemma 3.1. The first inequality is exactly Lemma 5.2 of [AS17], to which we refer for a proof. Concerning the second inequality, note that Proposition 1.5.10 of [Law91] shows in particular that for a positive constant c, for any $y \in \partial_i B_r(x)$,

$$\mathbb{P}_y\big(H_{\partial B_{2r}(x)} < H^+_{B_r(x)}\big) \ge \frac{c}{r}.$$

Now using that points of $\mathcal{C}$ are at distance at least $4r$ from each other, and applying the strong Markov property, we get

$$\mathbb{P}_y\big(H^+_{B_r(\mathcal{C})} = \infty\big) = \sum_{z \in \partial B_{2r}(x)} \mathbb{P}_y\big(S_{H_{\partial B_{2r}(x)}} = z,\ H_{\partial B_{2r}(x)} < H^+_{B_r(x)}\big) \cdot \mathbb{P}_z\big(H_{B_r(\mathcal{C})} = \infty\big).$$

Then Lemma 6.3.7 and Proposition 6.4.4. in [LL10] give for some constant $c' > 0$,

$$\mathbb{P}_y\big(H^+_{B_r(\mathcal{C})} = \infty\big) \geq \frac{c'}{r^{d-1}} \mathbb{P}_y\big(H_{\partial B_{2r}(x)} < H^+_{B_r(x)}\big) \cdot \sum_{z \in \partial B_{2r}(x)} \mathbb{P}_z\big(H_{B_r(\mathcal{C})} = \infty\big)$$
$$\geq \frac{cc'}{r^d} \sum_{z \in \partial B_{2r}(x)} \mathbb{P}_z\big(H_{B_r(\mathcal{C})} = \infty\big).$$

On the other hand, Harnack's inequality (see Theorem 6.3.9 in [LL10]) gives

$$\mathbb{P}\big(H^r_{\Lambda_r}(\gamma_x) = \infty\big) \leq \frac{C}{r^{d-1}} \sum_{z \in \partial B_{2r}(x)} \mathbb{P}_z\big(H_{B_r(\mathcal{C})} = \infty\big),$$

and altogether this concludes the proof of the lemma. □

4 Proof of Theorem 1.2

We start with a variant of Theorem 1.2 which deals with the event of visiting a certain number of times each site of a set, and connect the probability of such an event with the capacity of the set. Let $q := \mathbb{P}(H^+_{\{0\}} < \infty)$, and for a finite set $\Lambda \subset \mathbb{Z}^d$, and $z \in \Lambda$, let $q_z = q_{z,\Lambda} := \mathbb{P}_z(H^+_\Lambda < \infty)$.

Theorem 4.1. *Assume $d \geq 3$, and let Λ be a finite subset of $\mathbb{Z}^d$. Then, for any set of nonnegative integers $\{n_z\}_{z\in\Lambda}$,*

$$\mathbb{P}(\ell_\infty(z) \geq n_z \quad \forall z \in \Lambda) \leq \frac{\prod_{z\in\Lambda} q_z^{n_z}}{\min_{z\in\Lambda} q_z}. \tag{4.1}$$

In particular for all $t \geq 1$,

$$\mathbb{P}(\ell_\infty(z) \geq t \quad \forall z \in \Lambda) \leq \frac{1}{q} \exp\big(-t \cdot \mathrm{cap}(\Lambda)\big). \tag{4.2}$$

Both Theorems 1.2 and 4.1 use improvements of the proof of Proposition 1.7 of [AS17]. We present a simple self-contained proof of Theorem 4.1, and note that it works in fact for any random walk in a translation invariant setting.

4.1 Proof of Theorem 4.1

The proof proceeds by induction on $N := \sum_{z\in\Lambda} n_z$. Our induction hypothesis is that for any $y \in \mathbb{Z}^d$, any $\Lambda \subset \mathbb{Z}^d$ finite, and any family of nonnegative integers $\{n_z\}_{z\in\Lambda}$, we have

$$\mathbb{P}_y\big(\ell_\infty(z) \geq n_z,\ \forall z \in \Lambda\big) \leq \frac{\prod_{z\in\Lambda} q_z^{n_z}}{\min_{z\in\Lambda} q_z}.$$

Let us stress that it is important in the proof to allow some integers n_z to be equal to 0. If $N = 0$ or $N = 1$, there is nothing to prove, since the right-hand side of (4.1) is larger than or equal to 1 in this case. Assume now that the result is true for any family $\{n_z\}_{z\in\Lambda}$, with $\sum_{z\in\Lambda} n_z \leq N$, for some $N \geq 1$, and consider another family (which we still denote by $\{n_z\}_{z\in\Lambda}$) satisfying $\sum n_z = N + 1$.

If $0 \in \Lambda$, and $n_0 \geq 1$, we write (recalling that in the definition of local times, the time 0 is taken into account), with $\Lambda^* := \Lambda \setminus \{0\}$,

$$\begin{aligned}
\mathbb{P}_0(\ell_\infty(z) \geq n_z \ \forall z \in \Lambda) &= \sum_{y \in \Lambda \colon n_y \geq 1} \mathbb{P}_0(H_\Lambda^+ < \infty, S_{H_\Lambda^+} = y) \cdot \mathbb{P}_y(\ell_\infty(z) \geq n_z \ \forall z \in \Lambda^*, \ \ell_\infty(0) \geq n_0 - 1) \\
&\leq \sum_{y \in \Lambda \colon n_y \geq 1} \mathbb{P}_0(H_\Lambda^+ < \infty, S_{H_\Lambda^+} = y) \cdot \frac{(\prod_{z \in \Lambda^*} q_z^{n_z}) \cdot q_0^{n_0 - 1}}{\min_{z \in \Lambda} q_z} \\
&\leq \mathbb{P}_0(H_\Lambda^+ < \infty) \frac{(\prod_{z \in \Lambda^*} q_z^{n_z}) q_0^{n_0 - 1}}{\min_{z \in \Lambda} q_z} = \frac{\prod_{z \in \Lambda} q_z^{n_z}}{\min_{z \in \Lambda} q_z},
\end{aligned} \tag{4.3}$$

using the induction hypothesis at the second line. Now, when $0 \notin \Lambda$, or when $0 \in \Lambda$ and $n_0 = 0$, one can simply bound the probability on the left hand side above by the probability to hit a point $y \in \Lambda$ with $n_y \geq 1$, and then by translation-invariance and the Markov property, we are back to the previous situation. This concludes the proof of the first assertion (4.1).

The second assertion (4.2) follows immediately from (4.1), using that for any $z \in \Lambda$,

$$q_z = 1 - \mathbb{P}_z(H_\Lambda^+ = \infty) \leq \exp(-\mathbb{P}_z(H_\Lambda^+ = \infty)).$$

This concludes the proof. $\square$

4.2 Proof of Theorem 1.2

First, using Lemma 2.1, one can assume that all points of $\mathcal{C}$ are at distance at least $4r$ from each other, as stated in [AS17]. The proof of Proposition 1.7 of [AS17] then shows that for some positive constants κ, and K, for all $r \geq 1$, $\rho > 0$, and $\mathcal{C} \in \mathcal{X}_r$,

$$\mathbb{P}\big(\forall x \in \mathcal{C}, \ \ell_\infty(B_r(x)) > \rho r^d\big) \leq K \cdot 2^{|\mathcal{C}|} \exp\big(-\kappa \cdot \rho \cdot \mathrm{cap}(B_r(\mathcal{C}))\big).$$

Indeed, observe that the proof of (5.7) in [AS17] works as well if we take $n = \infty$, and consider a number of excursions larger than $\{n_z\}_{z \in \mathcal{C}}$ (instead of equal to it), as in (4.3) of the previous proof. Also, notice that all terms of the form $\binom{|\mathcal{C}|}{k}$ appearing in the former proof can be bounded by $2^{|\mathcal{C}|}$ rather than $|\mathcal{C}|!$. Note furthermore that the constant K above may be removed at the cost of taking the constant A large enough in the hypothesis $\rho \cdot r^{d-2} > A$ of Theorem 1.2, when we use (2.5). Then, Theorem 1.1 gives the existence of a subset $U \subseteq \mathcal{C}$, with $\mathrm{cap}(B_r(U))$ of the same order as both $r^{d-2} \cdot |U|$ and $\mathrm{cap}(B_r(\mathcal{C}))$, which entails (for some possibly different constant κ),

$$\mathbb{P}\big(\forall x \in \mathcal{C}, \ \ell_\infty(B_r(x)) > \rho r^d\big) \leq 2^{|U|} \exp\big(-\kappa \cdot \rho r^{d-2} \cdot |U|\big).$$

Now the hypothesis (1.4) allows to remove the term $2^{|U|}$ on the right-hand side of the above inequality (at the cost of taking a smaller constant κ if necessary), and this concludes the proof of the theorem, using again that $r^{d-2}|U|$ is of the same order as $\mathrm{cap}(B_r(\mathcal{C}))$. $\square$

5 Application to folding, and proof of Theorem 1.4

The proof of Theorem 1.4 is divided in two parts. In the first part (see Subsection 5.1 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 (1.9),

$$\mathbb{P}\left(|\mathcal{C}_n(r,\rho)| > L, \ \mathrm{cap}(\mathcal{V}_n(r,\rho)) > A|\mathcal{V}_n(r,\rho)|^{1-2/d}\right) \leq \exp(-\widetilde{\kappa} A \cdot \rho \cdot r^{d-2} L^{1-2/d}). \tag{5.1}$$

In the second part (see Subsection 5.2 below) we prove the lower bound in (1.10). Recall that a proof of the upper bound was already given in the introduction, see after (1.7), since with the notation thereof $\{|\mathcal{C}_n(r,\rho)| > L\} \subset \cup_{|\mathcal{C}|>L}\mathcal{F}_n(r,\rho,\mathcal{C})$. Note that altogether this gives (1.11) as well, and thus proves Theorem 1.4.

5.1 The upper bound: proof of (5.1)

We introduce the notation $Q_r(U)$ for $\cup_{x\in U}Q_r(x)$, and we use Theorem 1.1, with the condition (1.9) and then (1.5) as follows.

$$\mathbb{P}\big(|\mathcal{C}_n(r,\rho)| > L,\ \mathrm{cap}(\mathcal{V}_n(r,\rho)) \geq A\cdot|\mathcal{V}_n(r,\rho)|^{1-\frac{2}{d}}\big) \leq \sum_{k>L}\mathbb{P}\big(|\mathcal{C}_n(r,\rho)| = k, \mathrm{cap}(\mathcal{V}_n(r,\rho)) \geq A\cdot r^{d-2}k^{1-2/d}\big)$$

$$\leq \sum_{L<k\leq n}\mathbb{P}\big(\exists\,\mathcal{U}\subset[-n,n]^d:\ k\geq|\mathcal{U}|>\alpha Ak^{1-2/d},\ \mathrm{cap}(Q_r(\mathcal{U}))\geq\alpha r^{d-2}|\mathcal{U}|,\ \ell_n(Q_r(x))>\rho r^d\ \forall x\in\mathcal{U}\big)$$

$$\leq \sum_{L<k\leq n}\ \sum_{\alpha Ak^{1-2/d}<i\leq k}\mathbb{P}\big(\exists\,\mathcal{U}\subset[-n,n]^d:\ |\mathcal{U}|=i,\ \mathrm{cap}(Q_r(\mathcal{U}))\geq\alpha r^{d-2}i,\ \ell_n(Q_r(x))>\rho r^d\ \forall x\in\mathcal{U}\big)$$

$$\leq \sum_{L<k\leq n}\ \sum_{\alpha Ak^{1-2/d}<i\leq k}(2n+1)^{d\cdot i}\cdot\exp\big(-\kappa\alpha\rho r^{d-2}\cdot i\big)$$

$$\leq \sum_{k>L}\exp\big(-\widetilde{\kappa}\cdot\rho r^{d-2}k^{1-2/d}\big) \leq\ \exp\big(-\frac{\widetilde{\kappa}}{2}\cdot\rho r^{d-2}L^{1-2/d}\big),$$

for some constant $\widetilde{\kappa} > 0$. The combinatorial factor $(2n+1)^{d\cdot i}$ was swallowed after using the condition that $\rho r^{d-2} > C\log(n)$, and choosing C large enough.

5.2 Lower bound

In this subsection, we establish a result which slightly differs from the lower bound in (1.10), 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 $\rho\in[0,1]$,

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

Our result is as follows.

Proposition 5.1. *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

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

Note that this result implies the lower bound in (1.10), since $\widetilde{\mathcal{C}}_n(r,\rho)\subset\mathcal{C}_n(r,\rho)$. Note also that for the same reason, the upper bounds in (1.10), and in (1.11) hold as well for the set $\widetilde{\mathcal{C}}_n(r,\rho)$.

Remark 5.2. The hypothesis $\rho<1/2$ in Proposition 5.1 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 cube so that its occupation density is ρ. It is convenient to transform localization into a statement about excursions. To define properly an excursion, between ∂Q_{2R} and ∂Q_{4R}, we first define some stopping times. Let $\sigma_1 = \inf\{k \geq 0 : \ S_k \in \partial Q_{2R}\}$. Then, by induction, and as long as $\sigma_i < \infty$, define for $i \geq 1$,

$$\tau_i = \inf\{k \geq \sigma_i : \ S_k \in \partial Q_{4R}\}, \quad \text{and} \quad \sigma_{i+1} = \inf\{k \geq \tau_i : \ S_k \in \partial Q_{2R}\}.$$

On the other hand if $\sigma_i = \infty$, then the stopping times with larger indices are infinite as well. Now, the i-th excursion is the random walk trajectory with times within $[\sigma_i, \tau_i]$ when $\sigma_i < \infty$. The number of excursions before the walk escapes Q_{8R} is denoted $\mathcal{N}_R$ which reads

$$\mathcal{N}_R = \sup\{i : \ \sigma_i < \inf\{k : \ S_k \in \partial Q_{8R}\}\}.$$

We would like the cube Q_R to contain L cubes of side-length r each one filled with the random walk trace to a density above ρ. Also, we expect to localize the walk in Q_{8R} for a time T with

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

Consider the event

$$\mathcal{A}_1 := \{\mathcal{N}_R \geq N\}, \quad \text{with} \quad N = \lfloor C_2 \rho R^{d-2} \rfloor. \tag{5.2}$$

Now, for any C_2 (which tunes the desired number of excursions), one makes the event $\mathcal{A}_1$ typical by choosing C_1 large enough and imposing the walk to localize in Q_{8R} for a time T (which will lead to the desired conclusion by assuming also the time n to be larger than T). Then, a simple but key observation is that given $\mathbf{x} = (x_1, \dots, x_N) \in \partial Q_{2R}^N$, and $\mathbf{y} = (y_1, \dots, y_N) \in \partial Q_{4R}^N$, and conditioning the random walk on $\{S_{\sigma_i} = x_i \text{ and } S_{\tau_i} = y_i \text{ for } i = 1, \dots, N\}$, we have that the excursions $\{S_k, \ k \in \bigcup_{i \leq N} [\sigma_i, \tau_i]\}$ are independent from $\{S_k, \ k \in \bigcup_{i \leq N} [\tau_i, \sigma_{i+1}]\}$.

Let $\mathbb{P}_{\mathbf{x}}$ be the (product) law of N independent excursions starting from $\{x_i, \ i \leq N\}$, up to ∂Q_{4R}. We 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 Z to be the number of boxes of side-length r, whose fraction of visited sites, before the $\mathcal{N}_R$-th excursion, exceeds ρ. In other words, if $\mathcal{R}_{\tau_{\mathcal{N}_R}}$ is the set of visited sites before the $\mathcal{N}_R$-th excursion,

$$Z = \sum_{i=1}^{M} \mathbb{I}_{\{|\mathcal{R}_{\tau_{\mathcal{N}_R}} \cap Q_r(v_i)| > \rho \cdot r^d\}}.$$

We now define, in the space of N excursions starting from $\mathbf{x} = (x_1, \dots, x_N) \in \partial Q_{2R}^N$,

$$\mathcal{A}_2 := \left\{ \begin{array}{c} \text{Altogether the } N \text{ excursions visit at least } \rho r^d \text{ sites} \\ \text{of at least half of the boxes } \{Q_r(v_i)\}_{i \leq M} \end{array} \right\}.$$

With $\sigma = \inf\{n \geq 1 : S_n \in \partial Q_{2R} \cup \partial Q_{8R}\}$, and $\mathbf{y} = (y_1, \dots, y_N) \in \partial Q_{4R}^N$, we have

$$\mathbb{P}(\mathcal{A}_1 \cap \{Z \geq \frac{M}{2}\}, \forall i = 1, \dots, N, \ S_{\sigma_i} = x_i, \ S_{\tau_i} = y_i) \geq \prod_{i=1}^{N-1} \mathbb{P}_{y_i}\big(S_\sigma = x_{i+1}\big) \cdot \mathbb{P}_{\mathbf{x}}(\mathcal{A}_2, Y = \mathbf{y}). \tag{5.3}$$

Noting that $y \mapsto \mathbb{P}_y\big(S_\sigma = x\big)$ is harmonic, and using Harnack's inequality (see [LL10, Theorem 6.3.9]), we get for some constant $c_H > 0$, for any $x \in \partial Q_{2R}$,

$$\inf_y \mathbb{P}_y\big(S_\sigma = x\big) \geq c_H \mathbb{P}_{y^*}\big(S_\sigma = x\big), \quad \text{with} \quad y^* := (4R, 0, \dots, 0). \tag{5.4}$$

Assume for a moment that,

$$\forall \mathbf{x} \in (\partial Q_{2R})^N, \qquad \mathbb{P}_{\mathbf{x}}(\mathcal{A}_2) \geq 1/2. \tag{5.5}$$

Then, 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}, together with (5.3) and (5.4), we have $c_1 > 0$, such that

$$\begin{aligned}
\mathbb{P}(\mathcal{A}_1 \cap \{Z \geq \frac{M}{2}\}) &= \sum_{\mathbf{x}} \sum_{\mathbf{y}} \mathbb{P}(\mathcal{A}_1 \cap \{Z \geq \frac{M}{2}\}, (S_{\sigma_1}, \dots, S_{\sigma_N}) = \mathbf{x}, Y = \mathbf{y}) \\
&\geq \sum_{\mathbf{x}} c_H^{N-1} \prod_{i=1}^{N-1} \mathbb{P}_{y^*}\big(S_\sigma = x_{i+1}\big) \sum_{\mathbf{y}} \mathbb{P}_{\mathbf{x}}(\mathcal{A}_2, Y = \mathbf{y}) \\
&\geq c_H^{N-1} \inf_{\mathbf{x}} \mathbb{P}_{\mathbf{x}}(\mathcal{A}_2) \sum_{\mathbf{x}} \prod_{i=1}^{N-1} \mathbb{P}_{y^*}\big(S_\sigma = x_{i+1}\big) \\
&\geq \frac{c_H^{N-1}}{2} \prod_{i=1}^{N-1} \mathbb{P}_{y^*}\big(S_\sigma \in \partial Q_{2R}\big) \geq e^{-c_1 N}.
\end{aligned} \tag{5.6}$$

Finally, define

$$\mathcal{A}_3 := \{\text{The walk makes at least } N \text{ excursions from } \partial Q_{2R} \text{ to } \partial Q_{4R} \text{ before time } T\}.$$

Using that on the event $\mathcal{A}_1 \cap \mathcal{A}_3^c$, the walk spends a time at least T in Q_{8R}, we deduce that for some constant $c > 0$,

$$\mathbb{P}(\mathcal{A}_1 \cap \mathcal{A}_3^c) \leq \exp(-c\frac{T}{R^2}). \tag{5.7}$$

Then the proposition readily follows from (5.6) and (5.7), once we choose $C_1 c > 2C_2 c_1$ and use

$$\mathbb{P}\big(\mathcal{A}_1 \cap \{Z \geq \frac{M}{2}\} \cap \mathcal{A}_3\big) \geq \mathbb{P}\big(\mathcal{A}_1 \cap \{Z \geq \frac{M}{2}\}\big) - \mathbb{P}\big(\mathcal{A}_1 \cap \mathcal{A}_3^c\big).$$

It remains now to prove (5.5). 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 in (5.2) large enough, one has for any $i \leq M$,

$$\mathbb{P}_{\mathbf{x}}\big(|\mathcal{R}^N \cap Q_r(v_i)| > \rho r^d\big) \geq \frac{3}{4}. \tag{5.8}$$

Indeed, (5.8) shows that $\mathbb{E}[Z] \geq \frac{3}{4}M$, and using also that Z is bounded by M, it implies that $\mathbb{P}(Z \leq M/2) \leq 1/2$, as desired.

Thus, we are led to prove (5.8) 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 $\mathcal{V} \subseteq Q_r(v_i)$ 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 $\partial Q_{2r}(v_i)$, uniformly in its starting point, we have for some constant $c_0 > 0$, almost surely

$$\mathbb{E}_{\mathbf{x}}\big[|\mathcal{V}| \mid \mathcal{G}\big] \geq \Big(1 - (1 - \frac{c_0}{r^{d-2}})^{\mathcal{N}_r}\Big) \cdot |Q_r| \geq \Big(1 - \exp\big(-c_0 \frac{\mathcal{N}_r}{r^{d-2}}\big)\Big) \cdot |Q_r|, \tag{5.9}$$

using that $1 - u \leq e^{-u}$, for all $u \geq 0$. Since any excursion has a probability of order at least $(r/R)^{d-2}$ to hit $\partial Q_{2r}(v_i)$, we observe that for any fixed $K \geq 1$, it is possible to choose $C_2 = C_2(K)$ in (5.2) large enough so that

$$\mathbb{P}_{\mathbf{x}}\big(\mathcal{N}_r \geq K\rho r^{d-2}\big) \geq \sqrt{7/8}. \tag{5.10}$$

We fix K later, and we treat distinctly high and low densities.

High density. If ρ is such that $1 - \exp(-c_0\, K\rho) \geq \sqrt{7/8}$, then (5.9) and (5.10) imply that

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

Using that $\mathcal{V} \subseteq Q_r$, and as a consequence that $|\mathcal{V}|$ is bounded by $|Q_r|$, we obtain (5.8), writing

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

Low density. If ρ is such that $1 - \exp(-c_0\, K\rho) < \sqrt{7/8}$, then it means that ρK is bounded by some universal constant (only depending on c_0), and thus that for another univeral constant $c' > 0$, one has $1 - \exp(-c_0\, K\rho) \geq c'\rho K$. Then by (5.9), on the event $\{\mathcal{N}_r \geq K\rho r^{d-2}\}$,

$$\mathbb{E}_{\mathbf{x}}\big[|\mathcal{V}| \mid \mathcal{G}\big] \geq c'\rho K \cdot |Q_r|. \tag{5.11}$$

Our strategy now is to use a second (conditional) moment method, and show that the conditional variance of $\mathcal{V}$ is small. We denote with $\mathcal{V}_1$ the set of pairs of vertices $(y,z) \in \mathcal{V} \times \mathcal{V}$, for which there exists an excursion going through both y and z, and let $\mathcal{V}_2$ be the complement of $\mathcal{V}_1$ in $\mathcal{V} \times \mathcal{V}$ and note that $|\mathcal{V}|^2 = |\mathcal{V}_1| + |\mathcal{V}_2|$. 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 $c > 0$,

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

Since we have assumed that $\rho r^{d-2} \geq 1$ it follows using (5.11) that on the event $\{\mathcal{N}_r \geq K\rho r^{d-2}\}$ we have

$$\mathbb{E}_{\mathbf{x}}\big[|\mathcal{V}_1| \mid \mathcal{G}\big] \leq \frac{c}{c'K} \cdot \mathbb{E}_{\mathbf{x}}\big[|\mathcal{V}| \mid \mathcal{G}\big]^2. \tag{5.12}$$

We fix now K such that $K > 64c/c'$ (and fix accordingly C_2 and then C_1 as explained above). Then it remains to bound the conditional mean of $|\mathcal{V}_2|$ knowing $\mathcal{G}$. First we fix a constant $K' > 0$, such that

$$\mathbb{P}_{\mathbf{x}}(\mathcal{A}_4) \geq 7/8, \quad \text{for} \quad \mathcal{A}_4 := \Big\{K\rho r^{d-2} \leq \mathcal{N}_r \leq K'\rho r^{d-2}\Big\}.$$

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 $\mathcal{A}_4$, and for r large enough,

$$\begin{aligned}
\mathbb{P}_{\mathbf{x}}\left(z \in \bigcup_{k \notin \mathcal{I}_y} \mathcal{E}_k \;\middle|\; \mathcal{G}, \mathcal{I}_y\right) &= 1 - \prod_{k \notin \mathcal{I}_y} \Big(1 - \mathbb{P}_{\mathbf{x}}(z \in \mathcal{E}_k \mid \mathcal{G}, \mathcal{I}_y)\Big) \\
&\leq 1 - \prod_{k \notin \mathcal{I}_y} \Big(1 - \mathbb{P}_{\mathbf{x}}(z \in \mathcal{E}_k \mid \mathcal{G}) - \mathcal{O}(r^{2(2-d)})\Big) \\
&\leq 1 - \prod_{k \leq \mathcal{N}_r} \Big(1 - \mathbb{P}_{\mathbf{x}}(z \in \mathcal{E}_k \mid \mathcal{G})\Big) + \mathcal{O}\left(\frac{\mathcal{N}_r}{r^{2(d-2)}}\right) \\
&= \mathbb{P}_{\mathbf{x}}(z \in \mathcal{V} \mid \mathcal{G}) + \mathcal{O}\left(\frac{\mathcal{N}_r}{r^{2(d-2)}}\right),
\end{aligned} \tag{5.13}$$

where at the penultimate line we use that on $\mathcal{A}_4$, 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 $\mathcal{A}_4$, we get from (5.13),

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

Summing over $y,z\in Q_r$, we deduce from (5.11), that on the event $\mathcal{A}_4$,

$$\mathbb{E}_{\mathbf{x}}\big[|\mathcal{V}_2| \bigm| \mathcal{G}\big] \le \mathbb{E}_{\mathbf{x}}\big[|\mathcal{V}| \bigm| \mathcal{G}\big]^2(1+\mathcal{O}(r^{2-d})).$$

Combining this with (5.12), we get for r large enough,

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

Together with (5.11), it follows that for r large enough, on the event $\mathcal{A}_4$,

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

Finally, using that $\mathbb{P}_{\mathbf{x}}(\mathcal{A}_4)\ge 7/8$, we obtain the desired bound for r large enough,

$$\mathbb{P}_{\mathbf{x}}(|\mathcal{V}|\le\rho|Q_r|) \le \mathbb{P}_{\mathbf{x}}(\mathcal{A}_4^c) + \mathbb{P}_{\mathbf{x}}(|\mathcal{V}|\le\rho|Q_r|,\,\mathcal{A}_4) \le 1/4,$$

On the other hand, for small values of r, the result is immediate. This concludes the proof of (5.5) and the proposition. □

Acknowledgements: We warmly thank the two referees and the editors, for their careful reading and numerous comments/corrections which greatly improved the readability of the manuscript. We also acknowledge support from public grants overseen by the French National Research Agency, ANR SWiWS (ANR-17-CE40-0032-02) and ANR MALIN (ANR-16-CE93-0003).

References

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

[AS21a] A. Asselah; Br. Schapira. The two regimes of moderate deviations for the range of a transient walk. Probab. Theory Related Fields 180, (2021), 439–465.

[AS21b] A. Asselah; Br. Schapira. Deviations for the capacity of the range of a random walk. Electron. J. Probab. 25: 1–28 (2020).

[AS21c] A. Asselah; Br. Schapira. Large Deviations for Intersections of Random Walks, to appear in Comm. Pure Appl. Math.

[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), no. 2, 355–406.

[Chen09] X. Chen. Random Walk Intersections, Large Deviations and Related Topics. Mathematical Surveys and Monographs, Vol 157, 2009, AMS.

[DV75] M. D. Donsker; S. R. S. Varadhan. Asymptotic evaluation of certain Markov process expectations for large time. I.-II. Comm. Pure Appl. Math. 28 (1975), 1–47; ibid. 28 (1975), 279–301.

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

[IK60] K. Itô; H. P. Jr. McKean. Potentials and the random walk. Illinois J. Math. 4 (1960), 119–132.

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

[Law91] G. F. Lawler. Intersections of random walks. Reprint of the 1996 edition. Modern Birkhauser Classics. Birkhauser/Springer, New York, 2013. iv+223

[LL10] G. F. Lawler; V. Limic. Random walk: a modern introduction. Cambridge Studies in Advanced Mathematics, 123. Cambridge University Press, Cambridge, 2010.

[LSz15] X. Li; A.-S. Sznitman. Large deviations for occupation time profiles of random interlacements. Probab. Theory Related Fields 161 (2015), 309–350.

[LT21] G. Liddle; A. Turner. Scaling limits and fluctuations for random growth under capacity rescaling, Annales de l'Institut Henri Poincaré (B) 57, (2021), 980–1015.

[NSz20] M. Nitzschner; A.-S. Sznitman. Solidification of porous interfaces and disconnection. J. Eur. Math. Soc. (JEMS) 22, (2020), 2629–2672.

[Sz10] A.-S. Sznitman. Vacant set of random interlacements and percolation, Ann. of Math. (2) 171, (2010), 2039–2087.

[Sz12] A.-S. Sznitman. Topics in occupation times and Gaussian free fields. Zurich Lectures in Advanced Mathematics. European Mathematical Society (EMS), Zürich, 2012. viii+114 pp.

[Sz15] A.-S. Sznitman. Disconnection and level-set percolation for the Gaussian free field. J. Math. Soc. Japan 67 (2015), 1801–1843.

[Sz17] A.-S. Sznitman. Disconnection, random walks, and random interlacements. Probab. Theory Related Fields 167 (2017), 1–44.

[Sz20] A.-S. Sznitman. Excess deviations for points disconnected by random interlacements, Probability and Mathematical Physics, Vol. 2 (2021), No. 3, 563–611.

[Sz21] A.-S. Sznitman. On the cost of the bubble set for random interlacements, (2021), arXiv:2105.12110