Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2016, Vol. 52, No. 1, 58–83
DOI: 10.1214/14-AIHP629

ANNALES DE L'INSTITUT HENRI POINCARÉ
PROBABILITÉS ET STATISTIQUES
www.imstat.org/aihp

Fluctuations for internal DLA on the comb

Amine Asselah and Houda Rahmani

LAMA, Université Paris-Est Créteil, 61 avenue du général de Gaulle, 94010 Créteil cedex, France.
E-mail: amine.asselah@u-pec.fr; houda.rahmani@u-pec.fr

Received 27 January 2014; accepted 21 June 2014

Abstract. We study internal diffusion limited aggregation (DLA) on the two dimensional comb lattice. The comb lattice is a spanning tree of the Euclidean lattice, and internal DLA is a random growth model, where simple random walks, starting one at a time at the origin of the comb, stop when reaching the first unoccupied site. An asymptotic shape is suggested by a lower bound of Huss and Sava (*Electron. J. Probab.* **17** (2012) 30). We bound the fluctuations with respect to this shape.

Résumé. Nous étudions un modèle d'agrégation limitée par diffusion interne (DLA), sur le peigne bidimensionnel. Le peigne est un arbre couvrant du réseau cubique, et DLA interne est un modèle de croissance aléatoire : des marches simples, lancées une après l'autre à l'origine du peigne, s'arrêtent lorsqu'elles atteignent le premier sommet inexploré. Une forme asymptotique est suggérée par une borne inférieure de Huss et Sava (*Electron. J. Probab.* **17** (2012) 30). Nous étudions les fluctuations par rapport à cette forme asymptotique.

MSC: 60K35; 82B24; 60J45

Keywords: Diffusion limited aggregation; Cluster growth; Random walk; Shape theorem; Lattice comb

1. Introduction

The comb lattice, denoted $\mathcal{C}$, is an inhomogeneous spanning tree of $\mathbb{Z}^2$. Sites $z=(x,y)$ and $z'=(x',y')$ share an edge if either $x=x'$ and $|y-y'|=1$, or if $y=y'=0$ and $|x-x'|=1$; in this case, we say that z and z' are neighbors.

Internal DLA on $\mathcal{C}$ is a Markov chain on the finite subsets of the comb, with initial condition the empty set, and growing as follows. Assume we have obtained a cluster A. To build the cluster with one more site, launch a simple random walk on the comb starting at the origin. Stop the random walk when it exits A, say on site z. The new cluster is the union of A and z, the first visited site outside A. The random walk with the aggregation rule is called an explorer. We say that the explorer settles on z.

Internal DLA has been first studied on the cubic lattice $\mathbb{Z}^d$ in d dimension. Diaconis and Fulton [4] introduced it, as well as many variants, with a special emphasize on the invariance of the cluster with respect to the order in which the explorers are sent: the so-called *Abelian property*. Lawler, Bramson and Griffeath [12] established that the normalized asymptotic shape is the Euclidean sphere in dimension two or more.

Blachère and Brofferio [3] obtained a limiting shape when the graph is a finitely generated group with exponential growth. Huss [9] studied internal DLA for a large class of random walks on such graphs. Recently, internal DLA has been considered on the infinite percolation cluster, and the asymptotic shape is a Euclidean ball intersected with the infinite cluster: Shellef [13] obtained a bound on the inner fluctuations, and Duminil-Copin, Lucas, Yadin and Yehudayoff [5] obtained the corresponding bound on the outer fluctuations using the inner bound.

It is interesting to study internal DLA on the comb, since it is inhomogeneous, and distinct from a cubic lattice: a simple random walk is recurrent, however two random walks meet on the average a finite number of times [8]. Also,

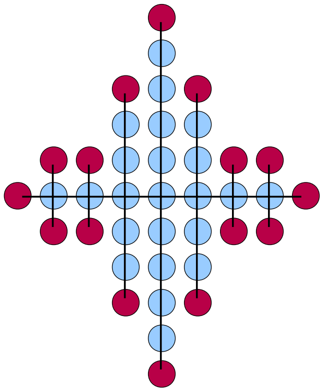

Fig. 1. For $\rho = 3.5$, $\mathcal{D}(\rho)$ and its boundary as balls.

the x and y axes play a different role: in a time n, a simple random walk on the comb, reaches a y-axis displacement of order $n^{1/2}$, and an x-axis displacement of order $n^{1/4}$. To discuss results on the comb, let us introduce some notations. For any real ρ, we define

$$D(\rho) = \left\{(x,y) \in \mathbb{R}^2 \colon |x| < \rho, |y| < \frac{1}{3}\big(\rho - |x|\big)^2\right\} \quad \text{and} \quad \mathcal{D}(\rho) = D(\rho) \cap \mathbb{Z}^2. \tag{1.1}$$

See Figure 1. For an integer n, the number of sites in $\mathcal{D}(n)$ is denoted $d(n)$, and the internal DLA cluster obtained by sending $d(n)$ explorers is denoted $A(n)$.

Recently, Huss and Sava [10] have characterized the limiting shape for a related model, the *divisible sandpile* introduced in [11], and shown a lower bound for the shape of the internal DLA cluster on the comb.

Theorem 1.1 (Theorem 4.2 of [10]). *For any $\varepsilon > 0$, with probability 1, we have for n large enough $\mathcal{D}(n - \varepsilon n) \subset A(n)$.*

Our main result here is the following improvement.

Theorem 1.2. *There is a positive constant a such that with probability 1, and n large enough*

$$\mathcal{D}\big(n - a\sqrt{\log(n)}\big) \subset A(n) \subset \mathcal{D}\big(n + a\sqrt{\log(n)}\big). \tag{1.2}$$

Remark 1.3. *This result does not mean that fluctuations are sub-logarithmic, but rather suggests that they are Gaussian. Indeed, a site $z = (x, y)$ on the boundary of $\mathcal{D}(n - a\sqrt{\log(n)})$ is at a distance of order $\frac{2}{3}a\sqrt{\log(n)}(n - |x|)$ from the boundary of $\mathcal{D}(n)$, whereas the tooth's length is of order $\frac{1}{3}(n - |x|)^2$. Thus, the fluctuations are similar to what would be observed in a collection of n independent segments whose lengths decrease quadratically. Diaconis and Fulton in [4] used an urn-representation to obtain a central limit theorem on $\mathbb{Z}$ for the right-end of the DLA cluster. The tip of $\mathcal{D}(n)$ has sublogarithmic fluctuations as in $\mathbb{Z}^d$, for $d \geq 3$.*

Theorem 1.2 follows a classical approach by Lawler, Bramson and Griffeath [12], and requires a study of the restricted Green's function on $\mathcal{D}$. It relies also on a deep connection with another cluster growth model, *the divisible sandpile*, which was discovered by Levine and Peres [11]. Finally, the limiting shape of the divisible sandpile cluster was shown to be $\mathcal{D}(n)$ on the comb in [10].

It is interesting to note that exit probabilities from the DLA cluster are not uniform, as it is the case for the cubic lattice, or for discrete groups having exponential growth [3], or for the layered square lattice [7]. To better appreciate the following estimate, note that for any $\rho > 0$, the size of the boundary of $\mathcal{D}(\rho)$ is of order ρ (see Figure 1).

Proposition 1.4. *For any real ρ, and $z = (x, y)$ in the boundary of $\mathcal{D}(\rho)$ with $x < \rho$, we have*

$$\frac{1}{2}\frac{\rho - |x|}{\rho^2 + 1/3} \leq P_0\big(\text{the walk exits } \mathcal{D}(\rho) \text{ at } z\big) \leq \frac{\rho + 1 - |x|}{\rho^2 + 1/3}. \tag{1.3}$$

However, one important property, which holds also for the cubic lattice, is the following *uniform hitting property*.

Proposition 1.5. *For each $\rho > 0$, there is a stopping time τ^*_ρ and two positive constants $\underline{\kappa}, \bar{\kappa}$, independent on ρ such that*

$$\forall z \in \mathcal{D}(\rho), \quad \frac{\underline{\kappa}}{|\mathcal{D}(\rho)|} \le P_0\big(S\big(\tau^*_\rho\big) = z\big) \le \frac{\bar{\kappa}}{|\mathcal{D}(\rho)|}. \tag{1.4}$$

Proposition 1.5 is crucial in proving some large deviation estimates about the cluster, which in turn, are key in controlling the outer error.

We now turn to some large deviations estimates which shed more light on the covering mechanism. Note that a general feature which emerges from all studies is that during the covering process, explorers do not leave *holes deep inside the bulk*. Our first lemma deals with the probability an explorer reaches site $(R,0)$, on the x-axis, without leaving the explored region $V \ni (R,0)$. The result, and its proof, are interesting on their own, and follow closely Lemma 1.6 of [2]. For a subset Λ in $\mathbb{Z}^2$, let $H(\Lambda)$ be the first time the random walk hits Λ.

Lemma 1.6. *Let R be a positive integer, and V a subset of $\mathbb{Z}^2$ containing $(0,0)$ and $(R,0)$. There are positive constants a_3 and κ_3, independent of R and V, such that*

$$P_0\big(H(R,0) < H\big(V^c\big)\big) \le \exp\left(a_3 - \kappa_3\sqrt{\frac{R^3}{|V|}}\right). \tag{1.5}$$

In other words, the random walk cannot reach $(R,0)$ inside V, unless V contains of the order of R^3 sites. As a corollary of Lemma 1.6, we have the following large deviation upper bound. Recall that $A(n)$ is the cluster obtained when sending $|\mathcal{D}(n)|$ explorers at the origin, and that the volume of $\mathcal{D}(n)$ is of order n^3. In other words, Lemma 1.6 quantifies the probability of making thin tentacles along the x-axis.

Corollary 1.7. *There is $\beta, \kappa_2 > 0$, such that if R and n are integers with $d(n) < \beta R^3$, then*

$$P\big((R,0) \in A(n)\big) \le \exp\big(-\kappa_2 R^2\big). \tag{1.6}$$

We wish now to explain how to find the asymptotic shape of the internal DLA cluster built with simple random walks all started at a distinguished vertex of a graph, say 0. A fundamental observation of Lawler, Bramson and Griffeath [12] yields a recipe: find an increasing family of subsets of the graph, say $\{\mathcal{D}(\rho), \rho \in \mathbb{R}\}$ containing 0, such that the discrete *mean value property holds* for harmonic functions. More precisely, h is harmonic if for any vertex x of the graph

$$\sum_{y \text{ neighbor of } x} \big(h(y) - h(x)\big) = 0. \tag{1.7}$$

Define now, for any subset Λ, and any function h on Λ, the centered average (recall that walks start from 0)

$$\mathrm{MV}(h, \Lambda) = \sum_{z \in \Lambda} \big(h(z) - h(0)\big). \tag{1.8}$$

Finally, we say that *the mean value property holds on Λ* when for any h harmonic, $\mathrm{MV}(h, \Lambda)$ is of smaller order than the volume of Λ times $h(0)$. Thus, we look for subsets $\{\mathcal{D}(\rho), \rho \in \mathbb{R}\}$, such that for any ρ, we can show that *the mean value property holds on $\mathcal{D}(\rho)$*.

The observation of [12] behind the connection between the DLA cluster and the mean value property is as follows. Each site of the DLA cluster is the settlement of exactly one explorer. Thus, paint green the explorers' trajectories until settlement, and add red independent random walks trajectories, starting one on each site of the cluster. The color-free trajectories, obtained by concatenating the end-point of a green strand with the red strand starting there, are independent random walks starting from 0. In short, green explorers glued to red walkers make independent random walks all started at 0. Now, if $\mathcal{D}(n)$ is the shape around which the DLA cluster fluctuates, then the probability a green explorer exits $\mathcal{D}(n)$ from site z in its boundary is small if *few deep holes are left* as $\mathcal{D}(n)$ gets covered. This probability is bounded by the difference between the expected number of random walks starting at 0 and exiting $\mathcal{D}(n)$ from z, and the expected number of red walkers exiting $\mathcal{D}(n)$ from z, with one starting on each site of $\mathcal{D}(n)$. This difference is $\mathrm{MV}(h_z, \mathcal{D}(n))$, where $h_z(y)$ is the probability of exiting $\mathcal{D}(n)$ from z when the initial position of the walk is y, and $y \mapsto h_z(y)$ is a harmonic function. The smaller is $\mathrm{MV}(h_z, \mathcal{D}(n))$, the better is the control of the fluctuation of the cluster (see (2.18)).

The discrete mean value property holds for spheres on $\mathbb{Z}^d$, as shown by Levine and Peres (Theorem 1.3 of [11] and Lemma 6 of [6]) who used the divisible sandpile for that purpose (this property can also be derived from unpublished estimates of Blachère, see the Appendix of [1]). On the comb, a mean value property for the domain $\mathcal{D}(n)$ is essentially contained in the study of Huss and Sava [10]. This is the starting point of our study.

Let us mention that it is delicate to estimate the shape of the divisible sandpile cluster. For instance on the comb of $\mathbb{Z}^3$ (where teeth stand on the two dimensional plane), we did not succeed in identifying the sandpile cluster.

The rest of the paper is organized as follows. We start with estimates on restricted Green's functions in Section 2. Estimates on the Green's function, Propositions 2.2 and 2.3 are the key technical novelties here. In Section 2.3, we recall the classical approach developed in [12]. The mean value property is proved in Section 2.4. The large deviations estimate Lemma 1.6 and Corollary 1.7 are proved in Section 3. Finally, inner and outer errors are respectively estimated in Sections 4 and 5. Finally, in the Appendix, we prove technical properties of the Green's function, most notably Proposition 2.2.

2. Preliminaries

2.1. *Notation*

The comb, denoted $\mathcal{C}$, is a tree rooted at the origin. Any nonzero site has a unique parent: that is its neighbor which is closer to the origin. It is convenient to call $\mathcal{A}(z)$ the parent of z.

The discrete boundary of $\mathcal{D}(\rho)$, denoted $\partial\mathcal{D}(\rho)$, consists of the sites of $\mathbb{Z}^2$ not in $\mathcal{D}(\rho)$, but at a distance 1 from $\mathcal{D}(\rho)$. The internal boundary of $\mathcal{D}(\rho)$, denoted $\partial_I\mathcal{D}(\rho)$, consists of sites of $\mathcal{D}(\rho)$ at a distance 1 from $\partial\mathcal{D}(\rho)$. The continuous boundary of $D(\rho)$ is denoted $\partial D(\rho)$, and is the curve $\{(x,y) \in \mathbb{R}^2\colon\ |x| \le \rho, |y| = \frac{1}{3}(\rho - |x|)^2\}$. The Euclidean ball of center 0, and radius R is denoted $B(R)$, and

$$\mathbb{B}(R) = B(R) \cap \mathbb{Z}^2, \quad \text{with } B(R) = \left\{(x,y) \in \mathbb{R}^2\colon\ x^2 + y^2 < R\right\}.$$

For a subset Λ of $\mathbb{Z}^2$, we denote by $H(\Lambda)$ the time at which a simple random walk on the comb first hits Λ, and we call Λ^+ the intersection of Λ with the positive quadrant.

2.2. *On Green's functions*

Henceforth, we consider a simple random walk on the comb $\mathcal{C}$. We establish many results on harmonic functions on the domain $\mathcal{D}(\rho)$. To ease the reading, their proofs are postponed to the Appendix.

For a subset Λ of $\mathbb{Z}^2$, let G_Λ be Green's function restricted to Λ. In other words, for $x, y \in \Lambda$, $G_\Lambda(x; y)$ is the expected number of visits to y before escaping Λ, when starting on x:

$$G_\Lambda(x; y) = E_x\left[\sum_{n=0}^{\infty} \mathbb{1}_{n < H(\Lambda^c)} \mathbb{1}_{\{S(n) = y\}}\right]. \tag{2.1}$$

We first approximate Green's function restricted to $\mathcal{D}(\rho)$. For a real x, $[x]$ denotes its integer part.

Lemma 2.1. *Let ρ be any positive real, and $z = (x, y) \in \mathcal{D}(\rho)$. Define h as*

$$h(z) = \frac{2(\rho - |x|)}{\rho^2 + 1/3}\left(\frac{(\rho - |x|)^2}{3} - |y|\right), \tag{2.2}$$

and h^+ as

$$h^+(z) = \frac{2([\rho]+1-|x|)}{([\rho]+1)^2+1/3+1}\left(\frac{([\rho]+1-|x|)^2}{3}+1-|y|\right). \tag{2.3}$$

Then

$$h(z) \le G_{\mathcal{D}(\rho)}(0;z) \le h^+(z). \tag{2.4}$$

Moreover, assume that $z' \in \partial\mathcal{D}(\rho)$, $z = \mathcal{A}(z')$. If $y \neq 0$, then $\rho - |x| > 1$ and we have

$$\frac{(\rho-|X_z|)}{\rho^2+1/3} \le G_{\mathcal{D}(\rho)}(0;z) \le \frac{2([\rho]+1-|X_z|)}{\rho^2+1/3}, \tag{2.5}$$

whereas if $y = 0$, then

$$\frac{1}{4}\frac{2(\rho+1-|X_z|)}{\rho^2+1/3} \le G_{\mathcal{D}(\rho)}(0;z) \le \frac{2([\rho]+1-|X_z|)}{\rho^2+1/3}. \tag{2.6}$$

For simplicity, we denote $\mathcal{E}(\rho) = H(\mathcal{D}^c(\rho))$ the exit time from $\mathcal{D}(\rho)$. Our second result is our main technical contribution in estimating hitting probabilities. This in turn allows us to establish accurate Green's function estimates.

Proposition 2.2. *For any positive real ρ, and any integer x, with $|x| \le \rho$, there is a constant $\kappa_a > 0$ independent of ρ*

$$P_0\big(H(x,0) < \mathcal{E}(\rho)\big) \ge \kappa_a\left(\frac{\rho-|x|}{\rho}\right)^2. \tag{2.7}$$

We now estimate the probability of exiting $\mathcal{D}(\rho)$ from $z \in \partial\mathcal{D}(\rho)$. This is equivalent to estimating Green's function $y \mapsto G_{\mathcal{D}(\rho)}(y, \mathcal{A}(z))$, since by a last passage decomposition

$$P_y\big(S\big(\mathcal{E}(\rho)\big) = z\big) = \frac{1}{\deg(\mathcal{A}(z))}G_{\mathcal{D}(\rho)}\big(y, \mathcal{A}(z)\big). \tag{2.8}$$

It is convenient for $w \in \mathbb{Z}^2$, to denote its two coordinates as X_w and Y_w. Also, let $L_\rho(w)$ denote the smallest integer larger or equal than $\frac{1}{3}(\rho - X_w)^2$, and $\mathrm{sg}(x)$ is the sign of x.

Proposition 2.3. *Assume $z \in \partial\mathcal{D}(\rho)$, and $w \in \mathcal{D}(\rho)$.*

(i) *When $0 \le X_w < X_z$ or $X_w = X_z$ but $\mathrm{sg}(Y_w) \neq \mathrm{sg}(Y_z)$, we have*

$$P_w\big(S\big(\mathcal{E}(\rho)\big) = z\big) \le \frac{4}{\kappa_a}\frac{L_\rho(w)-Y_w}{L_\rho(w)} \times \frac{(\rho+1-X_z)}{(\rho-X_w)^2}. \tag{2.9}$$

(ii) *When $0 \le X_z < X_w$, we have*

$$P_w\big(S\big(\mathcal{E}(\rho)\big) = z\big) \le \frac{4}{\kappa_a}\frac{L_\rho(w)-Y_w}{L_\rho(z)} \times \frac{(\rho-X_w)}{(\rho-X_z)^2}. \tag{2.10}$$

(iii) *When $X_w < 0 \le X_z$, there is a constant $\kappa > 0$ such that*

$$P_w\big(S\big(\mathcal{E}(\rho)\big) = z\big) \le \kappa\frac{L_\rho(w)-Y_w}{L_\rho(w)} \times \frac{(\rho-|X_w|)^3(\rho+1-X_z)}{\rho^5}. \tag{2.11}$$

(iv) *When $X_w = X_z$ and $\mathrm{sg}(Y_w) = \mathrm{sg}(Y_z)$, we have*

$$\frac{1}{2}\frac{Y_w}{L_\rho(w)} \le P_w\big(S\big(\mathcal{E}(\rho)\big) = z\big). \tag{2.12}$$

Finally, we have the following corollary of Proposition 2.3.

Corollary 2.4. *Assume that* $z \in \partial\mathcal{D}(\rho)$. *There are constants* $\underline{\kappa}, \bar{\kappa} > 0$ (*independent of* n *and* z) *such that*

$$\underline{\kappa}\big(\rho + 1 - |X_z|\big)^2 \le \sum_{w \in \mathcal{D}(\rho)} P_w\big(S\big(\mathcal{E}(\rho)\big) = z\big)^2 \le \bar{\kappa}\big(\rho + 1 - |X_z|\big)^2. \tag{2.13}$$

2.3. *On a classical approach*

Denote by $W(\eta, z)$ (resp. $M(\eta, z)$) the number of explorers (resp. random walkers) starting on configuration $\eta \in \mathbb{N}^{\mathcal{C}}$ which hit z. Two special initial configurations play a key role in internal DLA: we call $d(n)\mathbb{1}_0$ the configuration with $d(n)$ explorers at 0; when Λ is a subset of $\mathbb{Z}^2$, we still use Λ, rather than $\mathbb{1}_\Lambda$, to denote the configuration with one explorer on each site of Λ. The main observation of [12] yields the following inequality in law

$$W\big(d(n)\mathbb{1}_0, z\big) + M\big(A(n), z\big) \ge M\big(d(n)\mathbb{1}_0, z\big). \tag{2.14}$$

An important feature of (2.14) is that $W(d(n)\mathbb{1}_0, z)$ is expressed as a difference of two sums of Bernoulli variables. However, $A(n)$ is unknown, and as such (2.14) is of little use. Since we want to show that $A(n)$ is close to a deterministic region $\mathcal{D}(n)$, we first look for a region $\mathsf{I}(n) \subset \mathcal{D}(n)$ which is very likely covered by the cluster $A(n)$, when n is large. We even require that $\mathsf{I}(n)$ be covered by explorers not exiting $\mathsf{I}(n)$, and we call $A_{\mathsf{I}(n)}(n)$ the cluster made by these explorers. The possibility to discard trajectories exiting $\mathsf{I}(n)$ is made possible by a key observation of Diaconis and Fulton [4] named *the Abelian property*: the law of the cluster is independent on the order in which explorers are launched; this allows to obtain a smaller cluster if we discard some trajectories. The key point now is that by definition

$$A_{\mathsf{I}(n)}(n) \subset \mathsf{I}(n). \tag{2.15}$$

Now, for $z \in \mathsf{I}(n)$, $W_{\mathsf{I}(n)}(\eta, z)$ (resp. $M_{\mathsf{I}(n)}(\eta, z)$) denotes the number of explorers (resp. walkers starting on η) which hit z before exiting $\mathsf{I}(n)$. When $z \in \partial\mathsf{I}(n)$, $W_{\mathsf{I}(n)}(\eta, z)$ (resp. $M_{\mathsf{I}(n)}(\eta, z)$) still denotes the number of explorers (resp. walkers starting on η) which exit $\mathsf{I}(n)$ from z. The same idea leading to (2.14) yields for $z \in \mathsf{I}(n)$

$$W_{\mathsf{I}(n)}\big(d(n)\mathbb{1}_0, z\big) + M_{\mathsf{I}(n)}(A_{\mathsf{I}(n)}(n), z) \ge M_{\mathsf{I}(n)}\big(d(n)\mathbb{1}_0, z\big), \tag{2.16}$$

and this inequality becomes an equality when $z \in \partial\mathsf{I}(n)$. Using (2.15), we obtain for $z \in \mathsf{I}(n)$

$$W_{\mathsf{I}(n)}\big(d(n)\mathbb{1}_0, z\big) + M_{\mathsf{I}(n)}\big(A_{\mathsf{I}(n)}(n), z\big) \ge M_{\mathsf{I}(n)}(n\mathbb{1}_0, z). \tag{2.17}$$

The harmonic function $y \mapsto P_y(H(z) \le H(\partial\mathsf{I}(n)))$ is denoted $h_z(y)$. Taking the expectation of both sides of (2.17) allows a lower bound on the expectation of $W(d(n)\mathbb{1}_0, z)$

$$\begin{aligned} E\big[W\big(d(n)\mathbb{1}_0, z\big)\big] \ge E\big[W_{\mathsf{I}(n)}\big(d(n)\mathbb{1}_0, z\big)\big] \ge \mu(z) &:= E\big[M_{\mathsf{I}(n)}(n\mathbb{1}_0, z)\big] - E\big[M_{\mathsf{I}(n)}\big(\mathsf{I}(n), z\big)\big] \\ &= \big(d(n) - \big|\mathsf{I}(n)\big|\big) \times h_z(0) + \mathrm{MV}\big(h_z, \mathsf{I}(n)\big). \end{aligned} \tag{2.18}$$

If (2.17) were an equality, and using that $W_{\mathsf{I}(n)}(d(n)\mathbb{1}_0, z)$ and $M_{\mathsf{I}(n)}(\mathsf{I}(n), z)$ are independent, we would have the following bound for the variance of $W_{\mathsf{I}(n)}(d(n)\mathbb{1}_0, z)$ (we use that the M's are sums of Bernoulli)

$$\mathrm{var}\big(W_{\mathsf{I}(n)}\big(d(n)\mathbb{1}_0, z\big)\big) = \mathrm{var}\big(M_{\mathsf{I}(n)}(n\mathbb{1}_0, z)\big) - \mathrm{var}\big(M_{\mathsf{I}(n)}\big(\mathsf{I}(n), z\big)\big) \le \mu(z) + \sum_{y \in \mathsf{I}(n)} h_z^2(y). \tag{2.19}$$

Even though (2.19) is wrong, and that no bound on the variance is known, [1] shows that for a positive constant κ

$$P\big(W_{\mathsf{I}(n)}\big(d(n)\mathbb{1}_0, z\big) = 0\big) \le \exp\bigg(-\kappa \frac{\mu^2(z)}{\mu(z) + \sum_{y \in \mathsf{I}(n)} h_z^2(y)}\bigg). \tag{2.20}$$

Then, due to the tree structure of the comb $I(n)\not\subset A(n)$ implies that for some $z\in\partial I(n)$, $W_{I(n)}(z)=0$. We look for a subset $\mathsf{I}(n)$ in $\mathcal{D}(n)$ such that the following series on the right hand side converges.

$$\sum_{n\in\mathbb{N}} P\big(\mathsf{I}(n)\not\subset A(n)\big)\le\sum_{n\in\mathbb{N}}\sum_{z\in\partial\mathsf{I}(n)} P\big(W_{\mathsf{I}(n)}\big(d(n)\mathbb{1}_0,z\big)=0\big)$$
$$\le\sum_{n\in\mathbb{N}}\sum_{z\in\partial\mathsf{I}(n)}\exp\left(-\kappa\frac{\mu^2(z)}{\mu(z)+\sum_{y\in\mathsf{I}(n)}h_z^2(y)}\right).\tag{2.21}$$

Using Borel–Cantelli, (2.21) implies that almost surely, for n large enough, $\mathsf{I}(n)\subset A(n)$.

This approach can be implemented if we can estimate $\mu(z)$, and the sum of $y\mapsto h_z^2(y)$ over $\mathsf{I}(n)$. Note that $\mu(z)$ should be of order $(|\mathcal{D}(n)|-|\mathsf{I}(n)|)h_z(0)$ provided we can show that

$$\forall z\in\mathsf{I}(n),\quad \mathrm{MV}\big(h_z,\mathsf{I}(n)\big)\ll\big(\big|\mathcal{D}(n)\big|-\big|\mathsf{I}(n)\big|\big)h_z(0).\tag{2.22}$$

The divisible sandpile

Levine and Peres [11] have introduced a model, the divisible sandpile, whose cluster is a good candidate for $\mathcal{D}(n)$. In this model, we start with a mass n at the origin of our graph, and topple sand along some sequence of sites. We topple the sand at a site if its mass is above 1, and we transfer the total mass minus 1 equally to each nearest neighbor. The toppling sequence is arbitrary provided it covers each site of the graph infinitely often. We call $z\mapsto w_n(z)$ the final sand distribution, and we call $z\mapsto u_n(z)$ the odometer function: that is the amount of sand emitted from each site. The *sandpile cluster* is $\mathcal{S}_n=\{z\colon u_n(z)>0\}$. The key observation is that for any harmonic function on $h:\mathcal{S}_n\to\mathbb{R}$,

$$\sum_{z\in\mathbb{Z}^2} w_n(z)\big(h(z)-h(0)\big)=0.\tag{2.23}$$

When the graph is the comb $\mathcal{C}$, Huss and Sava obtain in [10] the following result.

Theorem 2.5 (Theorem 3.5 of [10]). *There is a positive constant R_{HS}, such that for n large enough*

$$\mathcal{D}(n-R_{\mathrm{HS}})\subset\mathcal{S}_n\subset\mathcal{D}(n+R_{\mathrm{HS}}).\tag{2.24}$$

This result (Theorem 3.5 of [10]) is not precise enough for our purpose, but the arguments in [10] yield easily the following stronger result. When A, B are subsets of $\mathbb{Z}^2$, it is handful to use the notation $A+B$ for Minkowski addition $\{z=x+y\colon x\in A, y\in B\}$, and $A-B=\{z=x-y\colon x\in A, y\in B\}$.

Lemma 2.6. *There is a constant $R_{\mathrm{HS}}>0$, such that for n large enough*

$$\mathcal{D}(n)-\mathbb{B}(R_{\mathrm{HS}})\subset\mathcal{S}_n\subset\mathcal{D}(n)+\mathbb{B}(R_{\mathrm{HS}}).\tag{2.25}$$

To prove the lemma, it is enough to check that on the (continuous) boundary of $\mathcal{D}(n)$, the obstacle function γ_n is bounded by a constant, independent of n. By the symmetry of $\mathcal{D}(n)$, it is enough to consider x,y both positive satisfying $x\le n$ and $y=\frac{1}{3}(n-x)^2$. We recall Huss and Sava's expression of γ_n with our normalizing of $\mathcal{D}(n)$ (that is if n' is their n, then $n^3=9n'/4$):

$$\gamma_n(x,y)=\frac{1}{2}\left(y-\frac{1}{2}\left(\frac{2}{3}x^2-tx+\frac{9}{24}t^2+\frac{1}{6}\right)\right)^2,\tag{2.26}$$

with (using the value for $T(n)$ after (3.12) of [10] with our n)

$$t=T-\frac{20}{27}\frac{1}{T}\quad\text{and}\quad T=\frac{4}{3}n+\mathrm{O}\left(\frac{1}{n^5}\right).\tag{2.27}$$

Note that

$$t = \frac{4}{3}n - \frac{5}{9}\frac{1}{n} + O\left(\frac{1}{n^5}\right) \quad \text{and} \quad t^2 = \left(\frac{4}{3}\right)^2 n^2 - \frac{40}{27} + O\left(\frac{1}{n^2}\right). \tag{2.28}$$

A simple computation yields for $0 \le x \le n$, and $y = \frac{1}{3}(n-x)^2$

$$\gamma_n(x,y) = \frac{1}{2}\left(-\frac{5}{18}\frac{x}{n} + \frac{7}{36} + O\left(\frac{x}{n^2}\right)\right)^2. \tag{2.29}$$

Thus, for a constant K independent of n,

$$\sup_{z \in \partial_c D(n)} \gamma_n(z) \le K. \tag{2.30}$$

Now, the obstacle function is a upper bound for the odometer u_n which decays by one unit as we move along a tooth (or along the x-axis), away from the origin. Thus, there is R_{HS} such that $\mathcal{S}_n \subset \mathcal{D}(n) + \mathbb{B}(R_{\mathrm{HS}})$. Note that on $\partial(\mathcal{D}(n) + \mathbb{B}(R_{\mathrm{HS}}))$ the odometer vanishes, whereas γ_n is bounded uniformly in n, say by $\tilde{K}$. Since $u_n - \gamma_n$ is superharmonic, it satisfies the minimum principle, and satisfies in $\mathcal{D}(n) + \mathbb{B}(R_{\mathrm{HS}})$ that $u_n \ge \gamma_n - \tilde{K}$. Since γ_n increases quadratically as we move toward the origin, this implies the lower estimate $\mathcal{D}(n) - \mathbb{B}(R_{\mathrm{HS}}) \subset \mathcal{S}_n$ for some constant R_{HS} independent of n.

2.4. *On the mean-value property in $\mathcal{D}(\rho)$*

Our main result in this section is the following mean value approximation, which relies on Lemma 2.6, where the constant R_{HS} appears. We consider $z \in \partial\mathcal{D}(\rho)$, and for $y \in \mathcal{D}(\rho)$ we set $h_z(y) = P_y(S(\mathcal{E}(\rho)) = z)$.

Lemma 2.7. *There is a constant $C_{\mathrm{MV}} > 0$, such that for any $\rho > 0$ and any $z \in \partial\mathcal{D}(\rho)$*

$$\left|\mathrm{MV}\left(h_z; \mathcal{D}(\rho)\right)\right| \le C_{\mathrm{MV}} R_{\mathrm{HS}}^2. \tag{2.31}$$

Remark 2.8. *For the outer fluctuation, we need a related and simpler result, that we present now. We consider $\rho' < \rho - R_{\mathrm{HS}}$, and have that for some positive constant C_{MV}*

$$\left|\mathrm{MV}\left(h_z; \mathcal{D}(\rho')\right)\right| \le C_{\mathrm{MV}} R_{\mathrm{HS}}. \tag{2.32}$$

We explain after the proof of Lemma 2.7 how to obtain this simpler statement.

Proof of Lemma 2.7. First, we extend $h_z : \mathcal{D}(\rho) \to [0,1]$ into a harmonic function on the smallest sandpile cluster, say $\mathcal{S}$, containing $\mathcal{D}(\rho)$. By Lemma 2.6, it is enough to extend it to $\mathcal{D}(\rho) + \mathbb{B}(R_{\mathrm{HS}})$ with the constant R_{HS} appearing there.

We set $\tilde{h} \equiv h_z$ on $\mathcal{D}(\rho) \cup \partial\mathcal{D}(\rho)$. Let $w \in \partial\mathcal{D}(\rho)$ with $|X_w| \le \rho - 1$. This implies that

$$h_z(w) = 0 \quad \text{and} \quad h_z\left(\mathcal{A}(w)\right) > 0.$$

Since w is not on the x-axis, there is a unique site w' such that $\mathcal{A}^k(w') = w$, we denote for simplicity $w' = \mathcal{A}^{-k}(w)$. Since teeth are one-dimensional, harmonicity of $\tilde{h}$ imposes that for any positive integer k

$$\tilde{h}\left(\mathcal{A}^{-k}(w)\right) - \tilde{h}(w) = k\left(h_z(w) - h_z\left(\mathcal{A}(w)\right)\right) = -kh_z\left(\mathcal{A}(w)\right),$$

so that if $w \in \partial\mathcal{D}(\rho)$ but w not on the x-axis,

$$\tilde{h}\left(\mathcal{A}^{-k}(w)\right) = -kh_z\left(\mathcal{A}(w)\right). \tag{2.33}$$

On the x-axis, we choose the following extension

$$\tilde{h}\big([\rho]+k,0\big)=-(k-1)h\big([\rho],0\big). \tag{2.34}$$

Now, and if $([\rho]-1,1)\notin\mathcal{D}(\rho)$ we set for $l\in\mathbb{Z}$, we set

$$\tilde{h}\big([\rho]-1,l\big)=-(l-1)h_z\big([\rho]-1,0\big).$$

Finally, we note that $([\rho],1)\notin\mathcal{D}(\rho)$, and we extend $\tilde{h}$ by linearity on each tooth rooted on $\{([\rho],k),k\in\mathbb{N}\}$, so that for integers $k\geq 0$, and $l\in\mathbb{Z}$, we have

$$\tilde{h}\big([\rho]+k,l\big)=(l-1)(k-1)h\big([\rho],0\big)\quad\text{and}\quad\tilde{h}\big(-[\rho]-k,l\big)=(l-1)(k-1)h\big(-[\rho],0\big). \tag{2.35}$$

Using a result of Levine and Peres (Theorem 1.3 of [11]), and Lemma 2.6 of [10], there exists a function $y\mapsto\omega(y)$ with value in $[0,1]$, which vanishes on $\mathcal{D}(\rho)+\mathbb{B}(R_{\mathrm{HS}})$, which equals 1 on $\mathcal{D}(\rho)-\mathbb{B}(R_{\mathrm{HS}})$, and which satisfies

$$\sum_{y\in\mathbb{Z}^2}\omega(y)\big(\tilde{h}(y)-\tilde{h}(0)\big)=0. \tag{2.36}$$

Thus, if we denote $\partial_R\mathcal{D}(\rho)$ the shell $(\mathcal{D}(\rho)+\mathbb{B}(R_{\mathrm{HS}}))\setminus(\mathcal{D}(\rho)-\mathbb{B}(R_{\mathrm{HS}}))$, we have

$$\begin{aligned}\Big|\sum_{y\in\mathcal{D}(\rho)}\tilde{h}(y)-\tilde{h}(0)\Big|&=\Big|\sum_{\mathcal{D}(\rho)\setminus(\mathcal{D}(\rho)-\mathbb{B}(R_{\mathrm{HS}}))}\big(1-\omega(y)\big)\big(\tilde{h}(y)-\tilde{h}(0)\big)-\sum_{(\mathcal{D}(\rho)+\mathbb{B}(R_{\mathrm{HS}}))\setminus\mathcal{D}(\rho)}\omega(y)\big(\tilde{h}(y)-\tilde{h}(0)\big)\Big|\\&\leq\sum_{y\in\partial_R\mathcal{D}(\rho)}\big|\tilde{h}(y)-\tilde{h}(0)\big|.\end{aligned} \tag{2.37}$$

This implies that for some positive constant C

$$\begin{aligned}\big|\mathrm{MV}\big(h_z(\cdot),\mathcal{D}(\rho)\big)\big|&=\Big|\sum_{y\in\mathcal{D}(\rho)}\big(\tilde{h}(y)-\tilde{h}(0)\big)\Big|\leq\sum_{y\in\partial_R\mathcal{D}(\rho)}\big|\tilde{h}(y)-\tilde{h}(0)\big|\\&\leq\big|\partial_R\mathcal{D}(\rho)\big|\times h_z(0)+CR_{\mathrm{HS}}^2\sum_{y\in\partial\mathcal{D}(\rho)}h_z\big(\mathcal{A}(y)\big)\\&\leq CR_{\mathrm{HS}}^2\bigg(\frac{\rho(\rho-X_z)}{\rho^2}+\sum_{y\in\partial\mathcal{D}(\rho)}h\big(\mathcal{A}(y)\big)\bigg).\end{aligned} \tag{2.38}$$

The following bound implies (2.31). It is a consequence of Proposition 2.3, after we decompose $\partial\mathcal{D}(\rho)$ into four regions to be dealt with estimates (2.11), (2.9), (2.12) and (2.10). Thus, there is a positive constant K such that

$$\begin{aligned}\sum_{y\in\partial\mathcal{D}(\rho)}h_z\big(\mathcal{A}(y)\big)&\leq\sum_{k=1}^{[\rho]}\frac{k^3(\rho+1-X_z)}{k^2\rho^5}+\sum_{k\geq\rho+1-X_z}\frac{(\rho+1-X_z)}{k^4}+1+\sum_{k=1}^{[\rho]+1-X_z}\frac{k}{(\rho+1-X_z)^4}\\&\leq K.\end{aligned} \tag{2.39}$$

This concludes the proof of Lemma 2.7. Finally, we wish to explain Remark 2.8. First, h_z is harmonic on the smallest sandpile cluster containing $\mathcal{D}(\rho')$, so there is no need to extend it as in the previous proof. The estimates (2.38) yields here

$$\big|\mathrm{MV}\big(h_z(\cdot),\mathcal{D}(\rho')\big)\big|\leq\sum_{y\in\partial_R\mathcal{D}(\rho')}\big|h_z(y)-h_z(0)\big|\leq\big|\partial_R\mathcal{D}(\rho')\big|\times h_z(0)+\sum_{y\in\partial_R\mathcal{D}(\rho')}h_z(y). \tag{2.40}$$

It is now enough to note that on each tooth intersecting $\partial_R\mathcal{D}(\rho')$ there are at most $2R_{\mathrm{HS}}$ sites, and that the estimates for $h_z(w)$ in Proposition 2.3 are worse when $Y_w=0$, and this yields the bound

$$\sum_{w\in\partial_R\mathcal{D}(\rho')}h_z(w)\leq 2R_{\mathrm{HS}}\sum_{|x|\leq\rho'+R_{\mathrm{HS}}}h_z(x,0)\leq C_{\mathrm{MV}}R_{\mathrm{HS}}. \tag{2.41}$$

□

3. Large deviations

Our aim in this section is to prove Proposition 1.5, Lemma 1.6 and Corollary 1.7.

3.1. *On the uniform hitting property*

For each $\rho > 0$, we build here a stopping time τ_ρ^* which satisfies (1.4) of Proposition 1.5. The time τ_ρ^* is called a *flashing time*. We set

$$g_\rho(r) = \frac{3r^2}{\rho^3} \quad \text{for } r \in [0, \rho] \quad \text{and} \quad \text{for } r > \rho \quad g_\rho(r) = 0. \tag{3.1}$$

The algorithm which defines τ_ρ^* is as follows.

- Draw R according to g_ρ.
- If $R < \frac{1}{2}$, then $\tau_\rho^* = 0$, and the walk flashes on its initial position, the origin.
- If $R > \frac{1}{2}$, then $\tau_\rho^* = \inf\{t > 0\colon\ S(t) \notin \mathcal{D}(R)\}$.

We need to estimate $P_0(S(\tau_\rho^*) = z)$ for $z \in \mathcal{D}(\rho)$. We have

$$P_0\big(S\big(\tau_\rho^*\big) = z\big) = P\left(R < \frac{1}{2}\right)\mathbb{1}_{z=0} + \int_{1/2}^{\rho} P_0\big(S\big(\mathcal{E}(r)\big) = z\big)g_\rho(r)\,\mathrm{d}r \times \mathbb{1}_{z\neq 0}.$$

First, $P(S(\tau_\rho^*) = 0) = P(R < \frac{1}{2}) = 1/(2\rho)^3$. Assume henceforth that $z \neq 0$, and note that $S(\mathcal{E}(r)) = z$ is possible only if $z \in \partial\mathcal{D}(r)$. Thus, we define $R(z) < \bar{R}(z)$ such that

$$z \in \partial\mathcal{D}(r) \quad \Longleftrightarrow \quad R(z) < r \le \bar{R}(z). \tag{3.2}$$

In other words, we define

$$\frac{1}{3}\big(\bar{R}(z) - |X_z|\big)^2 = |Y_z| \quad \text{and} \quad R(z) = \bar{R}\big(\mathcal{A}(z)\big). \tag{3.3}$$

We need to estimate $P_0(S(\mathcal{E}(r)) = z)$ for $R(z) \le r < \bar{R}(z)$. Upper and lower bounds are obtain for Green's function in Lemma 2.1, and hold for the exit distribution by the last passage decomposition (2.8). Now, the upper and the lower bound for $P_0(S(\tau_\rho^*) = z)$ are done in a similar way, and we write in details only the upper bound. Also, because of the symmetry of $\mathcal{D}(\rho)$, we can assume that $X_z \ge 0$ and $Y_z \ge 0$. We treat three cases: (i) when z is a nearest neighbor of the origin, (ii) when $\mathcal{A}(z) \neq 0$ and z is not on the x-axis, and (iii) when $\mathcal{A}(z) \neq 0$ and $Y_z = 0$.

Case (i): $\mathcal{A}(z) = 0$. Then, $\bar{R}(z) \le 2$ and $R(z) = 0$. We have

$$\begin{aligned} P_0\big(S\big(\tau_\rho^*\big) = z\big) &\le \int_{1/2}^{\bar{R}(z)} \frac{1}{\deg(z)} \frac{2(r + 1 - X_z)}{r^2} \frac{3r^2}{\rho}\,\mathrm{d}r \\ &\le \frac{3}{2\rho^3}\left(\bar{R}(z) - \frac{1}{2}\right)\big(\bar{R}(z) + 1 - X_z\big) \le \frac{27}{4\rho^3}. \end{aligned} \tag{3.4}$$

Case (ii): $\mathcal{A}(z) \neq 0$ *and* $Y_z \neq 0$. Note that $\mathcal{A}(z) = (X_z, Y_z - 1)$, $R(z) \ge 1$, and

$$\bar{R}(z) - R(z) = \sqrt{3Y_z} - \sqrt{3(Y_z - 1)} \le \frac{\sqrt{3}}{\sqrt{Y_z}}.$$

Then

$$\begin{aligned} P_0\big(S\big(\tau_\rho^*\big) = z\big) &\le \int_{R(z)}^{\bar{R}(z)} \frac{1}{\deg(z)} \frac{2(r + 1 - X_z)}{r^2} \frac{3r^2}{\rho}\,\mathrm{d}r \\ &\le \frac{3}{2\rho^3}\big(\bar{R}(z) + 1 - X_z\big)\big(\bar{R}(z) - R(z)\big) \le \frac{3}{2\rho^3}\sqrt{3Y_z}\frac{\sqrt{3}}{\sqrt{Y_z}} \le \frac{9}{2\rho^3}. \end{aligned} \tag{3.5}$$

Case (iii): $\mathcal{A}(z) \neq 0$ *and* $Y_z = 0$. Then $\mathcal{A}(z) = (X_z - 1, 0)$, $\bar{R}(z) = X_z$ and $R(z) = X_z - 1 \geq 1$. We have

$$P_0\big(S\big(\tau_\rho^*\big) = z\big) \leq \int_{X_z-1}^{X_z} \frac{1}{\deg(z)} \frac{2(r+1-X_z)}{r^2} \frac{3r^2}{\rho} \, \mathrm{d}r \leq \frac{3}{4\rho^3}. \tag{3.6}$$

We omit the similar estimates yielding the lower bound of (1.4).

3.2. *Proof of Lemma* 1.6

We consider here that the explored region is V, and estimate the probability an explorer reaches $(R, 0)$. To obtain (1.5) we can assume that the ratio $|V|/R^3$ is as small as we wish. Also, we can restrict to $|V| \geq R$, since an explorer reaching $(R, 0)$ has to visit all sites of $\{(x, 0), 0 \leq x \leq R\}$.

The proof makes use of the concept of *flashing explorer*, which was introduced in [1], and follows the arguments of the proof of Lemma 1.6 of [2]. A flashing explorer is a random walk which settles only if at some times, *the flashing times*, it is not on the explored region V.

If an explorer reaches $(R, 0)$ (without escaping V), then a flashing explorer following the same trajectory would reach $(R, 0)$ as well. Since Lemma 1.6 requires a bound from below on the crossing probability, it is enough to obtain an estimate for the flashing explorer.

We now define the flashing explorer associated with *the scale* h. Let h be a positive real smaller than $R/2$, and write M_h for the integer part of $R/(2h)$. We form M_h disjoint domains by translating $\mathcal{D}(h)$ so that they cover $[0, 2hM_h]$, see Figure 2, and we call $Z_1, \ldots, Z_{M_h}$ their centers. The flashing explorer associated with scale h is as follows.

- It performs a simple random walk on the comb, starting at 0.
- The first time the walk reaches Z_i, it draws one variable R_i according to g_h and a flashing time τ_i^* is constructed as in the previous section but around $\mathcal{D}(Z_i, h)$.
- It settles the first time $H(Z_i) + \tau_i^*$ that $S(H(Z_i) + \tau_i^*) \notin V$, for $i = 1, \ldots, M_h$.

We say that the domain $\mathcal{D}(Z_i, h)$ is *well-covered* when $|\mathcal{D}(Z_i, h) \cap V| > \beta|\mathcal{D}(Z_i, h)|$, for a positive $\beta < 1$ to be chosen later. We call Γ_h the set of well-covered domains:

$$\Gamma_h = \big\{i \in [1, M_h]\colon\ \big|\mathcal{D}(Z_i, h) \cap V\big| > \beta\big|\mathcal{D}(Z_i, h)\big|\big\}. \tag{3.7}$$

The reason we use a flashing explorer is that the probability that it settles in a *not well-covered* domain is easy to estimate. Indeed, by the *uniform hitting property*, it visits the domain $\mathcal{D}(Z_i, h)$ almost uniformly, and the probability

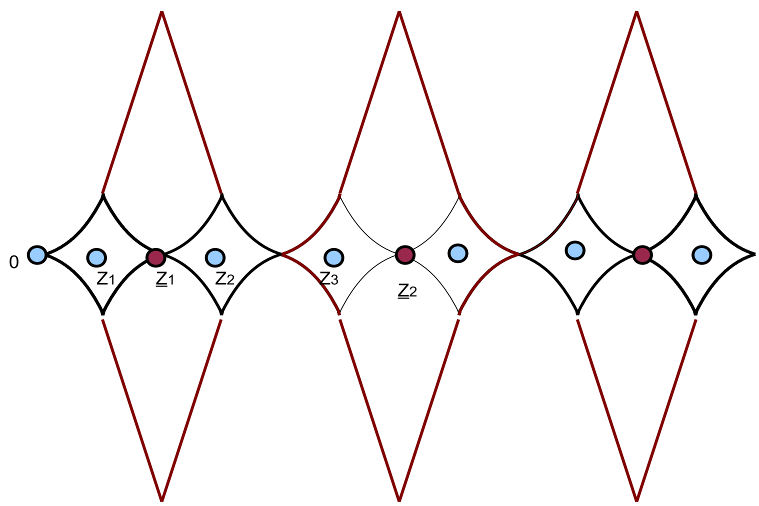

Fig. 2. Two scales h and $\underline{h} = 2h$, and corresponding centers.

it flashes on a site of V is less than $\kappa\beta$, for some positive constant $\kappa < 1$ (independent of β and h). We now choose β by requiring that when $h = R/2$ (and $M_h = 1$ and $Z_1 = (h, 0)$), then $\Gamma_h = \emptyset$. Thus,

$$P_0\big(\text{the flashing explorer reaches } (R,0)\big) \le \prod_{i \notin \Gamma_h} P_{Z_i}\big(S\big(H(Z_i) + \tau_i^*\big) \in V\big) \le (\kappa\beta)^{M_h - |\Gamma_h|}. \tag{3.8}$$

Now, $\beta|\mathcal{D}(h)||\Gamma_h| \le |V|$, and for some $\alpha > 0$ we have $|\mathcal{D}(h)| \ge \alpha h^3$. Replacing $|\Gamma_h|$ in (3.8) by an upper bound yields

$$\begin{aligned} P_0\big(\text{the flashing explorer reaches } (R,0)\big) &\le \exp\left(-\log\left(\frac{1}{\kappa\beta}\right)\left(M_h - \frac{|V|}{\beta|\mathcal{D}(h)|}\right)\right) \\ &\le \exp\left(-\log\left(\frac{1}{\kappa\beta}\right)\left(\frac{R}{2h} - 1 - \frac{|V|}{\beta\alpha h^3}\right)\right). \end{aligned} \tag{3.9}$$

We optimize now in h. The maximum of $R/(2h) - |V|/(\beta\alpha h^3)$ is reached for $h^* = \sqrt{6/\beta\alpha}\sqrt{|V|/R}$, and this choice completes the proof of (1.5).

3.3. *Proof of Corollary* 1.7

The proof relies on Lemma 1.6, and follows closely the arguments of Lemma 1.5 of [2] with $d = 3$. Indeed, $|\mathcal{D}(R)|$ is of order R^3.

The strategy of the proof is to built optimal disjoint random domains $\mathcal{D}(Z_0, h_0), \dots, \mathcal{D}(Z_L, h_L)$ inside $\mathcal{D}(0, R)$ in such a that if N_i is the number of settled explorers in $\mathcal{D}(Z_i, h_i)$, we have that $(R, 0) \in A(n)$ implies that for each j, $N_{j+1} + \cdots + N_L$ explorers have crossed $\mathcal{D}(Z_0, h_0) \cup \cdots \cup \mathcal{D}(Z_j, h_j)$. The randomness comes from $A(n)$.

We choose $h_0 = R/4 > 1$, and $Z_0 = (h_0, 0)$. We choose a positive (large) constant γ from a_3 and κ_3 of Lemma 1.6:

$$\gamma = \max\left(1, \left(\frac{2a_3}{\kappa_3}\right)^2\right).$$

The choice of γ will be clear later. We choose now β such that $\gamma d(n) < \gamma\beta R^3 < |\mathcal{D}(h_0)|$.

We now build by induction neighboring domains $\mathcal{D}(Z_i, h_i)$ for $i = 1, \dots, L$ such that

$$\sum_{i=0}^{L} \mathcal{D}(Z_i, h_i) \supset \big\{(0,k)\colon\ 0 \le k \le R, k \in \mathbb{N}\big\}.$$

Assume we have chosen $h_{i-1} \ge 1$ and $2(h_0 + \cdots + h_{i-1}) < R$. We choose h_i such that

$$\big|\mathcal{D}(h_i)\big| = \gamma\big|\mathcal{D}(Z_{i-1}, h_{i-1}) \cap A(n)\big| \quad \text{and} \quad Z_i = \left(2\sum_{j=0}^{i-1} h_j + h_i, 0\right). \tag{3.10}$$

Note that since $h_{i-1} \ge 1$, we have $N_{i-1} = |\mathcal{D}(Z_{i-1}, h_{i-1}) \cap A(n)| \ge 3$, and $h_i \ge 1$. Clearly, the induction stops with L domains, and $L \le R/2$.

For any choice of integers $l, n_0, \dots, n_l$, the event $\{L = l, N_0 = n_0, \dots, N_l = n_l\}$ implies that $n_1 + \cdots + n_l$ explorers have crossed $\mathcal{D}(Z_0, h_0)$ with an explored region made up of n_0 settled explorers, and $n_2 + \cdots + n_l$ explorers have crossed $\mathcal{D}(Z_1, h_1)$ with an explored region made up of n_1 settled explorers, and so on and so forth. We use now (1.5) of Lemma 1.6 to obtain

$$\begin{aligned} P\big(\text{an explorer reaches } (R,0)\big) &\le \sum_{l, n_0, \dots, n_l} \prod_{i>0} P\left(\sum_{k=i}^{l} n_k \text{ explorers cross } \mathcal{D}(Z_{i-1}, h_{i-1})\right) \\ &\le \big(R^3\big)^R \sup_{l, n_1, \dots, n_l} \exp\left(a_3 \sum_{i=1}^{l} i n_i - \kappa_3 \sum_{i=1}^{l} n_i \left(\sqrt{\frac{h_0^3}{n_0}} + \cdots + \sqrt{\frac{h_{i-1}^3}{n_{i-1}}}\right)\right). \end{aligned} \tag{3.11}$$

By the arithmetic–geometric inequality, for $1 \le i \le l$ (and using $h_i \le h_0$)

$$\frac{1}{i}\left(\sqrt{\frac{h_0^3}{n_0}} + \cdots + \sqrt{\frac{h_{i-1}^3}{n_{i-1}}}\right) \ge \left(\frac{h_0^3}{n_0} \times \cdots \times \frac{h_{i-1}^3}{n_{i-1}}\right)^{1/2i} = \left(\frac{h_0^3}{n_{i-1}}\gamma^{i-1}\right)^{1/2i} = \left(\frac{h_0^3}{h_i^3}\gamma^{i}\right)^{1/2i} \ge \frac{2a_3}{\kappa_3}. \tag{3.12}$$

Thus, from (3.11) and (3.12), we have

$$P\big((R,0)\in A(n)\big) \le R\big(\beta R^3\big)^{R+1} \max_{\substack{l\le R, n_0,n_1,\dots,n_l\le \beta R^3\\ \forall i, n_i\ge \lfloor h_i\rfloor}} \exp\left(-a_3\sum_{i=1}^{l} i n_i\right) \le R\big(\beta R^3\big)^{R+1} \max_{\substack{l\le R, n_0,n_1,\dots,n_l\le \beta R^3\\ \forall i, n_i\ge \lfloor h_i\rfloor}} \exp\left(-\frac{a_3}{\gamma}\sum_{i=1}^{l-1} i h_{i+1}^3\right). \tag{3.13}$$

Since $h_1 \le R/4$, we have $h_2 + \cdots + h_l \ge R/4$. By Hölder's inequality, and for a constant c_3, we have

$$\sum_{i=1}^{l-1} i h_{i+1}^3 \ge \frac{(\sum_{i=1}^{l-1} h_{i+1})^3}{(\sum_{i=1}^{l-1} 1/\sqrt{i})^2} \ge c_3\frac{R^3}{l} \ge c_3 R^2. \tag{3.14}$$

This concludes the proof.

4. Inner fluctuations

Our main result here is an inner estimate for the aggregate.

Proposition 4.1. *There is a positive constant κ_{in} (independent of n) such that for any positive a and integer n large enough*

$$P\big(\mathcal{D}\big(n - a\sqrt{\log(n)}\big) \not\subset A(n)\big) \le d(n)\exp\big(-\kappa_{in}a^2\log(n)\big). \tag{4.1}$$

Remark 4.2. *Since $d(n)$ is of order n^3, (4.1) implies the inner estimate (1.2) of Theorem 1.2. Our proof below establishes actually a stronger result than (4.1). Indeed, we only count explorers which remain in the domain $\mathcal{D}(n)$. This remark is used in the outer error bound.*

Proof of Proposition 4.1. The constant $A > 1$ will be chosen later. For any $\alpha > 0$, we set $a = A\sqrt{\alpha}$ and $L = \sqrt{\alpha\log(n)}$. Inequality (4.1) follows if we show that for $z \in \partial\mathcal{D}(n-L)$ with $|X_z| \le n - AL$, we have

$$P\big(W_{\mathcal{D}(n-L)}(z) = 0\big) \le \exp\big(-\kappa_{in}a^2\log(n)\big). \tag{4.2}$$

Indeed, since the comb is a tree, covering $\partial\mathcal{D}(n-L)\cap\{z\colon |X_z| < n - AL\}$ by the DLA cluster implies that $\mathcal{D}(n-AL)$ is entirely covered. Henceforth, consider $z \in \partial\mathcal{D}(n-L)$ with $|X_z| \le n - AL$. For $y \in \mathcal{D}(n-AL)$, define $h_z(y) = P_y(S(\mathcal{E}(n-L)) = z)$, and set

$$\mu(z) = E\big[M_{\mathcal{D}(n-L)}\big(d(n)\mathbb{1}_0; z\big)\big] - E\big[M_{\mathcal{D}(n-L)}\big(\mathcal{D}(n-L); z\big)\big] = \big(\big|\mathcal{D}(n)\big| - \big|\mathcal{D}(n-L)\big|\big)h_z(0) + \mathrm{MV}\big(h_z, \mathcal{D}(n-L)\big). \tag{4.3}$$

Using (2.5) of Lemma 2.1 and Lemma 2.7, we obtain when n is large enough, and for $c_1, c_2 > 0$ independent of n and z,

$$c_2 L\big(n - L - |X_z|\big) \le \mu(z) \le c_1 L\big(n - L - |X_z|\big). \tag{4.4}$$

Note also that by (2.13) of Corollary 2.4, there are κ_1, κ_2 independent of n and z such that

$$\kappa_2\big(n-L-|X_z|\big)^2 \le \sum_{y\in\mathcal{D}(n-L)} h_z^2(y) \le \kappa_1\big(n-L-|X_z|\big)^2.$$

In order to use Lemma C.1, we form the following partition of $\mathcal{D}(n-L)$:

$$\mathcal{B}=\big\{y\in\mathcal{D}(n-L)\colon\ X_z=X_y, Y_z\times Y_y\ge 0\big\} \quad\text{and}\quad \mathcal{A}=\mathcal{D}(n-L)\setminus\mathcal{B}. \tag{4.5}$$

We need to show that there is $\kappa>1$, independent of n such that for $y\in\mathcal{A}$, we have $h_z(y)<1-1/\kappa$. We choose here $\kappa=2$ for simplicity, and note that for $y\in\mathcal{A}$, any path joining y and z crosses $(X_z,0)$ so that $h_z(y)\le h_z(X_z,0)$. Note also that

$$h_z(X_z,0)=\frac{G_{\mathcal{D}(n-L)}((X_z,0);\mathcal{A}(z))}{2}=\frac{G_{\mathcal{D}(n-L)}(0;\mathcal{A}(z))}{2P_0(H(X_z,0)<\mathcal{E}(n-L))}. \tag{4.6}$$

Using now Lemma 2.1 and Proposition 2.2, we obtain readily that $h_z(X_z,0)$ can be made smaller than $1/2$ as soon as n is large enough.

For δ to be fixed later, we choose $A=1+\frac{1}{\delta}$ so that $\delta(n-L-|X_z|)\ge L$, and we choose λ as follows.

$$\lambda=\frac{\mu(z)}{2(16|\mathcal{B}|+(1/2)\sum_{y\in\mathcal{D}(n-L)}h_z^2(y))}\le\frac{c_1}{2(16/3+\kappa_2)}\frac{L(n-L-|X_z|)}{(n-L-|X_z|)^2}\le\frac{c_1}{2(16/3+\kappa_2)}\delta. \tag{4.7}$$

Since, we need $\lambda<\log(2)$ in Lemma C.1, the condition on δ is that

$$\delta<2\log(2)\frac{16/3+\kappa_2}{c_1}. \tag{4.8}$$

We use Lemma C.1, with $\xi=0$, $\kappa=1/2$ and $\lambda\le\log(2)$ to have

$$P\big(W_{\mathcal{D}(n-L)}(z)=0\big)\le\exp\left(-\lambda\mu(z)+\frac{\lambda^2}{2}\left(\mu(z)+16|\mathcal{B}|+\frac{1}{2}\sum_{y\in\mathcal{D}(n-L)}h_z^2(y)\right)\right).$$

Note that since $\lambda<1$, we have $\lambda\mu(z)-\lambda^2\mu(z)/2>\lambda\mu(z)/2$, and the choice of λ in (4.7) yields (4.2) with κ_{in} given by $c_2/(8A^2(16/3+\kappa_1))$. □

5. Outer fluctuations

We estimate the probability that the largest finger reaches $\partial\mathcal{D}(n+A\sqrt{\log(n)})$ for some large A. The analysis is distinct whether the finger protrudes in the *tip* of $\partial\mathcal{D}(n+A\sqrt{\log(n)})$, that is the region

$$\mathcal{T}=\left\{z\colon\ |X_z|>n+\frac{A}{2}L\right\}\cap\partial\mathcal{D}(n+AL)\quad\big(\text{here } L:=\big[\sqrt{\log(n)}\big]\big) \tag{5.1}$$

or in the complement of $\mathcal{T}$ in $\partial\mathcal{D}(n+AL)$ called the *bulk*, and made of the edges of long teeth. Indeed, the geometry of the graph is different on the tip, and on the edge of a long tooth. The goal is to show that the appearance of a long finger implies that a *narrow region* is crossed by *many explorers*. More precisely, when a finger reaches site z of the bulk, that is through a long tooth, it imposes that many explorers settle in the tooth: if $Y_z>2k$ and we set $\tilde{z}=\mathcal{A}^k(z)$, then in order to cover z, we need that k explorers cross $\tilde{z}$. Moreover, at least half of these explorers if they were random walks starting on z would very likely exit $\mathcal{D}(n+AL)$ from z. On the other hand, when a finger reaches a site z of the tip, say on $z=(n+AL,0)$, this imposes that site $(n,0)$ is crossed by AL explorers, but these explorers if they where random walks would have many ways to exit $\mathcal{D}(n+AL)$. We call $\mathcal{B}(z)$ the event that $A(n)\not\subset\mathcal{D}(n+AL)$, and z is the first site of $\partial\mathcal{D}(n+AL)$ to be covered by the aggregate. Note that

$$\big\{A(n)\not\subset\mathcal{D}(n+AL)\big\}=\bigcup_{z\in\partial\mathcal{D}(n+AL)}\mathcal{B}(z). \tag{5.2}$$

5.1. *The bulk region*

We assume for $z \in \partial\mathcal{D}(n+AL)$ with $|X_z| < n + AL/2$ that $\mathcal{B}(z)$ holds. This implies that explorers fill the tooth of z without escaping $\mathcal{D}(n+AL)$. For simplicity, we denote $\bar{D} = \mathcal{D}(n+AL)$. Let $\tilde{z}$ be the site of $\partial\mathcal{D}(n+3AL/4)$ on the same tooth as z. Necessarily, the number of explorers crossing $\tilde{z}$ before escaping $\bar{D}$ is larger than the length of the tooth to be covered

$$\xi(\tilde{z}) = \frac{1}{3}\big(n + AL - |X_z|\big)^2 - \frac{1}{3}\left(n + \frac{3AL}{4} - |X_z|\right)^2$$
$$\geq \frac{AL}{6}\left(n + \frac{3AL}{4} - |X_z|\right). \tag{5.3}$$

To exploit this information, we introduce an auxiliary process which proved useful in studying DLA [1,2]. The *flashing* process is a cluster growth, were explorers, called $*$-explorers settle less often than in DLA. We now build a flashing process adapted to our purpose.

- Inside $\mathcal{D}(n)$, $*$-explorers are just explorers.
- When a $*$-explorer exits $\mathcal{D}(n)$, it cannot settle in $\mathcal{D}(n)$ anymore.
- A $*$-explorer does not settle in $\bar{D} \setminus \mathcal{D}(n)$.
- Outside $\bar{D}$, $*$-explorers behave like explorers.

We call $A^*(n)$ the cluster made by $d(n)$ $*$-explorers sent at the origin. Note that by construction, the cluster made by $d(n)$ explorers before they exit $\mathcal{D}(n)$, denoted $A^*_{\mathcal{D}(N)}(n)$, is equal to $A_{\mathcal{D}(N)}(n)$.

The key fact, established in [1], is that this growth can be coupled with the internal DLA cluster in such a way that there are times $\{T_i, T_i^*, i = 1, \dots, d(n)\}$ with $T_i \leq T_i^*$ and such that for $d(n)$ independent random walks $S_1, \dots, S_{d(n)}$

$$A(n) = \big\{S_i(T_i), i = 1, \dots, d(n)\big\} \quad \text{and} \quad A^*(n) = \big\{S_i\big(T_i^*\big), i = 1, \dots, d(n)\big\}. \tag{5.4}$$

Thus, under the coupling of [1], if an explorer happens to visit $\tilde{z}$ before escaping $\bar{D}$, then this will be the case for the associated $*$-explorer. We add an index $*$ to denote objects linked with $*$-explorers. For instance, we denote by $W^*_{\Lambda}(\eta, z)$ the number of $*$-explorers which cross z before escaping Λ, and we drop the η dependence when $\eta = d(n)\mathbb{1}_0$. As a consequence of the coupling, we have

$$\mathcal{B}(z) \subset \big\{W^*_{\bar{D}}(\tilde{z}) > \xi(\tilde{z})\big\}. \tag{5.5}$$

Let us now estimate how many $*$-explorers exit $\bar{D}$ most likely from z.

Note first that when $|X_z| < n + AL/2$, then

$$\frac{Y_{\tilde{z}}}{Y_z} = \frac{(1/3)(n + 3AL/4 - |X_z|)^2}{(1/3)(n + AL - |X_z|)^2} \geq \frac{1}{4}, \quad \text{so } P_{\tilde{z}}\big(\text{walk hits } z \text{ before } (X_z, 0)\big) > \frac{1}{4}. \tag{5.6}$$

$W^*_{\bar{D}}(z)$ represents the number of $*$-explorers which exit $\bar{D}$ from z, out of $W^*_{\bar{D}}(\tilde{z})$ $*$-explorers at $\tilde{z}$. Inside $\bar{D}$, the $*$-explorers are just simple random walks, and by (5.6), we have that $E[W^*_{\bar{D}}(z) | W^*_{\bar{D}}(\tilde{z})] \geq \frac{1}{4} W^*_{\bar{D}}(\tilde{z})$. Thus, by Chernoff's inequality

$$P\left(W^*_{\bar{D}}(z) < \frac{1}{8} W^*_{\bar{D}}(\tilde{z}) \big| W^*_{\bar{D}}(\tilde{z}) > \xi(\tilde{z})\right) \leq \exp\left(-\frac{1}{4}\frac{\xi(\tilde{z})}{8}\right). \tag{5.7}$$

Second, note that since z is a bulk site $n + 3AL/4 - |X_z| > AL/4$, and on the event $\mathcal{B}(z)$ we have that $W_{\bar{D}}(\tilde{z}) > \xi(\tilde{z})$, which in turn implies $W^*_{\bar{D}}(\tilde{z}) > \xi(\tilde{z}) > (AL)^2/24$. Thus, we have

$$P\left(W^*_{\bar{D}}(z) < \frac{1}{8} W^*_{\bar{D}}(\tilde{z}) \big| \mathcal{B}(z)\right) \leq \exp\left(-\frac{1}{24 \times 32}(AL)^2\right). \tag{5.8}$$

Thus,

$$P\big(\mathcal{B}(z)\big) \le P\bigg(\mathcal{B}(z), W^*_{\bar D}(z) < \frac{1}{8} W^*_{\bar D}(\tilde z)\bigg) + P\bigg(\mathcal{B}(z), W^*_{\bar D}(z) \ge \frac{1}{8} W^*_{\bar D}(\tilde z)\bigg)$$
$$\le P\bigg(W^*_{\bar D}(z) < \frac{1}{8} W^*_{\bar D}(\tilde z) \big| \mathcal{B}(z)\bigg) + P\big(\mathcal{D}(n-a_0L) \not\subset A^*_{\mathcal{D}(n)}(n)\big)$$
$$+ P\bigg(W^*_{\bar D}(z) \ge \frac{\xi(\tilde z)}{8}, \mathcal{D}(n-a_0L) \subset A^*_{\mathcal{D}(n)}(n)\bigg). \tag{5.9}$$

The probability $\mathcal{D}(n-a_0L) \not\subset A^*_{\mathcal{D}(n)}(n)$ is actually estimated in Proposition 4.1 since $A^*_{\mathcal{D}(n)}(n) = A_{\mathcal{D}(n)}(n)$.

We explain now why $\{W^*_{\bar D}(z) \ge \xi\}$ is very unlikely, where we set for simplicity $\xi = \xi(\tilde z)/8$. Since our inner error estimate is also valid for $*$-explorers, we have the equality in law

$$W^*_{\bar D}(z) + M_{\bar D}\big(A^*_{\bar D}(n), z\big) = M_{\bar D}\big(d(n)\mathbb{1}_0, z\big). \tag{5.10}$$

This implies that

$$\mathbb{1}_{\mathcal{D}(n-a_0L)\subset A^*_{\bar D}(n)}\big(W^*_{\bar D}(z) + M_{\bar D}\big(\mathcal{D}(n-a_0L), z\big)\big) \le M_{\bar D}\big(d(n)\mathbb{1}_0, z\big). \tag{5.11}$$

Now, (5.11) allows us to estimate the probability that $W^*_{\bar D}(z)$ is large, through Lemma 2.5 of [2]: for $0 < \lambda < \log(2)$, and $\xi > \mu^*(z) := E[M_{\bar D}(d(n)\mathbb{1}_0, z)] - E[M_{\bar D}(\mathcal{D}(n-a_0L), z)]$,

$$P\big(W^*_{\bar D}(z) > \xi, \mathcal{D}(n-a_0L) \subset A^*_{\bar D}(n)\big) \le \exp\bigg(-\lambda\big(\xi - \mu^*(z)\big) + \lambda^2\bigg(\mu^*(z) + 4\sum_{y\in\bar D} h_z^2(y)\bigg)\bigg), \tag{5.12}$$

where $h_z(y)$ is the probability of exiting $\bar D$ from z, when a random walk starts on y. Note that the function $y \mapsto h_z(y)$ is harmonic on $\bar D$, and that since $(A-a_0)L \ge R_{\text{HS}}$ Remark 2.8 applies. There is a constant c^* such that (recall that z is in the bulk)

$$\mu^*(z) = \big(\big|\mathcal{D}(n)\big| - \big|\mathcal{D}(n-a_0L)\big|\big)h_z(0) + \text{MV}\big(h_z, \mathcal{D}(n-a_0L)\big)$$
$$\le c^* a_0 L\big(n + AL - |X_z|\big). \tag{5.13}$$

We choose A large enough, after a_0 is fixed, so that

$$\mu^*(z) \le c^* a_0 L\big(n + AL - |X_z|\big) \ll \frac{1}{4}\xi \le \frac{1}{32}\frac{AL}{4}\big(n + AL - |X_z|\big). \tag{5.14}$$

Also, by (2.13) of Corollary 2.4 we have

$$\sum_{y\in\bar D} h_z^2(y) \le \kappa_2\big(n + AL - |X_z|\big)^2. \tag{5.15}$$

In the bulk, the following choice of λ with the estimate (5.15) yields

$$\lambda = \frac{\mu^*(z)}{16\sum_{y\in\bar D} h_z^2(y)} \le \frac{a_0 L}{16\kappa_2(n + AL - |X_z|)} \le \frac{a_0}{8\kappa_2 A}. \tag{5.16}$$

One chooses A large enough so that $\lambda < \log(2)$. Note that since $\lambda < 1$, our choice of A in (5.14) is such that $\lambda(\xi - \mu^*(z)) - \lambda^2\mu \ge \lambda\xi/2$. Using (5.12) with the choice of λ in (5.16), and after simple algebra, one obtains for some constant κ

$$P\big(W^*_{\bar D}(z) > \xi, \mathcal{D}(n-a_0L) \subset A^*_{\bar D}(n)\big) \le \exp\big(-\kappa a_0 A L^2\big). \tag{5.17}$$

Finally, from the inner error, we know that for a_0 large enough most likely $\mathcal{D}(n-a_0L)\subset A_{\mathcal{D}(n)}(n)=A^*_{\mathcal{D}(n)}(n)\subset A^*_{\tilde{D}}(n)$. Combining the estimates for the three terms on the right hand side of (5.9), we obtain

$$P\big(\mathcal{B}(z)\big)\le \exp\left(-\frac{1}{24\times 32}A^2\log(n)\right)+\exp\big(-\kappa a_0 A\log(n)\big)+\exp\big(-\kappa_{in}a_0^2\log(n)\big).$$

We can choose a_0, and then A so that $P(\mathcal{B}(z))$ is smaller than any negative power of n.

5.2. *The tip*

Let us describe the additional idea needed to deal with the tip. A constant A large enough will be chosen later. Assume $A/4\in\mathbb{N}$, and define three points

$$B=\left(n+\frac{A}{2}L,0\right),\qquad \tilde{B}=\left(n+\frac{A}{4}L,0\right)\quad\text{and}\quad C\in\partial\mathcal{D}(n+AL)\quad\text{with } X_C=n+\frac{A}{4}L.$$

Assume in this section that a site of the tip is covered. This implies that B or $-B$ is in $A(n)$. Assume for instance that $B\in A(n)$. The internal DLA covering mechanism would say that $\tilde{B}$ is necessarily covered by $\frac{A}{4}L$ explorers. However, too small a fraction, of the order of $1/L$, of these explorers would exit $\mathcal{D}(\tilde{B},\frac{3}{4}AL)$ from site C. We first need to show that of the order of $(AL)^3$ explorers cross $\tilde{B}$, and secondly that it is very unlikely that of the order of $(AL)^2$ exit $\mathcal{D}(\tilde{B},\frac{3}{4}AL)$ from site C.

As in the previous section, we need to consider here the same $*$-explorers. An important property is the fact that the aggregate's law is independent of the order of the explorers we launch, or more generally, of stopping explorers in some region letting other explorers cover space before the stopped ones are eventually launched. Thus, we will realize the aggregate by sending two waves of exploration. We stop $*$-explorers on $\tilde{B}\cup\partial D(n+AL)$, and call ζ the configuration of stopped $*$-explorers, that is $\zeta:\{\tilde{B}\}\cup\partial\mathcal{D}(n+AL)\to\mathbb{N}$.

The event that B is covered, and $\zeta(\tilde{B})$ is less than $\beta(AL/4)^3$, is very unlikely by Corollary 1.7. Henceforth, assume that $\zeta(\tilde{B})>\beta(AL/4)^3$, where we recall that β is a constant independent of A,n. Assume that we launch the stopped $*$-explorers and stop them on $\partial\mathcal{D}(n+AL)$. It is very unlikely that less than $\kappa(AL)^2$ $*$-explorers exit $\partial\mathcal{D}(n+AL)$ from C for some positive constant κ. Indeed, let us call the latter number $W^*_{\tilde{D}}(\zeta(\tilde{B})\mathbb{1}_{\tilde{B}};C)$. Call for simplicity $\Lambda=\mathcal{D}(\tilde{B},\frac{3}{4}AL)$, and define $M_\Lambda(\zeta(\tilde{B})\mathbb{1}_{\tilde{B}};C)$ the number of random walks which exit Λ from C. Note the following obvious fact

$$\Lambda\subset\bar{D}\quad\Longleftrightarrow\quad W^*_{\tilde{D}}\big(\zeta(\tilde{B})\mathbb{1}_{\tilde{B}};C\big)\ge W^*_{\Lambda}\big(\zeta(\tilde{B})\mathbb{1}_{\tilde{B}};C\big).$$

Also, the Abelian property we mentioned says that (with equality in law)

$$W^*_{\tilde{D}}(C)=W^*_{\tilde{D}}\big(\zeta(\tilde{B})\mathbb{1}_{\tilde{B}};C\big).\tag{5.18}$$

Thus, to estimate the probability that $W^*_{\tilde{D}}(C)$ be small it is enough to estimate the probability that $W^*_\Lambda(\zeta(\tilde{B})\mathbb{1}_{\tilde{B}};C)$ be small. Note that $*$-explorers when starting in $\tilde{B}$ and staying in Λ have the same law as simple random walks, so that

$$W^*_\Lambda\big(\zeta(\tilde{B})\mathbb{1}_{\tilde{B}};C\big)=M_\Lambda\big(\zeta(\tilde{B})\mathbb{1}_{\tilde{B}};C\big).$$

M_Λ is a sum of independent Bernoulli, and from (2.5), there is a constant $\kappa>0$

$$E\left[M_\Lambda\big(\zeta(\tilde{B})\mathbb{1}_{\tilde{B}};C\big)\Big|\zeta(\tilde{B})\ge\beta\left(\frac{AL}{4}\right)^3\right]\ge\beta\left(\frac{AL}{4}\right)^3\frac{2}{3AL/4}\ge 2\kappa(AL)^2,$$

and therefore, using Chernoff's inequality on the event $\{\zeta(\tilde{B})\ge\beta(\frac{AL}{4})^3\}$

$$\begin{aligned}P\big(W^*_\Lambda\big(\zeta(\tilde{B})\mathbb{1}_{\tilde{B}};C\big)\le\kappa(AL)^2\big)&=P\big(M_\Lambda\big(\zeta(\tilde{B})\mathbb{1}_{\tilde{B}};C\big)\le\kappa(AL)^2\big)\\&\le\exp\big(-\kappa(AL)^2\big).\end{aligned}\tag{5.19}$$

We deal now with the event $\{W^*_{\bar{D}}(C) > \kappa(AL)^2\}$. Note that defining

$$\mu^*_{\bar{D}}(C) := E\big[M_{\bar{D}}\big(d(n)\mathbb{1}_0, C\big)\big] - E\big[M_{\bar{D}}\big(\mathcal{D}(n-a_0L), C\big)\big],$$

we have using our harmonic measure estimate, for some constant c^*

$$\mu^*_{\bar{D}}(C) \le c^* a_0 AL^2 \quad \text{and} \quad \sum_{y\in\mathcal{D}(n-a_0L)} h_C^2(y) \le \sum_{y\in\bar{D}} h_C^2(y) \le \kappa_2\left(\frac{3AL}{4}\right)^2. \tag{5.20}$$

Thus, for $\lambda < \log(2)$,

$$P\big(W^*_{\bar{D}}(C) > \kappa(AL)^2, \mathcal{D}(n-a_0L) \subset A^*_{\bar{D}}(n)\big) \le \mathrm{e}^{-\lambda(\kappa(AL)^2 - c^* a_0 AL^2) + \lambda^2(a_0AL^2 + \kappa_2(AL)^2)}. \tag{5.21}$$

We need now to choose A so large that $\kappa A \ge 2\kappa_2 c^* a_0$, and $\lambda = \min(\log(2), \kappa/(2\kappa_2))$, which gives finally that

$$P(\text{the tip is covered}) \le \exp\big(-cA^2L^2\big). \tag{5.22}$$

Appendix A: Proof of Lemma 2.1

Our goal in this section is to estimate precisely the restricted Green's function $z \mapsto G_{\mathcal{D}(\rho)}(0; z)$ for any positive real ρ, and $z \in \mathcal{D}(\rho)$. We use that the latter function is discrete harmonic on $\mathcal{D}(\rho) \setminus \{0\}$, vanishes on the (discrete) boundary of $\mathcal{D}(\rho)$, and satisfies $\Delta G_{\mathcal{D}(\rho)}(0, \cdot)|_0 = -1$.

We first find an explicit function, denoted $h : \mathbb{Z} \times \mathbb{R}$, discrete harmonic on the x-axis, real harmonic on each tooth of $\mathcal{D}(\rho) \setminus \{0\}$, vanishing on

$$\Sigma(\rho) = \left\{(x, y) \in \mathbb{Z} \times \mathbb{R}\colon\ |x| \le \rho, |y| = \frac{1}{3}\big(\rho - |x|\big)^2\right\} \quad \text{and} \quad \Delta h(0) = -1.$$

Since h is linear on each tooth of $\mathcal{D}(\rho)$, and can readily be extended to $\mathcal{D}(\rho) \cup \partial\mathcal{D}(\rho)$ with nonpositive values on $\partial\mathcal{D}(\rho)$, the maximum principle yields

$$\forall z \in \mathcal{D}(\rho), \quad G_{\mathcal{D}(\rho)}(0; z) \ge h(z). \tag{A.1}$$

Similarly we build h^+, positive and harmonic on a larger domain $\Sigma^+(\rho) \setminus \{0\}$ with

$$\Sigma^+(\rho) = \left\{(x, y) \in \mathbb{Z} \times \mathbb{R}\colon\ |x| \le [\rho] + 1, |y| = \frac{1}{3}\big([\rho] + 1 - |x|\big)^2\right\} \quad \text{and} \quad \Delta h^+(0) = -1.$$

We will have that $h^+ - G_{\mathcal{D}(\rho)}$ is harmonic on $\mathcal{D}(\rho)$ and nonnegative on $\partial\mathcal{D}(\rho)$. Again, by the maximum principle

$$\forall z \in \mathcal{D}(\rho), \quad h^+(z) \ge G_{\mathcal{D}(\rho)}(0; z). \tag{A.2}$$

The explicit expression of h and h^+, and estimates (A.1) and (A.2) are the desired results of this section.

Construction of h

By the symmetries of $\mathcal{D}(\rho)$, h is even in the x and y coordinates. Thus, we restrict the construction for $x \in [-\rho, 0]$. Also, it is convenient to shift $\mathcal{D}(\rho)$ by ρ units along the x-axis, so that $(-\rho, 0)$ becomes the origin, and $(0, 0)$ becomes $(0, \rho)$.

On each arm of the comb h is linear, and reads for $z = (x, y)$,

$$h(z) = a(x)y + b(x). \tag{A.3}$$

We set $Y(x)=\frac{1}{3}x^2$, and we impose

$$0=a(x)Y(x)+b(x) \quad \text{and} \quad b(0)=0. \tag{A.4}$$

We solve a set of equations: for $x_k=\rho-[\rho]+k$ with k integer in $\{1,\dots,[\rho]-1\}$,

$$4h(x_k,0)=h(x_k,1)+h(x_k,-1)+h(x_{k+1},0)+h(x_{k-1},0)=2h(x_k,1)+h(x_{k+1},0)+h(x_{k-1},0), \tag{A.5}$$

and a boundary equation

$$4h(\rho,0)=2h(\rho,1)+2h(\rho-1,0)+4. \tag{A.6}$$

When we choose $b(x)=\frac{1}{3}\alpha x^3$, (A.4) implies that $a(x)=-\alpha x$. In terms of a and b, (A.5) and (A.6) read for $k\in\{1,\dots,[\rho]-1\}$

$$2b(x_k)=2a(x_k)+b(x_{k+1})+b(x_{k-1}) \quad \text{and} \quad b(\rho)=a(\rho)+b(\rho-1)+2. \tag{A.7}$$

Solving (A.7), we find

$$\alpha=\frac{2}{\rho^2+1/3}.$$

Thus, we obtain a function $h:[0,\rho]\times\mathbb{R}$ given by

$$h(x,y)=\frac{2x}{\rho^2+1/3}\left(\frac{x^2}{3}-y\right). \tag{A.8}$$

Construction of h^+

Here, the domain $\mathcal{D}([\rho]+1)$ is shifted by $[\rho]+1$ units along the x-axis. We build a function $h^+(x,y)=a^+(x)y+b^+(x)$, with $h^+(0,0)=0$, and $h^+(x,Y^+(x))=0$ for

$$Y^+(x)=\frac{1}{3}x^2+1 \quad \text{and} \quad b^+(x)=\alpha x Y^+(x).$$

This implies that $a^+(x)=-\alpha x$. Now, a^+,b^+ solve (A.5) for x an integer from 1 to $[\rho]+1$. Also (A.6) holds with $[\rho]+1$ instead of ρ. This yields

$$\alpha=\frac{2}{([\rho]+1)^2+1/3+1}.$$

We obtain a function $h^+:[0,[\rho]+1]\times\mathbb{R}$ given by

$$h^+(x,y)=\frac{2x}{([\rho]+1)^2+1/3+1}\left(\frac{x^2}{3}+1-y\right). \tag{A.9}$$

Estimate on Green's function

We go back now to the usual coordinate system to obtain (2.4) with h and h^+ given in (2.2) and (2.3). One more relation is useful to obtain nondegenerate estimates when taking z close to the boundary of $\mathcal{D}(\rho)$. Recall that for $z\neq 0$,

$$\begin{aligned} G_{\mathcal{D}(\rho)}(0;z)&=E_0\left[\sum_{k=1}^{\mathcal{E}(\rho)-1}\mathbb{1}_{S(k)=z}\right]=E_0\left[\sum_{k=1}^{\mathcal{E}(\rho)-1}\mathbb{1}_{S(k-1)=\mathcal{A}(z),S(k)=z}\right]\\ &=G_{\mathcal{D}(\rho)}\big(0;\mathcal{A}(z)\big)\times p\big(\mathcal{A}(z);z\big). \end{aligned} \tag{A.10}$$

Thus, if $z=(x,0)\in\mathcal{D}(\rho)$, and $x+1>\rho$, note that $\mathcal{A}(z)=(x-1,0)$, and $p(\mathcal{A}(z);z)=1/4$ so that

$$G_{\mathcal{D}(\rho)}(0;z)=\frac{1}{4}G_{\mathcal{D}(\rho)}\big(0;\mathcal{A}(z)\big)\quad\text{and}\quad G_{\mathcal{D}(\rho)}(0;z)\ge\frac{h((x-1,0))}{4}. \tag{A.11}$$

Appendix B: Proof of Proposition 2.2

Proposition 2.2 uses the following lemma which we prove at the end of the section.

Lemma B.1. *For any* $\rho>0$,

$$P_0\big(H(1,0)<\mathcal{E}(\rho)\big)\ge\left(\frac{\rho}{\rho+1}\right)^3. \tag{B.1}$$

We assume that the integer y satisfies $0<y<\rho-1$, and denote for $k=0,\dots,y$

$$u(k)=P_{(k,0)}\big(H(y,0)<\mathcal{E}(\rho)\big) \tag{B.2}$$

and $L_\rho(k)$ denotes the height of the tooth of $\mathcal{D}(\rho)$ at site $(k,0)$, i.e. $L_\rho(k)=[\frac{1}{3}(\rho-k)^2]+1$. If we condition the event $\{H(y,0)<\mathcal{E}(\rho)\}$ on the first step of the random walk, then we obtain for $k=1,\dots,y-1$

$$u(k)=\frac{u(k+1)+u(k-1)}{4}+\frac{1}{2}\left(1-\frac{1}{L_\rho(k)}\right)u(k)\quad\text{and}\quad u(y)=1. \tag{B.3}$$

We rewrite (B.3) as

$$u(k)-\frac{u(k-1)}{\alpha(k-1)}=\frac{1}{\alpha(k-1)}\left(u(k+1)-\frac{u(k)}{\alpha(k)}\right), \tag{B.4}$$

and the $\{\alpha(k),k=0,\dots,y-1\}$ is a sequence obtained inductively with the constraint that

$$\forall k=1,\dots,y-1,\quad \alpha(k-1)+\frac{1}{\alpha(k)}=2+\frac{2}{L_\rho(k)}. \tag{B.5}$$

As we iterate (B.5), from $k=1$ to $k=y-1$, we find

$$u(1)-\frac{u(0)}{\alpha(0)}=\frac{1}{\alpha(0)}\times\cdots\times\frac{1}{\alpha(y-2)}\left(u(y)-\frac{u(y-1)}{\alpha(y-1)}\right). \tag{B.6}$$

Now, assume we have the following three relations

$$\text{(i)}\quad u(0)\ge u(1)\left(\frac{\rho}{\rho+1}\right)^3,\qquad \text{(ii)}\quad 1-\frac{u(y-1)}{\alpha(y-1)}\ge\frac{1}{\rho-y-1} \tag{B.7}$$

and

$$\text{(iii)}\quad 1\le\alpha(k)\le 1+\frac{3}{\rho-k-2}\quad\forall k<y.$$

Using (B.6) and (B.7), we obtain for ρ large enough and for a positive constant κ

$$\left(\left(\frac{\rho+1}{\rho}\right)^3-\frac{\rho-2}{\rho+1}\right)u(0)\ge\left(\prod_{k=0}^{y-2}\frac{\rho-k-2}{\rho-k+1}\right)\times\left(1-\frac{u(y-1)}{\alpha(y-1)}\right). \tag{B.8}$$

Now, some simple algebra yields

$$\left(\left(\frac{\rho+1}{\rho}\right)^3-\frac{\rho-2}{\rho+1}\right)\le\frac{6}{\rho}\left(1+\frac{1}{2\rho}+\frac{1}{6\rho^2}\right), \tag{B.9}$$

and

$$\prod_{k=0}^{y-2}\frac{\rho-k-2}{\rho-k+1}\ge\frac{(\rho-y)^3}{\rho^3}. \tag{B.10}$$

We deduce from (B.8), (B.9) and (B.10) that for some constant κ

$$u(0)\ge\kappa\left(\frac{\rho-y}{\rho}\right)^2. \tag{B.11}$$

We are left with showing the estimates of (B.7). Note that (i) is Lemma B.1.

We show (ii). When we start on $(y-1,0)$, one way to escape $\mathcal{D}(\rho)$ before reaching $(y,0)$ is to go up on one tooth and hit the boundary of $\mathcal{D}(\rho)$ before touching $(y-1,0)$. Thus, when $y<\rho-1$

$$1-u(y-1)=P_{(y-1,0)}\big(\mathcal{E}(\rho)<H(y,0)\big)\ge\frac{1}{2L_\rho(y-1)}\ge\frac{3/2}{(\rho-y+1)^2+1}\ge\frac{1}{(\rho-y+1)^2}. \tag{B.12}$$

This is equivalent to

$$1-\frac{u(y-1)}{1+1/(\rho-y+1)}\ge\frac{1}{\rho-y+1}. \tag{B.13}$$

Now, to produce a sequence satisfying (B.5), we choose $\alpha(y-1)$ as follows, and build $\alpha(k)$ by a backward induction:

$$\alpha(y-1)=1+\frac{3}{\rho-y+1}. \tag{B.14}$$

This implies, using (B.13), that

$$1-\frac{u(y-1)}{\alpha(y-1)}\ge1-\frac{u(y-1)}{1+1/(\rho-y+1)}\ge\frac{1}{\rho-y+1}. \tag{B.15}$$

We now show that (iii) is compatible with our choice (B.14). We do it by backward induction. First, it is obvious that $\alpha(k)>1$ implies that $\alpha(k-1)>1$. We assume now that $\alpha(k-1)>1+3/(\rho-k-1)$, and show that $\alpha(k)>1+3/(\rho-k-2)$. This in combination with (B.14) yields (iii). In view of (B.5) this is equivalent to checking that

$$\begin{aligned}
&1\ge\left(1+\frac{2}{L_\rho(k)}-\frac{3}{\rho-k-1}\right)\left(1+\frac{3}{\rho-k-2}\right)\\
&\iff\quad 1\ge1-\frac{9}{(\rho-k-1)(\rho-k-2)}+\left(\frac{3}{\rho-k-2}-\frac{3}{\rho-k-1}\right)+\frac{2}{L_\rho(k)}\left(1+\frac{3}{\rho-k-2}\right)\\
&\iff\quad L_\rho(k)\ge\frac{(\rho-k-1)(\rho-k+1)}{3}=\frac{1}{3}\big((\rho-k)^2-1\big).
\end{aligned} \tag{B.16}$$

The last inequality of (B.16) is true since $L_\rho(k)\ge\frac{1}{3}(\rho-k)^2$.

Proof of Lemma B.1. Calling $A=(1,0)$, we establish first,

$$P_0\big(H(A)<\mathcal{E}(\rho)\big)=\left(1+\frac{1}{L_\rho(0)}+\frac{2}{G_{\mathcal{D}(\rho)}(0;0)}\right)^{-1}. \tag{B.17}$$

For simplicity, we name $B = (0,1)$, $C = (-1,0)$ and $D = (0,-1)$. Then,

$$P_0\big(H(A) < \mathcal{E}(\rho)\big) = P_0\big(S(1) = A\big) + \sum_{z \in \{B,C,D\}} P_0\big(S(1) = z\big) P_z\big(H(0) < \mathcal{E}(\rho)\big) P_0\big(H(A) < \mathcal{E}(\rho)\big)$$
$$= \frac{1}{4} + P_0\big(H(A) < \mathcal{E}(\rho)\big)\left(\frac{P_B(H(0) < \mathcal{E}(\rho))}{2} + \frac{P_C(H(0) < \mathcal{E}(\rho))}{4}\right). \tag{B.18}$$

A classical gambler's ruin estimate yields

$$P_B\big(H(0) < \mathcal{E}(\rho)\big) = \frac{L_\rho(0) - 1}{L_\rho(0)}. \tag{B.19}$$

Also, by decomposing over the first step, we have for $H(0)^+$ the return time to 0,

$$1 - \frac{1}{G_{\mathcal{D}(\rho)}(0;0)} = P_0\big(H(0)^+ < \mathcal{E}(\rho)\big)$$
$$= \frac{1}{2} P_A\big(H(0) < \mathcal{E}(\rho)\big) + \frac{1}{2} P_B\big(H(0) < \mathcal{E}(\rho)\big). \tag{B.20}$$

Thus, using (B.19) and (B.20) in (B.18), we obtain (B.17). Now, by (2.4), we have

$$\frac{2}{G_{\mathcal{D}(\rho)}(0;0)} \le \frac{3}{\rho} + \frac{1}{\rho^3}. \tag{B.21}$$

Recalling that $L_\rho(0) \ge \rho^2/3$, we obtain the desired relation.

$$P_0\big(H(A) < \mathcal{E}(\rho)\big) \ge \left(1 + \frac{3}{\rho} + \frac{3}{\rho^2} + \frac{1}{\rho^3}\right)^{-1} = \left(\frac{\rho}{\rho+1}\right)^3. \tag{B.22}$$

□

Appendix C: Proof of Green's function estimates

C.1. *Proof of Proposition* 2.3

We prove (i). By symmetries of $\mathcal{D}(\rho)$, we can consider $0 < X_w < X_z$ or $X_w = X_z$ and $Y_w, Y_z \ge 0$. Note that the path joining w and z crosses $(X_w, 0)$, as well as the path joining 0 and z. Thus,

$$G_{\mathcal{D}(\rho)}(w;z) = P_w\big(H(X_w,0) < \mathcal{E}(\rho)\big) \times G_{\mathcal{D}(\rho)}\big((X_w,0);z\big) \quad \text{and}$$
$$G_{\mathcal{D}(\rho)}(0;z) = P_0\big(H(X_w,0) < \mathcal{E}(\rho)\big) \times G_{\mathcal{D}(\rho)}\big((X_w,0);z\big). \tag{C.1}$$

This implies that

$$G_{\mathcal{D}(\rho)}(w;z) = P_w\big(H(X_w,0) < \mathcal{E}(\rho)\big) \times \frac{G_{\mathcal{D}(\rho)}(0,z)}{P_0(H(X_w,0) < \mathcal{E}(\rho))}. \tag{C.2}$$

Since $P_w(H(X_w,0) < \mathcal{E}(\rho)) = (L_\rho(w) - Y_w)/L_\rho(w)$, (2.9) follows from Lemma 2.1 and Proposition 2.2.

Note that case (ii) follows from the previous argument by noting first that a reversible measure for the simple random walk assigns to a vertex its degree, and

$$G_{\mathcal{D}(\rho)}(w;z)=\frac{\deg(z)}{\deg(w)}G_{\mathcal{D}(\rho)}(z,w) \quad\Longrightarrow\quad G_{\mathcal{D}(\rho)}(w;z)\le 2G_{\mathcal{D}(\rho)}(z,w). \tag{C.3}$$

Then, we interchange in (C.2) the role of z and w. Note, however that z is at a distance 1 from the boundary of $\mathcal{D}(\rho)$ while w can be anywhere in $\mathcal{D}(\rho)$.

We prove now (iii). Note the two relations

$$G_{\mathcal{D}(\rho)}\big((X_w,0);z\big)=P_{(X_w,0)}\big(H(0)<\mathcal{E}(\rho)\big)\times G_{\mathcal{D}(\rho)}(0;z)$$

and

$$G_{\mathcal{D}(\rho)}\big(0;(X_w,0)\big)\le 2G_{\mathcal{D}(\rho)}\big((X_w,0);0\big)=P_{(X_w,0)}\big(H(0)<\mathcal{E}(\rho)\big)\times G_{\mathcal{D}(\rho)}(0,0). \tag{C.4}$$

Using Lemma 2.1 and Proposition 2.2, (C.4) yields

$$G_{\mathcal{D}(\rho)}\big((X_w,0);z\big)=\frac{G_{\mathcal{D}(\rho)}(0;(X_w,0))}{G_{\mathcal{D}(\rho)}(0;0)}\times G_{\mathcal{D}(\rho)}(0;z)\le\kappa\frac{(\rho-X_w)^3}{3\rho^2+1}\times\frac{3\rho^2+1}{2\rho^3}\times\frac{(\rho-X_z)}{\rho^2}. \tag{C.5}$$

We complete (C.5) with the gambler's ruin estimate to obtain (2.11).

Finally, we deal with (iv). Consider w,z with $X_w=X_z$. Then,

$$G_{\mathcal{D}(\rho)}(w;z)=P_w\big(H(z)<\mathcal{E}(\rho)\big)G_{\mathcal{D}(\rho)}(z;z)\le G_{\mathcal{D}(\rho)}(z;z). \tag{C.6}$$

On the other hand, by decomposing over the first step (and recalling that $z\in\partial_I\mathcal{D}(\rho)$)

$$G_{\mathcal{D}(\rho)}(z;z)=1+\frac{1}{2}G_{\mathcal{D}(\rho)}(z-1,z)\le 1+\frac{1}{2}G_{\mathcal{D}(\rho)}(z;z) \quad\Longrightarrow\quad G_{\mathcal{D}(\rho)}(z;z)\le 2. \tag{C.7}$$

This completes (2.12).

C.2. *Proof of Corollary* 2.4

It is enough to consider $z\in\partial\mathcal{D}^+(\rho)$, and to recall the last passage decomposition (2.8). Introduce now the following notation. For $\Lambda\subset\mathbb{Z}^2$,

$$\Gamma(\Lambda)=\sum_{w\in\Lambda}G^2_{\mathcal{D}(\rho)}\big(w,\mathcal{A}(z)\big), \tag{C.8}$$

and partition $\mathcal{D}(\rho)$ into four regions $\mathcal{D}_1,\dots,\mathcal{D}_4$ with

$$\mathcal{D}_1=\mathcal{D}(\rho)\cap\big(\{0\le x\le X_z\}\cup\{x=X_z,Y_z\cdot Y_w<0\}\big),\qquad \mathcal{D}_2=\mathcal{D}(\rho)\cap\{X_w=x,Y_z\cdot Y_w\ge 0\},$$
$$\mathcal{D}_3=\mathcal{D}(\rho)\cap\{x>X_z\}$$

and $\mathcal{D}_4$ the remaining part of $\mathcal{D}(\rho)$. Using κ_i to denote constants, whose meaning may change from line to line, we obtain using the estimates of Proposition 2.3. We set $n=[\rho]+1$, and

$$\Gamma(\mathcal{D}_1)\le\kappa(\rho+1-X_z)^2\sum_{k=n-X_z}^{n}\frac{1}{k^4}\sum_{i=1}^{k^2}\frac{i^2}{k^4}\le\kappa_1(\rho+1-X_z), \tag{C.9}$$

then,

$$\Gamma(\mathcal{D}_2)\le 2L_n(z)\le\frac{2(\rho+1-X_z)^2}{3}. \tag{C.10}$$

Also, note the lower bound

$$\Gamma(\mathcal{D}_2) \ge \kappa \sum_{i=1}^{[(\rho - X_{\mathcal{A}(z)})^2]} \frac{i^2}{(\rho - X_{\mathcal{A}(z)})^4} \ge \underline{\kappa}(\rho - X_{\mathcal{A}(z)})^2. \tag{C.11}$$

Now,

$$\Gamma(\mathcal{D}_3) \le \kappa \frac{1}{(\rho - X_z)^8} \sum_{k=1}^{n-X_z} k^5 \le \kappa_3 \frac{1}{(\rho + 1 - X_z)^2}. \tag{C.12}$$

Finally

$$\Gamma(\mathcal{D}_4) \le \frac{(\rho + 1 - X_z)^2}{\rho^{10}} \sum_{k=1}^{n} k^6 \sum_{i=1}^{k^2} \frac{i^2}{k^4} \le \kappa_3 \frac{(\rho + 1 - X_z)^2}{\rho}. \tag{C.13}$$

With our estimates, the dominant term in $\Gamma(\mathcal{D}(\rho))$ is $\Gamma(\mathcal{D}_2)$, and this concludes the proof.

C.3. On sums of Bernoulli

Let us recall Lemma 2.3 of [1]. Assume that for random variables W, M, and L we have

$$W + L \ge M, \tag{C.14}$$

and furthermore that L and M are sums of independent Bernoulli variables with $L = Y_1 + \cdots + Y_n$. Three hypotheses played a key role in [1]: $(\mathcal{H}_0)$ W is independent of L,

$$(\mathcal{H}_1) \quad \mu := E[M] - E[L] \ge 0 \quad \text{and} \quad (\mathcal{H}_2) \quad \text{for some } \kappa > 1 \quad \sup_i E[Y_i] < 1 - \frac{1}{\kappa}. \tag{C.15}$$

Then, Lemma 2.3 of [1] establishes that for $0 \le \xi < \mu$, any $\lambda \ge 0$,

$$P(W \le \xi) \le \exp\left(-\lambda(\mu - \xi) + \frac{\lambda^2}{2}\left(\mu + \kappa \sum_i E[Y_i]^2\right)\right). \tag{C.16}$$

In the inner estimate that we treat here, hypothesis $(\mathcal{H}_2)$ does not hold. Rather, we decompose the Bernoulli variables $\{Y_i, 1 \le i \le n\}$ into two subgroups, according to some $\kappa > 1$ as follows:

$$\mathcal{A} = \left\{ i \le n \colon E[Y_i] < 1 - \frac{1}{\kappa} \right\} \quad \text{and} \quad \mathcal{B} = \{1, \dots, n\} \setminus \mathcal{A}. \tag{C.17}$$

We show the following estimate.

Lemma C.1. *For $\{W, M, L\}$ satisfying $(\mathcal{H}_0)$ and $(\mathcal{H}_1)$, and any $0 \le \lambda \le \log(2)$, we have*

$$\text{for } \xi \ge 0, \quad P(W \le \xi) \le \exp\left(-\lambda(\mu - \xi) + \frac{\lambda^2}{2}\left(\mu + \frac{4}{\kappa^2}|\mathcal{B}| + \kappa \sum_{i \in \mathcal{A}} E[Y_i]^2\right)\right). \tag{C.18}$$

Proof. Using Chebychev's inequality with any $\lambda > 0$, and hypothesis $(\mathcal{H}_0)$

$$P(W \le \xi) \le \mathrm{e}^{\lambda\xi} \frac{E[\mathrm{e}^{-\lambda W}]E[\mathrm{e}^{-\lambda L}]}{E[\mathrm{e}^{-\lambda L}]} \le \mathrm{e}^{\lambda\xi} \frac{E[\mathrm{e}^{-\lambda M}]}{E[\mathrm{e}^{-\lambda L}]} \le \mathrm{e}^{-(\mu-\xi)\lambda} \frac{E[\mathrm{e}^{-\lambda(M - E[M])}]}{E[\mathrm{e}^{-\lambda(L - E[L])}]}.$$

We have now to estimate the Laplace transform of Bernoulli variables. The argument follows the proof of Lemma 2.3 of [1], with the following trick. When $i \in \mathcal{B}$, $\tilde{Y}_i = 1 - Y_i$ is again a Bernoulli variable, and

$$Y_i - E[Y_i] = -\big(\tilde{Y}_i - E[\tilde{Y}_i]\big). \tag{C.19}$$

Now, we recall two simple inequalities used in the proof of Lemma 2.3 of [1]: for $0 \le x \le 1$, we have $1 + x \ge \exp(x - x^2)$, whereas for $0 \le x \le 1 - 1/\kappa$, we have $1 - x \ge \exp(-x - \kappa x^2/2)$. Thus, using $\mathrm{e}^\lambda \le 2$ and the notation $f(t) = \mathrm{e}^t - 1 - t$, and $g(t) = (\mathrm{e}^t - 1)^2$, we have for $i \in \mathcal{B}$,

$$\begin{aligned} E\big[\exp\big(-\lambda\big(Y_i - E[Y_i]\big)\big)\big] &= E\big[\exp\big(\lambda\big(\tilde{Y}_i - E[\tilde{Y}_i]\big)\big)\big] = \mathrm{e}^{-E[\tilde{Y}_i]\lambda}\big(1 + E[\tilde{Y}_i]\big(\mathrm{e}^\lambda - 1\big)\big) \\ &\ge \exp\big(f(\lambda)E[\tilde{Y}_i] - g(\lambda)E[\tilde{Y}_i]^2\big). \end{aligned} \tag{C.20}$$

On the other hand, for $i \in \mathcal{A}$,

$$E\big[\exp\big(-\lambda\big(Y_i - E[Y_i]\big)\big)\big] \ge \exp\bigg(f(-\lambda)E[Y_i] - \frac{\kappa}{2}g(-\lambda)E[Y_i]^2\bigg). \tag{C.21}$$

Recall now that if $[\cdot]_+$ stands for the positive part

$$0 \le f(t) \le \frac{t^2}{2}\mathrm{e}^{[t]_+} \quad \text{and} \quad 0 \le g(t) \le t^2\mathrm{e}^{2[t]_+}.$$

Finally, we have (using also in the third line that for $\lambda \ge 0$, we have $f(\lambda) \ge f(-\lambda)$)

$$\begin{aligned} \frac{E[\mathrm{e}^{-\lambda(M-E[M])}]}{E[\mathrm{e}^{-\lambda(L-E[L])}]} &\le \exp\bigg(f(-\lambda)E[M] - f(-\lambda)\sum_{i\in\mathcal{A}} E[Y_i] - f(\lambda)\sum_{i\in\mathcal{B}} E[\tilde{Y}_i] \\ &\qquad + g(\lambda)\sum_{i\in\mathcal{B}} E[\tilde{Y}_i]^2 + \frac{\kappa}{2}g(-\lambda)\sum_{i\in\mathcal{A}} E[Y_i]^2\bigg) \\ &\le \exp\bigg(f(-\lambda)\mu + \big(f(-\lambda) - f(\lambda)\big)\sum_{i\in\mathcal{B}} E[\tilde{Y}_i] \\ &\qquad + g(\lambda)\sum_{i\in\mathcal{B}}\big(1 - E[Y_i]\big)^2 + \frac{\kappa}{2}g(-\lambda)\sum_{i\in\mathcal{A}} E[Y_i]^2\bigg) \\ &\le \exp\bigg(\frac{\lambda^2}{2}\bigg(\mu + \frac{4}{\kappa^2}|\mathcal{B}| + \frac{\kappa}{2}\sum_{i\in\mathcal{A}} E[Y_i]^2\bigg)\bigg). \end{aligned} \tag{C.22}$$

□

Acknowledgements

The authors thank an anonymous referee for careful reading, and suggestions which considerably improved the presentation.

References

[1] A. Asselah and A. Gaudillière. From logarithmic to subdiffusive polynomial fluctuations for internal DLA and related growth models. *Ann. Probab.* **41**(3A) (2013) 1115–1159. MR3098673

[2] A. Asselah and A. Gaudillière. Sub-logarithmic fluctuations for internal DLA. *Ann. Probab.* **41**(3A) (2013) 1160–1179. MR3098674

[3] S. Blachère and S. Brofferio. Internal diffusion limited aggregation on groups having exponential growth. *Probab. Theory Related Fields* **137**(3) (2007) 323–343. MR2278460

[4] P. Diaconis and W. Fulton. A growth model, a game, an algebra, Lagrange inversion, and characteristic classes. *Rend. Semin. Mat. Univ. Politec. Torino* **49**(1) (1991) 95–119. MR1218674

[5] H. Duminil-Copin, C. Lucas, A. Yadin and A. Yehudayoff. Containing internal diffusion limited aggregation. *Electron. Commun. Probab.* **18** (2013) 50. MR3078013

[6] D. Jerison, L. Levine and S. Sheffield. Logarithmic fluctuations for internal DLA. *J. Amer. Math. Soc.* **25**(1) (2012) 271–301. MR2833484

[7] W. Kager and L. Levine. Diamond aggregation. *Math. Proc. Cambridge Philos. Soc.* **149**(2) (2010) 351–372. MR2670220

[8] M. Krishnapour and Y. Peres. Recurrent graphs where two independent random walks collide finitely often. *Electron. Commun. Probab.* **9** (2004) 72–81. MR2081461

[9] W. Huss. Internal diffusion limited aggregation on non-amenable graphs. *Electron. Commun. Probab.* **13** (2008) 272–279. MR2415135

[10] W. Huss and E. Sava. Internal aggregation models on comb lattices. *Electron. J. Probab.* **17** (2012) 30. MR2915666

[11] L. Levine and Y. Peres. Strong spherical asymptotics for rotor-router aggregation and the divisible sandpile. *Potential Anal.* **30**(1) (2009) 1–27. MR2465710

[12] G. Lawler, M. Bramson and D. Griffeath. Internal diffusion limited aggregation. *Ann. Probab.* **20**(4) (1992) 2117–2140. MR1188055

[13] R. Shellef. IDLA on the supercritical percolation cluster. *Electron. J. Probab.* **15** (2010) 723–740. MR2650780