DIAMOND AGGREGATION

WOUTER KAGER AND LIONEL LEVINE

Abstract. Internal diffusion-limited aggregation is a growth model based on random walk in $\mathbb{Z}^d$. We study how the shape of the aggregate depends on the law of the underlying walk, focusing on a family of walks in $\mathbb{Z}^2$ for which the limiting shape is a diamond. Certain of these walks—those with a directional bias toward the origin—have at most logarithmic fluctuations around the limiting shape. This contrasts with the simple random walk, where the limiting shape is a disk and the best known bound on the fluctuations, due to Lawler, is a power law. Our walks enjoy a uniform layering property which simplifies many of the proofs.

1. Introduction and main results

Internal diffusion-limited aggregation (internal DLA) is a growth model proposed by Diaconis and Fulton [DF91]. In the original model on $\mathbb{Z}^d$, particles are released one by one from the origin o and perform simple symmetric discrete-time random walks. Starting from the set $A(1) = \{o\}$, the clusters $A(i+1)$ for $i \geq 1$ are defined recursively by letting the i-th particle walk until it first visits a site not in $A(i)$, then adding this site to the cluster. Lawler, Bramson and Griffeath [LBG92] proved that in any dimension $d \geq 2$, the asymptotic shape of the cluster $A(i)$ is a d-dimensional ball. Lawler [La95] subsequently showed that the fluctuations around a ball of radius r are at most of order $r^{1/3}$ up to logarithmic corrections. Moore and Machta [MM00] found experimentally that the fluctuations appear to be at most logarithmic in r, but there is still no rigorous bound to match their simulations. Other studies of internal DLA include [GQ00, BQR03, BB07, LP09b].

Here we investigate how the shape of an internal DLA cluster depends on the law of the underlying random walk. Perhaps surprisingly, small changes in the law can dramatically affect the limiting shape. Consider the walk in $\mathbb{Z}^2$ with the same law as simple random walk except on the x and y-axes, where steps toward the origin are reflected. For example, from a site $(x, 0)$ on the positive x-axis, the walk steps to $(x+1, 0)$ with probability $1/2$ and to each of $(x, \pm 1)$ with probability $1/4$; see Figure 1. It follows from Theorem 1, below, that when we rescale the resulting internal DLA cluster $A(i)$ to have

Date: November 10, 2009.

2000 *Mathematics Subject Classification.* 60K35.

Key words and phrases. asymptotic shape, growth model, internal diffusion-limited aggregation, uniform harmonic measure.

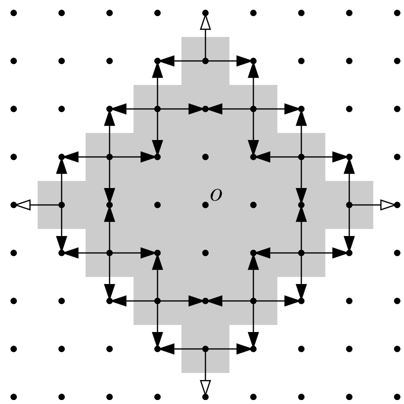

FIGURE 1. Example of a uniformly layered walk. The sites enclosed by the shaded area form the diamond $\mathcal{D}_3$. Only the transition probabilities from layer $\mathcal{L}_3$ are shown. Open-headed arrows indicate transitions that take place with probability 1/2; all the other transitions have probability 1/4.

area 2, its asymptotic shape as $i \to \infty$ is the diamond

$$\mathcal{D} = \{(x, y) \in \mathbb{R}^2 : |x| + |y| \leq 1\}.$$

In fact, a rather large family of walks produce this diamond as their limiting shape. The key property shared by the walks we will consider is that their position at any time t is distributed as a mixture of uniform distributions on diamond layers. To define these walks, for $k \geq 0$ let

$$\mathcal{L}_k := \{x \in \mathbb{Z}^2 : \|x\| = k\}$$

where for $x = (x_1, x_2)$ we write $\|x\| = |x_1| + |x_2|$. A *uniformly layered walk* is a discrete-time Markov chain on state space $\mathbb{Z}^2$ whose transition probabilities $Q(x, y)$ satisfy

(U1) $Q(x, y) = 0$ if $\|y\| > \|x\| + 1$;
(U2) For all $k \geq 0$ and all $x \in \mathcal{L}_k$, there exists $y \in \mathcal{L}_{k+1}$ with $Q(x, y) > 0$;
(U3) For all $k, \ell \geq 0$ and all $y, z \in \mathcal{L}_\ell$,

$$\sum_{x \in \mathcal{L}_k} Q(x, y) = \sum_{x \in \mathcal{L}_k} Q(x, z).$$

In order to state our main results, let us now give a more precise description of the aggregation rules. Set $A(1) = \{o\}$, and let $Y^i(t)$ $(i = 1, 2, \dots)$ be independent uniformly layered walks with the same law, started from the origin. For $i \geq 1$, define the stopping times σ^i and the growing cluster $A(i)$ recursively by setting

$$\sigma^i = \min\{t \geq 0 : Y^i(t) \notin A(i)\}$$

and

$$A(i + 1) = A(i) \cup \{Y^i(\sigma^i)\}.$$

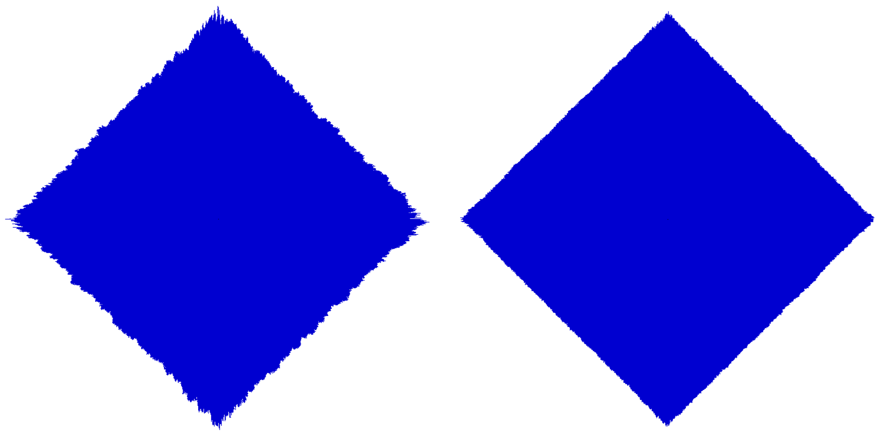

FIGURE 2. Internal DLA clusters in $\mathbb{Z}^2$ based on the uniformly layered walk with transition kernel $p\,Q_{\text{in}} + q\,Q_{\text{out}}$. Left: $p = 0$, walks are directed outward. Right: $p = 1/2$, walks have no directional bias. Each cluster is composed of $v_{350} = 245\,701$ particles.

Now for any real number $r \geq 0$, let

$$\mathcal{D}_r := \left\{x \in \mathbb{Z}^2 : \|x\| \leq r\right\}.$$

We call D_r the diamond of radius r in $\mathbb{Z}^2$. Note that $\mathcal{D}_r = D_{\lfloor r \rfloor}$. For integer $n \geq 0$, we have $\mathcal{D}_n = \bigcup_{k=0}^{n} \mathcal{L}_k$. Since $\#\mathcal{L}_k = 4k$ for $k \geq 1$, the volume of $\mathcal{D}_n$ is

$$v_n := \#\mathcal{D}_n = 2n(n+1) + 1.$$

Our first result says that the internal DLA cluster of v_n sites based on any uniformly layered walk is close to a diamond of radius n.

Theorem 1. *For any uniformly layered walk in $\mathbb{Z}^2$, the internal DLA clusters $A(v_n)$ satisfy*

$$\mathbb{P}\left(\mathcal{D}_{n-4\sqrt{n\log n}} \subset A(v_n) \subset \mathcal{D}_{n+20\sqrt{n\log n}} \text{ eventually}\right) = 1.$$

Here and throughout this paper *eventually* means "for all but finitely many n." Likewise, we will write *i.o.* or *infinitely often* to abbreviate "for infinitely many n."

Our proof of Theorem 1 in Section 5 follows the strategy of Lawler [La95]. The uniform layering property (U3) takes the place of the Green's function estimates used in that paper, and substantially simplifies some of the arguments.

Within the family of uniformly layered walks, we study how the law of the walk affects the fluctuations of the internal DLA cluster around the limiting diamond shape. A natural walk to start with is the outward-directed layered

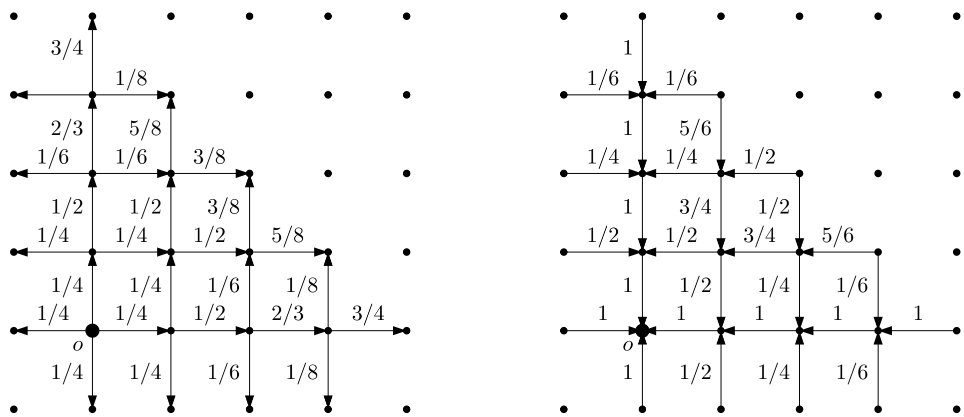

FIGURE 3. Left: transition probabilities of the outward directed kernel Q_{out}. Right: transition probabilities for the inward directed kernel Q_{in}. The origin o is near the lower-left corner.

walk $X(t)$ satisfying

$$\|X(t+1)\| = \|X(t)\| + 1$$

for all t. There is a unique such walk satisfying condition (U3) whose transition probabilities are symmetric with respect to reflection about the axes. It is defined in the first quadrant by

$$Q_{\text{out}}\big((x,y),(x,y+1)\big) = \frac{y+1/2}{x+y+1} \qquad \text{for } x,y = 1,2,\dots, \tag{1.1}$$

$$Q_{\text{out}}\big((x,y),(x+1,y)\big) = \frac{x+1/2}{x+y+1} \qquad \text{for } x,y = 1,2,\dots, \tag{1.2}$$

and on the positive horizontal axis by

$$Q_{\text{out}}\big((x,0),(x,\pm 1)\big) = \frac{1/2}{x+1} \qquad \text{for } x = 1,2,\dots, \tag{1.3}$$

$$Q_{\text{out}}\big((x,0),(x+1,0)\big) = \frac{x}{x+1} \qquad \text{for } x = 1,2,\dots. \tag{1.4}$$

In the other quadrants Q_{out} is defined by reflection symmetry, and at the origin we set $Q_{\text{out}}(o,z) = 1/4$ for all $z \in \mathbb{Z}^2$ with $\|z\| = 1$. See Figure 3.

Likewise one can construct a symmetric Markov kernel defining an inward directed random walk which remains uniformly distributed on diamond layers. This kernel is defined in the first quadrant by

$$Q_{\text{in}}\big((x,y),(x,y-1)\big) = \frac{y-1/2}{x+y-1} \qquad \text{for } x,y = 1,2,\dots, \tag{1.5}$$

$$Q_{\text{in}}\big((x,y),(x-1,y)\big) = \frac{x-1/2}{x+y-1} \qquad \text{for } x,y = 1,2,\dots, \tag{1.6}$$

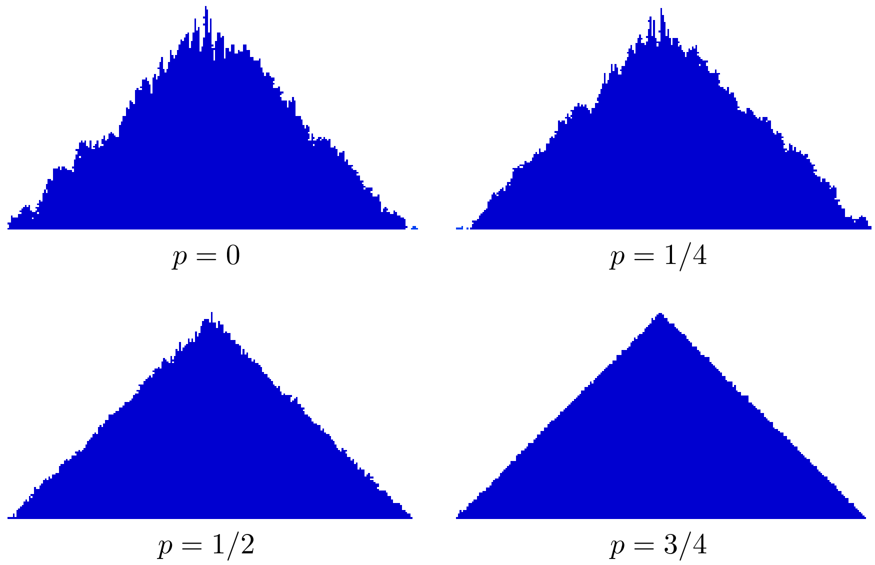

FIGURE 4. Closeups of the boundary of the diamond. Fluctuations decrease as the directional bias of the walk tends from outward ($p = 0$) to inward ($p = 1$).

and on the positive horizontal axis by

$$Q_{\text{in}}\big((x,0),(x-1,0)\big) = 1 \qquad \text{for } x = 1, 2, \ldots. \tag{1.7}$$

Again, the definition extends to the other quadrants by reflection symmetry, and is completed by making the origin an absorbing state: $Q_{\text{in}}(o,o) = 1$. See Figure 3.

We now choose a parameter $p \in [0,1)$, let $q = 1 - p$ and define the kernel $Q_p := p\,Q_{\text{in}} + q\,Q_{\text{out}}$. The parameter p allows us to interpolate between a fully outward directed walk at $p = 0$ and a fully inward directed walk at $p = 1$.

Theorem 1 shows that the fluctuations around the limit shape are at most of order $\sqrt{n \log n}$ for the entire family of walks Q_p. However, one may expect that the true size of the fluctuations depends on p. When p is large, particles tend to take a longer time to leave a diamond of given radius, affording them more opportunity to fill in unoccupied sites near the boundary of the cluster. Indeed, in simulations we find that the boundary becomes less ragged as p increases (Figure 4). Our next result shows that when $p > 1/2$, the boundary fluctuations are at most logarithmic in n.

Theorem 2. *For all $p \in (1/2, 1)$, we have*

$$\mathbb{P}\left(\mathcal{D}_{n-6\log_r n} \subset A(v_n) \subset \mathcal{D}_{n+6\log_r n} \text{ eventually}\right) = 1$$

where the base of the logarithm is $r = p/q$.

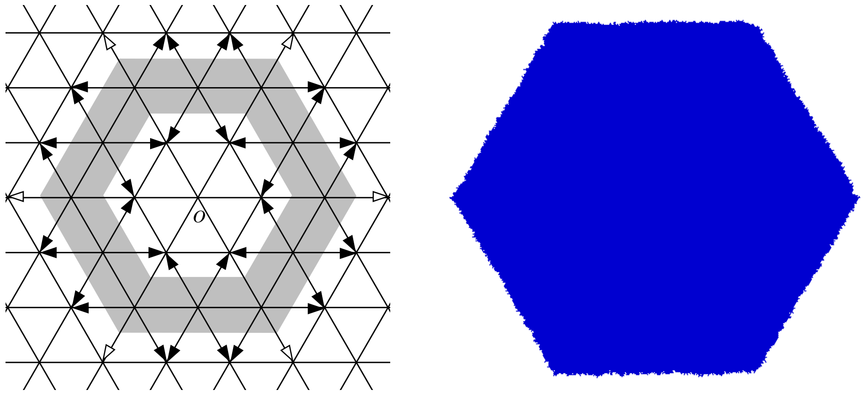

FIGURE 5. Left: Example of a uniformly layered walk on the triangular lattice with hexagonal layers. Only transitions from a single (shaded) layer are shown. Open-headed arrows indicate transitions that take place with probability 1/2; all the other transitions have probability 1/4. Right: An internal DLA cluster of 100 000 particles based on this uniformly layered walk.

We believe that for all $p \in [0, 1/2)$ the boundary fluctuations are of order $\sqrt{n}$ up to logarithmic corrections, and that therefore an abrupt change in the order of the fluctuations takes place at $p = 1/2$. At present, however, we are able to prove a lower bound on the order of fluctuations only in the case $p = 0$:

Theorem 3. *For $p = 0$ we have*

$$\mathbb{P}\left(\mathcal{D}_{n-(1-\varepsilon)\sqrt{2(n\log\log n)/3}} \not\subset A(v_n) \ i.o.\right) = 1 \qquad \forall \varepsilon > 0$$

and

$$\mathbb{P}\left(A(v_n) \not\subset \mathcal{D}_{n+(1-\varepsilon)\sqrt{2(n\log\log n)/3}} \ i.o.\right) = 1 \qquad \forall \varepsilon > 0.$$

Uniformly layered walks are closely related to the walks studied in [Du04, Ka07]. Indeed, the diamond shape of the layers does not play an important role in our arguments. A result similar to Theorem 1 will hold for walks satisfying (U1)–(U3) for other types of layers $\mathcal{L}_k$, provided the cardinality $\#\mathcal{L}_k$ grows at most polynomially in k. Figure 5 shows an example of a walk on the triangular lattice satisfying (U1)–(U3) for hexagonal layers. The resulting internal DLA clusters have the regular hexagon as their asymptotic shape. Blachère and Brofferio [BB07] study internal DLA based on uniformly layered walks for which $\#\mathcal{L}_k$ grows exponentially, such as simple random walk on a regular tree.

Given how sensitive the shape of an internal DLA cluster is to the law of the underlying walk, it is surprising how robust the shape is to other types of changes in the model. For example, the particles may perform deterministic rotor-router walks instead of simple random walks. These walks depend on an initial choice of rotors at each site in $\mathbb{Z}^d$, but for any such choice, the limiting shape is a ball. Another variant is the divisible sandpile model, which replaces the discrete particles by a continuous amount of mass at each lattice site. Its limiting shape is also a ball. These models are discussed in [LP09a].

The remainder of the paper is organized as follows. Section 2 explores the properties of uniformly layered walks, section 3 discusses an "abelian property" of internal DLA which is essential for the proof of Theorem 1, and section 4 collects the limit theorems we will use. Sections 5, 6 and 7 are devoted to the proofs of Theorems 1, 2 and 3, respectively.

2. Uniformly layered walks

Let $\{X(t)\}_{t\geq 0}$ be a uniformly layered walk, that is, a walk on $\mathbb{Z}^2$ satisfying properties (U1)–(U3) of the introduction. Write ν_k for the uniform measure on the sites of layer $\mathcal{L}_k$, and let $\mathbb{P}_k$ denote the law of the walk started from $X(0) \sim \nu_k$. Likewise, let $\mathbb{P}_x$ denote the law of the walk started from $X(0) = x$. Consider the stopping times

$$\begin{aligned} \tau_z &:= \min\{t \geq 0 : X(t) = z\} && \text{for } z \in \mathbb{Z}^2; \\ \tau_k &:= \min\{t \geq 0 : X(t) \in \mathcal{L}_k\} && \text{for } k \geq 0. \end{aligned}$$

The key to the diamond shape, as we shall see, is the fact that our random walks have the uniform distribution on diamond layers at all fixed times, and at the particular stopping times τ_k. The next lemma shows that under $\mathbb{P}_k$, conditionally on $\|X(s)\|$ for $s \leq t$, the distribution of $X(t)$ is uniform on $\mathcal{L}_{\|X(t)\|}$. We remark that the fact that this conditional distribution depends only on $\|X(t)\|$, and not on $\|X(s)\|$ for $s < t$, implies that $\|X(t)\|$ is a Markov chain under $\mathbb{P}_k$; see [RP81].

Lemma 4. *Fix $k \geq 0$. For all $t \geq 0$ and all sequences of nonnegative integers $k = \ell(0), \ldots, \ell(t)$ satisfying $\ell(s+1) \leq \ell(s) + 1$ for $s = 0, \ldots, t-1$, we have for all $z \in \mathcal{L}_{\ell(t)}$*

$$\begin{aligned} \mathbb{P}_k\big(X(t) = z \bigm| \|X(s)\| = \ell(s),\ 0 \leq s \leq t\big) &= \frac{1}{\#\mathcal{L}_{\ell(t)}} \\ &= \mathbb{P}_k\big(X(t) = z \bigm| \|X(t)\| = \ell(t)\big). \end{aligned}$$

Proof. We prove the first equality by induction on t. The base case $t = 0$ is immediate. Write

$$\mathcal{E}_t = \big\{\|X(s)\| = \ell(s),\, 0 \leq s \leq t\big\}.$$

By the Markov property and the inductive hypothesis, we have for $t \geq 1$ and any $y \in \mathcal{L}_{\ell(t)}$

$$\begin{aligned}\mathbb{P}_k(X(t) = y, \mathcal{E}_t) &= \sum_{x \in \mathcal{L}_{\ell(t-1)}} \mathbb{P}_k(X(t) = y, X(t-1) = x, \mathcal{E}_{t-1}) \\ &= \sum_{x \in \mathcal{L}_{\ell(t-1)}} Q(x,y) \cdot \mathbb{P}_k(X(t-1) = x, \mathcal{E}_{t-1}). \\ &= \sum_{x \in \mathcal{L}_{\ell(t-1)}} Q(x,y) \cdot \frac{1}{\#\mathcal{L}_{\ell(t-1)}} \cdot \mathbb{P}_k(\mathcal{E}_{t-1}).\end{aligned}$$

By property (U3), the right side does not depend on the choice of $y \in \mathcal{L}_{\ell(t)}$. It follows that

$$\mathbb{P}_k(X(t) = z \mid \mathcal{E}_t) = \frac{\mathbb{P}_k(X(t) = z, \mathcal{E}_t)}{\sum_{y \in \mathcal{L}_{\ell(t)}} \mathbb{P}_k(X(t) = y, \mathcal{E}_t)} = \frac{1}{\#\mathcal{L}_{\ell(t)}}.$$

By induction this holds for all $t \geq 0$ and all sequences $\ell(0), \ldots, \ell(t)$. Therefore, for fixed $\ell(t)$ and $z \in \mathcal{L}_{\ell(t)}$

$$\begin{aligned}\mathbb{P}_k\big(X(t) = z\big) &= \sum_{\ell(0),\ldots,\ell(t-1)} \mathbb{P}_k\big(X(t) = z, \|X(s)\| = \ell(s)\ \forall s \leq t\big) \\ &= \frac{1}{\#\mathcal{L}_{\ell(t)}} \sum_{\ell(0),\ldots,\ell(t-1)} \mathbb{P}_k\big(\|X(s)\| = \ell(s)\ \forall s \leq t\big) \\ &= \frac{1}{\#\mathcal{L}_{\ell(t)}} \mathbb{P}_k\big(\|X(t)\| = \ell(t)\big)\end{aligned}$$

which implies

$$\mathbb{P}_k\big(X(t) = z \,\big|\, \|X(t)\| = \ell(t)\big) = \frac{1}{\#\mathcal{L}_{\ell(t)}}. \qquad \square$$

As a consequence of Lemma 4, our random walks have the uniform distribution on layer ℓ at the stopping time τ_ℓ.

Lemma 5. *Fix integers $0 \leq k < \ell$. Then*

$$\mathbb{P}_k(X(\tau_\ell) = z) = \frac{1}{\#\mathcal{L}_\ell} = \frac{1}{4\ell} \qquad \textit{for every } z \in \mathcal{L}_\ell.$$

Proof. Note that property (U2) and Lemma 4 imply $\tau_\ell < \infty$ almost surely. For $t \geq 0$ we have

$$\{\tau_\ell = t\} = \bigcup_{\ell_0,\ldots,\ell_t} \big\{\|X(s)\| = \ell_s, 0 \leq s \leq t\big\},$$

where the union is over all sequences of nonnegative integers $\ell_0, \ell_1, \ldots, \ell_t$ with $\ell_0 = k$ and $\ell_t = \ell$, such that $\ell_{s+1} \leq \ell_s + 1$ and $\ell_s \neq \ell$ for all $s =$

$0, 1, \ldots, t-1$. Writing $\mathcal{E}_{\ell_0,\ldots,\ell_t}$ for the disjoint events in this union, it follows that

$$\begin{aligned}\mathbb{P}_k(X(\tau_\ell) = z) &= \sum_{t\geq 0} \mathbb{P}_k(X(t) = z,\, \tau_\ell = t) \\ &= \sum_{t\geq 0} \sum_{\ell_0,\ldots,\ell_t} \mathbb{P}_k(X(t) = z,\, \mathcal{E}_{\ell_0,\ldots,\ell_t}) \\ &= \sum_{t\geq 0} \sum_{\ell_0,\ldots,\ell_t} \mathbb{P}_k(X(t) = z \mid \mathcal{E}_{\ell_0,\ldots,\ell_t}) \mathbb{P}_k(\mathcal{E}_{\ell_0,\ldots,\ell_t}).\end{aligned}$$

Since $\sum_{t\geq 0} \mathbb{P}_k(\tau_\ell = t) = 1$, the result follows from Lemma 4. □

The previous lemmas show that one can view our random walks as walks that move from layer to layer on the lattice, while remaining uniformly distributed on these layers. This idea can be formalized in terms of an intertwining relation between our two-dimensional walks and a one-dimensional walk that describes the transitions between layers, an idea explored in [Du04, Ka07] for closely related random walks in wedges. This approach is particularly useful for computing properties of the Green's function.

Next we calculate some hitting probabilities for the walk with transition kernel $Q_p = p\,Q_{in} + q\,Q_{out}$ defined in the introduction; we will use these in the proof of Theorem 2. We start with the probability of visiting the origin before leaving the diamond of radius n. By the definition of Q_p, this probability depends only on the layer on which the walk is started, not on the particular starting point on that layer. That is, if $0 < \ell < n$, then $\mathbb{P}_x(\tau_o < \tau_n) = \mathbb{P}_\ell(\tau_o < \tau_n)$ for all $x \in \mathcal{L}_\ell$, since at every site except the origin, the probability to move inward is p and the probability to move outward is q. This leads to the following well-known gambler's ruin calculation (see, e.g., [Bi95, §7]).

Lemma 6. *Let $0 < \ell < n$ and $x \in \mathcal{L}_\ell$. If $p \neq q$, then*

$$\mathbb{P}_x(\tau_o < \tau_n) = \mathbb{P}_\ell(\tau_o < \tau_n) = \frac{r^n - r^\ell}{r^n - 1}$$

where $r = p/q$. If $p = q = 1/2$, then

$$\mathbb{P}_x(\tau_o < \tau_n) = \mathbb{P}_\ell(\tau_o < \tau_n) = \frac{n - \ell}{n}.$$

Next we bound the probability that the inward-biased walk ($p > 1/2$) exits the diamond $\mathcal{D}_{n-1}$ before hitting a given site $z \in \mathcal{D}_{n-1}$.

Lemma 7. *Write $r = p/q$. For $p \in (1/2, 1)$, if $z \in \mathcal{L}_k$ for $0 < k < n$, then*

$$\mathbb{P}_o(\tau_z \geq \tau_n) < (4k - 1)r^{k-n}.$$

Proof. Let $T_0 = 0$ and for $i \geq 1$ consider the stopping times

$$\begin{aligned}U_i &= \min\{t > T_{i-1} : X(t) \in \mathcal{L}_k\};\\ T_i &= \min\{t > U_i : X(t) = o\}.\end{aligned}$$

Let $M = \max\{i : U_i < \tau_n\}$. For any integer $m \geq 1$ and any $x_1, \ldots, x_m \in \mathcal{L}_k$, we have by the strong Markov property

$$\begin{aligned}
&\mathbb{P}_o(M = m,\ X(U_1) = x_1, \ldots, X(U_m) = x_m) \\
&\quad = \prod_{i=1}^{m-1} \left[\mathbb{P}_o(X(\tau_k) = x_i)\, \mathbb{P}_{x_i}(\tau_o < \tau_n)\right] \cdot \mathbb{P}_o(X(\tau_k) = x_m)\, \mathbb{P}_{x_m}(\tau_n < \tau_o).
\end{aligned}$$

By Lemma 5, $\mathbb{P}_o(X(\tau_k) = x_i) = 1/4k$ for each $x_i \in \mathcal{L}_k$. Moreover, by Lemma 6 we have for any $x \in \mathcal{L}_k$

$$\mathbb{P}_x(\tau_n < \tau_0) = \frac{r^k - 1}{r^n - 1} < r^{k-n},$$

where we have used the fact that $r = p/q > 1$. Hence

$$\mathbb{P}_o(M = m,\ X(U_i) \neq z\ \forall i \leq m) < r^{k-n} \left(1 - \frac{1}{4k}\right)^m.$$

Since the event $\{\tau_z \geq \tau_n\}$ is contained in the event $\{X(U_i) \neq z\ \forall i \leq M\}$, we conclude that

$$\begin{aligned}
\mathbb{P}_o(\tau_z \geq \tau_n) &= \sum_{m \geq 1} \mathbb{P}_o(M = m,\ \tau_z \geq \tau_n) \\
&\leq \sum_{m \geq 1} \mathbb{P}_o(M = m,\ X(U_i) \neq z\ \forall i \leq m) \\
&< \sum_{m \geq 1} r^{k-n} \left(1 - \frac{1}{4k}\right)^m \\
&= (4k - 1) r^{k-n}.
\end{aligned}$$

□

3. Abelian property

In this section we discuss an important property of internal DLA discovered by Diaconis and Fulton [DF91, Theorem 4.1], which gives some freedom in how the clusters $A(i)$ are constructed. We will use this property in the proof of Theorem 1. It was also used in [La95]. Instead of performing i random walks one at a time in sequence, start with i particles at the origin. At each time step, choose a site occupied by more than one particle, and let one particle take a single random walk step from that site. The abelian property says that regardless of these choices, the final set of i occupied sites has the same distribution as the cluster $A(i)$.

This property is not dependent on the law of the random walk, and in fact holds deterministically in a certain sense. Suppose that at each site $x \in \mathbb{Z}^2$ we place an infinite stack of cards, each labeled by a site in $\mathbb{Z}^2$. A *legal move* consists of choosing a site x which has at least two particles, burning the top card at x, and then moving one particle from x to the site labeled by the card just burned. A finite sequence of legal moves is *complete* if it results in a configuration in which each site has at most one particle.

Lemma 8 (Abelian property)**.** *For any initial configuration of particles on $\mathbb{Z}^2$, if there is a complete finite sequence of legal moves, then any sequence of legal moves is finite, and any complete sequence of legal moves results in the same final configuration.*

In our setting, the cards in the stack at x have i.i.d. labels with distribution $Q(x, \cdot)$. Starting with i particles at the origin, one way to construct a complete sequence of legal moves is to let each particle in turn perform a random walk until reaching an unoccupied site. The resulting set of occupied sites is the internal DLA cluster $A(i)$. By the abelian property, any other complete sequence of legal moves yields the same cluster $A(i)$.

For the proof of Theorem 1, it will be useful to define generalized internal DLA clusters for which not all walks start at the origin. Given a (possibly random) sequence $x_1, x_2, \ldots \in \mathbb{Z}^2$, we define the clusters $A(x_1, \ldots, x_i)$ recursively by setting $A(x_1) = \{x_1\}$, and

$$A(x_1, \ldots, x_{i+1}) = A(x_1, \ldots, x_i) \cup \{Y^i(\sigma^i)\}, \qquad i \geq 1,$$

where the Y^i are independent uniformly layered walks started from $Y^i(0) = x_{i+1}$, and

$$\sigma^i = \min\{t \geq 0 \,:\, Y^i(t) \notin A(x_1, \ldots, x_i)\}.$$

When $x_1 = \cdots = x_i = o$ we recover the usual cluster $A(i)$.

The next lemma gives conditions under which two such generalized clusters can be coupled so that one is contained in the other. Let $x_1, \ldots, x_r$ and $y_1, \ldots, y_s$ be random points in $\mathbb{Z}^2$. For $z \in \mathbb{Z}^2$, let

$$N_z = \#\{i \leq r \,:\, x_i = z\}$$
$$\tilde{N}_z = \#\{j \leq s \,:\, y_j = z\}$$

and consider the event

$$\mathcal{E} = \bigcap_{z \in \mathbb{Z}^2} \{N_z \leq \tilde{N}_z\}.$$

Lemma 9 (Monotonicity)**.** *There exists a random set A' with the same distribution as $A(y_1, \ldots, y_s)$, such that $\mathcal{E} \subset \big\{A(x_1, \ldots, x_r) \subset A'\big\}$.*

The proof follows directly from the abelian property: since the distribution of $A(y_1, \ldots, y_s)$ does not depend on the ordering of the points $y_1, \ldots, y_s$, we can take

$$A' = \begin{cases} A(y'_1, \ldots, y'_s) & \text{on } \mathcal{E} \\ A(y_1, \ldots, y_s) & \text{on } \mathcal{E}^c \end{cases}$$

where $y'_1, \ldots, y'_s$ is a (random) permutation of $y_1, \ldots, y_s$ such that $y'_i = x_i$ for all $i \leq r$.

4. Sums of independent random variables

We collect here a few standard results about sums of independent random variables. First we consider large deviation bounds for sums of independent indicators, which we will use several times in the proofs of Theorems 1 and 2. Let S be a finite sum of independent indicator random variables. We start with simple Chernoff-type bounds based on the inequality

$$\mathbb{P}(S \geq b) \leq e^{-tb}\,\mathbb{E}\left(e^{tS}\right).$$

There are various ways to give an upper bound on the right side when the summands of S are i.i.d. indicators; see for example [AS92, Appendix A]. These bounds extend to the case of independent but not necessarily identically distributed indicators by an application of Jensen's inequality, leading to the following bounds [Ja02, Theorems 1 and 2]:

Lemma 10 (Chernoff bounds)**.** *Let S be a finite sum of independent indicator random variables. For all $b \geq 0$,*

$$\mathbb{P}(S \geq \mathbb{E}S + b) \leq \exp\left(-\frac{1}{2}\frac{b^2}{\mathbb{E}S + b/3}\right),$$
$$\mathbb{P}(S \leq \mathbb{E}S - b) \leq \exp\left(-\frac{1}{2}\frac{b^2}{\mathbb{E}S}\right).$$

Next we consider limit theorems for sums of independent random variables, which we will use in the proof of Theorem 3. For $\{X_n\}_{n\geq 1}$ a sequence of independent random variables satisfying $\mathbb{E}|X_i|^3 < \infty$, we define

$$B_n = \sum_{1\leq i\leq n} \mathrm{Var}(X_i), \tag{4.1}$$

$$L_n = B_n^{-3/2} \sum_{1\leq i\leq n} \mathbb{E}|X_i - \mathbb{E}X_i|^3. \tag{4.2}$$

It is well known that the partial sums

$$S_n = \sum_{1\leq i\leq n} X_i \tag{4.3}$$

satisfy the Central Limit Theorem when $L_n \to 0$; this is a special case of Lyapunov's condition. We are interested in the rate of convergence. Let

$$\Delta_n = \sup_{x\in\mathbb{R}} \left|\mathbb{P}\left(S_n - \mathbb{E}S_n < x\sqrt{B_n}\right) - \Phi(x)\right|, \tag{4.4}$$

where Φ is the standard normal distribution function. Esseen's inequality (see [Es45, Introduction, equation (6)] and [PS00, Chapter I]) gives a bound on Δ_n in terms of L_n. This bound can be used to verify the conditions given by Petrov [Pe66, Theorem 1] (see also [PS00, Chapter I]), under which the partial sums S_n satisfy the Law of the Iterated Logarithm.

Lemma 11 (Esseen's inequality)**.** *Let $X_1, \dots, X_n$ be independent and such that $\mathbb{E}|X_i|^3 < \infty$, and define B_n, L_n, S_n and Δ_n by (4.1)–(4.4). Then*

$$\Delta_n \leq 7.5 \cdot L_n.$$

Lemma 12 (Petrov's theorem)**.** *Let $\{X_i\}_{i\geq 1}$ be a sequence of independent random variables with finite variances, and define B_n, S_n and Δ_n by (4.1), (4.3) and (4.4). If, as $n \to \infty$,*

$$B_n \to \infty, \quad \frac{B_{n+1}}{B_n} \to 1 \quad \text{and} \quad \Delta_n = O\left(\frac{1}{(\log B_n)^{1+\delta}}\right) \text{ for some } \delta > 0,$$

then

$$\mathbb{P}\left(\limsup_{n\to\infty} \frac{S_n - \mathbb{E}S_n}{\sqrt{2B_n \log\log B_n}} = 1\right) = 1.$$

5. Proof of Theorem 1

We control the growth of the cluster $A(i)$ by relating it to two modified growth processes, the *stopped process* $S(i)$ and the *extended process* $E(i)$. In the stopped process, particles stop walking when they hit layer $\mathcal{L}_n$, even if they have not yet found an unoccupied site. More formally, let $S(1) = \{o\}$, and define the stopping times σ_S^i and clusters $S(i)$ for $i \geq 1$ recursively by

$$\sigma_S^i = \min\{t \geq 0 : Y^i(t) \in \mathcal{L}_n \cup S(i)^c\}$$

and

$$S(i+1) = S(i) \cup \{Y^i(\sigma_S^i)\}. \tag{5.1}$$

Here $Y^i(t)$ for $i = 1, 2, \dots$ are independent uniformly layered walks started from the origin in $\mathbb{Z}^2$, all having the same law. Note that $S(i+1) = S(i)$ on the event that the walk Y^i hits layer $\mathcal{L}_n$ before exiting the cluster $S(i)$. By the abelian property, Lemma 8, we have $S(i) \subset A(i)$. Indeed, $A(i)$ can be obtained from $S(i)$ by letting all but one of the particles stopped at each site in $\mathcal{L}_n$ continue walking until reaching an unoccupied site.

The extended process $E(i)$ is defined by starting with every site in the diamond $\mathcal{D}_n$ occupied, and letting each of i additional particles in turn walk from the origin until reaching an unoccupied site. More formally, let $E(0) = \mathcal{D}_n$, and for $i \geq 0$ define

$$\sigma_E^i = \min\{t \geq 0 : Y^{v_n+i}(t) \notin E(i)\}$$

and

$$E(i+1) = E(i) \cup \{Y^{v_n+i}(\sigma_E^i)\}.$$

An outline of the proof of Theorem 1 runs as follows. We first show in Lemma 13 that the stopped cluster $S(v_n)$ contains a large diamond with high probability. Since the stopped cluster is contained in $A(v_n)$, the inner bound of Theorem 1 follows. The proof of the outer bound proceeds in three steps. Lemma 14 shows that the particles that stop in layer $\mathcal{L}_n$ during the stopped process cannot be too bunched up at any single site $z \in \mathcal{L}_n$. We then use this to argue in Lemma 15 that with high probability, $A(v_n)$ is contained in a suitable extended cluster $E(m)$. Finally, Lemma 16 shows that this extended cluster is contained in a slightly larger diamond.

A notable feature of the argument (also present in [La95]) is that the proof of the outer bound relies on the inner bound: Lemma 13 is used in the proof of Lemma 14. This dependence is responsible for the larger constant in the outer bound of Theorem 1. It would be interesting to know whether this asymmetry is merely an artifact of the proof, or whether the outer fluctuations are really larger than the inner fluctuations.

We introduce an auxiliary collection of walks that will appear in the proofs. Let $\{Y^x(t) : x \in \mathbb{Z}^2\}$ be independent walks with the same transition probabilities as Y^1, which are independent of the Y^i, and which start from $Y^x(0) = x$. Now for $i = 1, \dots, v_n - 1$ define

$$X^i(t) = \begin{cases} Y^i(t) & \text{for } 0 \le t \le \sigma_S^i, \\ Y^{Y^i(\sigma_S^i)}(t - \sigma_S^i) & \text{for } t > \sigma_S^i. \end{cases}$$

Note that replacing the walks Y^i with X^i in (5.1) has no effect on the clusters $S(i)$. Finally, for $i \ge v_n$ we set $X^i(t) = Y^i(t)$ for all $t \ge 0$.

We associate the following stopping times with the auxiliary walks $Y^x(t)$:

$$\begin{aligned} \tau_z^x &:= \min\{t \ge 0 : Y^x(t) = z\} && \text{for } z \in \mathbb{Z}^2; \\ \tau_k^x &:= \min\{t \ge 0 : Y^x(t) \in \mathcal{L}_k\} && \text{for } k \ge 0. \end{aligned}$$

Likewise, let

$$\begin{aligned} \tau_z^i &:= \min\{t \ge 0 : X^i(t) = z\} && \text{for } z \in \mathbb{Z}^2; \\ \tau_k^i &:= \min\{t \ge 0 : X^i(t) \in \mathcal{L}_k\} && \text{for } k \ge 0. \end{aligned}$$

Lemma 13. *There exists n_0 such that for all uniformly layered walks and all $n \ge n_0$*

$$\mathbb{P}\left(\mathcal{D}_{n-4\sqrt{n \log n}} \not\subset S(v_n)\right) < 6n^{-2}. \tag{5.2}$$

Remark. To avoid referring to too many unimportant constants, for the rest of this section we will take the phrase "for sufficiently large n," and its variants, to mean that a single bound on n applies to all uniformly layered walks.

Proof. For $z \in \mathcal{D}_{n-1}$, write

$$\mathcal{E}_z(v_n) = \bigcap_{i=1}^{v_n - 1} \{\sigma_S^i < \tau_z^i\}$$

for the event that the site z does not belong to the stopped cluster $S(v_n)$. We want to show that $\mathbb{P}(\mathcal{E}_z(v_n))$ is very small when z is taken too deep inside $\mathcal{D}_n$. To this end, let $\ell = \|z\|$, and consider the random variables

$$N_z = \sum\nolimits_{0<i<v_n} \mathbb{1}\{\tau_z^i \le \sigma_S^i\},$$
$$M_z = \sum\nolimits_{0<i<v_n} \mathbb{1}\{\tau_z^i = \tau_\ell^i\},$$
$$L_z = \sum\nolimits_{0<i<v_n} \mathbb{1}\{\sigma_S^i < \tau_z^i = \tau_\ell^i\}.$$

Then $\mathcal{E}_z(v_n) = \{N_z = 0\}$. Since $N_z \ge M_z - L_z$, we have for any real number a

$$\begin{aligned}\mathbb{P}\big(\mathcal{E}_z(v_n)\big) = \mathbb{P}(N_z = 0) &\le \mathbb{P}(M_z \le a \text{ or } L_z \ge a)\\ &\le \mathbb{P}(M_z \le a) + \mathbb{P}(L_z \ge a).\end{aligned} \tag{5.3}$$

Our choice of a will be made below. Note that M_z is a sum of i.i.d. indicator random variables, and by Lemma 5,

$$\mathbb{E}M_z = 2n(n+1)\,\mathbb{P}_o(X(\tau_\ell) = z) = \frac{1}{2}\,\frac{n(n+1)}{\ell}. \tag{5.4}$$

The summands of L_z are not independent. Following [LBG92], however, we can dominate L_z by a sum of independent indicators as follows. By property (U1), a uniformly layered walk cannot exit the diamond $\mathcal{D}_{\ell-1}$ without passing through layer $\mathcal{L}_\ell$, so the event $\{\sigma_S^i < \tau_z^i = \tau_\ell^i\}$ is contained in the event $\{X^i(\sigma_S^i) \in \mathcal{D}_{\ell-1}\}$. Hence

$$\begin{aligned}L_z &= \sum_{0<i<v_n} \mathbb{1}\left\{X^i(\sigma_S^i) \in \mathcal{D}_{\ell-1},\ \tau_z^{X^i(\sigma_S^i)} = \tau_\ell^{X^i(\sigma_S^i)}\right\}\\ &\le \sum_{x\in\mathcal{D}_{\ell-1}-\{o\}} \mathbb{1}\{\tau_z^x = \tau_\ell^x\} =: \tilde{L}_z\end{aligned}$$

where we have used the fact that the locations $X^i(\sigma_S^i)$ inside $D_{\ell-1}$ where particles attach to the cluster are distinct. Note that $\tilde{L}_z$ is a sum of independent indicator random variables. To compute its expectation, note that for every $0 < k < \ell$, by Lemma 5

$$\sum_{x\in\mathcal{L}_k} \mathbb{P}_x(X(\tau_\ell) = z) = 4k\,\mathbb{P}_k(X(\tau_\ell) = z) = \frac{k}{\ell},$$

hence

$$\mathbb{E}\tilde{L}_z = \sum_{k=1}^{\ell-1} \frac{k}{\ell} = \frac{\ell-1}{2}. \tag{5.5}$$

Now set $a = \frac{1}{2}(\mathbb{E}M_z + \mathbb{E}\tilde{L}_z)$, and let

$$b = \frac{\mathbb{E}M_z - \mathbb{E}\tilde{L}_z}{2} > \frac{n^2-\ell^2}{4\ell}$$

where the inequality follows from (5.4) and (5.5). Since $a = \mathbb{E}M_z - b = \mathbb{E}\tilde{L}_z + b$, we have by Lemma 10

$$\mathbb{P}(\tilde{L}_z \geq a) \leq \exp\left(-\frac{1}{2}\frac{b^2}{\mathbb{E}\tilde{L}_z + b/3}\right) \leq \exp\left(-\frac{1}{2}\frac{b^2}{\mathbb{E}M_z}\right)$$

and

$$\begin{aligned}\mathbb{P}(M_z \leq a) &\leq \exp\left(-\frac{1}{2}\frac{b^2}{\mathbb{E}M_z}\right)\\ &< \exp\left(-\frac{1}{2}\frac{(n^2-\ell^2)^2}{16\ell^2}\frac{2\ell}{n(n+1)}\right)\\ &\leq \exp\left(-\frac{1}{16}\frac{(n^2-\ell^2)^2}{n^3}\right)\end{aligned}$$

where in the last line we have used $\ell \leq n-1$. Since $L_z \leq \tilde{L}_z$, we obtain from (5.3)

$$\begin{aligned}\mathbb{P}\big(\mathcal{E}_z(v_n)\big) &\leq \mathbb{P}(M_z \leq a) + \mathbb{P}(\tilde{L}_z \geq a)\\ &< 2\exp\left(-\frac{1}{16}\frac{(n^2-\ell^2)^2}{n^3}\right).\end{aligned}$$

Writing $\ell = n - \rho$, with $\rho \geq \lceil 4\sqrt{n\log n}\,\rceil$, we obtain for sufficiently large n

$$\begin{aligned}\mathbb{P}\big(\mathcal{E}_z(v_n)\big) &< 2\exp\left(-\frac{1}{16}\frac{\rho^2(2n-\rho)^2}{n^3}\right)\\ &\leq 2\exp\left(-\frac{\rho^2}{4n} + \frac{\rho^3}{4n^2}\right)\\ &\leq 3n^{-4}.\end{aligned}$$

We conclude that for n sufficiently large

$$\mathbb{P}\left(\mathcal{D}_{n-4\sqrt{n\log n}} \not\subset S(v_n)\right) \leq \sum_{z\in\mathcal{D}_{n-4\sqrt{n\log n}}} \mathbb{P}\big(\mathcal{E}_z(v_n)\big) < 6n^{-2}. \qquad \square$$

Turning to the outer bound of Theorem 1, the first step is to bound the number

$$N_z := \sum\nolimits_{0<i<v_n} \mathbb{1}\{\sigma_S^i = \tau_z^i\} \tag{5.6}$$

of particles stopping at each site $z \in \mathcal{L}_n$ in the course of the stopped process. To get a rough idea of the order of N_z, note that according to Lemma 13, with high probability, at least $v_{n-4\sqrt{n\log n}}$ of the v_n particles find an occupied site before hitting layer $\mathcal{L}_n$. The number of particles remaining is of order $n^{3/2}\sqrt{\log n}$. If these remaining particles were spread evenly over $\mathcal{L}_n$, then there would be order $\sqrt{n\log n}$ particles at each site $z \in \mathcal{L}_n$. The following lemma shows that with high probability, all of the N_z are at most of this order.

Lemma 14. *If n is sufficiently large, then*

$$\mathbb{P}\left(\bigcup_{z\in\mathcal{L}_n}\left\{N_z > 7\sqrt{n\log n}\right\}\right) < 13n^{-5/4}.$$

Proof. For $z \in \mathcal{L}_n$, define

$$M_z = \sum\nolimits_{0<i<v_n} \mathbb{1}\{\tau_z^i = \tau_n^i\},$$
$$L_z = \sum\nolimits_{0<i<v_n} \mathbb{1}\{\sigma_S^i < \tau_z^i = \tau_n^i\},$$

so that $N_z = M_z - L_z$. Write $\eta = \sqrt{n\log n}$ and $\rho = \lceil 4\eta \rceil$, and let

$$\tilde{L}_z = \sum_{y\in\mathcal{D}_{n-\rho}-\{o\}} \mathbb{1}\{\tau_z^y = \tau_n^y\}.$$

Note that $\tilde{L}_z \le L_z$ on the event $\{\mathcal{D}_{n-\rho} \subset S(v_n)\}$. Therefore,

$$\mathbb{P}\left(\bigcup_{z\in\mathcal{L}_n}\{N_z > 7\eta\}\right) = \mathbb{P}\left(\bigcup_{z\in\mathcal{L}_n}\{M_z - L_z > 7\eta\}\right)$$
$$\le \sum_{z\in\mathcal{L}_n}\mathbb{P}(M_z - \tilde{L}_z > 7\eta) + \mathbb{P}\left(\mathcal{D}_{n-4\sqrt{n\log n}} \not\subset S(v_n)\right). \quad (5.7)$$

To obtain a bound on $\mathbb{P}(M_z - \tilde{L}_z > 7\eta)$, note that

$$\mathbb{E}M_z = 2n(n+1)\,\mathbb{P}_o(X(\tau_n) = z) = \frac{n+1}{2}.$$

Moreover, by Lemma 5

$$\sum_{y\in\mathcal{L}_k}\mathbb{P}(\tau_z^y = \tau_n^y) = 4k\,\mathbb{P}_k(\tau_z = \tau_n) = \frac{k}{n},$$

hence

$$\mathbb{E}\tilde{L}_z = \sum_{k=1}^{n-\rho}\frac{k}{n} = \frac{n+1}{2} - \rho + \frac{\rho(\rho-1)}{2n}.$$

In particular, $\mathbb{E}M_z - \mathbb{E}\tilde{L}_z < \rho - 1 \le 4\eta$ for large enough n, so that

$$\begin{aligned}\mathbb{P}(M_z - \tilde{L}_z > 7\eta) &\le \mathbb{P}(M_z - \tilde{L}_z > \mathbb{E}M_z - \mathbb{E}\tilde{L}_z + 3\eta)\\ &\le \mathbb{P}\big(M_z > \mathbb{E}M_z + \tfrac{3}{2}\eta \quad \text{or} \quad \tilde{L}_z < \mathbb{E}\tilde{L}_z - \tfrac{3}{2}\eta\big) \qquad (5.8)\\ &\le \mathbb{P}\big(M_z > \mathbb{E}M_z + \tfrac{3}{2}\eta\big) + \mathbb{P}\big(\tilde{L}_z < \mathbb{E}\tilde{L}_z - \tfrac{3}{2}\eta\big).\end{aligned}$$

By Lemma 10,

$$\mathbb{P}\big(\tilde{L}_z < \mathbb{E}\tilde{L}_z - \tfrac{3}{2}\eta\big) \le \exp\left(-\frac{1}{2}\frac{(3\eta/2)^2}{\mathbb{E}\tilde{L}_z}\right) < \exp\left(-\frac{9}{8}\frac{\eta^2}{n/2}\right) = n^{-9/4}.$$

Likewise, for sufficiently large n

$$\mathbb{P}\big(M_z > \mathbb{E}M_z + \tfrac{3}{2}\eta\big) \le \exp\left(-\frac{1}{2}\frac{(3\eta/2)^2}{\mathbb{E}M_z + \eta/2}\right)$$
$$= \exp\left(-\frac{9}{4}\frac{n\log n}{n+1+\sqrt{n\log n}}\right)$$
$$< 2n^{-9/4}.$$

Combining (5.7), (5.8) and Lemma 13 yields for sufficiently large n

$$\mathbb{P}\left(\bigcup_{z\in\mathcal{L}_n}\{N_z > 7\eta\}\right) < 3n^{-9/4}\#\mathcal{L}_n + 6n^{-2} < 13n^{-5/4}. \qquad \square$$

Given random sets $A, B \subset \mathbb{Z}^2$, we write $A \stackrel{d}{=} B$ to mean that A and B have the same distribution.

Lemma 15. *Let $m = \lceil 29n\sqrt{n\log n}\rceil$. For all sufficiently large n, there exist random sets $A' \stackrel{d}{=} A(v_n)$ and $E' \stackrel{d}{=} E(m)$ such that*

$$\mathbb{P}\big(A' \not\subset E'\big) < 14n^{-5/4}.$$

Proof. By the abelian property, Lemma 8, we can obtain $A(v_n)$ from the stopped cluster $S(v_n)$ by starting N_z particles at each $z \in \mathcal{L}_n$, and letting all but one of them walk until finding an unoccupied site. More formally, let $x_1 = o$ and $x_{i+1} = Y^i(\sigma_S^i)$ for $0 < i < v_n$. Then

$$\#\{i \le v_n : x_i = z\} = \begin{cases} N_z, & z \in \mathcal{L}_n \\ 1, & z \in S(v_n) - \mathcal{L}_n \\ 0, & \text{else} \end{cases}$$

and

$$A(v_n) \stackrel{d}{=} A(x_1, \dots, x_{v_n}).$$

To build up the extended cluster $E(m)$ in a similar fashion, let $s = v_n + m$, and let $y_1, \dots, y_s \in \mathbb{Z}^2$ be such that $\{y_1, \dots, y_{v_n}\} = \mathcal{D}_n$, and

$$y_{v_n+i} = Y^{v_n+i-1}(\tau_n^{v_n+i-1}), \qquad i = 1, 2, \dots, m.$$

By Lemma 8, we have

$$E(m) \stackrel{d}{=} A(y_1, \dots, y_s).$$

For each $z \in \mathcal{L}_n$, let

$$\tilde{N}_z = \sum\nolimits_{0\le i<m} \mathbb{1}\left\{\tau_z^{v_n+i} = \tau_n^{v_n+i}\right\}$$

be the number of extended particles that first hit layer $\mathcal{L}_n$ at z. Then

$$\#\{i \le s : y_i = z\} = \begin{cases} \tilde{N}_z, & z \in \mathcal{L}_n \\ 1, & z \in \mathcal{D}_{n-1} \\ 0, & \text{else.} \end{cases}$$

Now let $A' = A(x_1, \dots, x_{v_n})$ and consider the event

$$\mathcal{E} = \bigcap_{z \in \mathcal{L}_n} \{N_z \le \tilde{N}_z\}.$$

By Lemma 9, on the event $\mathcal{E}$ there exists a random set $E' \overset{d}{=} A(y_1, \dots, y_s)$ such that $A' \subset E'$. Therefore, to finish the proof it suffices to show that $\mathbb{P}(\mathcal{E}^c) < 14n^{-5/4}$. Note that $\tilde{N}_z$ is a sum of independent indicators, and

$$\mathbb{E}\tilde{N}_z = \frac{m}{4n} \ge \frac{29}{4}\eta$$

where $\eta := \sqrt{n \log n}$. Setting $b = \eta/4$ in Lemma 10 yields for sufficiently large n

$$\mathbb{P}\left(\tilde{N}_z \le 7\eta\right) \le \exp\left(-\frac{1}{2}\frac{b^2}{\mathbb{E}\tilde{N}_z}\right) = \exp\left(-\frac{1}{232}\sqrt{n \log n}\right) < \frac{1}{4}n^{-9/4},$$

hence by Lemma 14

$$\begin{aligned}\mathbb{P}(\mathcal{E}^c) &\le \mathbb{P}\Bigg(\bigcup_{z \in \mathcal{L}_n} \{N_z > 7\eta \quad \text{or} \quad \tilde{N}_z \le 7\eta\}\Bigg) \\ &\le \mathbb{P}\Bigg(\bigcup_{z \in \mathcal{L}_n} \{N_z > 7\eta\}\Bigg) + \sum_{z \in \mathcal{L}_n} \mathbb{P}\big(\tilde{N}_z \le 7\eta\big) \\ &< 14n^{-5/4}.\end{aligned}$$

□

To finish the argument it remains to show that with high probability, the extended cluster $E(m)$ is contained in a slightly larger diamond. Here we follow the strategy used in the proof of the outer bound in [LBG92].

Lemma 16. *Let* $m = \lceil 29n\sqrt{n \log n} \rceil$. *For all sufficiently large* n,

$$\mathbb{P}\left(E(m) \not\subset \mathcal{D}_{n+20\sqrt{n \log n}}\right) < n^{-2}.$$

Proof. For $j, k \ge 1$, let

$$Z_k(j) = \#\big(E(j) \cap \mathcal{L}_{n+k}\big)$$

and let $\mu_k(j) = \mathbb{E}Z_k(j)$. Then $\mu_k(j)$ is the expected number of particles that have attached to the cluster in layer $\mathcal{L}_{n+k}$ after the first j extended particles have aggregated. Note that

$$\mu_k(i+1) - \mu_k(i) = \mathbb{P}\big(Y^{v_n+i+1}(\sigma_E^{i+1}) \in \mathcal{L}_{n+k}\big).$$

By property (U1), in order for the $(i+1)^{\text{th}}$ extended particle to attach to the cluster in layer $\mathcal{L}_{n+k}$, it must be inside the cluster $E(i)$ when it first reaches layer $\mathcal{L}_{n+k-1}$. Therefore, by Lemma 5,

$$\begin{aligned}\mu_k(i+1) - \mu_k(i) &\le \mathbb{P}\big(Y^{v_n+i+1}(\tau^{v_n+i+1}_{n+k-1}) \in E(i)\big) \\ &= \sum_{y\in\mathcal{L}_{n+k-1}} \mathbb{P}_o\big(Y^{v_n+i+1}(\tau^{v_n+i+1}_{n+k-1}) = y\big)\cdot \mathbb{P}\big(y \in E(i)\big) \\ &= \frac{1}{4(n+k-1)}\cdot \mu_{k-1}(i) \le \frac{\mu_{k-1}(i)}{4n}.\end{aligned}$$

Since $\mu_k(0) = 0$, summing over i yields

$$\mu_k(j) \le \frac{1}{4n}\sum_{i=1}^{j-1} \mu_{k-1}(i).$$

Since $\mu_1(j) \le j$ and $\sum_{i=1}^{j-1} i^{k-1} \le j^k/k$, we obtain by induction on k

$$\mu_k(j) \le 4n\left(\frac{j}{4n}\right)^k \frac{1}{k!} \le 4n\left(\frac{je}{4nk}\right)^k,$$

where in the last equality we have used the fact that $k! \ge k^k e^{-k}$. Since $29e/80 < 1$, setting $j = m$ and $k = \lfloor 20\sqrt{n\log n}\rfloor$ we obtain

$$\mu_{k+1}(m) \le 4n\left(\frac{\lceil 29n\sqrt{n\log n}\rceil\, e}{4n\cdot 20\sqrt{n\log n}}\right)^{k+1} < n^{-2}$$

for sufficiently large n. To complete the proof, note that

$$\mathbb{P}(E(m) \not\subset \mathcal{D}_{n+k}) = \mathbb{P}(Z_{k+1}(m) \ge 1) \le \mu_{k+1}(m). \qquad \square$$

Proof of Theorem 1. Write $\eta = \sqrt{n\log n}$. Since $S(v_n) \subset A(v_n)$, we have by Lemma 13

$$\sum_{n\ge1} \mathbb{P}\big(\mathcal{D}_{n-4\eta} \not\subset A(v_n)\big) \le \sum_{n\ge1} \mathbb{P}\big(\mathcal{D}_{n-4\eta} \not\subset S(v_n)\big) < \infty.$$

Likewise, by Lemmas 15 and 16

$$\begin{aligned}\sum_{n\ge1} \mathbb{P}\big(A(v_n) \not\subset \mathcal{D}_{n+20\eta}\big) &\le \sum_{n\ge1} \mathbb{P}\big(A(v_n) \not\subset E(m)\big) + \sum_{n\ge1} \mathbb{P}\big(E(m) \not\subset \mathcal{D}_{n+20\eta}\big) \\ &< \infty.\end{aligned}$$

By Borel-Cantelli we obtain Theorem 1. $\square$

6. The inward directed case

Proof of Theorem 2. Write $\ell = n - \lceil 6\log_r n\rceil$, and denote by

$$\mathcal{A}_n = \bigcap_{0<i<v_n} \bigcap_{z\in\mathcal{D}_\ell} \{\tau^i_z < \tau^i_n\}$$

the event that each of the first $v_n - 1$ walks visits every site $z \in \mathcal{D}_\ell$ before hitting layer $\mathcal{L}_n$. Since $\#\mathcal{D}_{n-1} < v_n$, at least one of the first $v_n - 1$ particles must exit $\mathcal{D}_{n-1}$ before aggregating to the cluster: $\sigma^i \ge \tau^i_n$ for some $i < v_n$.

On the event $\mathcal{A}_n$, this particle visits every site $z \in \mathcal{D}_\ell$ before aggregating to the cluster, so $\mathcal{D}_\ell \subset A(i) \subset A(v_n)$. Hence

$$\mathbb{P}\big(\mathcal{D}_\ell \not\subset A(v_n)\big) \le \mathbb{P}(\mathcal{A}_n^c) \le \sum_{0<i<v_n} \sum_{z\in\mathcal{D}_\ell} \mathbb{P}(\tau_z^i \ge \tau_n^i)$$

By Lemma 7,

$$\begin{aligned}\mathbb{P}\big(\mathcal{D}_\ell \not\subset A(v_n)\big) &< 2n(n+1)\sum_{k=1}^{\ell} 4k(4k-1)r^{k-n} \\ &\le 32n^3(n+1)\frac{r^{\ell+1}-r}{r^n(r-1)} \\ &\le \frac{32r}{r-1}n^3(n+1)\cdot n^{-6},\end{aligned}$$

and by Borel-Cantelli we conclude that $\mathbb{P}(\mathcal{D}_\ell \subset A(v_n) \text{ eventually}) = 1$.

Likewise, writing $m = n + \lfloor 6\log_r n \rfloor$, let

$$\mathcal{B}_n = \bigcap_{0<i<v_n} \bigcap_{z\in\mathcal{D}_n} \{\tau_z^i < \tau_m^i\}$$

be the event that each of the first $v_n - 1$ walks visits every site $z \in \mathcal{D}_n$ before hitting layer $\mathcal{L}_m$. Since the occupied cluster $A(v_n-1)$ has cardinality $v_n - 1 = \#\mathcal{D}_n - 1$, there is at least one site $z \in \mathcal{D}_n$ belonging to $A(v_n-1)^c$. On the event $\mathcal{B}_n$, each of the first $v_n - 1$ particles visits z before hitting layer $\mathcal{L}_m$, so

$$\sigma^i \le \tau_z^i < \tau_m^i, \qquad i = 1, \dots, v_n - 1.$$

Therefore,

$$\mathbb{P}\big(A(v_n) \not\subset \mathcal{D}_m\big) \le \mathbb{P}(\mathcal{B}_n^c) \le \sum_{0<i<v_n} \sum_{z\in\mathcal{D}_n} \mathbb{P}(\tau_z^i \ge \tau_m^i).$$

By Lemma 7,

$$\begin{aligned}\mathbb{P}\big(A(v_n) \not\subset \mathcal{D}_m\big) &< 2n(n+1)\sum_{k=1}^{n} 4k(4k-1)r^{k-m} \\ &\le 32n^3(n+1)\frac{r^{n+1}-r}{r^m(r-1)} \\ &\le \frac{32r^2}{r-1}n^3(n+1)\cdot n^{-6},\end{aligned}$$

and by Borel-Cantelli we conclude that $\mathbb{P}(A(v_n) \subset \mathcal{D}_m \text{ eventually}) = 1$. □

7. The outward directed case

To prove Theorem 3 we make use of a specific property of the uniformly layered walks for $p = 0$. Recall that these walks have transition kernel Q_{out}. By (1.1)–(1.4), such a walk can only reach the site $(m, 0)$ for $m \ge 1$ by visiting the sites $(0,0), (1,0), \dots, (m,0)$ in turn. We can use this to find the exact growth rate of the clusters $A(i)$ along the x-axis.

Suppose that we count time according to the number of particles we have added to the growing cluster, and for $m \geq 1$ set

$$T_m := \min\{n \geq 0 : (m, 0) \in A(n+1)\}.$$

Then we can interpret T_m as the time it takes before the site $(m, 0)$ becomes occupied. The following lemma gives the exact order of the fluctuations in T_m as $m \to \infty$.

Lemma 17. *For $p = 0$ we have that*

$$\mathbb{P}\left(\limsup_{m\to\infty} \frac{T_m - 2m(m+1)}{\sqrt{32(m^3 \log\log m)/3}} = 1\right) = 1$$

and

$$\mathbb{P}\left(\liminf_{m\to\infty} \frac{T_m - 2m(m+1)}{\sqrt{32(m^3 \log\log m)/3}} = -1\right) = 1.$$

Proof. Set $X_1 = T_1$ and $X_m = T_m - T_{m-1}$ for $m > 1$. Consider the aggregate at time T_{m-1} when $(m-1, 0)$ gets occupied. Since a walk must follow the x-axis to reach the site $(m-1, 0)$, we know that at time T_{m-1} all sites $\{(i, 0) : i = 0, 1, \ldots, m-1\}$ are occupied and all sites $\{(i, 0) : i \geq m\}$ are vacant. Now consider the additional time $X_m = T_m - T_{m-1}$ taken before the site $(m, 0)$ becomes occupied. Each walk visits $(m, 0)$ if and only if it passes through the sites $(1, 0), (2, 0), \ldots, (m, 0)$ during the first m steps, which happens with probability $1/4m$. Thus X_m has the geometric distribution with parameter $1/4m$. Moreover, the X_i are independent. Hence T_m is a sum of independent geometric random variables X_i.

Since $\mathbb{E}X_i = 4i$, $\operatorname{Var} X_i = 16i^2 - 4i$ and $\mathbb{E}X_i^3 = 384i^3 - 96i^2 + 4i$,

$$B_m = \sum_{1\leq i\leq m} \operatorname{Var} X_i = \frac{16}{3}m^3 + O(m^2)$$

and

$$\sum_{1\leq i\leq m} \mathbb{E}\big(|X_i - \mathbb{E}X_i|^3\big) \leq \sum_{1\leq i\leq m} \big(\mathbb{E}X_i^3 + (\mathbb{E}X_i)^3\big) = O(m^4).$$

By Lemma 11, $\Delta_m = O(m^{-1/2})$, which shows that Petrov's conditions of Lemma 12 are satisfied. Therefore,

$$\mathbb{P}\left(\limsup_{m\to\infty} \frac{T_m - \mathbb{E}T_m}{\sqrt{2B_m \log\log B_m}} = 1\right) = 1.$$

Since $\mathbb{E}T_m = \sum_{i=1}^m 4i = 2m(m+1)$ and $B_m = 16m^3/3 + O(m^2)$, this proves the first statement in Lemma 17. The second statement is obtained by applying Lemma 12 to $-T_m = \sum_{i=1}^m (-X_i)$. □

Proof of Theorem 3. Fix $\varepsilon > 0$, set $\eta := \sqrt{2(n \log\log n)/3}$ and let $\rho = \lceil (1-\varepsilon)\eta \rceil$. If we write $m = n - \rho$, then

$$2m(m+1) = 2n(n+1) - 4n\rho + o(n),$$
$$\sqrt{32(m^3 \log\log m)/3} = 4n\eta + o(n^{5/4} \log\log n).$$

Hence, setting $m = n - \rho$ in Lemma 17 gives

$$\mathbb{P}\left(\limsup_{n\to\infty} \frac{T_{n-\rho} - 2n(n+1) + 4n\rho}{4n\eta} = 1 \right) = 1.$$

Since $\{(n-\rho, 0) \notin A(v_n)\} = \{T_{n-\rho} > v_n - 1\}$ and $v_n - 1 = 2n(n+1)$, this implies

$$\mathbb{P}\big((n-\rho,0) \notin A(v_n) \text{ i.o.}\big) = 1.$$

Likewise, setting $m = n + \rho$ in Lemma 17 gives

$$\mathbb{P}\left(\liminf_{n\to\infty} \frac{T_{n+\rho} - 2n(n+1) - 4n\rho}{4n\eta} = -1 \right) = 1,$$

hence

$$\mathbb{P}\big((n+\rho,0) \in A(v_n) \text{ i.o.}\big) = 1. \qquad \square$$

8. Concluding Remarks

Lawler, Bramson and Griffeath [LBG92] discovered a key property of the Euclidean ball that characterizes it as the limiting shape of internal DLA clusters based on simple random walk in $\mathbb{Z}^d$: for simple random walk killed on exiting the ball, any point z sufficiently far from the boundary of the ball is visited more often in expectation by a walk started at the origin than by a walk started at a uniform point in the ball. Uniformly layered walks have an analogous property with respect to the diamond: the Green's function $g(y,\cdot)$ for a walk started at y and killed on exiting $\mathcal{D}_n$ satisfies

$$g(o,z) \geq \frac{1}{\#\mathcal{D}_n} \sum_{y\in\mathcal{D}_n} g(y,z)$$

for all $z \in \mathcal{D}_n$. Indeed, both the walk started at o and the walk started at a uniform point in $\mathcal{D}_n$ are uniformly distributed on layer $\mathcal{L}_{\|z\|}$ at the time $\tau_{\|z\|}$ when they first hit this layer, so the expected number of visits to z after time $\tau_{\|z\|}$ is the same for both walks. The inequality comes from the fact that a walk started at the origin must hit layer $\mathcal{L}_{\|z\|}$ before exiting $\mathcal{D}_n$.

We conclude with two questions. The first concerns uniformly layered walks started from a point other than the origin. Figure 6 shows internal DLA clusters for six different starting points in the first quadrant of $\mathbb{Z}^2$. These clusters are all contained in the first quadrant. Our simulations indicate that a limiting shape exists for each starting point, and that no two starting points have the same limiting shape; but we do not know of any explicit characterization of the shapes arising in this way.

The second question is, do there exist walks with bounded increments having uniform harmonic measure on L^1 spheres in $\mathbb{Z}^d$ for $d \geq 3$?

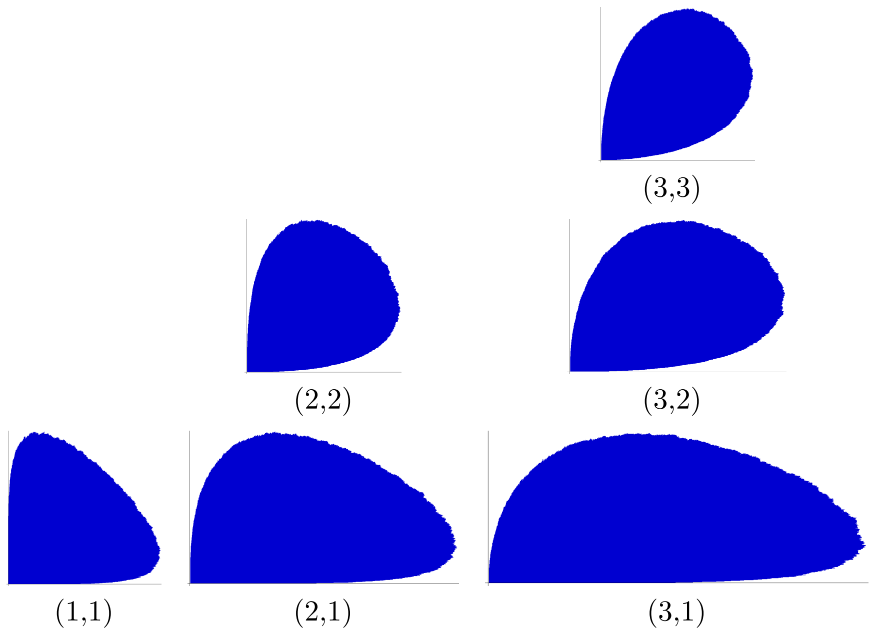

FIGURE 6. Internal DLA clusters in the first quadrant of $\mathbb{Z}^2$ based on the outward-directed layered walk Q_{out} started from a point other than the origin. For example, the cluster on the lower left is formed from 405 900 particles started at the point $(1, 1)$.

Acknowledgement

We thank Ronald Meester for fruitful discussions.

References

[AS92] N. Alon and J. H. Spencer. *The probabilistic method.* John Wiley & Sons Inc., New York (1992).

[BQR03] G. Ben Arous, J. Quastel and A. F. Ramírez. Internal DLA in a random environment. *Ann. Inst. H. Poincaré Probab. Statist.* **39**(2): 301–324 (2003).

[Bi95] P. Billingsley. *Probability and Measure.* 3rd ed. John Wiley & Sons (1995).

[BB07] S. Blachère and S. Brofferio. Internal diffusion limited aggregation on discrete groups having exponential growth. *Probab. Theory Related Fields* **137**(3-4): 323–343 (2007).

[DF91] P. Diaconis and W. Fulton. A growth model, a game, an algebra, Lagrange inversion, and characteristic classes. *Rend. Sem. Mat. Univ. Politec. Torino* **49**(1): 95–119 (1991).

[Du04] J. Dubédat. Reflected planar Brownian motions, intertwining relations and crossing probabilities. *Ann. Inst. H. Poincaré Probab. Statist.* **40**: 539–552 (2004).

[Es45] C.-G. Esseen. Fourier analysis of distribution functions. A mathematical study of the Laplace-Gaussian law. *Acta Math.* **77**: 1–125 (1945).

[GQ00] J. Gravner and J. Quastel. Internal DLA and the Stefan problem. *Ann. Probab.* **28**(4): 1528–1562 (2000).

[Ja02] S. Janson. On concentration of probability. In: Contemporary combinatorics. *Bolyai Soc. Math. Stud.* **10**: 289–301, János Bolyai Math. Soc., Budapest (2002).

[Ka07] W. Kager. Reflected Brownian motion in generic triangles and wedges. *Stoch. Process. Appl.* **117**(5): 539–549 (2007).

[LBG92] G. F. Lawler, M. Bramson and D. Griffeath. Internal diffusion limited aggregation. *Ann. Probab.* **20**(4): 2117–2140 (1992).

[La95] G. F. Lawler. Subdiffusive fluctuations for internal diffusion limited aggregation. *Ann. Probab.* **23**(1): 71–86 (1995).

[LP09a] L. Levine and Y. Peres. Strong spherical asymptotics for rotor-router aggregation and the divisible sandpile. *Potential Anal.* **30**: 1–27 (2009).

[LP09b] L. Levine and Y. Peres. Scaling limits for internal aggregation models with multiple sources. *J. d'Analyse Math.*, to appear. http://arxiv.org/abs/0712.3378

[MM00] C. Moore and J. Machta. Internal diffusion-limited aggregation: parallel algorithms and complexity. *J. Statist. Phys.* **99**(3–4): 661–690 (2000).

[Pe66] V. V. Petrov. On a relation between an estimate of the remainder in the central limit theorem and the law of iterated logarithm. *Theor. Probab. Appl.* **11**(3): 454–458 (1966).

[PS00] Yu. V. Prokhorov and V. Statulevičius (Eds.). Limit theorems of probability theory. Springer-Verlag, Berlin (2000).

[RP81] L. C. G. Rogers and J. W. Pitman Markov functions. *Ann. Probab.* **9**(4): 573–582 (1981).

Wouter Kager, Department of Mathematics, VU University Amsterdam, De Boelelaan 1081a, 1081 HV Amsterdam, The Netherlands, http://www.few.vu.nl/~wkager

Lionel Levine, Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139, http://math.mit.edu/~levine