Internal DLA in Higher Dimensions

David Jerison* Lionel Levine† Scott Sheffield‡

June 22, 2012

Abstract

Let $A(t)$ denote the cluster produced by internal diffusion limited aggregation (internal DLA) with t particles in dimension $d \geq 3$. We show that $A(t)$ is approximately spherical, up to an $O(\sqrt{\log t})$ error.

In the process known as internal diffusion limited aggregation (internal DLA) one constructs for each integer time $t \geq 0$ an **occupied set** $A(t) \subset \mathbb{Z}^d$ as follows: begin with $A(0) = \emptyset$ and $A(1) = \{0\}$. Then, for each integer $t > 1$, form $A(t+1)$ by adding to $A(t)$ the first point at which a simple random walk from the origin hits $\mathbb{Z}^d \setminus A(t)$. Let $B_r \subset \mathbb{R}^d$ denote the ball of radius r centered at 0, and write $\mathbf{B}_r := B_r \cap \mathbb{Z}^d$. Let ω_d be the volume of the unit ball in $\mathbb{R}^d$. Our main result is the following.

Theorem 1. *Fix an integer $d \geq 3$. For each γ there exists an $a = a(\gamma, d) < \infty$ such that for all sufficiently large r,*

$$\mathbb{P}\left\{\mathbf{B}_{r-a\sqrt{\log r}} \subset A(\omega_d r^d) \subset \mathbf{B}_{r+a\sqrt{\log r}}\right\}^c \leq r^{-\gamma}.$$

We treated the case $d = 2$ in [JLS12] (see also the overview in [JLS09]), where we obtained a similar statement with $\log r$ in place of $\sqrt{\log r}$. Together with a Borel-Cantelli argument, these results in particular imply the following: let $D(r)$ be the Hausdorff distance between the ball B_r and the set $A(\omega_d r^d) + [-\frac{1}{2}, \frac{1}{2}]^d$ centered at points of the internal DLA cluster. Then

Corollary 2. *For each $d \geq 2$ there is a constant $a = a(d)$ such that*

$$\mathbb{P}\left\{D(r) \leq a(\log r)^\alpha \text{ eventually}\right\} = 1$$

*Partially supported by NSF grant DMS-1069225.
†Partially supported by NSF grants DMS-0803064 and DMS-1243606.
‡Partially supported by NSF grant DMS-0645585.

where

$$\alpha = \begin{cases} 1, & d = 2 \\ \frac{1}{2}, & d \geq 3. \end{cases}$$

These results show that internal DLA in dimensions $d \geq 3$ is extremely close to a perfect sphere: when the cluster $A(t)$ has the same size as a ball of radius r, its fluctuations around that ball are confined to the $\sqrt{\log r}$ scale (versus $\log r$ in dimension 2). A recent result of Asselah and Gaudillière [AG13c] shows that Theorem 1 is sharp in the sense that $\sqrt{\log r}$ cannot be replaced by any function that is $o(\sqrt{\log r})$.

In [JLS12] we explained that our method for $d = 2$ would also apply in dimensions $d \geq 3$ with the $\log r$ replaced by $\sqrt{\log r}$. We outlined the changes needed in higher dimensions (stating that the full proof would follow in this paper) and included a key step: Lemma A, which bounds the probability of "thin tentacles" in the internal DLA cluster in all dimensions. The purpose of this note is to carry out the adaptation of the $d = 2$ argument of [JLS12] to higher dimensions. We remark that in [JLS12] we used an estimate from [LBG92] to start this iteration, while here we have modified the argument slightly so that this a priori estimate is no longer required.

One way for $A(\omega_d r^d)$ to deviate from the radius r sphere is for it to have a single "tentacle" extending beyond the sphere. The thin tentacle estimate [JLS12, Lemma A] essentially says that in dimensions $d \geq 3$, the probability that there is a tentacle of length m and volume less than a small constant times m^d (near a given location) is at most e^{-cm^2}. By summing over all locations, one may use this to show that the length of the longest "thin tentacle" produced before time t is $O(\sqrt{\log t})$. To complete the proof of Theorem 1, we will have to show that other types of deviations from the radius r sphere are also unlikely.

Lemma A of [JLS12] was also proved for $d = 2$, albeit with e^{-cm^2} replaced by $e^{-cm^2/\log m}$. However, when $d = 2$ there appear to be other more "global" fluctuations that swamp those produced by individual tentacles. (Indeed, we expect, but did not prove, that the $\log r$ fluctuation bound is tight when $d = 2$.) We bound these other fluctuations in higher dimensions via the same scheme introduced in [JLS09, JLS12], which involves constructing and estimating certain martingales related to the growth of $A(t)$. It turns out the quadratic variations of these martingales are, with high probability, of order $\log t$ when $d = 2$ and of constant order when $d \geq 3$, closely paralleling what one obtains for the discrete Gaussian free field (as outlined in more detail in [JLS12]). The connection to the Gaussian free field is made more explicit in [JLS13].

Section 1 proves Theorem 1 by iteratively applying higher dimensional analogues of the two main lemmas of [JLS12]. The lemmas themselves are proved in Section 3, which is the heart of the argument. Section 2 contains preliminary estimates about random walks that are used in Section 3.

A brief history of internal DLA fluctuation bounds

In 1986, Meakin and Deutch [MD86] defined a closely related process which they termed *diffusion limited annihilation*. In numerical experiments, they found that the average fluctuation (as opposed to the "worst case" fluctuation bounded by Theorem 1) was of order $\sqrt{\log r}$ in dimension 2 and of constant order in dimension 3. Diaconis and Fulton proposed internal DLA in its modern form in 1991 [DF91]. In 1992, Lawler, Bramson, and Griffeath proved that the limit shape of internal DLA from a point is the ball in all dimensions [LBG92]. In 1995 Lawler gave a more quantitative proof, showing that the fluctuations of $A(\omega_d r^d)$ from the ball of radius r are at most of order $O(r^{1/3}\log^4 r)$ [Law95]. In December 2009, the present authors announced the bound $O(\log r)$ on fluctuations in dimension $d = 2$ [JLS09] and gave an overview of the argument, making clear that the details remained to be written. In April 2010, Asselah and Gaudillière [AG10] gave a proof, using different methods from [JLS09], of the bound $O(r^{1/(d+1)})$ in all dimensions, improving the Lawler bound for all $d \geq 3$. In September 2010, Asselah and Gaudillière improved this to $O((\log r)^2)$ in all dimensions $d \geq 2$ with an $O(\log r)$ bound on "inner" errors [AG13a]. In October 2010 the present authors proved the $O(\log r)$ bounds (announced in December 2009) for dimension $d = 2$ and outlined the proof of the $O(\sqrt{\log r})$ bound for dimensions $d \geq 3$ [JLS12]. In November 2010, Asselah and Gaudillière gave a second proof of the $O(\sqrt{\log r})$ bound [AG13b]. Their proof uses methods from [AG13a] along with Lemma A of [JLS12] to bound "outer" errors and a new large deviation bound (in some sense symmetric to Lemma A) to bound "inner" errors.

More references and a more general discussion of internal DLA history appear in [JLS12].

1 Proof of Theorem 1

We recall the overall structure of the proof in [JLS12]. The first step is to quantify how early or late each point joins the cluster $A(T)$. Lemma 3, below, then says that an early point is unlikely unless there is also a comparably late point. Lemma 4 says that a late point is unlikely unless there is

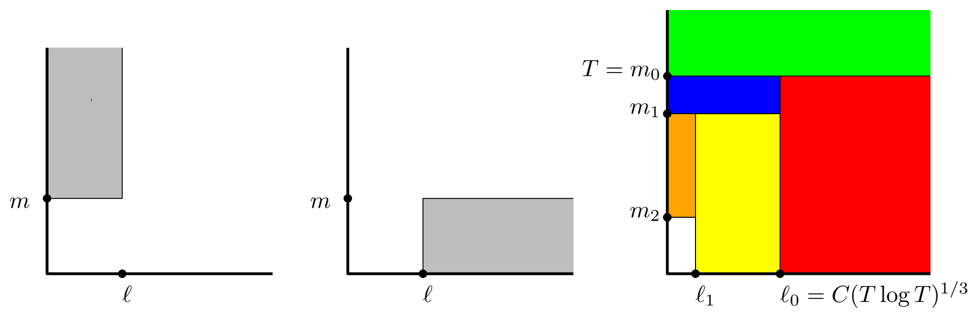

Figure 1: Let m^T be the largest m' for which $A(T)$ contains an m' early point. Let ℓ^T be the largest ℓ' for which some point of $B_{(T/\omega_d)^{1/d}-\ell'}$ is ℓ'-late. By Lemma 3, the pair of random variables (ℓ^T, m^T) is unlikely to belong to the semi-infinite rectangle in the left figure if $\ell \le m/C_0$. By Lemma 4, (ℓ^T, m^T) is unlikely to belong to the semi-infinite rectangle in the second figure if $\ell \ge C_1((\log T)m)^{1/3}$. Theorem 1 will follow because the event $\{m^T > T\}$ is impossible and the other colored rectangles on the right are all (by Lemmas 3 and 4) unlikely.

also a *significantly* earlier point. Since $A(T)$ is a connected set of T lattice sites in $\mathbb{Z}^d$, we have $A(T) \subset \mathbf{B}_T$, which gives an upper bound on how early any point can be. Thus the region $\{m > T\}$ at the top right of Figure 1 has probability 0. The other colored rectangles in Figure 1 are unlikely by Lemmas 3 and 4, so with high probability there are no very early or late points.

To make the above outline more precise, let m and ℓ be positive real numbers. We say that $x \in \mathbb{Z}^d$ is m-early if

$$x \in A(\omega_d(|x| - m)^d),$$

where $|x| = \left(\sum_{i=1}^d x_i^2\right)^{1/2}$, and ω_d is the volume of the unit ball in $\mathbb{R}^d$. Likewise, we say that x is ℓ-late if

$$x \notin A(\omega_d(|x| + \ell)^d).$$

Let $\mathcal{E}_m[T]$ be the event that some point of $A(T)$ is m-early. Let $\mathcal{L}_\ell[T]$ be the event that some point of $\mathbf{B}_{(T/\omega_d)^{1/d}-\ell}$ is ℓ-late. These events correspond to "outer" and "inner" deviations of $A(T)$ from circularity.

Lemma 3. (Early points imply late points) *Fix a dimension $d \geq 3$. For each $\gamma \geq 1$, there is a constant $C_0 = C_0(\gamma, d)$, such that for all sufficiently large T, if $m \geq C_0\sqrt{\log T}$ and $\ell \leq m/C_0$, then*

$$\mathbb{P}(\mathcal{E}_m[T] \cap \mathcal{L}_\ell[T]^c) < T^{-10\gamma}.$$

Lemma 4. (Late points imply early points) *Fix a dimension $d \geq 3$. For each $\gamma \geq 1$, there is a constant $C_1 = C_1(\gamma, d)$ such that for all sufficiently large T, if $m \geq \ell \geq C_1\sqrt{\log T}$ and $\ell \geq C_1((\log T)m)^{1/3}$, then*

$$\mathbb{P}(\mathcal{E}_m[T]^c \cap \mathcal{L}_\ell[T]) \leq T^{-10\gamma}.$$

We now proceed to derive Theorem 1 from Lemmas 3 and 4. The lemmas themselves will be proved in Section 3. Let $C = \max(C_0, C_1)$. We start with

$$m_0 = T.$$

Note that $A(T) \subset \mathbf{B}_T$, so $\mathbb{P}(\mathcal{E}_T[T]) = 0$. Next, for $j \geq 0$ we let

$$\ell_j = \max(C((\log T)m_j)^{1/3}, C\sqrt{\log T})$$

and

$$m_{j+1} = C\ell_j.$$

By induction on j, we find

$$\begin{aligned}\mathbb{P}(\mathcal{E}_{m_j}[T]) &< 2jT^{-10\gamma}\\ \mathbb{P}(\mathcal{L}_{\ell_j}[T]) &< (2j+1)T^{-10\gamma}.\end{aligned}$$

To estimate the size of ℓ_j, let $K = C^4 \log T$ and note that $\ell_j \leq \ell'_j$, where

$$\ell'_0 = (KT)^{1/3}; \quad \ell'_{j+1} = \max((K\ell'_j)^{1/3}, K^{1/2}).$$

Then

$$\ell'_j \leq \max(K^{1/3+1/9+\cdots+1/3^j}T^{1/3^j}, K^{1/2})$$

so choosing $J = \log T$ we have

$$T^{1/3^J} < 2$$

and

$$\ell_J \leq 2K^{1/2} \leq C\sqrt{\log T}.$$

Setting $T = \omega_d r^d$, $\ell = \ell_J$ and $m = m_J$, the event $\mathcal{E}_m[T] \cup \mathcal{L}_\ell[T]$ has probability at most

$$(4J+1)T^{-10\gamma} < T^{-9\gamma} < r^{-\gamma}.$$

We conclude that if a is sufficiently large, then

$$\mathbb{P}\left\{\mathbf{B}_{r-a\sqrt{\log r}} \subset A(\omega_d r^d) \subset \mathbf{B}_{r+a\sqrt{\log r}}\right\} \leq \mathbb{P}(\mathcal{E}_m[T] \cup \mathcal{L}_\ell[T]) < r^{-\gamma}$$

which completes the proof of Theorem 1.

2 Green function estimates on the grid

This section assembles several Green function estimates that we need to prove Lemmas 3 and 4. The reader who prefers to proceed to the heart of the argument may skip this section on a first read and refer to the lemma statements as necessary. Fix $d \geq 3$ and consider the d-dimensional grid

$$\mathcal{G} = \{(x_1, \ldots, x_d) \in \mathbb{R}^d \,:\, \text{at most one } x_i \notin \mathbb{Z}\}.$$

In many of the estimates below, we will assume that a positive integer k and a $y \in \mathbb{Z}^d$ have been fixed. We write

$$\Omega = \Omega(y,k) := \mathcal{G} \cap B_{|y|+k} \backslash \{y\}.$$

For $x \in \Omega \cup \partial\Omega$, let

$$P(x) = P_{y,k}(x)$$

be the probability that a Brownian motion on the grid $\mathcal{G}$ (defined in the obvious way; see [JLS12]) starting at x reaches y before exiting $B_{|y|+k}$. Note that P is **grid harmonic** in Ω (i.e., P is linear on each segment of $\Omega \setminus \mathbb{Z}^d$, and for each $x \in \Omega \cap \mathbb{Z}^d$, the sum of the slopes of P on the $2d$ directed edge segments starting at x is zero). Boundary conditions are given by $P(y) = 1$ and $P(x) = 0$ for $x \in (\partial\Omega) \setminus \{y\}$.

The point y plays the role that ζ played in [JLS12], and $P_{y,k}$ plays the role of the discrete harmonic function H_ζ. One difference from [JLS12] is that we will sometimes take $k > 1$ so that y lies inside the ball instead of on the boundary. As we explain in Section 3, this extra parameter k (in particular, the gain of a factor of k in the lower bound of Lemma 9(a)) is what enables the improved arithmetic in Lemma 4 which results in fluctuation bounds of order $\sqrt{\log T}$ instead of $\log T$ in Theorem 1.

To estimate P we use the discrete Green function $g(x)$, defined as the expected number of visits to x by a simple random walk started at the origin in $\mathbb{Z}^d$. The well-known asymptotic estimate for g is [Uch98]

$$\left| g(x) - a_d |x|^{2-d} \right| \leq C|x|^{-d} \tag{1}$$

for dimensional constants a_d and C (i.e., constants depending only on the dimension d). We extend g to a function, also denoted g, defined on the grid $\mathcal{G}$ by making g linear on each segment between lattice points. Note that g is grid harmonic on $\mathcal{G} \setminus \{0\}$.

Throughout we use C to denote a large positive dimensional constant, and c to denote a small positive dimensional constant, whose values may change from line to line.

Lemma 5. *There is a dimensional constant C such that*

(a) $P(x) \leq C/(1 + |x-y|^{d-2})$.

(b) $P(x) \leq Ck(|y| + k + 1 - |x|)/|x-y|^d$, *for* $|x-y| \geq k/2$.

(c) $\max_{x \in \mathbf{B}_r} P(x) \leq Ck/(|y| - r - k)^{d-1}$ *for* $r < |y| - 2k$.

Proof. The maximum principle (for grid harmonic functions) implies $Cg(x-y) \geq P(x)$ on Ω, which gives part (a).

For part (b), let y^* be one of the lattice points nearest to $(1 + (2k + C_1)/|y|)y$. By (1) we can choose a dimensional constant C_1 large enough so that $g(x-y) \geq g(x-y^*)$ for all $x \in \partial B_{|y|+k}$. By the maximum principle it follows that for $x \in \Omega$ we have

$$P(x) \leq C(g(x-y) - g(x-y^*)) \tag{2}$$

where $C = (g(0) - g(y-y^*))^{-1}$. Indeed, both sides are grid harmonic on Ω, and the right side is nonnegative on $\partial B_{|y|+k}$.

Combining (1) and (2) yields the bound

$$P(x) \leq \frac{Ck}{|x-y|^{d-1}}, \quad \text{for } |x-y| \geq 2k.$$

Next, let $z \in \partial B_{|y|+k}$ be such that $|z-y| = 2L$, with $L \geq 2k$. The bound above implies

$$P(x) \leq \frac{Ck}{L^{d-1}}, \quad \text{for } x \in B_L(z)$$

Let z^* be one of the lattice points nearest to $(|y| + k + L + C_1)z/|z|$. Then

$$F(x) = a_d L^{2-d} - g(x - z^*)$$

is comparable to L^{2-d} on $\partial B_{2L}(z^*)$ and positive outside the ball $B_L(z^*)$ (for a large enough dimensional constant C_1 — in fact, we can also do this with $C_1 = 1$ with L large enough). It follows that

$$P(x) \leq C(k/L^{d-1})(L^{d-2})F(x)$$

on $\partial(B_{2L}(z^*) \cap \Omega)$ and hence by the maximum principle on $B_{2L}(z^*) \cap \Omega$. Moreover,

$$F(x) \leq C(|y| + k + 1 - |x|)/L^{d-1}$$

for x a multiple of z and $|y| + k - L \leq |x| \leq |y| + k$. Thus for these values of x,

$$P(x) \leq C(k/L)F(x) \leq Ck(|y| + k + 1 - |x|)/L^d$$

We have just confirmed the bound of part (b) for points x collinear with 0 and z, but z was essentially arbitrary. To cover the cases $|x - y| \leq 2k$ one has to use exterior tangent balls of radius, say $k/2$, but actually the upper bound in part (a) will suffice for us in the range $|x - y| \leq Ck$.

Part (c) of the lemma follows from part (b). □

The mean value property (as typically stated for continuum harmonic functions) holds only approximately for discrete harmonic functions. There are two choices for where to put the approximation: one can show that the average of a discrete harmonic function h over the discrete ball $\mathbf{B}_r$ is approximately $h(0)$, or one can find an approximation w_r to the discrete ball $\mathbf{B}_r$ such that averaging h with respect to w_r yields *exactly* $h(0)$. The divisible sandpile model of [LP09] accomplishes the latter. In particular, the following discrete mean value property follows from Theorem 1.3 of [LP09]. For the sake of completeness we include a proof in dimensions $d \geq 3$. (Although we only use the $d \geq 3$ result here, the proof below also applies in dimension 2 after replacing the Green function $g(x)$ by $-a(x)$ where a is the recurrent potential kernel for $\mathbb{Z}^2$.)

Lemma 6. (Exact mean value property on an approximate ball) *For each real number $r > 0$, there is a function $w = w_r : \mathbb{Z}^d \to [0, 1]$ such that*

(i) $w(x) = 1$ *for all* $x \in \mathbf{B}_{r-c}$, *for a constant* c *depending only on* d.

(ii) $w(x) = 0$ *for all* $x \notin \mathbf{B}_r$.

(iii) *For any function h that is discrete harmonic on $\mathbf{B}_r$,*

$$\sum_{x\in\mathbb{Z}^d} w(x)(h(x)-h(0)) = 0.$$

Proof. Let $m = \omega_d(r-a)^d$, for a constant a to be chosen below. Let $\mathcal{F}$ be the set of all functions $f : \mathbb{Z}^d \to \mathbb{R}$ satisfying

$$\begin{aligned} f &\geq 0 && \text{on } \mathbb{Z}^d \\ \Delta f &\leq 1 - m\delta_0 && \text{on } \mathbb{Z}^d. \end{aligned}$$

Here $\Delta f(x) = \frac{1}{2d}\sum_{y\sim x}(f(y)-f(x))$ denotes the discrete Laplacian of f, where the sum is over the $2d$ lattice neighbors y of x; and δ_0 denotes the function that is 1 at the origin and 0 elsewhere.

Let $u : \mathbb{Z}^d \to \mathbb{R}$ be defined by

$$u(x) = \inf_{f\in\mathcal{F}} f(x)$$

and let $w = m\delta_0 + \Delta u$. (Intuitively, w is the result of starting with mass m at the origin and spreading it out by discrete balayage until every site in $\mathbb{Z}^d$ has mass at most 1.)

It is straightforward to show that $u \in \mathcal{F}$ and hence $w \leq 1$ on $\mathbb{Z}^d$. Next we show $w \geq 1_U$ where $U = \{x \in \mathbb{Z}^d | u(x) > 0\}$ is the support of u. Indeed, if for some $x \in \mathbb{Z}^d$ we have $w(x) < 1_U(x)$, then for small enough $\epsilon > 0$ we would have $u - \epsilon\delta_x \in \mathcal{F}$, contradicting the minimality of u.

To prove items (i) and (ii) we express u in terms of an obstacle problem for the discrete Laplacian. Consider the "obstacle" $\gamma : \mathbb{Z}^d \to \mathbb{R}$ given by

$$\gamma(x) = -|x|^2 - mg(x)$$

where g is the discrete Green function for $\mathbb{Z}^d$. Let Φ be the set of all functions $\phi : \mathbb{Z}^d \to \mathbb{R}$ satisfying

$$\begin{aligned} \phi &\geq \gamma && \text{on } \mathbb{Z}^d \\ \Delta\phi &\leq 0 && \text{on } \mathbb{Z}^d. \end{aligned}$$

Let

$$s(x) = \inf_{\phi\in\Phi} \phi(x).$$

Since $\Delta\gamma = -1 + m\delta_0$, a simple argument using the maximum principle shows that $u = s - \gamma$.

By the Green function estimate (1), we have

$$\gamma(x) = \Gamma(|x|) + O((r/|x|)^d)$$

where $\Gamma(t) := -t^2 - \frac{2}{d-2}(r-a)^d t^{2-d}$, and the constant in the error term depends only on the dimension d. In particular, there is a constant C depending only on d, such that for all $t \in [r/2, 2r]$ and all $x, y \in \partial\mathbf{B}_t$ we have $|\gamma(x) - \gamma(y)| < C$. We choose $a = 3C$.

Since $s \geq \gamma \geq \Gamma(r) - C$ on $\partial\mathbf{B}_r$ and s is superharmonic, we have $s \geq \Gamma(r) - C$ on $\mathbf{B}_r$ by the minimum principle. Since Γ is maximized at $t = r-a$, it follows that $s > \gamma$ on $\mathbf{B}_{r-a-b}$ for a constant b depending only on d. Hence for all $x \in \mathbf{B}_{r-a-b}$ we have $u(x) > 0$ and hence $w(x) = 1$, which proves (i).

To prove (ii), note that the constant function $\phi(x) \equiv \Gamma(r-a) + C$ belongs to Φ. Hence $s \leq \Gamma(r-a) + C$, which shows that $u \leq 2C$ on $\partial\mathbf{B}_{r-a}$. We will show this implies u is supported in $\mathbf{B}_{r-a+2C}$. For each $x \in U - \{0\}$ the equality $\Delta u(x) = 1$ implies that at least one neighbor y of x has $u(y) \geq u(x) + 1$; hence there is a path $x = x_0, x_1, \ldots, x_k = 0$ such that $u(x_i) > i$. If $|x| > r - a$ then this path must pass through $\partial\mathbf{B}_{r-a}$, which shows that $|x| \leq r - a + 2C$, proving (ii).

To prove (iii), let $h : \mathbb{Z}^d \to \mathbb{R}$ be discrete harmonic on $\mathbf{B}_r$ and let $H(x) = h(x) - h(0)$. Since $w = m\delta_0 + \Delta u$ is supported on $\mathbf{B}_r$, we have by summation by parts

$$\sum_{x\in\mathbb{Z}^d} w(x)H(x) = \sum_{x\in\mathbf{B}_r} \Delta u(x)H(x) = \sum_{x\in\mathbf{B}_r} u(x)\Delta H(x) = 0. \quad \square$$

The next lemma bounds sums of $P = P_{y,k}$ over discrete spherical shells and discrete balls.

Lemma 7. *There is a dimensional constant C such that*

(a) $\displaystyle\sum_{x\in\mathbf{B}_{r+1}\setminus\mathbf{B}_r} P(x) \leq Ck$ *for all* $r \leq |y| + k$.

(b) $\displaystyle\left|\sum_{x\in\mathbf{B}_r} (P(x) - P(0))\right| \leq Ck$ *for all* $r \leq |y|$.

(c) $\displaystyle\left|\sum_{x\in\mathbf{B}_{|y|+k}} (P(x) - P(0))\right| \leq Ck^2.$

Proof. Part (a) follows from Lemma 5: Take the worst shell, when $r = |y|$. Then the sum over x satisfying $|x-y| \le k$ and $|y| \le |x| \le |y|+1$ is bounded by Lemma 5(a)

$$\int_0^k s^{2-d} s^{d-2} ds = k$$

(volume element on disk with thickness 1 and radius k in $\mathbb{Z}^{d-1}$ is $s^{d-2}ds$.) For the remaining portion of the shell, Lemma 5(b) has numerator $k(|y| + k - |y|) = k^2$, so that

$$\int_k^\infty k^2 s^{-d} s^{d-2} ds = k.$$

Next, for part (b), let w_r be as in Lemma 6. Since P is discrete harmonic in $\mathbf{B}_{|y|}$, we have for $r \le |y|$

$$\sum_{x \in \mathbb{Z}^d} w_r(x)(P(x) - P(0)) = 0.$$

Since w_r equals the indicator $\mathbf{1}_{\mathbf{B}_r}$ except on the annulus $\mathbf{B}_r \setminus \mathbf{B}_{r-c}$, and $|w_r| \le 1$, we obtain

$$\begin{aligned} \left| \sum_{x \in \mathbf{B}_r} (P(x) - P(0)) \right| &\le \sum_{x \in \mathbf{B}_r \setminus \mathbf{B}_{r-c}} |w_r(x)| \, |P(x) - P(0)| \\ &\le \sum_{x \in \mathbf{B}_r \setminus \mathbf{B}_{r-c}} (P(x) + P(0)) \\ &\le Ck. \end{aligned}$$

In the last step we have used part (a) to bound the first term; the second term is bounded by Lemma 5(b), which says that $P(0) \le Ck/|y|^{d-1}$.

Part (c) follows by splitting the sum over $\mathbf{B}_{|y|+k}$ into k sums over spherical shells $\mathbf{B}_{|y|+j} \setminus \mathbf{B}_{|y|+j-1}$ for $j = 1, \ldots, k$, each bounded by part (a), plus a sum over the ball $\mathbf{B}_{|y|}$, bounded by part (b). □

Fix $\alpha > 0$, and consider the level set

$$U = \{x \in \mathcal{G} \mid g(x) > \alpha\}.$$

For $x \in \partial U$, let $p(x)$ be the probability that a Brownian motion started at the origin in $\mathcal{G}$ first exits U at x.

Lemma 8. *Choose α so that ∂U does not intersect $\mathbb{Z}^d$. For each $x \in \partial U$, the quantity $p(x)$ equals the directional derivative of $g/2d$ along the directed edge in U starting at x.*

Proof. We use a discrete form of the divergence theorem

$$\int_U \operatorname{div} V = \sum_{\partial U} \nu_U \cdot V. \tag{3}$$

where V is a vector-valued function on the grid, and the integral on the left is a one-dimensional integral over the grid. The dot product $\nu_U \cdot V$ is defined as $e_j \cdot V(x - 0e_j)$, where e_j is the unit vector pointing toward x along the unique incident edge in U. To define the divergence, for $z = x + te_j$, where $0 \le t < 1$ and $x \in \mathbb{Z}^d$, let

$$\operatorname{div} V(z) := \frac{\partial}{\partial x_j} e_j \cdot V(z) + \delta_x(z) \sum_{j=1}^{d} (e_j \cdot V(x + 0e_j) - e_j \cdot V(x - 0e_j)).$$

If f is a continuous function on U that is C^1 on each connected component of $U - \mathbb{Z}^d$, then the gradient of f is the vector-valued function

$$V = \nabla f = (\partial f/\partial x_1, \partial f/\partial x_2, \ldots, \partial f/\partial x_d)$$

with the convention that the entry $\partial f/\partial x_j$ is 0 if the segment is not pointing in the direction x_j. Note that ∇f may be discontinuous at points of $\mathbb{Z}^d$.

Let $G = -g/2d$, so that $\operatorname{div} \nabla G = \delta_0$. If u is grid harmonic on U, then $\operatorname{div} \nabla u = 0$ and

$$\operatorname{div} (u \nabla G - G \nabla u) = u(0) \delta_0.$$

Indeed, on each segment this is the same as $(uG' - u'G)' = u'G' - u'G' + uG'' - u''G = 0$ because u and G are linear on segments. At lattice points u and G are continuous, so the divergence operation commutes with the factors u and G and gives exactly one nonzero delta term, the one indicated.

Let $u(y)$ be the probability that Brownian motion on U started at y first exits U at x. Then $p(x) = u(0)$. Since u is grid-harmonic on U, we have $\operatorname{div} \nabla u = 0$ on U, hence by the divergence theorem

$$u(0) = \int_U \operatorname{div} (u \nabla G - G \nabla u) = \sum_{\partial U} u \, \nu_U \cdot \nabla G.$$

Since u vanishes on $\partial U \setminus \{x\}$, the only nonzero term in the sum on the right side is $\nu_U \cdot \nabla G(x)$. Since ∂U does not intersect $\mathbb{Z}^d$, this term equals the directional derivative of $g/2d$ along the directed edge in U starting at x. $\square$

Next we establish some lower bounds for P.

Lemma 9. *There is a dimensional constant $c > 0$ such that* *(a)* $P(0) \geq ck/|y|^{d-1}$.

(b) *Let* $k = 1$, *and* $z = (1 - \frac{2m}{|y|})y$ *for* $0 < m < |y|/2$. *Then*

$$\min_{x \in \mathbf{B}(z,m)} P(x) \geq c/m^{d-1}.$$

Proof. By the maximum principle, there is a dimensional constant $c > 0$ such that

$$P(x) \geq c(g(x-y) - a_d(k/2)^{2-d})$$

for $x \in B_{k/2}(y)$. In particular,

$$P(x) \geq ck^{2-d} \quad \text{for all} \quad |x-y| \leq k/4$$

Now consider the region

$$U = \{x \in \mathcal{G} : g(x) > a_d s^{2-d}\}$$

where s is chosen so that $|s - (|y| - k/8)| < 1/2$ and all of the boundary points of U are non-lattice points. (A generic value of s in the given range will suffice.)

By (1), this set is within unit distance of the ball of radius $|y| - k/8$. Let $p(z)$ represent the probability that a Brownian motion on the grid starting from the origin first exits U at $z \in \partial U$. Thus

$$u(0) = \sum_{z \in \partial U} u(z)p(z) \tag{4}$$

for all grid harmonic functions u in U.

Take any boundary point of $z \in \partial U$. Take the nearest lattice point z^*. Let z_j be a coordinate of z largest in absolute value. Then $|z_j| \geq |z|/d$. The rate of change of $|x|^{2-d}$ in the jth direction near z has size $\geq 1/d|z|^{d-1}$, which is much larger than the error term $C|z|^{-d}$ in (1). It follows that on the segment in that direction, where the function $g(x) - a_d(|y| - k/8)^{2-d}$ changes sign, its derivative is bounded below by $1/2d|z|^{d-1}$. In other words, by Lemma 8, within distance 2 of every boundary point of $z \in \partial U$ there is a point $z' \in \partial U$ for which $p(z') \geq c/|y|^{d-1}$. There are at least ck^{d-1} such points in the ball $\mathbf{B}_{k/4}(y)$ where the lower bound for P was ck^{2-d}, so

$$P(0) \geq ck^{2-d}k^{d-1}/|y|^{d-1} = ck/|y|^{d-1}.$$

Next, the argument for Lemma 9(b) is nearly the same. We are only interested in $k = 1$. It is obvious that for points x within constant distance of y (and unit distance from the boundary at radius $|y|+1$) the values of $P(x)$ are bounded below by a positive constant. We then bound $P((1-2m/|y|)y)$ from below using the same argument as above, but with Green's function for a ball of radius comparable to m. Finally, Harnack's inequality says that the values of $P(x)$ for x in the whole ball of size m around this point $(1-2m/|y|)y$ are comparable. □

3 Proofs of main lemmas

The proofs in this section make use of the martingale

$$M(t) = M_{y,k}(t) := \sum_{x \in A_{y,k}(t)} (P_{y,k}(x) - P_{y,k}(0))$$

where $P_{y,k}$ is the grid harmonic function defined in Section 2, and $A_{y,k}(t)$ is the modified internal DLA cluster in which particles are stopped if they exit Ω.

As in [JLS12], we take the time parameter t to be real-valued: starting at each integer time n, a particle is released from the origin and performs Brownian motion on the grid $\mathcal{G}$ until reaching a point in $(\mathcal{G} \setminus \Omega) \cap (\mathbb{Z}^d \setminus A(n))$. By applying a deterministic time change to the Brownian motion we can ensure that this happens before time $n+1$, so only one particle is active at any given time. The choice of continuous time is convenient for applying the martingale representation theorem, but it is not essential for the argument: One can embed the discrete time martingale $M(n)$ into a Brownian motion using Skorohod's theorem, and estimate the elapsed time $(M(n+1) - M(n))^2$.

We view $A_{y,k}(t)$ as a multiset: points on the boundary of Ω where many stopped particles accumulate are counted with multiplicity in the sum defining M. In addition to these stopped particles, the set $A_{y,k}(t)$ contains one more point, the location of the currently active particle performing Brownian motion on $\mathcal{G}$.

Recall that $P = P_{y,k}$ and $M = M_{y,k}$ depend on k, which is the distance from y to the boundary of Ω. We will choose $k = 1$ for the proof of Lemma 3, and $k = a\ell$ for a small constant a in the proof of Lemma 4. Taking $k > 1$ is one of the main differences from the argument in [JLS12]. The factor of k in the lower bound of Lemma 9(a) results in the bound $M(T_1) \leq -k^2$ on the event that y is ℓ-late, and consequently the weaker hypothesis $\ell \geq C_1((\log T)m)^{1/3}$ suffices at the end of the proof of Lemma 4 (compare to [JLS12, Lemma 13] where the power was $1/2$ instead of $1/3$).

Proof of Lemma 3. The proof follows the same method as [JLS12, Lemma 12]. We highlight here the changes needed in dimensions $d \geq 3$. We use the discrete harmonic function $P(x)$ with $k = 1$. Fix $z \in \mathbb{Z}^d$, let $r = |z|$ and $y = (r + 2m)z/r$. Let

$$T_1 = \lceil \omega_d (r - m)^d \rceil$$

where ω_d is the volume of the unit ball in $\mathbb{R}^d$. If z is m-early, then $z \in A(T_1)$; in particuar, this means that $r \geq m$, so that $r+m$, $r+2m$ are all comparable to r. Since $k = 1$, we have by Lemmas 5(c) and 9(a)

$$P(0) \approx 1/r^{d-1},$$

where $\approx$ denotes equivalence up to a constant factor depending only on d.

First we control the quadratic variation

$$S(t) = \lim_{\substack{0=t_0 \leq \ldots \leq t_N = t \\ \max(t_i - t_{i-1}) \to 0}} \sum_{i=1}^{N} (M(t_i) - M(t_{i-1}))^2$$

on the event $\mathcal{E}_{m+1}[T]^c$ that there are no $(m+1)$-early points by time T. As in [JLS12, Lemma 9], there are independent standard Brownian motions $\widetilde{B}^0, \widetilde{B}^1, \ldots$ such that each increment $(S(n+1) - S(n))\mathbf{1}_{\mathcal{E}_{m+1}[T]^c}$ is bounded above by the first exit time of $\widetilde{B}^n$ from the interval $[-a_n, b_n]$, where

$$a_n = P(0) \approx \frac{1}{r^{d-1}}$$

$$b_n = \max_{|x| \leq (n/\omega_d)^{1/d} + m + 1} P(x) \leq \frac{1}{[r + 2m - ((n/\omega_d)^{1/d} + m + 1)]^{d-1}}.$$

Here we have used Lemma 5(b) in the bound on b_n.

Unlike in dimension 2, we will use the large deviation bound for Brownian exit times [JLS12, Lemma 5] with $\lambda = cm^2$ instead of $\lambda = 1$. Here c is a constant depending only on d. Note that $b_n \leq 1/m^{d-1}$, for all $n \leq T_1$, so this is a valid choice of λ in all dimensions $d \geq 3$ (that is, the hypothesis

$\sqrt{\lambda}(a_n + b_n) \le 3$ of [JLS12, Lemma 5] holds). We obtain

$$\begin{aligned}\log \mathbb{E}\left[e^{\lambda S(T_1)} 1_{\mathcal{E}_{m+1}[T]^c}\right] &\le \sum_{n=1}^{T_1} 10\lambda a_n b_n \\ &\le \int_1^{T_1} \lambda \frac{C}{r^{d-1}} \frac{1}{(r+m-(n/\omega_d)^{1/d}-1)^{d-1}} dn \\ &\le \int_1^{r} \lambda \frac{C}{r^{d-1}} \frac{1}{(r+m-j-1)^{d-1}} j^{d-1} dj \\ &\le \int_1^{r} \frac{C\lambda\, dj}{(r+m-j-1)^{d-1}} \le C\lambda/m^{d-2}.\end{aligned}$$

Note that the last step uses $d \ge 3$. Taking $\lambda = cm^2$ for small enough c we obtain

$$\mathbb{E}\left[e^{cm^2 S(T_1)} 1_{\mathcal{E}_{m+1}[T]^c}\right] \le e^{m^2/m^{d-2}} \le e^m.$$

Therefore, by Markov's inequality,

$$\mathbb{P}(\{S(T_1) > 1/c\} \cap \mathcal{E}_{m+1}[T]^c) \le e^{m-m^2} < T^{-20\gamma}. \tag{5}$$

Fix $z \in \mathbf{B}_T$ and $t \in \{1, \dots, T\}$, and let $Q_{z,t}$ be the event that $z \in A(t) \setminus A(t-1)$ and z is m-early and no point of $A(t-1)$ is m-early. This event is empty unless $(t/\omega_d)^{1/d} + m \le |z| \le (t/\omega_d)^{1/d} + m + 1$; in particular, the first inequality implies $t \le T_1$. We will bound from below the martingale $M(t)$ on the event $Q_{z,t} \cap \mathcal{L}_\ell[T]^c$. With no ℓ-late point, the ball $\mathbf{B}_{r-m-\ell-1}$ is entirely filled by time t. Lemma 7(b) shows that the sites in this ball contribute at most a constant to $M(t)$ (recall that $k = 1$). The thin tentacle estimate [JLS12, Lemma A] says that except for an event of probability e^{-cm^2}, there are order m^d sites in $A(t)$ within the ball $\mathbf{B}(z, m)$. By Lemma 9(b), P is bounded below by c/m^{d-1} on this ball, so these sites taken together contribute order m to $M(t)$. Each of the remaining terms in the sum defining $M(t)$ is bounded below by $-P(0)$, and there are at most ℓr^{d-1} sites in $A(t) \setminus \mathbf{B}_{r-m-\ell-1}$. So these terms contribute at least

$$-\ell r^{d-1}(1/r^{d-1}) = -\ell \ge -m/C$$

which cannot overcome the order m term. Thus

$$\mathbb{P}(Q_{z,t} \cap \{M_\zeta(t) < m/C\} \cap \mathcal{L}_\ell[t]^c) < e^{-cm^2}. \tag{6}$$

We conclude that

$$\begin{aligned}\mathbb{P}(Q_{z,t} \cap \mathcal{L}_\ell[T]^c) \le\ & \mathbb{P}(Q_{z,t} \cap \{S(t) > 1/c\}) \\ &+ \mathbb{P}(Q_{z,t} \cap \{M(t) < m/C\} \cap \mathcal{L}_\ell[t]^c) \\ &\quad + \mathbb{P}(\{S(t) \le 1/c\} \cap \{M(t) \ge m/C\}).\end{aligned}$$

The first two terms are bounded by (5) and (6). Since $M(t) = B(S(t))$ for a standard Brownian motion B, the final term is bounded by

$$\mathbb{P}\left\{\sup_{0\le s\le 1/c} B(s) \ge m/C\right\} < e^{-c(m/C)^2/2} < T^{-20\gamma}. \quad \square$$

Proof of Lemma 4. Fix $y \in \mathbb{Z}^d$, and let $L[y]$ be the event that y is ℓ-late. Set $k = a\ell$ in the definition of $P = P_{y,k}$. Here $a > 0$ is a small dimensional constant chosen below. Note that the hypotheses on m and ℓ imply that ℓ is at least of order $\sqrt{\log T}$; after choosing a, we take the constant C_1 appearing in the statement of the lemma large enough so that $k^2 > 1000\gamma \log T$.

Case 1. $1 \le |y| \le 2k$. Then $P(0) \approx 1/|y|^{d-2}$. Let

$$T_1 = \lfloor \omega_d(|y| + \ell)^d \rfloor$$

With $a_n = P(0)$ and $b_n = 1$, we have $S(n+1) - S(n) \le \tau_n$, where τ_n is the first exit time of the Brownian motion $\widetilde{B}^n$ from the interval $[-a_n, b_n]$. (Note that because we take $b_n = 1$, the indicator $\mathbf{1}_{\mathcal{E}_{m+1}[T]^c}$ is not needed here as it was in the proof of Lemma 3.) We obtain

$$\log \mathbb{E}e^{S(T_1)} \le \sum_{t=1}^{T_1} \log \mathbb{E}e^{\tau_n} \le T_1 P(0).$$

Let $Q = T_1 P(0)$. By Markov's inequality, $\mathbb{P}(S(T_1) > 2Q) \le e^{-Q}$.

On the event $L[y]$, the site y is still not occupied at time T_1. Accordingly, the largest $M(T_1)$ can be is if $A_{y,k}(T_1)$ fills the whole ball $\mathbf{B}_{|y|+k}$ (except for y), and then the rest of the particles will have to collect on the boundary where P is zero. The contribution from $\mathbf{B}_{|y|+k}$ is at most Ck^2 by Lemma 7(c). The number of particles stopped on the boundary is at least

$$T_1 - 2\omega_d(|y| + k)^d \ge \frac{T_1}{2}.$$

Therefore, on the event $L[y]$ we have

$$M(T_1) \le Ck^2 - \frac{T_1}{2}P(0). \tag{7}$$

Note that $Q := T_1 P(0) \approx (|y| + \ell)^d/|y|^{d-2} \ge \ell^d/(k/2)^{d-2}$, so by taking $a = k/\ell$ sufficiently small, we can ensure that the right side of (7) is at most

$-Q/4$. Also, $Q \geq \ell^2 \geq 1000\gamma \log T$. Since $M(t) = B(S(t))$ for a standard Brownian motion B, we conclude that

$$\begin{aligned}\mathbb{P}(L[y]) &\leq \mathbb{P}(S(T_1) > 2Q) + \mathbb{P}\left\{\inf_{0\leq s\leq 2Q} B(s) \leq -Q/4\right\} \\ &\leq e^{-Q} + e^{-(Q/4)^2/4Q} \\ &< T^{-20\gamma}.\end{aligned}$$

Case 2. $|y| \geq 2k$. Then by Lemma 5(c) with $r = 1$, and Lemma 9(a), we have $P(0) \approx k/s^{d-1}$. First take

$$T_0 = \lfloor \omega_d(|y| + k - 3m)^d \rfloor$$

(or $T_0 = 0$ if $|y| + k - 3m \leq 0$). As in the previous lemma (but taking $\lambda = 1$ instead of $\lambda = cm^2$) we have

$$\log \mathbb{E}\left[e^{S(T_0)} 1_{\mathcal{E}_m[T]^c}\right] \leq C \frac{k}{|y|^{d-1}} \int_0^{T_0} \frac{dn}{\left(|y| + k - (n/\omega_d)^{1/d}\right)^{d-1}} \leq Ck/m^{d-2}.$$

Since $d \geq 3$ and $m \geq k/a$, the right side is $\leq C$. By Markov's inequality,

$$\mathbb{P}(\{S(T_0) > C + k^2\} \cap \mathcal{E}_m[T]^c) < e^{-k^2} < T^{-20\gamma}.$$

Now since

$$(T_1 - T_0)P(0) \approx m|y|^{d-1}(k/|y|^{d-1}) = km$$

we have

$$\log \mathbb{E} e^{S(T_1)-S(T_0)} \leq Ckm.$$

Thus (since $km \geq k^2$)

$$\mathbb{P}(\{S(T_1) > 2Ckm\} \cap \mathcal{E}_m[T]^c) < 2T^{-20\gamma}. \tag{8}$$

As in case 1, the martingale $M(T_1)$ is largest if the ball $\mathbf{B}_{|y|+k}$ is completely filled, and in that case the total contribution of sites in this ball is at most Ck^2. On the event $L[y]$, the number of particles stopped on the boundary of Ω at time T_1 is at least

$$T_1 - \#\mathbf{B}_{|y|+k} \geq \omega_d((|y| + \ell)^d - (|y| + k + C)^d) \approx \ell|y|^{d-1}.$$

Each such particle contributes $-P(0) \approx -k/|y|^{d-1}$ to $M(T_1)$, for a total contribution of order $-k\ell = -k^2/a$. Taking a sufficiently small we obtain $M(T_1) \leq Ck^2 - k^2/a \leq -k^2$. We conclude that

$$\mathbb{P}(L[y] \cap \mathcal{E}_m[T]^c) \leq \mathbb{P}(\{S(T_1) > 2Ckm\} \cap \mathcal{E}_m[T]^c) + \\ + \mathbb{P}(\{S(T_1) \leq 2Ckm\} \cap \{M(T_1) \leq -k^2\}).$$

The first term is bounded above by (8), and the second term is bounded above by

$$\mathbb{P}\left\{ \inf_{s \leq 2Ckm} B(s) \leq -k^2 \right\} \leq e^{-k^4/4Ckm} < T^{-20\gamma}.$$

Hence $\mathbb{P}(L[y] \cap \mathcal{E}_m[T]^c) < 3T^{-20\gamma}$. Since $\mathcal{L}_\ell[T]$ is the union of the events $L[y]$ for $y \in \mathcal{B} := \mathbf{B}_{(T/\omega_d)^{1/d} - \ell}$, summing over $y \in \mathcal{B}$ completes the proof. □

References

[AG10] A. Asselah and A. Gaudillière, A note on the fluctuations for internal diffusion limited aggregation. arXiv:1004.4665

[AG13a] A. Asselah and A. Gaudillière, From logarithmic to subdiffusive polynomial fluctuations for internal DLA and related growth models. *Ann. Probab.* **41**:1115–1159, 2013. arXiv:1009.2838

[AG13b] A. Asselah and A. Gaudillière, Sub-logarithmic fluctuations for internal DLA. *Ann. Probab.* **41**:1160–1179, 2013. arXiv:1011.4592

[AG13c] A. Asselah and A. Gaudillière, Lower bounds on fluctuations for internal DLA, *Probab. Theory Related Fields*, to appear. arXiv:1111.4233

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

[JLS09] D. Jerison, L. Levine and S. Sheffield, Internal DLA: slides and audio. *Midrasha on Probability and Geometry: The Mathematics of Oded Schramm.* http://iasmac31.as.huji.ac.il:8080/groups/midrasha_14/weblog/855d7, 2009.

[JLS12] D. Jerison, L. Levine and S. Sheffield, Logarithmic fluctuations for internal DLA, *J. Amer. Math. Soc.* **25**: 271–301, 2012. arXiv:1010.2483

[JLS13] D. Jerison, L. Levine and S. Sheffield, Internal DLA and the Gaussian free field. *Duke Math. J.*, to appear. arXiv:1101.0596

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

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

[Law95] G. F. Lawler, Subdiffusive fluctuations for internal diffusion limited aggregation, *Ann. Probab.* **23**(1):71–86, 1995.

[LP09] L. Levine and Y. Peres, Strong spherical asymptotics for rotor-router aggregation and the divisible sandpile, *Potential Anal.* **30**:1–27, 2009. arXiv:0704.0688

[MD86] P. Meakin and J. M. Deutch, The formation of surfaces by diffusion-limited annihilation, *J. Chem. Phys.* **85**:2320, 1986.

[Uch98] K. Uchiyama, Green's functions for random walks on $\mathbb{Z}^N$, *Proc. London Math. Soc.* **77** (1998), no. 1, 215–240.