Potential Anal (2009) 30:1–27
DOI 10.1007/s11118-008-9104-6

Strong Spherical Asymptotics for Rotor-Router Aggregation and the Divisible Sandpile

Lionel Levine · Yuval Peres

Received: 22 November 2007 / Accepted: 7 October 2008 / Published online: 29 October 2008

Abstract The rotor-router model is a deterministic analogue of random walk. It can be used to define a deterministic growth model analogous to internal DLA. We prove that the asymptotic shape of this model is a Euclidean ball, in a sense which is stronger than our earlier work (Levine and Peres, Indiana Univ Math J 57(1):431–450, 2008). For the shape consisting of $n = \omega_d r^d$ sites, where ω_d is the volume of the unit ball in $\mathbb{R}^d$, we show that the inradius of the set of occupied sites is at least $r - O(\log r)$, while the outradius is at most $r + O(r^\alpha)$ for any $\alpha > 1 - 1/d$. For a related model, the *divisible sandpile*, we show that the domain of occupied sites is a Euclidean ball with error in the radius a constant independent of the total mass. For the classical abelian sandpile model in two dimensions, with $n = \pi r^2$ particles, we show that the inradius is at least $r/\sqrt{3}$, and the outradius is at most $(r + o(r))/\sqrt{2}$. This improves on bounds of Le Borgne and Rossin. Similar bounds apply in higher dimensions, improving on bounds of Fey and Redig.

Keywords Abelian sandpile · Asymptotic shape · Discrete Laplacian · Divisible sandpile · Growth model · Internal diffusion limited aggregation · Rotor-router model

Mathematics Subject Classifications (2000) Primary 60G50 · Secondary 35R35

Lionel Levine is supported by an NSF Graduate Research Fellowship, and NSF grant DMS-0605166.

Yuval Peres is partially supported by NSF grant DMS-0605166.

L. Levine (✉)
University of California, Berkeley, Berkeley, CA, USA
e-mail: levine@math.berkeley.edu

Y. Peres
Microsoft Research, Redmond, WA, USA
e-mail: peres@microsoft.com

1 Introduction

Rotor-router walk is a deterministic analogue of random walk, first studied by Priezzhev et al. [18] under the name "Eulerian walkers." It was later independently discovered by Jim Propp, and has also been called the "Propp machine" [3]. At each site in the integer lattice $\mathbb{Z}^2$ is a *rotor* pointing north, south, east or west. A particle starts at the origin; during each time step, the rotor at the particle's current location is rotated clockwise by 90 degrees, and the particle takes a step in the direction of the newly rotated rotor. In rotor-router aggregation, introduced by Jim Propp, we start with n particles at the origin; each particle in turn performs rotor-router walk until it reaches a site not occupied by any other particles. Let A_n denote the resulting region of n occupied sites. For example, if all rotors initially point north, the sequence will begin $A_1 = \{(0,0)\}$, $A_2 = \{(0,0),(1,0)\}$, $A_3 = \{(0,0),(1,0),(0,-1)\}$. The region $A_{1,000,000}$ is pictured in Fig. 1. In higher dimensions, the model can be defined analogously by repeatedly cycling the rotors through an ordering of the $2d$ cardinal directions in $\mathbb{Z}^d$.

Jim Propp observed from simulations in two dimensions that the regions A_n are extraordinarily close to circular, and asked why this was so [7, 13, 19]. Despite the impressive empirical evidence for circularity, the best result known until now [14] says only that if A_n is rescaled to have unit volume, the volume of the symmetric difference of A_n with a ball of unit volume tends to zero as a power of n, as $n \uparrow \infty$. The main outline of the argument is summarized in [15]. Fey and Redig [5] also show that A_n contains a diamond. In particular, these results do not rule out the possibility of "holes" in A_n far from the boundary or of long tendrils extending far beyond the boundary of the ball, provided the volume of these features is negligible compared to n.

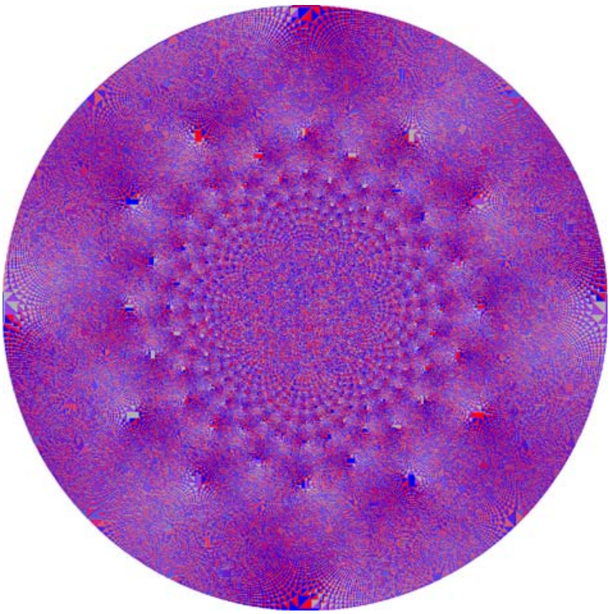


Fig. 1 Rotor-router aggregate of one million particles in $\mathbb{Z}^2$. Each site is colored according to the direction of its rotor

Our main result is the following, which rules out the possibility of holes far from the boundary or of long tendrils in the rotor-router shape. For $r \geq 0$ let

$$B_r = \left\{x \in \mathbb{Z}^d \ : \ |x| < r\right\}.$$

Theorem 1.1 *Let A_n be the region formed by rotor-router aggregation in $\mathbb{Z}^d$ starting from n particles at the origin and any initial rotor state. There exist constants c, c' depending only on d, such that*

$$B_{r - c\log r} \subset A_n \subset B_{r(1+c'r^{-1/d}\log r)}$$

where $r = (n/\omega_d)^{1/d}$, and ω_d is the volume of the unit ball in $\mathbb{R}^d$.

We remark that the same result holds when the rotors evolve according to stacks of bounded discrepancy; see the remark following Lemma 5.1.

Internal diffusion limited aggregation ("internal DLA") is an analogous growth model defined using random walks instead of rotor-router walks. Starting with n particles at the origin, each particle in turn performs simple random walk until it reaches an unoccupied site. Lawler et al. [10] showed that for internal DLA in $\mathbb{Z}^d$, the occupied region A_n, rescaled by a factor of $n^{1/d}$, converges with probability one to a Euclidean ball in $\mathbb{R}^d$ as $n \to \infty$. Lawler [11] estimated the rate of convergence. By way of comparison with Theorem 1.1, if I_n is the internal DLA region formed from n particles started at the origin, the best known bounds [11] are (up to logarithmic factors)

$$B_{r-r^{1/3}} \subset I_n \subset B_{r+r^{1/3}}$$

for all sufficiently large n, with probability one.

We also study another model which is slightly more difficult to define, but much easier to analyze. In the *divisible sandpile*, each site $x \in \mathbb{Z}^d$ starts with a quantity of "mass" $\nu_0(x) \in \mathbb{R}_{\geq 0}$. A site *topples* by keeping up to mass 1 for itself, and distributing the excess (if any) equally among its neighbors. Thus if x has mass $m > 1$, then each of the $2d$ neighboring sites gains mass $(m-1)/2d$ when we topple x, and x is left with mass 1; if $m \leq 1$, then no mass moves when we topple x.

Note that individual topplings do not commute; however, the divisible sandpile is "abelian" in the following sense.

Proposition 1.2 *Let $x_1, x_2, \ldots \in \mathbb{Z}^d$ be a sequence with the property that for any $x \in \mathbb{Z}^d$ there are infinitely many terms $x_k = x$. Let*

$$u_k(x) = \text{total mass emitted by } x \text{ after toppling } x_1, \ldots, x_k;$$

$$\nu_k(x) = \text{amount of mass present at } x \text{ after toppling } x_1, \ldots, x_k.$$

Then $u_k \uparrow u$ and $\nu_k \to \nu \leq 1$. Moreover, the limits u and ν are independent of the sequence $\{x_k\}$.

The abelian property can be generalized as follows: after performing some topplings, we can add some additional mass and then continue toppling. The resulting limits u and ν will be the same as in the case when all mass was initially present. For a further generalization, see [16].

The limiting function u in Proposition 1.2 is the *odometer function* for the divisible sandpile. This function plays a central role in our analysis. The limit ν represents the final mass distribution. Sites $x \in \mathbb{Z}^d$ with $\nu(x) = 1$ are called *fully occupied*. Proposition 1.2 is proved in Section 3, along with the following.

Theorem 1.3 *For $m \geq 0$ let $D_m \subset \mathbb{Z}^d$ be the domain of fully occupied sites for the divisible sandpile formed from a pile of mass m at the origin. There exist constants c, c' depending only on d, such that*

$$B_{r-c} \subset D_m \subset B_{r+c'},$$

where $r = (m/\omega_d)^{1/d}$ and ω_d is the volume of the unit ball in $\mathbb{R}^d$.

The divisible sandpile is similar to the "oil game" studied by Van den Heuvel [22]. In the terminology of [5], it also corresponds to the $h \to -\infty$ limit of the classical abelian sandpile (defined below), that is, the abelian sandpile started from the initial condition in which every site has a very deep "hole."

In the classical abelian sandpile model [1], each site in $\mathbb{Z}^d$ has an integer number of grains of sand; if a site has at least $2d$ grains, it *topples*, sending one grain to each neighbor. If n grains of sand are started at the origin in $\mathbb{Z}^d$, write S_n for the set of sites that are visited during the toppling process; in particular, although a site may be empty in the final state, we include it in S_n if it was occupied at any time during the evolution to the final state.

Until now the best known constraints on the shape of S_n in two dimensions were due to Le Borgne and Rossin [12], who proved that

$$\left\{x \in \mathbb{Z}^2 \mid x_1 + x_2 \leq \sqrt{n/12} - 1\right\} \subset S_n \subset \left\{x \in \mathbb{Z}^2 \mid x_1, x_2 \leq \sqrt{n}/2\right\}.$$

Fey and Redig [5] proved analogous bounds in higher dimensions, and extended these bounds to arbitrary values of the height parameter h. This parameter is discussed in Section 4.

The methods used to prove the near-perfect circularity of the divisible sandpile shape in Theorem 1.3 can be used to give constraints on the shape of the classical abelian sandpile, improving on the bounds of [5] and [12].

Theorem 1.4 *Let S_n be the set of sites that are visited by the classical abelian sandpile model in $\mathbb{Z}^d$, starting from n particles at the origin. Write $n = \omega_d r^d$. Then for any $\epsilon > 0$ we have*

$$B_{c_1 r - c_2} \subset S_n \subset B_{c_1' r + c_2'}$$

where

$$c_1 = (2d-1)^{-1/d}, \qquad c_1' = (d-\epsilon)^{-1/d}.$$

The constant c_2 depends only on d, while c_2' depends only on d and ϵ.

Note that Theorem 1.4 does not settle the question of the asymptotic shape of S_n, and indeed it is not clear from simulations whether the asymptotic shape in two dimensions is a disc or perhaps a polygon (Fig. 2). To our knowledge even the existence of an asymptotic shape is not known (Fig. 3).

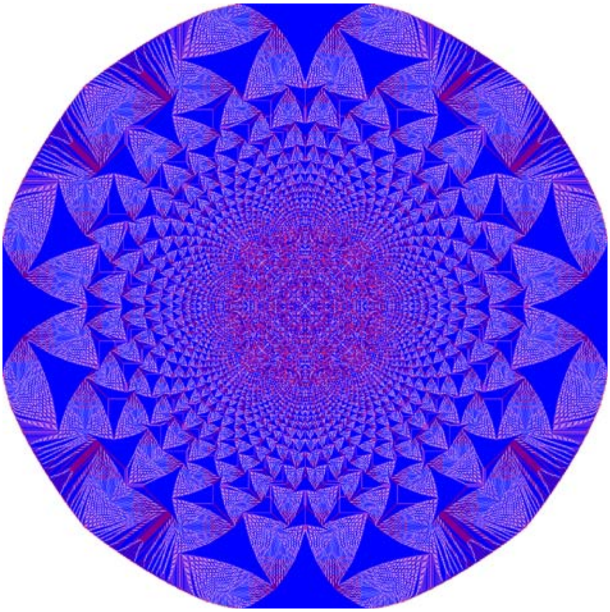


Fig. 2 Classical abelian sandpile aggregate of one million particles in $\mathbb{Z}^2$. *Colors* represent the number of grains at each site

The rest of the paper is organized as follows. In Section 2, we derive the basic Green's function estimates that are used in the proofs of Theorems 1.1, 1.3 and 1.4. In Section 3 we prove Proposition 1.2 and Theorem 1.3 for the divisible sandpile. In

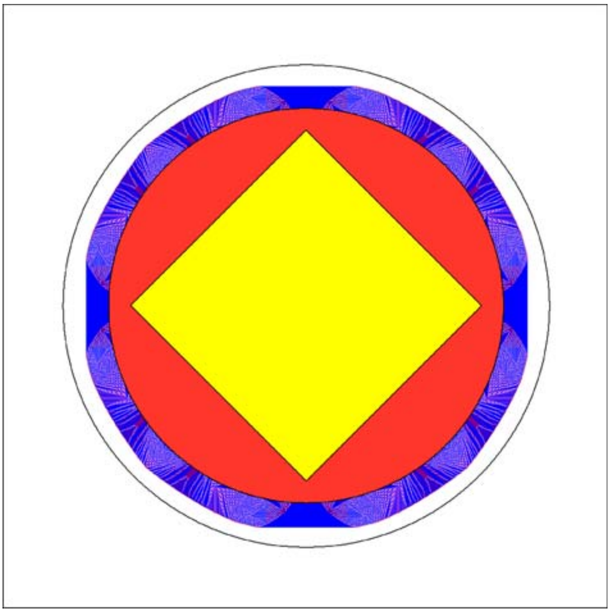


Fig. 3 Known bounds on the shape of the classical abelian sandpile in $\mathbb{Z}^2$. The *inner diamond* and *outer square* are due to Le Borgne and Rossin [12]; the *inner* and *outer circles* are those in Theorem 1.4

Section 4 we adapt the methods of the previous section to prove Theorem 1.4 for the classical abelian sandpile model. Section 5 is devoted to the proof of Theorem 1.1.

2 Basic Estimate

Write $(X_k)_{k\geq 0}$ for simple random walk in $\mathbb{Z}^d$, and for $d \geq 3$ denote by

$$g(x) = \mathbb{E}_o\#\{k | X_k = x\}$$

the expected number of visits to x by simple random walk started at the origin. This is the *discrete harmonic Green's function* in $\mathbb{Z}^d$; it satisfies $\Delta g(x) = 0$ for $x \neq o$, and $\Delta g(o) = -1$, where Δ is the discrete Laplacian

$$\Delta g(x) = \frac{1}{2d}\sum_{y\sim x} g(y) - g(x).$$

The sum is over the $2d$ lattice neighbors y of x. In dimension $d = 2$, simple random walk is recurrent, so the expectation defining $g(x)$ is infinite. Here we define the *potential kernel*

$$g(x) = \lim_{n\to\infty} g_n(x) - g_n(o) \tag{1}$$

where

$$g_n(x) = \mathbb{E}_o\#\{k \leq n | X_k = x\}.$$

The limit defining $g(x)$ in Eq. 1 is finite [9, 20], and $g(x)$ has Laplacian $\Delta g(x) = 0$ for $x \neq o$, and $\Delta g(o) = -1$. Note that Eq. 1 is the negative of the usual definition of the potential kernel; we have chosen this sign convention so that g has the same Laplacian in dimension two as in higher dimensions.

Fix a real number $m > 0$ and consider the function on $\mathbb{Z}^d$

$$\widetilde{\gamma}_d(x) = |x|^2 + mg(x). \tag{2}$$

Let r be such that $m = \omega_d r^d$, and let

$$\gamma_d(x) = \widetilde{\gamma}_d(x) - \widetilde{\gamma}_d(\lfloor r \rfloor e_1) \tag{3}$$

where e_1 is the first standard basis vector in $\mathbb{Z}^d$. The function γ_d plays a central role in our analysis. To see where it comes from, recall the divisible sandpile *odometer function* of Proposition 1.2

$$u(x) = \text{total mass emitted from } x.$$

Let $D_m \subset \mathbb{Z}^d$ be the domain of fully occupied sites for the divisible sandpile formed from a pile of mass m at the origin. For $x \in D_m$, since each neighbor y of x emits an equal amount of mass to each of its $2d$ neighbors, we have

$$\begin{aligned}\Delta u(x) &= \frac{1}{2d}\sum_{y\sim x} u(y) - u(x) \\ &= \text{mass received by } x - \text{mass emitted by } x \\ &= 1 - m\delta_{ox}.\end{aligned}$$

Moreover, $u = 0$ on ∂D_m. By construction, the function γ_d obeys the same Laplacian condition: $\Delta\gamma_d = 1 - m\delta_o$; and as we will see shortly, $\gamma_d \approx 0$ on ∂B_r. Since we expect the domain D_m to be close to the ball B_r, we should expect that $u \approx \gamma_d$. In fact, we will first show that u is close to γ_d, and then use this to conclude that D_m is close to B_r.

We will use the following estimates for the Green's function [6, 21]; see also [9, Theorems 1.5.4 and 1.6.2].

$$g(x) = \begin{cases} -\frac{2}{\pi}\log|x| + \kappa + O\left(|x|^{-2}\right), & d = 2 \\ a_d|x|^{2-d} + O\left(|x|^{-d}\right), & d \geq 3. \end{cases} \tag{4}$$

Here $a_d = \frac{2}{(d-2)\omega_d}$, where ω_d is the volume of the unit ball in $\mathbb{R}^d$, and κ is a constant whose value we will not need to know. For $x \in \mathbb{Z}^d$ we use $|x|$ to denote the Euclidean norm of x. Here and throughout the paper, constants in error terms denoted $O(\cdot)$ depend only on d.

We will need an estimate for γ_d near the boundary of the ball B_r. We first consider dimension $d = 2$. From Eq. 4 we have

$$\widetilde{\gamma}_2(x) = \phi(x) - \kappa m + O\left(m|x|^{-2}\right), \tag{5}$$

where

$$\phi(x) = |x|^2 - \frac{2m}{\pi}\log|x|.$$

In the Taylor expansion of ϕ about $|x| = r$

$$\phi(x) = \phi(r) - \phi'(r)(r - |x|) + \frac{1}{2}\phi''(t)\,(r - |x|)^2 \tag{6}$$

the linear term vanishes, leaving

$$\gamma_2(x) = \left(1 + \frac{m}{\pi t^2}\right)(r - |x|)^2 + O\left(m|x|^{-2}\right) \tag{7}$$

for some t between $|x|$ and r.

In dimensions $d \geq 3$, from Eq. 4 we have

$$\widetilde{\gamma}_d(x) = |x|^2 + a_d m|x|^{2-d} + O\left(m|x|^{-d}\right).$$

Setting $\phi(x) = |x|^2 + a_d m|x|^{2-d}$, the linear term in the Taylor expansion Eq. 6 of ϕ about $|x| = r$ again vanishes, yielding

$$\gamma_d(x) = \left(1 + (d-1)(r/t)^d\right)(r - |x|)^2 + O\left(m|x|^{-d}\right)$$

for t between $|x|$ and r. Together with Eq. 7, this yields the following estimates in all dimensions $d \geq 2$.

Lemma 2.1 *Let γ_d be given by Eq. 3. For all $x \in \mathbb{Z}^d$ we have*

$$\gamma_d(x) \geq (r - |x|)^2 + O\left(\frac{r^d}{|x|^d}\right). \tag{8}$$

Lemma 2.2 *Let γ_d be given by Eq. 3. Then uniformly in r,*

$$\gamma_d(x) = O(1), \qquad x \in B_{r+1} - B_{r-1}.$$

The following lemma is useful for x near the origin, where the error term in Eq. 8 blows up.

Lemma 2.3 *Let γ_d be given by Eq. 3. Then for sufficiently large r, we have*

$$\gamma_d(x) > \frac{r^2}{4}, \qquad \forall x \in B_{r/3}.$$

Proof Since $\gamma_d(x) - |x|^2$ is superharmonic, it attains its minimum in $B_{r/3}$ at a point z on the boundary. Thus for any $x \in B_{r/3}$

$$\gamma_d(x) - |x|^2 \geq \gamma_d(z) - |z|^2,$$

hence by Lemma 2.1

$$\gamma_d(x) \geq (2r/3)^2 - (r/3)^2 + O(1) > \frac{r^2}{4}. \qquad \square$$

Lemmas 2.1 and 2.3 together imply the following.

Lemma 2.4 *Let γ_d be given by Eq. 3. There is a constant a depending only on d, such that $\gamma_d \geq -a$ everywhere.*

3 Divisible Sandpile

Let μ_0 be a nonnegative function on $\mathbb{Z}^d$ with finite support. We start with mass $\mu_0(y)$ at each site y. The operation of *toppling* a vertex x yields the mass distribution

$$T_x\mu_0 = \mu_0 + \alpha(x)\Delta\delta_x$$

where $\alpha(x) = \max(\mu_0(x) - 1, 0)$ and Δ is the discrete Laplacian on $\mathbb{Z}^d$. Thus if $\mu_0(x) \leq 1$ then $T_x\mu_0 = \mu_0$ and no mass topples; if $\mu_0(x) > 1$ then the mass in excess of 1 is distributed equally among the neighbors of x.

Let $x_1, x_2, \ldots \in \mathbb{Z}^d$ be a sequence with the property that for any $x \in \mathbb{Z}^d$ there are infinitely many terms $x_k = x$. Let

$$\mu_k(y) = T_{x_k} \ldots T_{x_1}\mu_0(y).$$

be the amount of mass present at y after toppling the sites $x_1, \ldots, x_k$ in succession. The total mass emitted from y during this process is

$$u_k(y) := \sum_{j \leq k: x_j = y} \mu_{j-1}(y) - \mu_j(y) = \sum_{j \leq k: x_j = y} \alpha_j(y) \tag{9}$$

where $\alpha_j(y) = \max(\mu_j(y) - 1, 0)$.

Lemma 3.1 *As $k \uparrow \infty$ the functions u_k and μ_k tend to limits $u_k \uparrow u$ and $\mu_k \to \mu$. Moreover, these limits satisfy*

$$\mu = \mu_0 + \Delta u \leq 1.$$

Proof Write $M = \sum_x \mu_0(x)$ for the total starting mass, and let $B \subset \mathbb{Z}^d$ be a ball centered at the origin containing all points within L^1-distance M of the support of μ_0. Note that if $\mu_k(x) > 0$ and $\mu_0(x) = 0$, then x must have received its mass from a neighbor, so $\mu_k(y) \geq 1$ for some $y \sim x$. Since $\sum_x \mu_k(x) = M$, it follows that μ_k is supported on B. Let R be the radius of B, and consider the *quadratic weight*

$$Q_k = \sum_{x \in \mathbb{Z}^d} \mu_k(x)|x|^2 \leq MR^2.$$

Since $\mu_k(x_k) - \mu_{k-1}(x_k) = -\alpha_k(x_k)$ and for $y \sim x_k$ we have $\mu_k(y) - \mu_{k-1}(y) = \frac{1}{2d}\alpha_k(x_k)$, we obtain

$$Q_k - Q_{k-1} = \alpha_k(x_k)\left(\frac{1}{2d}\sum_{y \sim x_k} |y|^2 - |x_k|^2\right) = \alpha_k(x_k).$$

Summing over k we obtain from Eq. 9

$$Q_k = Q_0 + \sum_{x \in \mathbb{Z}^d} u_k(x).$$

Fixing x, the sequence $u_k(x)$ is thus increasing and bounded above, hence convergent.

Given neighboring vertices $x \sim y$, since y emits an equal amount of mass to each of its $2d$ neighbors, it emits mass $u_k(y)/2d$ to x up to time k. Thus x receives a total mass of $\frac{1}{2d}\sum_{y \sim x} u_k(y)$ from its neighbors up to time k. Comparing the amount of mass present at x before and after toppling, we obtain

$$\mu_k(x) = \mu_0(x) + \Delta u_k(x).$$

Since $u_k \uparrow u$ we infer that $\mu_k \to \mu := \mu_0 + \Delta u$. Note that if $x_k = x$, then $\mu_k(x) \leq 1$. Since for each $x \in \mathbb{Z}^d$ this holds for infinitely many values of k, the limit satisfies $\mu \leq 1$. □

A function s on $\mathbb{Z}^d$ is *superharmonic* if $\Delta s \leq 0$. Given a function γ on $\mathbb{Z}^d$ the *least superharmonic majorant* of γ is the function

$$s(x) = \inf\{f(x) \mid f \text{ is superharmonic and } f \geq \gamma\}.$$

The study of the least superharmonic majorant is a classical topic in analysis and PDE; see, for example, [8]. Note that if f is superharmonic and $f \geq \gamma$ then

$$f(x) \geq \frac{1}{2d}\sum_{y \sim x} f(y) \geq \frac{1}{2d}\sum_{y \sim x} s(y).$$

Taking the infimum on the left side we obtain that s is superharmonic.

Lemma 3.2 *The limit u in Lemma 3.1 is given by $u = s + \gamma$, where*

$$\gamma(x) = |x|^2 + \sum_{y \in \mathbb{Z}^d} g(x-y)\mu_0(y)$$

and s is the least superharmonic majorant of $-\gamma$.

Proof By Lemma 3.1 we have

$$\Delta u = \mu - \mu_0 \leq 1 - \mu_0.$$

Since $\Delta\gamma = 1 - \mu_0$, the difference $u - \gamma$ is superharmonic. As u is nonnegative, it follows that $u - \gamma \geq s$. For the reverse inequality, note that $s + \gamma - u$ is superharmonic on the domain $D = \{x \mid \mu(x) = 1\}$ of fully occupied sites and is nonnegative outside D, hence nonnegative inside D as well. □

As a corollary of Lemmas 3.1 and 3.2, we obtain the abelian property of the divisible sandpile, Proposition 1.2, which was stated in the introduction.

We now turn to the case of a point source mass m started at the origin: $\mu_0 = m\delta_o$. More general starting distributions are treated in [16], where we identify the scaling limit of the divisible sandpile model and show that it coincides with that of internal DLA and of the rotor-router model. In the case of a point source of mass m, the natural question is to identify the shape of the resulting domain D_m of fully occupied sites, i.e. sites x for which $\mu(x) = 1$. According to Theorem 3.3, D_m is extremely close to a ball of volume m; in fact, the error in the radius is a constant independent of m. As before, for $r \geq 0$ we write

$$B_r = \left\{x \in \mathbb{Z}^d \ : \ |x| < r\right\}$$

for the lattice ball of radius r centered at the origin.

Theorem 3.3 *For $m \geq 0$ let $D_m \subset \mathbb{Z}^d$ be the domain of fully occupied sites for the divisible sandpile formed from a pile of size m at the origin. There exist constants c, c' depending only on d, such that*

$$B_{r-c} \subset D_m \subset B_{r+c'},$$

where $r = (m/\omega_d)^{1/d}$ and ω_d is the volume of the unit ball in $\mathbb{R}^d$.

The idea of the proof is to use Lemma 3.2 along with the basic estimates on γ, Lemmas 2.1 and 2.2, to obtain estimates on the odometer function

$$u(x) = \text{total mass emitted from } x.$$

We will need the following simple observation.

Lemma 3.4 *For every point $x \in D_m - \{o\}$ there is a path $x = x_0 \sim x_1 \sim \ldots \sim x_k = o$ in D_m with $u(x_{i+1}) \geq u(x_i) + 1$.*

Proof If $x_i \in D_m - \{o\}$, let x_{i+1} be a neighbor of x_i maximizing $u(x_{i+1})$. Then $x_{i+1} \in D_m$ and

$$\begin{aligned} u(x_{i+1}) &\geq \frac{1}{2d} \sum_{y \sim x_i} u(y) \\ &= u(x_i) + \Delta u(x_i) \\ &= u(x_i) + 1, \end{aligned}$$

where in the last step we have used the fact that $x_i \in D_m$. $\square$

Proof of Theorem 3.3 We first treat the inner estimate. Let γ_d be given by Eq. 3. By Lemma 3.2 the function $u - \gamma_d$ is superharmonic, so its minimum in the ball B_r is attained on the boundary. Since $u \geq 0$, we have by Lemma 2.2

$$u(x) - \gamma_d(x) \geq -C, \qquad x \in \partial B_r$$

for a constant C depending only on d. Hence by Lemma 2.1,

$$u(x) \geq (r - |x|)^2 - C' r^d / |x|^d, \qquad x \in B_r. \tag{10}$$

for a constant C' depending only on d. It follows that there is a constant c, depending only on d, such that $u(x) > 0$ whenever $r/3 \leq |x| < r - c$. Thus $B_{r-c} - B_{r/3} \subset D_m$. For $x \in B_{r/3}$, by Lemma 2.3 we have $u(x) \geq r^2/4 - C > 0$, hence $B_{r/3} \subset D_m$.

For the outer estimate, note that $u - \gamma_d$ is harmonic on D_m. By Lemma 2.4 we have $\gamma_d \geq -a$ everywhere, where a depends only on d. Since u vanishes on ∂D_m it follows that $u - \gamma_d \leq a$ on D_m. Now for any $x \in D_m$ with $r - 1 < |x| \leq r$, we have by Lemma 2.2

$$u(x) \leq \gamma_d(x) + a \leq c'$$

for a constant c' depending only on d. Lemma 3.4 now implies that $D_m \subset B_{r+c'+1}$. $\square$

4 Classical Sandpile

We consider a generalization of the classical abelian sandpile, proposed by Fey and Redig [5]. Each site in $\mathbb{Z}^d$ begins with a "hole" of depth H. Thus, each site absorbs the first H grains it receives, and thereafter functions normally, toppling once for each additional $2d$ grains it receives. If H is negative, we can interpret this as saying that every site starts with $h = -H$ grains of sand already present. Aggregation is only well-defined in the regime $h \leq 2d - 2$, since for $h = 2d - 1$ the addition of a single grain already causes every site in $\mathbb{Z}^d$ to topple infinitely often.

Let $S_{n,H}$ be the set of sites that are visited if n particles start at the origin in $\mathbb{Z}^d$. Fey and Redig [5, Theorem 4.7] prove that

$$\lim_{H \to \infty} \limsup_{n \to \infty} \frac{H}{n} \# \left(S_{n,H} \triangle B_{H^{-1/d} r} \right) = 0,$$

where $n = \omega_d r^d$, and $\triangle$ denotes symmetric difference. The following theorem strengthens this result.

Theorem 4.1 *Fix an integer $H \geq 2 - 2d$. Let $S_n = S_{n,H}$ be the set of sites that are visited by the classical abelian sandpile model in $\mathbb{Z}^d$, starting from n particles at the origin, if every lattice site begins with a hole of depth H. Write $n = \omega_d r^d$. Then*

$$B_{c_1 r - c_2} \subset S_{n,H}$$

where

$$c_1 = (2d - 1 + H)^{-1/d}$$

and c_2 is a constant depending only on d. Moreover if $H \geq 1 - d$, then for any $\epsilon > 0$ we have

$$S_{n,H} \subset B_{c_1' r + c_2'}$$

where

$$c_1' = (d - \epsilon + H)^{-1/d}$$

and c_2' is independent of n but may depend on d, H and ϵ.

Note that the ratio $c_1/c_1' \uparrow 1$ as $H \uparrow \infty$. Thus, the classical abelian sandpile run from an initial state in which each lattice site starts with a deep hole yields a shape very close to a ball. Intuitively, one can think of the classical sandpile with deep holes as approximating the divisible sandpile, whose limiting shape is a ball by Theorem 3.3. Following this intuition, we can adapt the proof of Theorem 3.3 to prove Theorem 4.1; just one additional averaging trick is needed, which we explain below.

Consider the odometer function for the abelian sandpile

$$u(x) = \text{total number of grains emitted from } x.$$

Let $T_n = \{x | u(x) > 0\}$ be the set of sites which topple at least once. Then

$$T_n \subset S_n \subset T_n \cup \partial T_n.$$

In the final state, each site which has toppled retains between 0 and $2d - 1$ grains, in addition to the H that it absorbed. Hence

$$H \leq \Delta u(x) + n\delta_{ox} \leq 2d - 1 + H, \qquad x \in T_n. \tag{11}$$

We can improve the lower bound by averaging over a small box. For $x \in \mathbb{Z}^d$ let

$$Q_k(x) = \left\{ y \in \mathbb{Z}^d \,:\, ||x - y||_\infty \leq k \right\}$$

be the box of side length $2k + 1$ centered at x, and let

$$u^{(k)}(x) = (2k + 1)^{-d} \sum_{y \in Q_k(x)} u(y).$$

Write

$$T_n^{(k)} = \{x \mid Q_k(x) \subset T_n\}.$$

Le Borgne and Rossin [12] observe that if T is a set of sites all of which topple, the number of grains remaining in T is at least the number of edges internal to T: indeed, for each internal edge, the endpoint that topples last sends the other a grain which never moves again. Since the box $Q_k(x)$ has $2dk(2k+1)^{d-1}$ internal edges, we have

$$\Delta u^{(k)}(x) \geq \frac{2k}{2k+1} d + H - \frac{n}{(2k+1)^d} \mathbf{1}_{Q_k(o)}(x), \qquad x \in T_n^{(k)}. \tag{12}$$

The following lemma is analogous to Lemma 3.4.

Lemma 4.2 *For every point $x \in T_n$ adjacent to ∂T_n there is a path $x = x_0 \sim x_1 \sim \ldots \sim x_m = o$ in T_n with $u(x_{i+1}) \geq u(x_i) + 1$.*

Proof By Eq. 11 we have

$$\frac{1}{2d} \sum_{y \sim x_i} u(y) \geq u(x_i).$$

Since $u(x_{i-1}) < u(x_i)$, some term $u(y)$ in the sum above must exceed $u(x_i)$. Let $x_{i+1} = y$. □

Proof of Theorem 4.1 Let

$$\widetilde{\xi}_d(x) = (2d - 1 + H)|x|^2 + ng(x),$$

and let

$$\xi_d(x) = \widetilde{\xi}_d(x) - \widetilde{\xi}_d(\lfloor c_1 r \rfloor e_1).$$

Taking $m = n/(2d - 1 + H)$ in Lemma 2.2, we have

$$u(x) - \xi_d(x) \geq -\xi_d(x) \geq -C(2d - 1 + H), \qquad x \in \partial B_{c_1 r} \tag{13}$$

for a constant C depending only on d. By Eq. 11, $u - \xi_d$ is superharmonic, so $u - \xi_d \geq -C(2d - 1 + H)$ in all of $B_{c_1 r}$. Hence by Lemma 2.1 we have for $x \in B_{c_1 r}$

$$u(x) \geq (2d - 1 + H) \left((c_1 r - |x|)^2 - C'(c_1 r)^d / |x|^d\right), \tag{14}$$

where C' depends only on d. It follows that u is positive on $B_{c_1 r - c_2} - B_{c_1 r/3}$ for a suitable constant c_2 depending only on d. For $x \in B_{c_1 r/3}$, by Lemma 2.3 we have $u(x) > (2d - 1 + H)(c_1^2 r^2/4 - C) > 0$. Thus $B_{c_1 r - c_2} \subset T_n \subset S_n$.

For the outer estimate, let

$$\hat{\psi}_d(x) = (d - \epsilon + H)|x|^2 + ng(x).$$

Choose k large enough so that $\frac{2k}{2k+1} d \geq d - \epsilon$, and define

$$\widetilde{\psi}_d(x) = (2k+1)^{-d} \sum_{y \in Q_k(x)} \hat{\psi}_d(y).$$

Finally, let

$$\psi_d(x) = \widetilde{\psi}_d(x) - \widetilde{\psi}_d\left(\lfloor c_1' r \rfloor e_1\right).$$

By Eq. 12, $u^{(k)} - \psi_d$ is subharmonic on $T_n^{(k)}$. Taking $m = n/(d - \epsilon + H)$ in Lemma 2.4, there is a constant a depending only on d, such that $\psi_d \geq -a(d + H)$ everywhere.

Since $u^{(k)} \le (2d+H)^{(d+1)k}$ on $\partial T_n^{(k)}$ it follows that $u^{(k)} - \psi_d \le a(d+H) + (2d+H)^{(d+1)k}$ on $T_n^{(k)}$. Now for any $x \in S_n$ with $c_1'r - 1 < |x| \le c_1'r$ we have by Lemma 2.2

$$u^{(k)}(x) \le \psi_d(x) + a(d+H) + (2d+H)^{(d+1)k} \le \tilde{c}_2$$

for a constant $\tilde{c}_2$ depending only on d, H and ϵ. Then $u(x) \le c_2' := (2k+1)^d \tilde{c}_2$. Lemma 4.2 now implies that $T_n \subset B_{c_1'r+c_2'}$, and hence

$$S_n \subset T_n \cup \partial T_n \subset B_{c_1'r+c_2'+1}. \qquad \square$$

We remark that the crude bound of $(2d+H)^{(d+1)k}$ used in the proof of the outer estimate can be improved to a bound of order $k^2 H$, and the final factor of $(2k+1)^d$ can be replaced by a constant factor independent of k and H, using the fact that a nonnegative function on $\mathbb{Z}^d$ with bounded Laplacian cannot grow faster than quadratically; see [16, sec. 2].

5 Rotor-Router Model

Given a function f on $\mathbb{Z}^d$, for a directed edge (x, y) write

$$\nabla f(x, y) = f(y) - f(x).$$

Given a function s on directed edges in $\mathbb{Z}^d$, write

$$\operatorname{div} s(x) = \frac{1}{2d} \sum_{y \sim x} s(x, y).$$

The discrete Laplacian of f is then given by

$$\Delta f(x) = \operatorname{div} \nabla f = \frac{1}{2d} \sum_{y \sim x} f(y) - f(x).$$

5.1 Inner Estimate

Fixing $n \ge 1$, consider the odometer function for rotor-router aggregation

$$u(x) = \text{total number exits from } x \text{ by the first } n \text{ particles.}$$

We learned the idea of using the odometer function to study the rotor-router shape from Matt Cook [2].

Lemma 5.1 *For a directed edge (x, y) in $\mathbb{Z}^d$, denote by $\kappa(x, y)$ the net number of crossings from x to y performed by the first n particles in rotor-router aggregation. Then*

$$\nabla u(x, y) = -2d\kappa(x, y) + R(x, y) \tag{15}$$

for some edge function R which satisfies

$$|R(x, y)| \le 4d - 2$$

for all edges (x, y).

Remark In the more general setting of rotor stacks of bounded discrepancy, the $4d - 2$ will be replaced by a different constant here.

Proof Writing $N(x, y)$ for the number of particles routed from x to y, we have

$$\frac{u(x) - 2d + 1}{2d} \leq N(x, y) \leq \frac{u(x) + 2d - 1}{2d}$$

hence

$$\begin{aligned} |\nabla u(x, y) + 2d\kappa(x, y)| &= |u(y) - u(x) + 2dN(x, y) - 2dN(y, x)| \\ &\leq 4d - 2. \end{aligned}$$ □

In what follows, $C_0, C_1, \ldots$ denote constants depending only on d.

Lemma 5.2 *Let $\Omega \subset \mathbb{Z}^d - \{o\}$ with $2 \leq \#\Omega < \infty$. Then*

$$\sum_{y \in \Omega} |y|^{1-d} \leq C_0 \operatorname{Diam}(\Omega).$$

Proof For each positive integer k, let

$$\mathcal{S}_k = \left\{ y \in \mathbb{Z}^d \, : \, k \leq |y| < k + 1 \right\}.$$

Then

$$\sum_{y \in \mathcal{S}_k} |y|^{1-d} \leq k^{1-d} \# \mathcal{S}_k \leq C_0'$$

for a constant C_0' depending only on d. Since Ω can intersect at most $\operatorname{Diam}(\Omega) + 1 \leq 2 \operatorname{Diam}(\Omega)$ distinct sets $\mathcal{S}_k$, taking $C_0 = 2C_0'$ the proof is complete. □

Lemma 5.3 *Let $G = G_{B_r}$ be the Green's function for simple random walk in $\mathbb{Z}^d$ stopped on exiting B_r. For any $\rho \geq 1$ and $x \in B_r$,*

$$\sum_{\substack{y \in B_r \\ |x-y| \leq \rho}} \sum_{z \sim y} |G(x, y) - G(x, z)| \leq C_1 \rho. \tag{16}$$

Proof Let $(X_t)_{t \geq 0}$ denote simple random walk in $\mathbb{Z}^d$, and let T be the first exit time from B_r. For fixed y, the function

$$A(x) = g(x - y) - \mathbb{E}_x g(X_T - y) \tag{17}$$

has Laplacian $\Delta A(x) = -\delta_{xy}$ in B_r and vanishes on ∂B_r, hence $A(x) = G(x, y)$.

Let $x, y \in B_r$ and $z \sim y$. From Eq. 4 we have

$$|g(x - y) - g(x - z)| \leq \frac{C_2}{|x - y|^{d-1}}, \qquad y, z \neq x.$$

Using the triangle inequality together with Eq. 17, we obtain

$$\begin{aligned}|G(x,y)-G(x,z)| &\le |g(x-y)-g(x-z)| + \mathbb{E}_x|g(X_T-y)-g(X_T-z)| \\ &\le \frac{C_2}{|x-y|^{d-1}} + \sum_{w\in\partial B_r} H_x(w)\frac{C_2}{|w-y|^{d-1}},\end{aligned}$$

where $H_x(w) = \mathbb{P}_x(X_T = w)$.

Write $D = \{y \in B_r : |x-y| \le\}$. Then

$$\sum_{\substack{y\in D\\ y\neq x}}\sum_{\substack{z\sim y\\ z\neq x}} |G(x,y)-G(x,z)| \le C_3\rho + C_2\sum_{w\in\partial B_r} H_x(w)\sum_{y\in D}|w-y|^{1-d}. \tag{18}$$

Taking $\Omega = w - D$ in Lemma 5.2, the inner sum on the right is at most $C_0\mathrm{Diam}(D) \le 2C_0\rho$, so the right side of Eq. 18 is bounded above by $C_1\rho$ for a suitable C_1.

Finally, the terms in which y or z coincides with x make a negligible contribution to the sum in Eq. 16, since for $y \sim x \in \mathbb{Z}^d$

$$|G(x,x)-G(x,y)| \le |g(o)-g(x-y)| + \mathbb{E}_x|g(X_T-x)-g(X_T-y)| \le C_4. \qquad \square$$

Lemma 5.4 *Let H_1, H_2 be linear half-spaces in $\mathbb{Z}^d$, not necessarily parallel to the coordinate axes. Let T_i be the first hitting time of H_i. If $x \notin H_1 \cup H_2$, then*

$$\mathbb{P}_x(T_1 > T_2) \le \frac{5}{2}\frac{h_1+1}{h_2}\left(1+\frac{1}{2h_2}\right)^2$$

where h_i is the distance from x to H_i.

Proof If one of H_1, H_2 contains the other, the result is vacuous. Otherwise, let $\widetilde{H}_i$ be the half-space shifted parallel to H_i^c by distance $2h_2$ in the direction of x (Fig. 4), and let $\widetilde{T}_i$ be the first hitting time of $H_i \cup \widetilde{H}_i$. Let $(X_t)_{t\ge 0}$ denote simple random walk in $\mathbb{Z}^d$, and write M_t for the (signed) distance from X_t to the hyperplane defining the

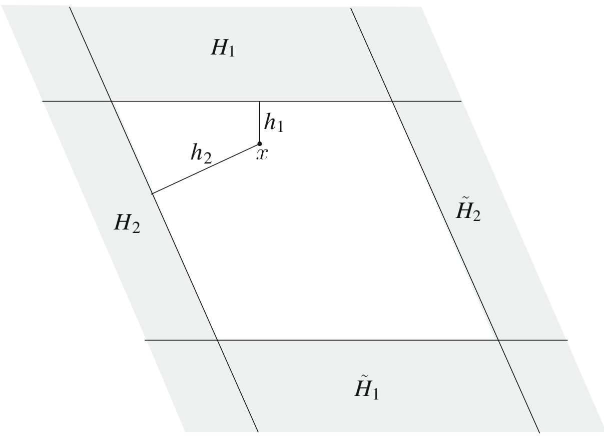


Fig. 4 Diagram for the Proof of Lemma 5.4

boundary of H_1, with $M_0 = h_1$. Then M_t is a martingale with bounded increments. Since $\mathbb{E}_x \widetilde{T}_1 < \infty$, we obtain from optional stopping

$$h_1 = \mathbb{E}_x M_{\widetilde{T}_1} \geq 2h_2 \, \mathbb{P}_x \left(X_{\widetilde{T}_1} \in \widetilde{H}_1 \right) - \mathbb{P}_x \left(X_{\widetilde{T}_1} \in H_1 \right),$$

hence

$$\mathbb{P}_x \left(X_{\widetilde{T}_1} \in \widetilde{H}_1 \right) \leq \frac{h_1 + 1}{2h_2}. \tag{19}$$

Likewise, $dM_t^2 - t$ is a martingale with bounded increments, giving

$$\begin{aligned} \mathbb{E}_x \, \widetilde{T}_1 &\leq d \, \mathbb{E}_x \, M^2_{\widetilde{T}_1} \\ &\leq d(2h_2 + 1)^2 \, \mathbb{P}_x \left(X_{\widetilde{T}_1} \in \widetilde{H}_1 \right) \\ &\leq d(h_1 + 1)(2h_2 + 1) \left(1 + \frac{1}{2h_2} \right). \end{aligned} \tag{20}$$

Let $T = \min(\widetilde{T}_1, \widetilde{T}_2)$. Denoting by D_t the distance from X_t to the hyperplane defining the boundary of H_2, the quantity

$$N_t = \frac{d}{2} \left(D_t^2 + (2h_2 - D_t)^2 \right) - t$$

is a martingale. Writing $p = \mathbb{P}_x(T = \widetilde{T}_2)$ we have

$$\begin{aligned} dh_2^2 = \mathbb{E} N_0 = \mathbb{E} N_T &\geq p \frac{d}{2} (2h_2)^2 + (1 - p) dh_2^2 - \mathbb{E}_x T \\ &\geq (1 + p) dh_2^2 - \mathbb{E}_x T \end{aligned}$$

hence by Eq. 20

$$p \leq \frac{\mathbb{E}_x T}{dh_2^2} \leq 2 \frac{h_1 + 1}{h_2} \left(1 + \frac{1}{2h_2} \right)^2.$$

Finally by Eq. 19

$$\mathbb{P}(T_1 > T_2) \leq p + \mathbb{P} \left(X_{\widetilde{T}_1} \in \widetilde{H}_1 \right) \leq \frac{5}{2} \frac{h_1 + 1}{h_2} \left(1 + \frac{1}{2h_2} \right)^2. \qquad \square$$

Lemma 5.5 *Let $x \in B_r$ and let $\rho = r + 1 - |x|$. Let*

$$\mathcal{S}_k^* = \left\{ y \in B_r \ : \ 2^k \rho < |x - y| \leq 2^{k+1} \rho \right\}. \tag{21}$$

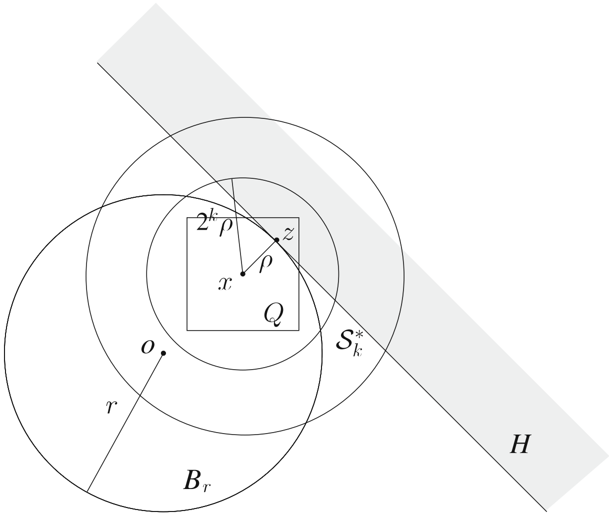
Let τ_k be the first hitting time of $\mathcal{S}_k^*$, and T the first exit time from B_r. Then

$$\mathbb{P}_x(\tau_k < T) \leq C_2 2^{-k}.$$

Proof Let H be the outer half-space tangent to B_r at the point $z \in \partial B_r$ closest to x. Let Q be the cube of side length $2^k \rho / \sqrt{d}$ centered at x (Fig. 5). Then Q is disjoint from $\mathcal{S}_k^*$, hence

$$\mathbb{P}_x(\tau_k < T) \leq \mathbb{P}_x(T_{\partial Q} < T) \leq \mathbb{P}_x(T_{\partial Q} < T_H)$$

Fig. 5 Diagram for the proof of Lemma 5.5



where $T_{\partial Q}$ and T_H are the first hitting times of ∂Q and H. Let $H_1, \ldots, H_{2d}$ be the outer half-spaces defining the faces of Q, so that $Q = H_1^c \cap \ldots \cap H_{2d}^c$. By Lemma 5.4 we have

$$\begin{aligned}\mathbb{P}_x(T_{\partial Q} < T_H) &\leq \sum_{i=1}^{2d} \mathbb{P}_x(T_{H_i} < T_H) \\ &\leq \frac{5}{2}\sum_{i=1}^{2d} \frac{\operatorname{dist}(x, H)+1}{\operatorname{dist}(x, H_i)}\left(1+\frac{1}{2\operatorname{dist}(x, H_i)}\right)^2.\end{aligned}$$

Since $\operatorname{dist}(x, H) = |x - z| \leq \rho$ and $\operatorname{dist}(x, H_i) = 2^{k-1}\rho/\sqrt{d}$, and $\rho \geq 1$, taking $C_2 = 20\, d^{3/2}(1+\sqrt{d})^2$ completes the proof. □

Lemma 5.6 *Let $G = G_{B_r}$ be the Green's function for random walk stopped on exiting B_r. Let $x \in B_r$ and let $\rho = r + 1 - |x|$. Then*

$$\sum_{y \in B_r} \sum_{z \sim y} |G(x, y) - G(x, z)| \leq C_3 \rho \log \frac{r}{\rho}.$$

Proof Let $\mathcal{S}_k^*$ be given by Eq. 21, and let

$$W = \left\{w \in \partial\left(\mathcal{S}_k^* \cup \partial\mathcal{S}_k^*\right) \ : \ |w - x| < 2^k\rho\right\}$$

be the portion of the boundary of the enlarged spherical shell $\mathcal{S}_k^* \cup \partial\mathcal{S}_k^*$ lying closer to x. Let τ_W be the first hitting time of W, and T the first exit time from B_r. For $w \in W$ let

$$H_x(w) = \mathbb{P}_x(X_{\tau_W \wedge T} = w).$$

For any $y \in \mathcal{S}_k^*$ and $z \sim y$, simple random walk started at x must hit W before hitting either y or z, hence

$$|G(x,y) - G(x,z)| \le \sum_{w \in W} H_x(w) |G(w,y) - G(w,z)|.$$

For any $y \in \mathcal{S}_k^*$ and any $w \in W$ we have

$$|y - w| \le |y - x| + |w - x| \le 3 \cdot 2^k \rho.$$

Lemma 5.3 yields

$$\sum_{y \in \mathcal{S}_k^*} \sum_{z \sim y} |G(x,y) - G(x,z)| \le 3C_1 2^k \rho \sum_{w \in W} H_x(w).$$

By Lemma 5.5 we have $\sum_{w \in W} H_x(w) \le C_2 2^{-k}$, so the above sum is at most $3C_1C_2\rho$. Since the union of shells $\mathcal{S}_0^*, \mathcal{S}_1^*, \ldots, \mathcal{S}_{\lceil \log_2(r/\rho) \rceil}^*$ covers all of B_r except for those points y within distance ρ of x, and $\sum_{|y-x| \le \rho} \sum_{z \sim y} |G(x,y) - G(x,z)| \le C_1 \rho$ by Lemma 5.3, the result follows. □

Proof of Theorem 1.1, Inner Estimate Let κ and R be defined as in Lemma 5.1. Since the net number of particles to enter a site $x \neq o$ is at most one, we have $2d \operatorname{div} \kappa(x) \ge -1$. Likewise $2d \operatorname{div} \kappa(o) = n - 1$. Taking the divergence in Eq. 15, we obtain

$$\Delta u(x) \le 1 + \operatorname{div} R(x), \qquad x \neq o; \tag{22}$$

$$\Delta u(o) = 1 - n + \operatorname{div} R(o). \tag{23}$$

Let T be the first exit time from B_r, and define

$$f(x) = \mathbb{E}_x u(X_T) - \mathbb{E}_x T + n\, \mathbb{E}_x \#\{j < T | X_j = 0\}.$$

Then $\Delta f(x) = 1$ for $x \in B_r - \{o\}$ and $\Delta f(o) = 1 - n$. Moreover $f \ge 0$ on ∂B_r. It follows from Lemma 2.2 with $m = n$ that $f \ge \gamma - C_4$ on B_r for a suitable constant C_4.

We have

$$u(x) - \mathbb{E}_x u(X_T) = \sum_{k \ge 0} \mathbb{E}_x \Big(u(X_{k \wedge T}) - u(X_{(k+1) \wedge T}) \Big).$$

Each summand on the right side is zero on the event $\{T \le k\}$, hence

$$\mathbb{E}_x \Big(u(X_{k \wedge T}) - u(X_{(k+1) \wedge T}) \mid \mathcal{F}_{k \wedge T} \Big) = -\Delta u(X_k) 1_{\{T > k\}}.$$

Taking expectations and using Eqs. 22 and 23, we obtain

$$\begin{aligned} u(x) - \mathbb{E}_x u(X_T) &\ge \sum_{k \ge 0} \mathbb{E}_x \left[1_{\{T>k\}} \left(n 1_{\{X_k = o\}} - 1 - \operatorname{div} R(X_k) \right) \right] \\ &= n\, \mathbb{E}_x \#\{k < T | X_k = o\} - \mathbb{E}_x T - \sum_{k \ge 0} \mathbb{E}_x \left[1_{\{T>k\}} \operatorname{div} R(X_k) \right], \end{aligned}$$

hence

$$u(x) - f(x) \ge -\frac{1}{2d} \sum_{k \ge 0} \mathbb{E}_x \left[1_{\{T>k\}} \sum_{z \sim X_k} R(X_k, z) \right]. \tag{24}$$

Since random walk exits B_r with probability at least $\frac{1}{2d}$ every time it reaches a site adjacent to the boundary ∂B_r, the expected time spent adjacent to the boundary before time T is at most $2d$. Since $|R| \le 4d$, the terms in Eq. 24 with $z \in \partial B_r$ contribute at most $16d^3$ to the sum. Thus

$$u(x) - f(x) \ge -\frac{1}{2d} \sum_{k \ge 0} \mathbb{E}_x \left[\sum_{\substack{y,z \in B_r \\ y \sim z}} 1_{\{T>k\} \cap \{X_k = y\}} R(y, z) \right] - 8d^2.$$

For $y \in B_r$ we have $\{X_k = y\} \cap \{T > k\} = \{X_{k \wedge T} = y\}$, hence

$$u(x) - f(x) \ge -\frac{1}{2d} \sum_{k \ge 0} \sum_{\substack{y,z \in B_r \\ y \sim z}} \mathbb{P}_x(X_{k \wedge T} = y) R(y, z) - 8d^2. \tag{25}$$

Write $p_k(y) = \mathbb{P}_x(X_{k \wedge T} = y)$. Note that since ∇u and κ are antisymmetric, R is antisymmetric. Thus

$$\begin{aligned} \sum_{\substack{y,z \in B_r \\ y \sim z}} p_k(y) R(y, z) &= - \sum_{\substack{y,z \in B_r \\ y \sim z}} p_k(z) R(y, z) \\ &= \sum_{\substack{y,z \in B_r \\ y \sim z}} \frac{p_k(y) - p_k(z)}{2} R(y, z). \end{aligned}$$

Summing over k and using the fact that $|R| \le 4d$, we conclude from Eq. 25 that

$$u(x) \ge f(x) - \sum_{\substack{y,z \in B_r \\ y \sim z}} |G(x, y) - G(x, z)| - 8d^2,$$

where $G = G_{B_r}$ is the Green's function for simple random walk stopped on exiting B_r. By Lemma 5.6 we obtain

$$u(x) \ge f(x) - C_3(r + 1 - |x|) \log \frac{r}{r + 1 - |x|} - 8d^2.$$

Using the fact that $f \ge \gamma - C_4$, we obtain from Lemma 2.1

$$u(x) \ge (r - |x|)^2 - C_3(r + 1 - |x|) \log \frac{r}{r + 1 - |x|} + O\left(\frac{r^d}{|x|^d}\right).$$

The right side is positive provided $r/3 \le |x| < r - C_5 \log r$. For $x \in B_{r/3}$, by Lemma 2.3 we have $u(x) > r^2/4 - C_3 r \log \frac{3}{2} > 0$, hence $B_{r - C_5 \log r} \subset A_n$. □

5.2 Outer Estimate

The following result is due to Holroyd and Propp (unpublished); we include a proof for the sake of completeness. Notice that the bound in Eq. 26 does not depend on the number of particles.

Proposition 5.7 *Let Γ be a finite connected graph, and let $Y \subset Z$ be subsets of the vertex set of Γ. Let s be a nonnegative integer-valued function on the vertices of Γ. Let $H_w(s, Y)$ be the expected number of particles stopping in Y if $s(x)$ particles start at each* *vertex x and perform independent simple random walks stopped on first hitting Z. Let $H_r(s, Y)$ be the number of particles stopping in Y if $s(x)$ particles start at each vertex x and perform rotor-router walks stopped on first hitting Z. Let $H(x) = H_w(1_x, Y)$. Then*

$$|H_r(s, Y) - H_w(s, Y)| \leq \sum_{u \notin Z} \sum_{v \sim u} |H(u) - H(v)| \tag{26}$$

independent of s and the initial positions of the rotors.

Proof For each vertex $u \notin Z$, arbitrarily choose a neighbor $\eta(u)$. Order the neighbors $\eta(u) = v_1, v_2, \ldots, v_d$ of u so that the rotor at u points to v_{i+1} immediately after pointing to v_i (indices mod d). We assign *weight* $w(u, \eta(u)) = 0$ to a rotor pointing from u to $\eta(u)$, and weight $w(u, v_i) = H(u) - H(v_i) + w(u, v_{i-1})$ to a rotor pointing from u to v_i. These assignments are consistent since H is a harmonic function: $\sum_i (H(u) - H(v_i)) = 0$. We also assign weight $H(u)$ to a particle located at u. The sum of rotor and particle weights in any configuration is invariant under the operation of routing a particle and rotating the corresponding rotor. Initially, the sum of all particle weights is $H_w(s, Y)$. After all particles have stopped, the sum of the particle weights is $H_r(s, Y)$. Their difference is thus at most the change in rotor weights, which is bounded above by the sum in Eq. 26. □

For $\rho \in \mathbb{Z}$ let

$$\mathcal{S}_\rho = \{x \in \mathbb{Z}^d \ : \ \rho \leq |x| < \rho + 1\}. \tag{27}$$

Then

$$B_\rho = \{x \in \mathbb{Z}^d \ : \ |x| < \rho\} = \mathcal{S}_0 \cup \ldots \cup \mathcal{S}_{\rho-1}.$$

Note that for simple random walk started in B_ρ, the first exit time of B_ρ and first hitting time of $\mathcal{S}_\rho$ coincide. Our next result is a modification of Lemma 5(b) of [10].

Lemma 5.8 *Fix $\rho \geq 1$ and $y \in \mathcal{S}_\rho$. For $x \in B_\rho$ let $H(x) = \mathbb{P}_x(X_T = y)$, where T is the first hitting time of $\mathcal{S}_\rho$. Then*

$$H(x) \leq \frac{J}{|x - y|^{d-1}} \tag{28}$$

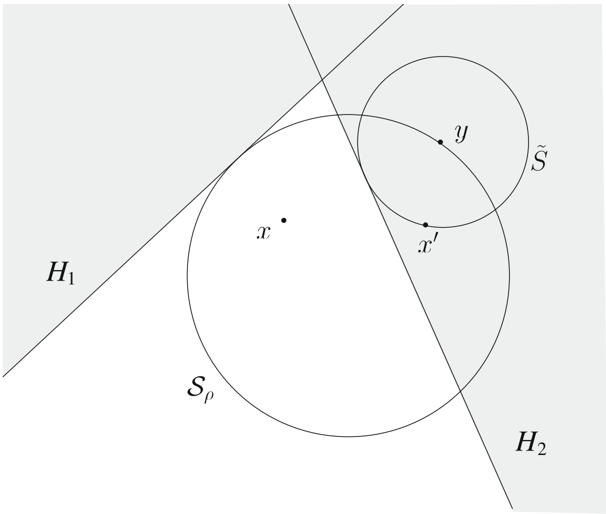
for a constant J depending only on d.

Proof We induct on the distance $|x - y|$, assuming the result holds for all x' with $|x' - y| \leq \frac{1}{2}|x - y|$; the base case can be made trivial by choosing J sufficiently large. By Lemma 5(b) of [10], we can choose J large enough so that the result holds provided $|y| - |x| \geq 2^{-d-3}|x - y|$. Otherwise, let H_1 be the outer half-space tangent to $\mathcal{S}_\rho$ at the point of $\mathcal{S}_\rho$ closest to x, and let H_2 be the inner half-space tangent to the ball $\widetilde{S}$ of radius $\frac{1}{2}|x - y|$ about y, at the point of $\widetilde{S}$ closest to x (Fig. 6). By Lemma 5.4 applied to these half-spaces, the probability that random walk started at x reaches $\widetilde{S}$ before hitting $\mathcal{S}_\rho$ is at most 2^{1-d}. Writing $\widetilde{T}$ for the first hitting time of $\widetilde{S} \cup \mathcal{S}_\rho$, we have

$$H(x) \leq \sum_{x' \in \widetilde{S}} \mathbb{P}_x \left(X_{\widetilde{T}} = x' \right) H\left(x'\right) \leq 2^{1-d} J \cdot \left(\frac{|x - y|}{2} \right)^{1-d}$$

where we have used the inductive hypothesis to bound $H(x')$. □

Fig. 6 Diagram for the proof of Lemma 5.8



The *lazy* random walk in $\mathbb{Z}^d$ stays in place with probability $\frac{1}{2}$, and moves to each of the $2d$ neighbors with probability $\frac{1}{4d}$. We will need the following standard result, which can be derived e.g. from the estimates in [17], section II.12; we include a proof for the sake of completeness.

Lemma 5.9 *Given $u \sim v \in \mathbb{Z}^d$, lazy random walks started at u and v can be coupled with probability $1 - C/R$ before either reaches distance R from u, where C depends only on d.*

Proof Let i be the coordinate such that $u_i \neq v_i$. To define a step of the coupling, choose one of the d coordinates uniformly at random. If the chosen coordinate is different from i, let the two walks take the same lazy step so that they still agree in this coordinate. If the chosen coordinate is i, let one walk take a step while the other stays in place. With probability $\frac{1}{2}$ the walks will then be coupled. Otherwise, they are located at points u', v' with $|u' - v'| = 2$. Moreover, $\mathbb{P}\big(|u - u'| \geq \frac{R}{2\sqrt{d}}\big) < \frac{C'}{R}$ for a constant C' depending only on d. From now on, whenever coordinate i is chosen, let the two walks take lazy steps in opposite directions.

Let

$$H_1 = \left\{ x \,\middle|\, x_i = \frac{u'_i + v'_i}{2} \right\}$$

be the hyperplane bisecting the segment $[u', v']$. Since the steps of one walk are reflections in H_1 of the steps of the other, the walks couple when they hit H_1. Let Q be the cube of side length $R/\sqrt{d} + 2$ centered at u, and let H_2 be a hyperplane defining one of the faces of Q. By Lemma 5.4 with $h_1 = 1$ and $h_2 = R/4\sqrt{d}$, the probability that one of the walks exits Q before the walks couple is at most $2d \cdot \frac{5}{2}\frac{h_1+1}{h_2}\left(1+\frac{1}{2h_2}\right)^2 \leq 40\, d^{3/2}\big(1+2\sqrt{d}\big)^2/R$. □

Lemma 5.10 *With H defined as in Lemma* 5.8*, we have*

$$\sum_{u\in B_\rho}\sum_{v\sim u}|H(u)-H(v)| \leq J'\log\rho$$

for a constant J' depending only on d.

Proof Given $u \in B_\rho$ and $v \sim u$, by Lemma 5.9, lazy random walks started at u and v can be coupled with probability $1 - 2C/|u-y|$ before either reaches distance $|u-y|/2$ from u. If the walks reach this distance without coupling, by Lemma 5.8 each has still has probability at most $J/|u-y|^{d-1}$ of exiting B_ρ at y. By the strong Markov property it follows that

$$|H(u)-H(v)| \leq \frac{2CJ}{|u-y|^d}.$$

Summing in spherical shells about y, we obtain

$$\sum_{u\in B_\rho}\sum_{v\sim u}|H(u)-H(v)| \leq \sum_{t=1}^{\rho} d\omega_d t^{d-1}\frac{2CJ}{t^d} \leq J'\log\rho. \qquad \square$$

We remark that Lemma 5.10 could also be inferred from Lemma 5.8 using [9, Thm. 1.7.1] in a ball of radius $|u-y|/2$ about u.

To prove the outer estimate of Theorem 1.1, we will make use of the *abelian property* of rotor-router aggregation. Fix a finite set $\Gamma \subset \mathbb{Z}^d$ containing the origin. Starting with n particles at the origin, at each time step, choose a site $x \in \Gamma$ with more than one particle, rotate the rotor at x, and move one particle from x to the neighbor the rotor points to. After a finite number of such choices, each site in Γ will have at most one particle, and all particles that exited Γ will be on the boundary $\partial\Gamma$. The abelian property says that *the final configuration of particles and the final configuration of rotors do not depend on the choices*. For a proof, see [4, Prop. 4.1].

In our application, we will fix $\rho \geq r$ and stop each particle in rotor-router aggregation either when it reaches an unoccupied site or when it reaches the spherical shell $\mathcal{S}_\rho$. Let N_ρ be the number of particles that reach $\mathcal{S}_\rho$ during this process. Note that at some sites in $\mathcal{S}_\rho$, more than one particle may have stopped. If we let each of these extra particles in turn continue performing rotor-router walk, stopping either when it reaches an unoccupied site or when it hits the larger shell $\mathcal{S}_{\rho+h}$, then by the abelian property, the number of particles that reach $\mathcal{S}_{\rho+h}$ will be $N_{\rho+h}$. We will show that when h is order $r^{1-1/d}$, a constant fraction of the particles that reach $\mathcal{S}_\rho$ find unoccupied sites before reaching $\mathcal{S}_{\rho+h}$.

Proof of Theorem 1.1, Outer Estimate Fix integers $\rho \geq r$ and $h \geq 1$. In the setting of Proposition 5.7, let Γ be the lattice ball $B_{\rho+h+1}$, and let $Z = \mathcal{S}_{\rho+h}$. Fix $y \in \mathcal{S}_{\rho+h}$ and let $Y = \{y\}$. For $x \in \mathcal{S}_\rho$, let $s(x)$ be the number of particles stopped at x if each particle in rotor-router aggregation is stopped either when it reaches an unoccupied site or when it reaches $\mathcal{S}_\rho$. Write

$$H(x) = \mathbb{P}_x(X_T = y)$$

where T is the first hitting time of $\mathcal{S}_{\rho+h}$. By Lemma 5.8 we have

$$H_w(s, y) = \sum_{x \in \mathcal{S}_\rho} s(x) H(x) \leq \frac{J N_\rho}{h^{d-1}} \tag{29}$$

where

$$N_\rho = \sum_{x \in \mathcal{S}_\rho} s(x)$$

is the number of particles that ever visit the shell $\mathcal{S}_\rho$.

By Lemma 5.10 the sum in Eq. 26 is at most $J' \log h$, hence from Proposition 5.7 and Eq. 29 we have

$$H_r(s, y) \leq \frac{J N_\rho}{h^{d-1}} + J' \log h. \tag{30}$$

Let $\rho(0) = r$, and define $\rho(i)$ inductively by

$$\rho(i+1) = \min\left\{\rho(i) + N_{\rho(i)}^{2/(2d-1)},\ \min\{\rho > \rho(i) | N_\rho \leq N_{\rho(i)}/2\}\right\}. \tag{31}$$

Fixing $h < \rho(i+1) - \rho(i)$, we have

$$h^{d-1} \log h \leq N_{\rho(i)}^{\frac{2d-2}{2d-1}} \log N_{\rho(i)} \leq N_{\rho(i)};$$

so Eq. 30 with $\rho = \rho(i)$ simplifies to

$$H_r(s, y) \leq \frac{C N_{\rho(i)}}{h^{d-1}} \tag{32}$$

where $C = J + J'$.

Since all particles that visit $\mathcal{S}_{\rho(i)+h}$ during rotor-router aggregation must pass through $\mathcal{S}_{\rho(i)}$, we have by the abelian property

$$N_{\rho(i)+h} \leq \sum_{y \in \mathcal{S}_{\rho(i)+h}} H_r(s, y). \tag{33}$$

Let $M_k = \#(A_n \cap \mathcal{S}_k)$. There are at most $M_{\rho(i)+h}$ nonzero terms in the sum on the right side of Eq. 33, and each term is bounded above by Eq. 32, hence

$$M_{\rho(i)+h} \geq N_{\rho(i)+h} \frac{h^{d-1}}{C N_{\rho(i)}} \geq \frac{h^{d-1}}{2C}$$

where the second inequality follows from $N_{\rho(i)+h} \geq N_{\rho(i)}/2$. Summing over h, we obtain

$$\sum_{\rho=\rho(i)+1}^{\rho(i+1)-1} M_\rho \geq \frac{1}{2dC} (\rho(i+1) - \rho(i) - 1)^d. \tag{34}$$

The left side is at most $N_{\rho(i)}$, hence

$$\rho(i+1) - \rho(i) \leq \left(2dCN_{\rho(i)}\right)^{1/d} \leq N_{\rho(i)}^{2/(2d-1)}$$

provided $N_{\rho(i)} \geq C' := (2dC)^{2d-1}$. Thus the minimum in Eq. 31 is not attained by its first argument. It follows that $N_{\rho(i+1)} \leq N_{\rho(i)}/2$, hence $N_{\rho(a \log r)} < C'$ for a sufficiently large constant a.

By the inner estimate, since the ball $B_{r-c\log r}$ is entirely occupied, we have

$$\sum_{\rho \geq r} M_\rho \leq \omega_d r^d - \omega_d (r - c \log r)^d$$
$$\leq cd\omega_d r^{d-1} \log r.$$

Write $x_i = \rho(i+1) - \rho(i) - 1$; by Eq. 34 we have

$$\sum_{i=0}^{a \log r} x_i^d \leq cd\omega_d r^{d-1} \log r,$$

By Jensen's inequality, subject to this constraint, $\sum x_i$ is maximized when all x_i are equal, in which case $x_i \leq C'' r^{1-1/d}$ and

$$\rho(a \log r) = r + \sum x_i \leq r + C'' r^{1-1/d} \log r. \tag{35}$$

Since $N_{\rho(a\log r)} < C'$ we have $N_{\rho(a \log r)+C'} = 0$; that is, no particles reach the shell $\mathcal{S}_{\rho(a\log r)+C'}$. Taking $c' = C' + C''$, we obtain from Eq. 35

$$A_n \subset B_{r(1+c'r^{-1/d}\log r)}. \qquad \square$$

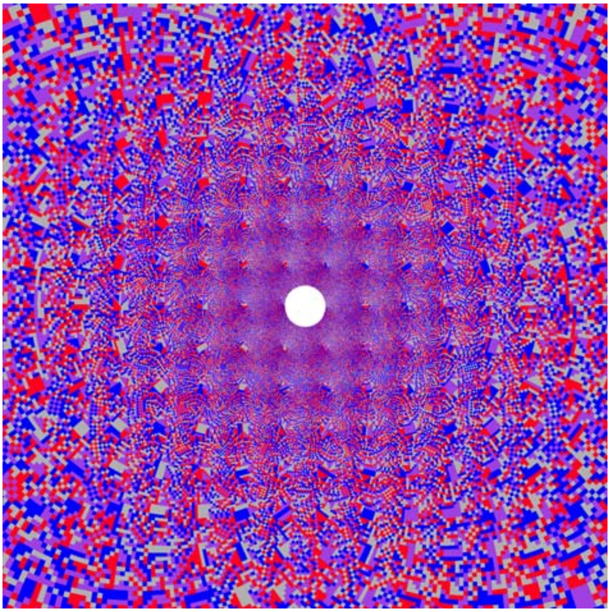


Fig. 7 Image of the rotor-router aggregate of one million particles under the map $z \mapsto 1/z^2$. The *colors* represent the rotor directions. The *white disc* in the center is the image of the complement of the occupied region

6 Concluding Remarks

A number of intriguing questions remain unanswered. Although we have shown that the asymptotic shape of the rotor-router model is a ball, the near perfect circularity found in Fig. 1 remains a mystery. In particular, we do not know whether an analogue of Theorem 1.3 holds for the rotor-router model, with constant error in the radius as the number of particles grows.

Equally mysterious are the patterns in the rotor directions evident in Fig. 1. The rotor directions can be viewed as values of the odometer function mod $2d$, but our control of the odometer is not fine enough to provide useful information about the patterns. If the rescaled occupied region $\sqrt{\pi/n}\,A_n$ is viewed as a subset of the complex plane, it appears that the monochromatic regions visible in Fig. 1, in which all rotors point in the same direction, occur near points of the form $(1+2z)^{-1/2}$, where $z = a + bi$ is a Gaussian integer (i.e. $a, b \in \mathbb{Z}$). We do not even have a heuristic explanation for this phenomenon. Figure 7 shows the image of $A_{1,000,000}$ under the map $z \mapsto 1/z^2$; the monochromatic patches in the transformed region occur at lattice points.

László Lovász (personal communication) has asked whether the occupied region A_n is simply connected, i.e. whether its complement is connected. While Theorem 1.1 shows that A_n cannot have any holes far from the boundary, we cannot answer his question at present.

A final question is whether our methods could be adapted to internal DLA to show that if $n = \omega_d r^d$, then with high probability $B_{r-c\log r} \subset I_n$, where I_n is the internal DLA cluster of n particles. The current best bound is due to Lawler [11], who proves that with high probability $B_{r-r^{1/3}(\log r)^2} \subset I_n$.

Acknowledgements The authors thank Jim Propp for bringing the rotor-router model to our attention and for many fruitful discussions. We also had useful discussions with Oded Schramm, Scott Sheffield and Misha Sodin. We thank Wilfried Huss for pointing out an error in an earlier draft. Yelena Shvets helped draw some of the figures.

References

1. Bak, P., Tang, C., Wiesenfeld, K.: Self-organized criticality: an explanation of the $1/f$ noise. Phys. Rev. Lett. **59**(4), 381–384 (1987)
2. Cook, M.: The tent metaphor. http://paradise.caltech.edu/~cook/Warehouse/ForPropp/ (2002)
3. Cooper, J.N., Spencer, J.: Simulating a random walk with constant error. Combin. Probab. Comput. **15**, 815–822 (2006). http://www.arxiv.org/abs/math.CO/0402323
4. Diaconis, P., Fulton, W.: A growth model, a game, an algebra, Lagrange inversion, and characteristic classes. Rend. Sem. Mat. Univ. Politec. Torino **49**(1), 95–119 (1991)
5. Fey, A., Redig, F.: Limiting shapes for deterministic centrally seeded growth models. J. Statist. Phys. **130**(3), 579–597 (2008). http://arxiv.org/abs/math.PR/0702450
6. Fukai, Y., Uchiyama, K.: Potential kernel for two-dimensional random walk. Ann. Probab. **24**(4), 1979–1992 (1996)
7. Kleber, M.: Goldbug variations. Math. Intelligencer **27**(1), 55–63 (2005)
8. Koosis, P.: La plus petite majorante surharmonique et son rapport avec l'existence des fonctions entières de type exponentiel jouant le rôle de multiplicateurs. Ann. Inst. Fourier (Grenoble) **33** fasc. (1), 67–107 (1983)
9. Lawler, G.: Intersections of Random Walks. Birkhäuser, Boston (1996)
10. Lawler, G., Bramson, M., Griffeath, D.: Internal diffusion limited aggregation. Ann. Probab. **20**(4), 2117–2140 (1992)

11. Lawler, G.: Subdiffusive fluctuations for internal diffusion limited aggregation. Ann. Probab. **23**(1), 71–86 (1995)
12. Le Borgne, Y., Rossin, D.: On the identity of the sandpile group. Discrete Math. **256**, 775–790 (2002)
13. Levine, L.: The rotor-router model. Harvard University senior thesis. http://arxiv.org/abs/math/0409407 (2002)
14. Levine, L., Peres, Y.: Spherical asymptotics for the rotor-router model in $\mathbb{Z}^d$. Indiana Univ. Math. J. **57**(1), 431–450 (2008). http://arxiv.org/abs/math/0503251
15. Levine, L., Peres, Y.: The rotor-router shape is spherical. Math. Intelligencer **27**(3), 9–11 (2005)
16. Levine, L., Peres, Y.: Scaling limits for internal aggregation models with multiple sources. http://arxiv.org/abs/0712.3378 (2007)
17. Lindvall, T.: Lectures on the Coupling Method. Wiley, New York (1992)
18. Priezzhev, V.B., Dhar, D., Dhar, A., Krishnamurthy, S.: Eulerian walkers as a model of self-organised criticality. Phys. Rev. Lett. **77**, 5079–82 (1996)
19. Propp, J.: Three lectures on quasirandomness. http://faculty.uml.edu/jpropp/berkeley.html (2004)
20. Spitzer, F.: Principles of Random Walk. Springer, New York (1976)
21. Uchiyama, K.: Green's functions for random walks on $\mathbb{Z}^N$. Proc. London Math. Soc. **77**(1), 215–240 (1998)
22. Van den Heuvel, J.: Algorithmic aspects of a chip-firing game. Combin. Probab. Comput. **10**(6), 505–529 (2001)