Random Growth Models

by

Laura Florescu

a dissertation submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Department of Computer Science

New York University

May, 2017

Joel Spencer

NYU

I dedicate this thesis to the love of my life, Wojciech Zaremba.

Acknowledgments

I wish to say thanks to many people. First of all, my advisor, Joel Spencer has been extremely supportive throughout my five years. He has motivated when there was a need and encouraged everytime. I will have a happy, successful life if all my mentors will be like him.

I thank my committee, Oded Regev, Yevgeniy Dodis, Prasad Tetali, Igor Shinkar for agreeing to review my dissertation and participate in my defense.

Many thanks go to my co-authors, from whom I have learned a lot: Yuval Peres, Lionel Levine, Will Perkins, Shirshendu Ganguly, Prasad Tetali, Miklos Racz. Much of this work has been completed at Microsoft Research, and I thank Yuval Peres for inviting me there many times and the Theory group for their hospitality and stimulating environment. Part of this work was also done at the Institute of Mathematics at University of Minnesota. I wish to thank them for a great, collaborating environment. I also thank Prasad Tetali for inviting me to Georgia Tech, I have learned a great deal from my visits.

I wish to thank some of my professors at Reed College and National Computer Science College “Tudor Vianu”, whose training, encouragement and role-model example has been invaluable: David Perkinson, Jerry Shurman, Rao Potluri, Iolanda Podeanu, Steliana Serban.

I also thank the NYU Computer Science department for so much support and help, in particular Rosemary Amico, Santiago Pizzini, Oded Regev. Their help with whatever issues has been critical and I am extremely grateful for that.

A lot of my motivation, happiness and encouragement came from some of my best friends: Sei Howe, Matt Steele, Kylie Byrd, Amy Goldsmith, Jess Gordon, Ronen Eldan, Riddhi Basu. I thank friends from Courant for inspiring me and for their friendship: Sasha Golovnev, Huck Bennett, Shravas Rao, Sid Krishna, Noah Stephens-Davidowitz, Chaya Ganesh.

I thank my family, and especially my father for inspiring me to study mathematics and science. Any project I take on is inspired by his memory.

I thank Wojciech for inspiring, motivating and challenging me every day. I love you.

Abstract

This work explores variations of randomness in networks, and more specifically, how drastically the dynamics and structure of a network change when a little bit of information is added to "chaos". On one hand, I investigate how much determinism in diffusions de-randomizes the process, and on the other hand, I look at how superposing "planted" information on a random network changes its structure in such a way that the "planted" structure can be recovered.

The first part of the dissertation is concerned with rotor-router walks, a deterministic counterpart to random walk, which is the mathematical model of a path consisting of a succession of random steps. I study and show results on the volume ("the range") of the territory explored by the random rotor-router model, confirming an old prediction of physicists.

The second major part in the dissertation consists of two constrained diffusion problems. The questions in this model are to understand the long-term behavior of the models, as well as how the boundary of the processes evolves in time.

The third part is detecting communities in, or more generally, clustering networks. This is a fundamental problem in mathematics, machine learning, biology and economics, both for its theoretical foundations as well as for its practi-

cal implications. This problem can be viewed as “planting” some structure in a random network; for example, in cryptography, a code can be viewed as hiding some integers in a random sequence. For such a model with two communities, I show both information theoretic thresholds when it is impossible to recover the communities based on the density of the edges “planted” between the communities, as well as thresholds for when it is computationally possible to recover the communities.

Contents

List of Figures

1

Introduction

The unifying theme of this dissertation is the interplay between randomness and planted information. More specifically, it is composed of a few problems in deterministic random walks, diffusions on graphs, stochastic block models. This research explores variations of randomness in complex systems, and more specifically, how drastically the dynamics and structure of a network change when a little bit of information is added to "chaos". On one hand, I investigate how much determinism in diffusions (the random spread of various objects in media) de-randomizes the process, and on the other hand, I look at how superposing "planted" information on a random network changes its structure in such a way that the “planted” structure can be recovered. For these types of problems, a great deal of inspiration, motivation and intuition comes from statistical physics.

The first main part of my dissertation is concerned with rotor-router walks[80], a deterministic counterpart to random walk, which is the mathematical model of a path consisting of a succession of random steps. Around the time of the initiation of the study of deterministic walks (published under the name ”Eulerian walkers as a model of self-organized criticality” by Priezzhev, Dhar, Dhar and Krishnamurthy[80]), there was, and still is, much interest in the study of complex systems exhibiting self-organized criticality[7]. Different models have been proposed for these type of systems such as sandpiles[7], earthquakes[87], forest-fires[27], and biological evolution[6]. These models involve a slowly driven system, in which an external disturbance propagates in the random medium following a random or deterministic rule.

In a rotor-router walk, each node in a network remembers its neighbors in a specific cyclical order, and each time the node is visited by a particle, it sends it to its neighbors in that order. For example, on the line, the “exits” of a particle from a site would alternate between left and right. This model has been employed in studying fundamental questions such as optimal transport[61], information spreading in networks[25], load balancing in distributed computing[38], condensed matter[78]. Deep connections with famous statistical mechanics also arise through the concept of “self-organized criticality“, which is a property of dynamical systems to stabilize without outside intervention[32].

More formally, let $G = (V, E)$ be a finite or infinite directed graph. For $v \in V$ let $E_v \subset E$ be the set of outbound edges from v, and let $\mathcal{C}_v$ be the set of all cyclic permutations of E_v. A *rotor configuration* on G is a choice of an outbound edge $\rho(v) \in E_v$ for each $v \in V$. A *rotor mechanism* on G is a choice of cyclic permutation $m(v) \in \mathcal{C}_v$ for each $v \in V$. Given ρ and m, the *simple rotor walk* started at X_0 is a sequence of vertices $X_0, X_1, \ldots \in \mathbb{Z}^d$ and rotor configurations $\rho = \rho_0, \rho_1, \ldots$ such that for all integer times $t \geq 0$

$$\rho_{t+1}(v) = \begin{cases} m(v)(\rho_t(v)), & v = X_t \\ \rho_t(v), & v \neq X_t \end{cases}$$

and $X_{t+1} = \rho_{t+1}(X_t)^+$, where e^+ denotes the target of the directed edge e. In words, the rotor at X_t "rotates" to point to a new neighbor of X_t and then the walker steps to that neighbor.

For this model, in chapter 2 I show results on the range of the walk and bounds on escape rates. This chapter is composed of two papers,[32] and[35]

In the next chapter, I study two models of constrained diffusions, one on "frozen random walk" in which particles far away from the origin are not allowed to move, and one on a controlled diffusion model on various graphs. This chapter is composed of two papers,[33] and[36].

The last chapter is concerned with a version of the stochastic block model (SBM), the bipartite stochastic block model. Although the initial motivation comes from community detection, this model comes as a reduction from planted constraint satisfaction problems (CSPs). This chapter is composed of one paper,[37].

2

Deterministic random walks

2.1 Escape rates

This chapter is based on papers[32] and[35].

In a *rotor walk* on a graph, the successive exits from each vertex follow a prescribed periodic sequence. For instance, in the square grid $\mathbb{Z}^2$, successive exits could repeatedly cycle through the sequence North, East, South West. Such walks were first studied in[88] as a model of mobile agents exploring a territory, and in[78] as a model of self-organized criticality. In a lecture at Microsoft in 2003[81], Jim Propp proposed rotor walk as a deterministic analogue of random walk, which naturally invited the question of whether rotor walk is recurrent in dimension 2 and transient in dimensions 3 and higher. One direction was settled immediately by Oded Schramm, who showed that rotor walk is “at least as recurrent” as random walk. Schramm’s elegant argument, which we recall below, applies to any initial rotor configuration ρ.

The other direction is more subtle because it depends on ρ. We say that ρ is *recurrent* if the rotor walk started at the origin with initial configuration ρ returns to the origin infinitely often; otherwise, we say that ρ is *transient*. Angel and Holroyd[3] showed that for all d there exist initial rotor configurations on $\mathbb{Z}^d$ such that rotor walk is recurrent. These special configurations are primed to send particles initially back toward the origin. Here, we analyze the case $\rho = \uparrow$ when all rotors send their first particle in the same direction. To measure how transient this configuration is, we run n rotor walks starting from the origin and record whether each returns to the origin or escapes to infinity. We show that the number of escapes is of order n in dimensions $d \geq 3$, and of order $n/\log n$ in dimension 2.

To give the formal definition of a rotor walk, write $\mathcal{E} = \{\pm e_1, \ldots, \pm e_d\}$ for the set of $2d$ cardinal directions in $\mathbb{Z}^d$, and let $\mathcal{C}$ be the set of cyclic permutations of $\mathcal{E}$. A *rotor mechanism* is a map $m : \mathbb{Z}^d \to \mathcal{C}$, and a *rotor configuration* is a map $\rho : \mathbb{Z}^d \to \mathcal{E}$. A *rotor walk* started at x_0 with initial configuration ρ is a sequence of vertices $x_0, x_1, \ldots \in \mathbb{Z}^d$ and rotor configurations $\rho = \rho_0, \rho_1, \ldots$ such that for all $n \geq 0$

$$x_{n+1} = x_n + \rho_n(x_n).$$

and

$$\rho_{n+1}(x_n) = m(x_n)(\rho_n(x_n))$$

and $\rho_{n+1}(x) = \rho_n(x)$ for all $x \neq x_n$.

For example in $\mathbb{Z}^2$, each rotor $\rho(x)$ points North, South, East or West. An example of a rotor mechanism is the permutation North $\mapsto$ East $\mapsto$ South $\mapsto$ West $\mapsto$ North at all $x \in \mathbb{Z}^2$. The resulting rotor walk in $\mathbb{Z}^2$ has the following description: A particle repeatedly steps in the direction indicated by the rotor at its current location, and then this rotor turns 90 degrees clockwise. Note that this "prospective" convention — move the particle before updating the rotor — differs from the "retrospective" convention of past works such as[3,47]. In the prospective convention, $\rho(x)$ indicates where the next particle will step from x, instead of where the previous particle stepped. The prospective convention is often more convenient when studying questions of recurrence and transience.

Here, we fix once and for all a rotor mechanism m on $\mathbb{Z}^d$. Now depending on the initial rotor configuration ρ, one of two things can happen to a rotor walk started from the origin:

1. The walk eventually returns to the origin; or
2. The walk never returns to the origin, and visits each vertex in $\mathbb{Z}^d$ only finitely often.

Indeed, if any site were visited infinitely often, then each of its neighbors must be visited infinitely often, and so the origin itself would be visited infinitely often. In case 2 we say that the walk "escapes to infinity." Note that after the walk has either returned to the origin or escaped to infinity, the rotors are in a new configuration.

To quantify the degree of transience of an initial configuration ρ, consider the following experiment: let each of n particles in turn perform rotor walk starting from the origin until either returning to the origin or escaping to infinity. Denote by $I(\rho, n)$ the number of walks that escape to infinity. (Importantly, we do not reset the rotors in between trials!)

Schramm[85] proved that for any ρ,

$$\limsup_{n\to\infty} \frac{I(\rho, n)}{n} \leq \alpha_d \tag{2.1}$$

where α_d is the probability that simple random walk in $\mathbb{Z}^d$ does not return to the origin. Our first result gives a corresponding lower bound for the initial configuration $\uparrow$ in which all rotors start pointing in the same direction: $\uparrow(x) = e_d$ for all $x \in \mathbb{Z}^d$.

Theorem 1. For the rotor walk on $\mathbb{Z}^d$ with $d \geq 3$ with all rotors initially aligned $\uparrow$, a positive fraction of particles escape to infinity; that is,

$$\liminf_{n\to\infty} \frac{I(\uparrow, n)}{n} > 0.$$

One cannot hope for such a result to hold for an arbitrary ρ: Angel and Holroyd[3] prove that in all dimensions there exist rotor configurations ρ_{rec} such that $I(\rho_{\text{rec}}, n) = 0$ for all n. Reddy first proposed such a configuration in dimension 3 on the basis of numerical simulations[82].

Our next result concerns the fraction of particles that escape in dimension 2: for any rotor configuration ρ this fraction is at most $\frac{\pi/2}{\log n}$, and for the initial configuration $\uparrow$ it is at least $\frac{c}{\log n}$ for some $c > 0$.

Theorem 2. For rotor walk in $\mathbb{Z}^2$ with any rotor configuration ρ, we have

$$\limsup_{n\to\infty} \frac{I(\rho, n)}{n/\log n} \leq \frac{\pi}{2}.$$

Moreover, if all rotors are initially aligned $\uparrow$, then

$$\liminf_{n\to\infty} \frac{I(\uparrow, n)}{n/\log n} > 0.$$

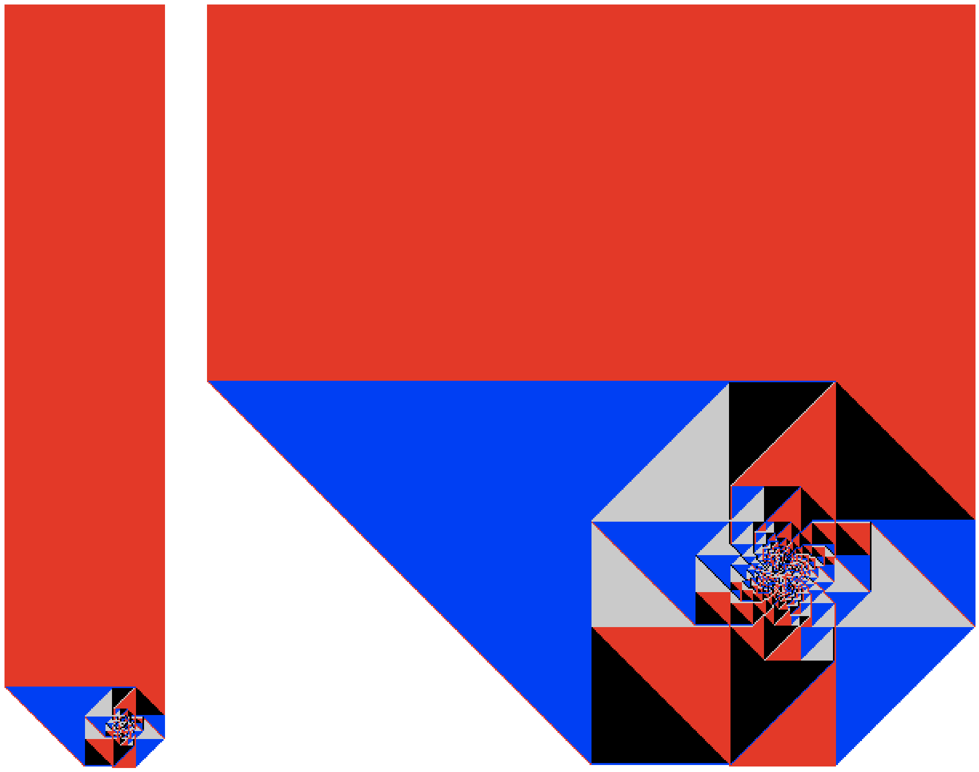


Figure 2.1: The configuration of rotors in $\mathbb{Z}^2$ after n particles started at the origin have escaped to infinity, with initial configuration $\uparrow$ (that is, all rotors send their first particle North). Left: $n = 100$; Right: $n = 480$. Each non-white pixel represents a point in $\mathbb{Z}^2$ that was visited at least once, and its color indicates the direction of its rotor.

2.1.1 Schramm's argument

One way to estimate the number of escapes to infinity of a rotor walk is to look at how many particles exit a large ball before returning to the origin. Let

$$\mathcal{B}_r = \{x \in \mathbb{Z}^d : |x| < r\}$$

be the set of lattice points in the open ball of radius r centered at the origin. Here $|x| = (x_1^2 + \cdots + x_d^2)^{1/2}$ denotes the Euclidean norm of x. Consider rotor walk started from the origin and stopped on hitting the boundary

$$\partial\mathcal{B}_r = \{y \in \mathbb{Z}^d \,:\, y \notin r \text{ and } y \sim x \text{ for some } x \in r\}.$$

Since $\mathcal{B}_r$ is a finite connected graph, this walk stops in finitely many steps.

Starting from initial rotor configuration ρ, let each of n particles in turn perform rotor walk starting from the origin until either returning to the origin or exiting the ball r. Denote by $I_r(\rho, n)$ the number of particles that exit r. The next lemmas give convergence and monotonicity of this quantity.

Lemma 1. [48] Lemma 18 For any rotor configuration ρ and any $n \in \mathbb{N}$, we have $I_r(\rho, n) \to I(\rho, n)$ as $r \to \infty$.

Proof. Let $w_n(x)$ be the number of exits from x by n rotor walks started at o and stopped if they return to o. Then $I(\rho, n)$ is determined by the values $w_n(x)$ for neighbors x of o.

Let $w_n^r(x)$ be the number of exits from x by n rotor walks started at o and stopped on hitting $\partial\mathcal{B}_r \cup \{o\}$. Then $I_r(\rho, n)$ is determined by the values $w_n^r(x)$ for neighbors x of o.

We first show that $w_n^r \leq w_n$ pointwise. Let $w_n^{r,t}(y)$ be the number of exits from y before time t if the walks are stopped on hitting $\partial\mathcal{B}_r \cup \{o\}$. If $w_n^r \not\leq w_n$, then choose t minimal such that $w_n^{r,t} \not\leq w_n$. Then there is a single point y such that $w_n^{r,t}(y) > w_n(y)$. Note that $y \neq o$, because $w_n^r(o) = w_n(o) = n$. Since $w_n^{r,t}(x) \leq w_n(x)$ for all $x \neq y$, at time t in the finite experiment the site y has received at most as many particles as it ever receives in the infinite experiment. But y has emitted strictly more particles in the finite experiment than it ever emits in the infinite experiment, so the number of particles at y at time t is < 0, a contradiction.

Now we induct on n to show that $w_n^r \uparrow w_n$ pointwise. Assume that $w_{n-1}^r \uparrow w_{n-1}$. Fix $s > 0$. There exists $R = R(s)$ such that $w_{n-1}^r = w_{n-1}$ on $\mathcal{B}_s$ for all $r \geq R$. If the nth walk returns to o then it does so without exiting $\mathcal{B}_S$ for some S; in this case $w_n^r = w_n$ on $\mathcal{B}_s$ for all $r \geq \max(R, S)$.

If the nth walk escapes to infinity, then there is some radius S such that after exiting $\mathcal{B}_S$ the walk never returns to $\mathcal{B}_s$. Now choose R' such that $w_{n-1}^{R'} = w_{n-1}$ on $\mathcal{B}_S$. Then we claim $w_n^r = w_n$ on $\mathcal{B}_s$ for all $r \geq R'$. Denote by ρ_{n-1}^r (resp. ρ_{n-1}) the rotor configuration after $n-1$ walks started at the origin have stopped on $\partial\mathcal{B}_r \cup \{o\}$ (resp. stopped if they return to o). If $r \geq R'$ then the rotor walks started at o with initial conditions ρ_{n-1}^r and ρ_{n-1} agree until they exit $\mathcal{B}_S$. Thereafter the latter walk never returns to $\mathcal{B}_s$, hence $w_n^r \geq w_n$ on $\mathcal{B}_s$. Since also $w_n^r \leq w_n$ everywhere, the inductive step is complete. $\square$

For the next lemma we recall the *abelian property* of rotor walk[47] Lemma 3.9. Let A be a finite subset of $\mathbb{Z}^d$. In an experiment of the form "run n rotor walks from prescribed starting points until they exit A," suppose that we repeatedly choose a particle in A and ask it to take a rotor walk step. Regardless of our choices, all particles will exit A in finitely many steps; for each $x \in A^c$, the number of particles that stop at x does not depend on the choices; and for each $x \in A$, the number of times we pointed to a particle at x does not depend on the choices.

Lemma 2. [48] Lemma 19 For any rotor configuration ρ, any $n \in \mathbb{N}$ and any $r < R$, we have $I_R(\rho, n) \leq I_r(\rho, n)$.

Proof. By the abelian property, we may compute $I_R(\rho, n)$ in two stages. First stop particles when they reach $\partial r \cup \{o\}$, where $o \in \mathbb{Z}^d$ is the origin, and then let the $I_r(\rho, n)$ particles stopped on ∂r continue walking until they reach $\partial R \cup \{o\}$. Therefore at most $I_r(\rho, n)$ particles stop in ∂R. □

Oded Schramm's upper bound (2.1) begins with the observation that if $2dm$ particles at a single site $x \in \mathbb{Z}^d$ each take a single rotor walk step, the result will be that m particles move to each of the $2d$ neighbors of x. Fix $r, m \in \mathbb{N}$ and consider $N = (2d)^r m$ particles at the origin. Let each particle take a single rotor walk step. Then repeat $r - 1$ times the following operation: let each particle that is not at the origin take a single rotor walk step. The result is that for each path $(\gamma_0, \ldots, \gamma_\ell)$ of length $\ell \leq r$ with $\gamma_0 = \gamma_\ell = o$ and $\gamma_i \neq o$ for all $1 \leq i \leq \ell - 1$, exactly $(2d)^{-\ell} N$ particles traverse this path. Denoting the set of such paths by $\Gamma(r)$ and the length of a path γ by $|\gamma|$, the number of particles now at the origin is

$$N \sum_{\gamma \in \Gamma(r)} (2d)^{-|\gamma|} = Np$$

where $p = \mathbb{P}(T_o^+ \leq r)$ is the probability that simple random walk returns to the origin by time r.

Now letting each particle that is not at the origin continue performing rotor walk until hitting $\partial r \cup \{o\}$, the number of particles that stop in ∂r is at most $N(1 - p)$, so

$$\frac{I_r(\rho, N)}{N} \leq 1 - p.$$

This holds for every N which is an integer multiple of $(2d)^r$. For general n, let N be the smallest multiple of $(2d)^r$ that is $\geq n$. Then

$$\frac{I_r(\rho, n)}{n} \leq \frac{I_r(\rho, N)}{N - (2d)^r}$$

The right side is at most $(1-p)(1 + 2(2d)^r/N)$, so

$$\limsup_{n\to\infty} \frac{I(\rho, n)}{n} \leq \limsup_{n\to\infty} \frac{I_r(\rho, n)}{n} \leq 1 - p = \mathbb{P}(T_o^+ > r).$$

As $r \to \infty$ the right side converges to α_d, completing the proof of (2.1).

See Holroyd and Propp[48] Theorem 10 for an extension of Schramm's argument to a general irreducible Markov chain with rational transition probabilities.

2.1.2 An odometer estimate for balls in all dimensions

To estimate $I_r(\rho, n)$, consider now a slightly different experiment. Let each of n particles started at the origin perform rotor walk until hitting ∂r. (The difference is that we do not stop the particles on returning to the origin!) Define the *odometer function* u_n^r by

$$u_n^r(x) = \text{total number of exits from } x \text{ by } n \text{ rotor walks stopped on hitting } \partial\mathcal{B}_r.$$

Note that $u_n^r(x)$ counts the total number of exits (as opposed to the net number).

Now we relate the two experiments.

Lemma 3. For any $r > 0$ and $n \in \mathbb{N}$ and any initial rotor configuration ρ, we have

$$I_r(\rho, u_n^r(o)) = n.$$

Proof. Starting with $N = u_n^r(o)$ particles at the origin, consider the following two experiments:

1. Let n of the particles in turn perform rotor walk until hitting $\partial\mathcal{B}_r$.

2. Let N of the particles in turn perform rotor walk until hitting $\partial\mathcal{B}_r \cup \{o\}$.

By the definition of u_n^r, in the first experiment the total number of exits from the origin is exactly N. Therefore the two experiments have exactly the same outcome: n particles reach ∂r and $N - n$ remain at the origin. $\square$

Our next task is to estimate u_n^r. We begin by introducing some notation. Given a function f on $\mathbb{Z}^d$, its *gradient* is the function on directed edges given by

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

Given a function κ on directed edges of $\mathbb{Z}^d$, its *divergence* is the function on vertices given by

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

where the sum is over the $2d$ nearest neighbors of x. The *discrete Laplacian* of f is the function

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

We recall some results from[59].

Lemma 4. [59] 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 by n rotor walks started at the origin and stopped on exiting r. Then

$$\nabla u_n^r(x, y) = -2d\,\kappa(x, y) + R(x, y)$$

for some edge function R satisfying $|R(x, y)| \leq 4d - 2$ for all edges (x, y).

Denote by $(X_j)_{j \geq 0}$ the simple random walk in $\mathbb{Z}^d$, whose increments are independent and uniformly distributed on $\mathcal{E} = \{\pm e_1, \ldots, \pm e_d\}$. Let $T = \min\{j : X_j \notin r\}$ be the first exit time from the ball of radius r. For $x, y \in r$, let

$$G_r(x, y) = \mathbb{E}_x \#\{j < T | X_j = y\}$$

be the expected number of visits to y by a simple random walk started at x before time T. The following well known estimates can be found in[57] Prop. 1.5.9, Prop. 1.6.7: for a constant a_d depending only on d,

$$G_r(x, o) = \begin{cases} a_d(|x|^{2-d} - r^{2-d}) + O(|x|^{1-d}), & d \geq 3 \\ \frac{2}{\pi}(\log r - \log |x|) + O(|x|^{-1}), & d = 2. \end{cases} \tag{2.2}$$

We will also use[57] Theorem 1.6.6 the fact that in dimension 2,

$$G_r(o,o) = \frac{2}{\pi}\log r + O(1). \tag{2.3}$$

(As usual, we write $f(n) = \Theta(g(n))$ (respectively, $f(n) = O(g(n))$) to mean that there is a constant $0 < C < \infty$ such that $1/C < f(n)/g(n) < C$ (respectively, $f(n)/g(n) < C$) for all sufficiently large n. Here and in what follows, the constants implied in $O()$ and $\Theta()$ notation depend only on the dimension d.)

The next lemma bounds the L^1 norm of the discrete gradient of the function $G_r(x,\cdot)$. It appears in[59] Lemma 5.6 with the factor of 2 omitted (this factor is needed for x close to the origin). The proof given there actually shows the following.

Lemma 5. Let $x \in \mathcal{B}_r$ and let $\rho = r + 1 - |x|$. Then for some C depending only on d,

$$\sum_{y\in\mathcal{B}_r}\sum_{z\sim y}|G_r(x,y) - G_r(x,z)| \le C\rho\log\frac{2r}{\rho}.$$

The next lemma is proved in the same way as the inner estimate of[59] Theorem 1.1. Let $f(x) = nG_r(x,o)$.

Lemma 6. In $\mathbb{Z}^d$, let $x \in \mathcal{B}_r$ and $\rho = r + 1 - |x|$. Then,

$$|u_n^r(x) - f(x)| \le C\rho\log\frac{2r}{\rho} + 4d.$$

where u_n^r is the odometer function for n particles performing rotor walk stopped on exiting $\mathcal{B}_r$, and C is the constant in Lemma 5.

Proof. If we consider the rotor walk stopped on exiting $\mathcal{B}_r$, all sites that have positive odometer value have been hit by particles. Using notation of Lemma 4, we notice that since the net number of particles to enter a site $x \neq o$ not on the boundary is zero, we have $2d \operatorname{div} \kappa(x) = 0$. For the origin, $2d \operatorname{div} \kappa(o) = n$. Also, the odometer function vanishes on the boundary, since the boundary does not emit any particles.

Write $u = u_n^r$. Using the definition of κ in Lemma 4, we see that

$$\Delta u(x) = \operatorname{div} R(x),\ \ x \neq o, \tag{2.4}$$

$$\Delta u(o) = -n + \operatorname{div} R(o). \tag{2.5}$$

Then $\Delta f(x) = 0$ for $x \in \mathcal{B}_r \setminus \{o\}$ and $\Delta f(o) = -n$ and f vanishes on $\partial\mathcal{B}_r$.

Since $u(X_T)$ is equal to 0, we have

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

Also, since the k^{th} term in the sum is zero when $T \leq k$

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

where $\mathcal{F}_j = \sigma(X_0, \ldots, X_j)$ is the standard filtration for the random walk.

Taking expectation of the conditional expectations and using (2.4) and (2.5), we get

$$u(x) = \sum_{k \geq 0} \mathbb{E}_x \left[1_{\{T>k\}}(n1_{\{X_k=o\}} - \operatorname{div} R(X_k))\right]$$

$$= \quad n\,\mathbb{E}_x\#\{k<T|X_k=0\}-\sum_{k\geq 0}\mathbb{E}_x\left[1_{\{T>k\}}\mathrm{div}\,R(X_k)\right].$$

So,

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

Let $N(y)$ be the number of edges joining y to $\partial\mathcal{B}_r$. Since $\mathbb{E}_x\sum_{k\geq 0}1_{\{T>k\}}N(X_k)=2d$, and $|R|\leq 4d$, the terms with $z\in\partial\mathcal{B}_r$ contribute at most $8d^2$ to the sum. Thus,

$$|u(x)-f(x)|\leq\frac{1}{2d}\left|\sum_{k\geq 0}\mathbb{E}_x\left[\sum_{\substack{y,z\in\mathcal{B}_r\\ y\sim z}}1_{\{T>k\}\cap\{X_k=y\}}\,R(y,z)\right]\right|+4d. \tag{2.6}$$

Note that for $y\in\mathcal{B}_r$ we have $\{X_k=y\}\cap\{T>k\}=\{X_{k\wedge T}=y\}$. Considering $p_k(y)=\mathbb{P}_x(X_{k\wedge T}=y)$, and noting that R is antisymmetric (because of antisymmetry in Lemma 4), we see that

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

Summing over k in (2.6) and using the fact that $|R|\leq 4d$, we conclude that

$$|u(x)-f(x)|\leq\sum_{\substack{y,z\in B_r\\ y\sim z}}|G(x,y)-G(x,z)|+4d.$$

The result now follows from the estimate of the gradient of Green's function in Lemma 5. ∎

Now we make our choice of radius, $r = n^{1/(d-1)}$. The next lemma shows that for this value of r, the support of the odometer function contains a large sphere.

Lemma 7. There exists a sufficiently small $\beta > 0$ depending only on d, such that for any initial rotor configuration and $r = n^{1/(d-1)}$ we have $u_n^r(x) > 0$ for all $x \in \partial\mathcal{B}_{\beta r}$.

Proof. For $x \in \partial\mathcal{B}_{\beta r}$ we have $\beta r \leq |x| \leq \beta r + 1$. By Lemma 6 we have

$$|u_n^r(x) - f(x)| \leq C'(1-\beta)r \log \frac{1}{1-\beta}$$

for a constant C' depending only on d. To lower bound $f(x)$ we use (2.2): in dimensions $d \geq 3$ we have

$$\begin{aligned} f(x) = nG_r(x,o) &\geq n(a_d(|x|^{2-d} - r^{2-d}) - K|x|^{1-d}) \\ &= a_d(\beta^{2-d} - 1)nr^{2-d} - K\beta^{1-d} \end{aligned}$$

for a constant K depending only on d. Since $r = nr^{2-d}$, we can take $\beta > 0$ sufficiently small so that

$$a_d(\beta^{2-d} - 1)nr^{2-d} - K\beta^{1-d} > 2C'(1-\beta)r \log \frac{1}{1-\beta}$$

for all sufficiently large n. Hence $u_n^r(x) > 0$.

In dimension 2, we have $r = n$ and $nG_n(x,o) \geq n\frac{2}{\pi}\log\frac{1}{\beta} - \frac{K}{\beta}$, by (2.2). So for β small enough, we have that

$$nG_n(x,o) = n\frac{2}{\pi}\log\frac{1}{\beta} - \frac{K}{\beta} > C'(1-\beta)n\log\frac{1}{1-\beta}$$

for all sufficiently large n. Hence $u_n^n(x) > 0$. $\square$

Identify $\mathbb{Z}^d$ with $\mathbb{Z}^{d-1}\times\mathbb{Z}$ and call each set of the form $(x_1,\ldots,x_{d-1})\times\mathbb{Z}$ a "column." Starting n particles at the origin and letting them each perform rotor walk until exiting $\mathcal{B}_r$ where $r = n^{1/(d-1)}$, let $\mathrm{col}(\rho,n)$ be the number of distinct columns that are visited. That is,

$$\mathrm{col}(\rho,n) = \#\{(x_1,\ldots,x_{d-1}) : u_n^r(x_1,x_2,\ldots,x_d) > 0 \text{ for some } x_d\in\mathbb{Z}\}.$$

By Lemma 7, every site of $\partial\beta r$ is visited at least once, so

$$\begin{aligned}\mathrm{col}(\rho,n) &\geq \#\{(x_1,\ldots,x_{d-1}) :\ (x_1,x_2,\ldots,x_d)\in\partial\mathcal{B}_{\beta r} \text{ for some } x_d\in\mathbb{Z}\}\\ &\geq C(\beta r)^{d-1} = \Theta(n). \qquad (2.7)\end{aligned}$$

All results so far have not made any assumptions on the initial configuration. The next lemma assumes the initial rotor configuration to be $\uparrow$. The important property of this initial condition for us is that the first particle to visit a given column travels straight along that column in direction e_d thereafter.

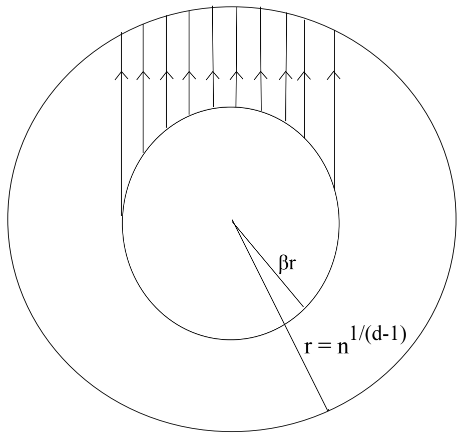


Figure 2.2: Diagram for the proof of Lemma 8. The first visit to each column results in an escape along that column, so at least $\mathrm{col}(\uparrow, n)$ particles escape.

Lemma 8. In $\mathbb{Z}^d$ with initial rotor configuration $\uparrow$, we have

$$I_R(\uparrow, u_n^r(o)) \geq \mathrm{col}(\uparrow, n)$$

for all $R \geq r$.

Proof. By the abelian property of rotor walk, we may compute $I_R(\rho, u_n^r(o))$ in two stages. First we stop the particles when they first hit $\partial\mathcal{B}_r \cup \{o\}$. Then we let all the particles stopped on $\partial\mathcal{B}_r$ continue walking until they hit $\partial\mathcal{B}_R \cup \{o\}$. By Lemma 3, exactly n particles stop on ∂r during the first stage, and therefore $\mathrm{col}(\uparrow, n)$ distinct columns are visited during the first stage. Because the initial rotors are $\uparrow$, the first particle to visit a given column travels straight along that column to hit ∂R (Figure 2.2). Therefore the number of particles stopping in ∂R is at least $\mathrm{col}(\uparrow, n)$. $\square$

2.1.3 The transient case: Proof of Theorem 1

In this section we consider $\mathbb{Z}^d$ for $d \geq 3$. We will prove Theorem 1 by comparing the number of escapes $I(\uparrow, n)$ with $\mathrm{col}(\uparrow, n)$.

Let $r = n^{1/(d-1)}$ and $N = u_n^r(o)$. By the transience of simple random walk in $\mathbb{Z}^d$ for $d \geq 3$ we have

$$f(o) = nG_r(o, o) = \Theta(n).$$

By Lemma 6 we have $|N - f(o)| = O(r)$ and hence $N = \Theta(n)$. By Lemmas 1 and 8 we have $I(\uparrow, N) \geq \mathrm{col}(\uparrow, n)$. Recalling (2.7) that $\mathrm{col}(\uparrow, n) = \Theta(n)$ and that $I(\uparrow, n)$ is nondecreasing in n, we conclude that there is a constant $c > 0$ depending only on d such that for all sufficiently large n

$$\frac{I(\uparrow, n)}{n} > c$$

which completes the proof.

2.2 The recurrent case: Proof of Theorem 2

In this section we work in $\mathbb{Z}^2$ and take $r = n$. We start by estimating the odometer function at the origin for the rotor walk stopped on exiting $\mathcal{B}_n$.

Lemma 9. For any initial rotor configuration in $\mathbb{Z}^2$ we have

$$u_n^n(o) = \frac{2}{\pi} n \log n + O(n).$$

Proof. By (2.3), we have $f(o) = nG_n(o,o) = n(\frac{2}{\pi}\log n + O(1))$, and $|u_n^n(o) - f(o)| = O(n)$ by Lemma 6. $\square$

Turning to the proof of the upper bound in Theorem 2, let $N = u_n^n(o)$. By Lemmas 1 and 2, $I(\rho, N) \le I_n(\rho, N)$. By Lemma 3, $I_n(\rho, N) = n$. Now by Lemma 9, $\frac{N}{\log N} = \frac{(2/\pi)n\log n + O(n)}{\log n + O(\log\log n)} = (\frac{2}{\pi} + o(1))n$, hence

$$\frac{I(\rho, N)}{N/\log N} \le \frac{n}{(\frac{2}{\pi} + o(1))n} = \frac{\pi}{2} + o(1).$$

Since $I(\rho, n)$ is nondecreasing in n, the desired upper bound follows.

To show the lower bound for $\uparrow$ we use lemmas 1 and 8 along with (2.7)

$$I(\uparrow, N) = \lim_{R\to\infty} I_R(\uparrow, N) \ge \mathrm{col}(\uparrow, n) \ge \beta n = \Theta(\frac{N}{\log N}).$$

Since $I(\rho, n)$ is nondecreasing in n the desired lower bound follows.

Remark 1. The proofs of the lower bounds in Theorems 1 and 2 apply to a slightly more general class of rotor configurations than $\uparrow$. Given a rotor configuration ρ, the *forward path* from x is the path $x = x_0, x_1, x_2, \ldots$ defined by $x_{k+1} = x_k + \rho(x_k)$ for $k \ge 0$. Let us say that $x \in \partial r$ has a *simple path to infinity* if the forward path from x is simple (that is, all x_k are distinct) and $x_k \notin \partial r$ for all $k \ge 1$. The proofs we have given for $\uparrow$ remain valid for ρ as long as there is a constant C and a sequence of radii $r_1, r_2, \ldots$ with $r_{i+1}/r_i < C$, such that for each i, at least r_i^{d-1}/C sites on ∂r_i have disjoint simple paths to infinity. For instance, the rotor configuration

$$\rho(x) = \begin{cases} \alpha, & x_d \geq 0 \\ \beta, & x_d < 0 \end{cases}$$

satisfies this condition as long as $(\alpha, \beta) \neq (-e_d, +e_d)$.

2.2.1 Some open questions

We conclude this chapter with a few natural questions.

- When is Schramm's bound attained? In $\mathbb{Z}^d$ for $d \geq 3$ with rotors initially aligned in one direction, is the escape rate for rotor walk asymptotically equal to the escape probability of the simple random walk? Theorem 1 shows that the escape rate is positive.

- If random walk on a graph is transient, must there be a rotor configuration ρ for which a positive fraction of particles escape to infinity, that is, $\liminf_{n\to\infty} \frac{I(\rho,n)}{n} > 0$?

- Let us choose initial rotors $\rho(x)$ for $x \in \mathbb{Z}^d$ independently and uniformly at random from $\{\pm e_1, \ldots, \pm e_d\}$. Is the resulting rotor walk recurrent in dimension 2 and transient in dimensions $d \geq 3$? Angel and Holroyd[3] Corollary 6 prove that two initial configurations differing in only a finite number of rotors are either both recurrent or both transient. Hence the set of recurrent ρ is a tail event and consequently has probability 0 or 1.

- Starting from initial rotor configuration $\uparrow$ in $\mathbb{Z}^2$, let ρ_n be the rotor configuration after n particles have escaped to infinity. Does $\rho_n(nx, ny)$ have a limit as $n \to \infty$? Figure 2.1 suggests that the answer is yes.

- Consider rotor walk in $\mathbb{Z}^2$ with a drift to the north: each rotor mechanism is period 5 with successive exits cycling through North, North, East, South, West. Is this walk transient for all initial rotor configurations?

Angel and Holroyd resolved many of these questions when $\mathbb{Z}^d$ is replaced by an arbitrary rooted tree: if only finitely many rotors start pointing toward the root (recall we use the prospective convention), then the escape rate for rotor walk started at the root equals the escape probability $\mathcal{E}$ for random walk started at the root[4] Theorem 3. On the other hand if *all* rotors start pointing toward the root, then the rotor walk is recurrent[4] Theorem 2(iii). On the regular b-ary tree, the i.i.d. uniformly random initial rotor configuration has escape rate $\mathcal{E} = 1/b$ for $b \geq 3$ but is recurrent for $b = 2$[4] Theorem 6. In the latter case particles travel extremely far[4] Theorem 7: There is a constant $c > 0$ such that with probability tending to 1 as $n \to \infty$, one of the first n particles reaches distance $e^{e^{cn}}$ from the root before returning!

2.3 Range of rotor walk

Imagine walking your dog on an infinite square grid of city streets. At each intersection, your dogs tugs you one block further North, East, South or West. After you've been dragged in this fashion down t blocks, how many distinct intersections have you seen?

The answer depends of course on your dog's algorithm. If she makes a beeline for the North then every block brings you to a new intersection, so you see $t + 1$ distinct intersections. At the opposite extreme, she could pull you back and forth repeatedly along her favorite block so that you see only ever see 2 distinct intersections.

In the *clockwise rotor walk* each intersection has a signpost pointing the way when you first arrive there. But your dog likes variety, and she has a capacious memory. If you come back to an intersection you have already visited, your dog chooses the direction 90° clockwise from the direction you went the last time you were there. We can formalize the city grid as the infinite graph $\mathbb{Z}^2$. The intersections are all the points (x, y) in the plane with integer coordinates, and the city blocks are the line segments from (x, y) to $(x \pm 1, y)$ and $(x, y \pm 1)$. More generally, we can consider a d-dimensional city $\mathbb{Z}^d$ or even an arbitrary graph, but the 90° clockwise rule will have to be replaced by something more abstract (a rotor mechanism, defined below).

In a *rotor walk* on a graph, the exits from each vertex follow a prescribed periodic sequence. Such walks were first studied in[88] as a model of mobile agents exploring a territory, and in[79] as a model of self-organized criticality. Propp proposed rotor walk as a deterministic analogue of random walk, a perspective explored in[22,32,48]. This section is concerned with the following questions. How much territory does a rotor walk cover in a fixed number of steps? Conversely, how many steps does it take for a rotor walk to completely explore a given finite graph?

Let $G = (V, E)$ be a finite or infinite directed graph. For $v \in V$ let $E_v \subset E$ be the set of outbound edges from v, and let $\mathcal{C}_v$ be the set of all cyclic permutations of E_v. A *rotor configuration* on G is a choice of an outbound edge $\rho(v) \in$

E_v for each $v \in V$. A *rotor mechanism* on G is a choice of cyclic permutation $m(v) \in \mathcal{C}_v$ for each $v \in V$. Given ρ and m, the *simple rotor walk* started at X_0 is a sequence of vertices $X_0, X_1, \ldots \in \mathbb{Z}^d$ and rotor configurations $\rho = \rho_0, \rho_1, \ldots$ such that for all integer times $t \geq 0$

$$\rho_{t+1}(v) = \begin{cases} m(v)(\rho_t(v)), & v = X_t \\ \rho_t(v), & v \neq X_t \end{cases}$$

and

$$X_{t+1} = \rho_{t+1}(X_t)^+$$

where e^+ denotes the target of the directed edge e. In words, the rotor at X_t "rotates" to point to a new neighbor of X_t and then the walker steps to that neighbor.

We have chosen the retrospective rotor convention—each rotor at an already visited vertex indicates the direction of the most recent exit from that vertex—because it makes a few of our results such as Lemma 11 easier to state.

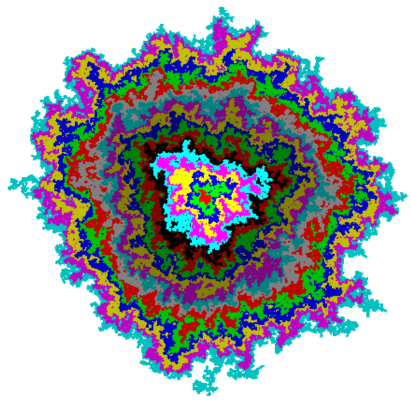


Figure 2.3: The range of a clockwise uniform rotor walk on $\mathbb{Z}^2$ after 80 returns to the origin. The mechanism m cycles through the four neighbors in clockwise order (North, East, South, West), and the initial rotors $\rho(v)$ were oriented independently North, East, South or West, each with probability $1/4$. Colors indicate the first twenty excursion sets $A_1, \dots, A_{20}$, defined in §2.3.2.

The *range* of rotor walk at time t is the set

$$R_t = \{X_1, \dots, X_t\}.$$

We investigate the size of the range, $\#R_t$, in terms of the growth rate of balls in the underlying graph G. Fix an origin $o \in V$ (the starting point of our rotor walk). For $r \in \mathbb{N}$ the *ball* of radius r centered at o, denoted $B(o,r)$, is the set of vertices reachable from o by a directed path of length $\leq r$. Suppose that there are constants $d, k > 0$ such that

$$\#B(o,r) \geq kr^d \tag{2.8}$$

for all $r \geq 1$. Intuitively, this condition says that G is "at least d-dimensional."

A directed graph is called *Eulerian* if each vertex has as many incoming as outgoing edges. In particular, any undirected graph can be made Eulerian by converting each edge into a pair of oppositely oriented directed edges.

Theorem 3. For any Eulerian graph G of bounded degree satisfying (2.8), the number of distinct sites visited by a rotor walk started at o in t steps satisfies

$$\#R_t \geq ct^{d/(d+1)}.$$

for a constant $c > 0$ depending only on G (and not on ρ or m).

Priezzhev et al.[79] and Povolotsky et al.[77] gave a heuristic argument that $\#R_t$ has order $t^{2/3}$ for the clockwise rotor walk on $\mathbb{Z}^2$ with uniform random initial rotors. Theorem 3 gives a lower bound of this order, and our proof is directly inspired by their argument.

The upper bound promises to be more difficult because it depends on the initial rotor configuration ρ. Indeed, the next theorem shows that for certain ρ, the number of visited sites $\#R_t$ grows linearly in t (which we need not point out is much faster than $t^{2/3}$!). Rotor walk is called *recurrent* if $X_t = X_0$ for infinitely many t, and *transient* otherwise.

Theorem 4. For any Eulerian graph G and any mechanism m, if the initial rotor configuration ρ has an infinite path directed toward o, then rotor walk started at o is transient and

$$\#R_t \geq \frac{t}{\Delta},$$

where Δ is the maximal degree of a vertex in G.

Theorems 3 and 4 are proved in §2.3.3. But enough about the size of the range; what about its shape? Each pixel in 2.3 corresponds to a vertex of $\mathbb{Z}^2$, and R_t is the set of all colored pixels (the different colors correspond to *excursions* of the rotor walk, defined in §2.3.2); the mechanism m is clockwise, and the initial rotors ρ independently point North, East, South, or West with probability 1/4 each. Although the set R_t of Figure 2.3 looks far from round, Kapri and Dhar have conjectured that for very large t it becomes nearly a circular disk! From now on, by **uniform rotor walk** we will always mean that the initial rotors $\{\rho(v)\}_{v \in V}$ are independent and uniformly distributed on E_v.

Conjecture 1 (Kapri-Dhar[53]). The set of sites R_t visited by the clockwise uniform rotor walk in $\mathbb{Z}^2$ is asymptotically a disk. There exists a constant c such that for any $\epsilon > 0$,

$$\mathbb{P}\{\mathcal{D}_{(c-\epsilon)t^{1/3}} \subset R_t \subset \mathcal{D}_{(c+\epsilon)t^{1/3}}\} \to 1$$

as $t \to \infty$, where $\mathcal{D}_r = \{(x, y) \in \mathbb{Z}^2 \, : \, x^2 + y^2 < r^2\}$.

We are a long way from proving anything like Conjecture 1, but we can show that an analogous shape theorem holds on a much simpler graph, the *comb* obtained from $\mathbb{Z}^2$ by deleting all horizontal edges except those along the x-axis (Figure 2.4).

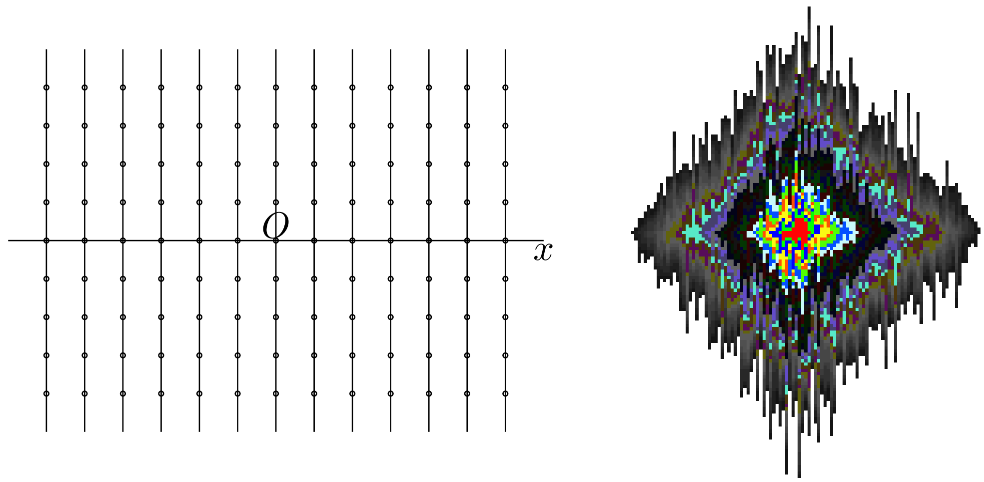
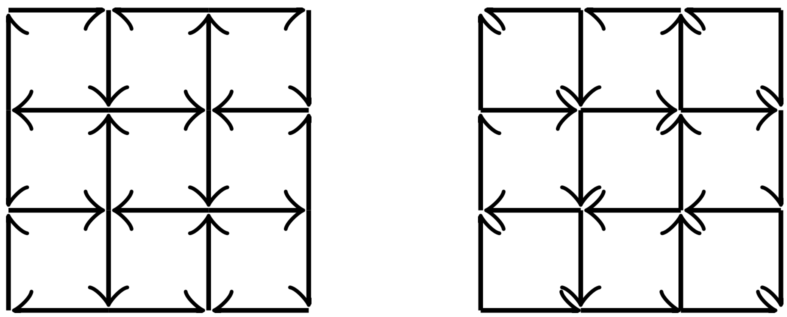


Figure 2.4: A piece of the comb graph (left) and the set of sites visited by a uniform rotor walk on the comb graph in 10000 steps.

Theorem 5. For uniform rotor walk on the comb graph, $\#R_t$ has order $t^{2/3}$ and the asymptotic shape of R_t is a diamond.

For the precise statement, see §2.3.4. This result contrasts with random walk on the comb, for which the expected number of sites visited is only on the order of $t^{1/2}\log t$ as shown by Pach and Tardos[73]. Thus the uniform rotor walk explores the comb more efficiently than random walk. (On the other hand, it is conjectured to explore $\mathbb{Z}^2$ *less* efficiently than random walk!)

The main difficulty in proving upper bounds for $\#R_t$ lies in showing that the uniform rotor walk is recurrent. This seems to be a difficult problem in $\mathbb{Z}^2$, but we can show it for two different directed graphs obtained by orienting the edges of $\mathbb{Z}^2$: the Manhattan lattice and the F-lattice, pictured in Figure 2.5. The F-lattice has two outgoing horizontal edges at every odd node and two outgoing vertical edges at every even node (we call (x, y) odd or even according to whether $x + y$ is odd or even). The Manhattan lattice is full of one-way streets: rows alternate pointing left and right, while columns alternate pointing up and down.



(a) F-lattice (b) Manhattan lattice

Figure 2.5: Two different periodic orientations of the square grid with indegree and outdegree 2.

Theorem 6. Uniform rotor walk is recurrent on both the F-lattice and the Manhattan lattice.

The proof uses a connection to the mirror model and critical bond percolation on $\mathbb{Z}^2$; see §2.3.5.

Theorems 3-6 bound the rate at which rotor walk explores various infinite graphs. In §2.4 we bound the time it takes a rotor walk to completely explore a given finite graph.

2.3.1 Related work

By comparing to a branching process, Angel and Holroyd[4] showed that uniform rotor walk on the infinite b-ary tree is transient for $b \geq 3$ and recurrent for $b = 2$. In the latter case the corresponding branching process is critical, and the distance traveled by rotor walk before returning n times to the root is doubly exponential in n. They also studied rotor walk on a singly infinite comb with the "most transient" initial rotor configuration ρ. They showed that if n particles start at the origin, then order $\sqrt{n}$ of them escape to infinity (more generally, order $n^{1-2^{1-d}}$ for a d-dimensional analogue of the comb).

In *rotor aggregation*, each of n particles starting at the origin performs rotor walk until reaching an unoccupied site, which it then occupies. For rotor aggregation in $\mathbb{Z}^d$, the asymptotic shape of the set of occupied sites is a Euclidean ball[59]. For the layered square lattice ($\mathbb{Z}^2$ with an outward bias along the x- and y-axes) the asymptotic shape becomes a diamond[51]. Huss and Sava[49] studied rotor aggregation on the 2-dimensional comb with the "most recurrent" initial rotor configuration. They showed that at certain times the boundary of the set of occupied sites is composed of four segments of exact parabolas. It is interesting to compare their result with Theorem 5: The asymptotic shape, and even the scaling, is different.

2.3.2 Excursions

Let $G = (V, E)$ be a connected Eulerian graph. In this section G can be either finite or infinite, and the rotor mechanism m can be arbitrary. The main idea of the proof of Theorem 3 is to decompose rotor walk on G into a sequence of excursions. This idea was also used in[3] to construct recurrent rotor configurations on $\mathbb{Z}^d$ for all d, and in[8,12,89] to bound the cover time of rotor walk on a finite graph (about which we say more in §2.4). For a vertex $o \in V$ we write $\deg(o)$

for the number of outgoing edges from o, which equals the number of incoming edges since G is Eulerian.

Definition 1. An *excursion* from o is a rotor walk started at o and run until it returns to o exactly $\deg(o)$ times.

More formally, let $(X_t)_{t\geq 0}$ be a rotor walk started at $X_0 = o$. For $t \geq 0$ let

$$u_t(x) = \#\{1 \leq s \leq t \,:\, X_s = x\}.$$

For $n \geq 0$ let

$$T(n) = \min\{t \geq 0 \,:\, u_t(o) \geq n \deg(o)\},$$

be the time taken for the rotor walk to complete n excursions from o (with the convention that min of the empty set is ∞). For all $n \geq 1$ such that $T(n-1) < \infty$, define

$$e_n \equiv u_{T(n)} - u_{T(n-1)}$$

so that $e_n(x)$ counts the number of visits to x during the nth excursion. To make sense of this expression when $T(n) = \infty$, we write $u_\infty(x) \in \mathbb{N} \cup \{\infty\}$ for the increasing limit of the sequence $u_t(x)$.

Our first lemma says that each $x \in V$ is visited at most $\deg(x)$ times per excursion. The assumption that G is Eulerian is crucial here.

Lemma 10. [3] Lemma 8;[12] §4.2 For any initial rotor configuration ρ,

$$e_1(x) \leq \deg(x) \qquad \forall x \in V.$$

Proof. If the rotor walk never traverses the same directed edge twice, then $u_t(x) \leq \deg(x)$ for all t and x, so we are done. Otherwise, consider the smallest t such that $(X_s, X_{s+1}) = (X_t, X_{t+1})$ for some $s < t$. By definition, rotor walk reuses an outgoing edge from X_t only after it has used all of the outgoing edges from X_t. Therefore, at time t the vertex X_t has been visited $\deg(X_t) + 1$ times, but by the minimality of t each incoming edge to X_t has been traversed at most once. Since G is Eulerian it follows that $X_t = X_0 = o$ and $t = T(1)$.

Therefore every directed edge is used at most once during the first excursion, so each $x \in V$ is visited at most $\deg(x)$ times during the first excursion. □

Lemma 11. If $T(1) < \infty$ and there is a directed path of <u>initial</u> rotors from x to o, then

$$e_1(x) = \deg(x).$$

Proof. Let y be the first vertex after x on the path of initial rotors from x to o. By induction on the length of this path, y is visited exactly $\deg(y)$ times in an excursion from o. Each incoming edge to y is traversed at most once by Lemma 10, so in fact each incoming edge to y is traversed exactly once. In particular, the edge (x, y) is traversed. Since $\rho(x) = (x, y)$, the edge (x, y) is the last one traversed out of x, so x must be visited at least $\deg(x)$ times. □

If G is finite, then $T(n) < \infty$ for all n, since by Lemma 10 the number of visits to a vertex is at most or equal to the degree of that vertex. If G is infinite, then depending on the rotor mechanism m and initial rotor configuration ρ, rotor walk may or may not complete an excursion from o. In particular, Lemma 11 implies the following.

Corollary 6.1. If ρ has an infinite path directed toward o, then $T(1) = \infty$.

Now let

$$A_n = \{x \in V \,:\, e_n(x) > 0\}$$

be the set of sites visited during the nth excursion. We also set $e_0 = \delta_o$ (where, as usual, $\delta_o(x) = 1$ if $x = o$ and 0 otherwise) and $A_0 = \{o\}$. For a subset $A \subset V$, define its outer boundary ∂A as the set

$$\partial A := \{y \notin A \,:\, (x, y) \in E \text{ for some } x \in A\}.$$

Lemma 12. For each $n \geq 0$, if $T(n+1) < \infty$ then

(i) $e_{n+1}(x) \leq \deg(x)$ for all $x \in V$,

(ii) $e_{n+1}(x) = \deg(x)$ for all $x \in A_n$,

(iii) $A_{n+1} \supseteq A_n \cup \partial A_n$.

Proof. Part (i) is immediate from Lemma 10.

Part (ii) follows from Lemma 11 and the observation that in the rotor configuration $\rho_{T(n)}$, the rotor at each $x \in A_n$ points along the edge traversed most recently from x, so for each $x \in A_n$ there is a directed path of rotors in $\rho_{T(n)}$ leading to $X_{T(n)} = o$.

Part (iii) follows from (ii): the $(n+1)$st excursion traverses each outgoing edge from each $x \in A_n$, so in particular it visits each vertex in $A_n \cup \partial A_n$. $\square$

Note that the balls $B(o, n)$ can be defined inductively by $B(o, 0) = \{o\}$ and

$$B(o, n + 1) = B(o, n) \cup \partial B(o, n)$$

for each $n \geq 0$. Inducting on n using Lemma 12(iii), we obtain the following.

Corollary 6.2. For each $n \geq 1$, if $T(n) < \infty$, then $B(o, n) \subseteq A_n$.

Rotor walk is called *recurrent* if $T(n) < \infty$ for all n. Consider the rotor configuration $\rho_{T(n)}$ at the end of the nth excursion. By Lemma 12, each vertex in $x \in A_n$ is visited exactly $\deg(x)$ times during the Nth excursion for each $N \geq n + 1$, so we obtain the following.

Corollary 6.3. For a recurrent rotor walk, $\rho_{T(N)}(x) = \rho_{T(n)}(x)$ for all $x \in A_n$ and all $N \geq n$.

The following proposition is a kind of converse to Lemma 12 in the case of undirected graphs.

Proposition 1. [8] Lemma 3;[3] Prop. 11 Let $G = (V, E)$ be an undirected graph. For a sequence $S_1, S_2, \ldots \subset V$ of sets inducing connected subgraphs such that $S_{n+1} \supseteq S_n \cup \partial S_n$ for all $n \geq 1$, and any vertex $o \in S_1$, there exists a rotor mechanism m and initial rotors ρ such that the nth excursion for rotor walk started at o traverses each edge incident to S_n exactly once in each direction, and no other edges.

2.3.3 Lower bound on the range

In this section $G = (V, E)$ is an infinite connected Eulerian graph. Fix an origin $o \in V$ and let $v(n)$ be the number of directed edges incident to the ball $B(o, n)$. Let $W(m) = \sum_{n=0}^{m-1} v(n)$. Write $W^{-1}(t) = \min\{m \in \mathbb{N} : W(m) > t\}$.

Fix a rotor mechanism m and an initial rotor configuration ρ on G. For $x \in V$ let $u_t(x)$ be the number of times x is visited by a rotor walk started at o and run for t steps. In the proof of the next theorem, our strategy for lower bounding the size of the range

$$R_t = \{x \in V : u_t(x) > 0\}$$

will be to (i) upper bound the number of excursions completed by time t, in order to (ii) upper bound the number of times each vertex is visited, so that (iii) many distinct vertices must be visited.

Theorem 7. For any rotor mechanism m, any initial rotor configuration ρ on G, and any time $t \geq 0$, the following bounds hold.

(i) $\frac{u_t(o)}{\deg(o)} < W^{-1}(t)$.

(ii) $\frac{u_t(x)}{\deg(x)} \leq \frac{u_t(o)}{\deg(o)} + 1$ for all $x \in V$.

(iii) Let $\Delta_t = \max_{x \in B(o,t)} \deg(x)$. Then

$$\#R_t \geq \frac{t}{\Delta_t(W^{-1}(t) + 1)}. \tag{2.9}$$

Before proving this theorem, let us see how it implies Theorem 3. The volume growth condition (2.8) implies $v(r) \geq kr^d$, so $W(r) \geq k'r^{d+1}$ for a constant k', so $W^{-1}(t) \leq (t/k')^{1/(d+1)}$. Now if G has bounded degree, then the right side of (2.9) is at least $ct^{d/(d+1)}$ for a constant c (which depends only on k and the maximal degree).

Proof of Theorem 7. We first argue that the total length $T(m)$ of the first m excursions is at least $W(m)$. By Corollary 6.2, the nth excursion visits every site in the ball $B(o,n)$. Therefore, by Lemma 12(ii), the $(n+1)$st excursion visits every site $x \in B(o,n)$ exactly $\deg(x)$ times, so the $(n+1)$st excursion traverses each directed edge incident to $B(o,n)$. The length $T(n+1) - T(n)$ of the $(n+1)$st excursion is therefore at least $v(n)$. Summing over $n < m$ yields the desired inequality $T(m) \geq W(m)$. Now let $m = W^{-1}(t)$. Since $t < W(m)$, the rotor walk has not yet completed its mth excursion at time t, so $u_t(o) < m \deg(o)$, which proves (i).

Part (ii) now follows from Lemma 10, since $e_1(x) = u_{T(1)}(x) \leq \deg(x)$. During each completed excursion, the origin o is visited $\deg(o)$ times while x is visited at most $\deg(x)$ times. The $+1$ accounts for the possibility that time t falls in the middle of an excursion.

Part (iii) follows from the fact that $t = \sum_{x \in B(o,t)} u_t(x)$. By parts (i) and (ii), each term in the sum is at most $\Delta_t(W^{-1}(t)+1)$, so there are at least $t/(\Delta_t(W^{-1}(t)+1))$ nonzero terms. $\square$

Pausing to reflect on the proof, we see that an essential step was the inclusion $B(o,n) \subseteq A_n$ of Corollary 6.2. Can this inclusion ever be an equality? Yes!

By Proposition 1, if G is undirected then there exists a rotor walk (that is, a particular m and ρ) for which

$$A_n = B(o,n) \qquad \text{for all } n \geq 1.$$

If $G = \mathbb{Z}^d$ (or any undirected graph satisfying (2.8) along with its upper bound counterpart, $\#B(o,n) \leq Kn^d$ for a constant K) then the range of this particular rotor walk satisfies $R_{W(n)} = B(o,n)$ and hence

$$\#R_t \leq \#B(o, W^{-1}(t)) \leq Ct^{d/(d+1)}$$

for a constant C. So in this case the exponent in Theorem 3 is best possible. We derived this upper bound just for a *particular* rotor walk, by choosing a rotor mechanism m and initial rotors ρ. For example, when $G = \mathbb{Z}^2$ the rotor mechanism is clockwise and the initial rotors are shown in Figure 2.6. Next we are going to see that by varying ρ we can make $\#R_t$ a lot larger.

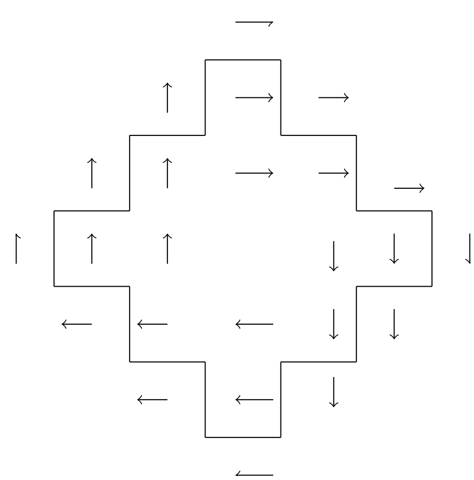


Figure 2.6: Minimal range rotor configuration for $\mathbb{Z}^2$. The excursion sets are diamonds.

Part (i) of the next theorem gives a sufficient condition for rotor walk to be transient. Parts (i) and (ii) together prove Theorem 4. Part (iii) shows that on a graph of bounded degree, the number of visited sites $\#R_t$ of a transient rotor walk grows linearly in t.

Theorem 8. On any Eulerian graph, the following hold:

(i) If ρ has an infinite path of initial rotors directed toward the origin o, then $u_t(o) < \deg(o)$ for all $t \geq 1$.

(ii) If $u_t(o) < \deg(o)$, then $\#R_t \geq t/\Delta_t$ where $\Delta_t = \max_{x \in B(o,t)} \deg(x)$.

(iii) If rotor walk is transient, then there is a constant $C = C(m, \rho)$ such that

$$\#R_t \geq \frac{t}{\Delta_t} - C$$

for all $t \geq 1$.

Proof. (i) By Corollary 6.1, if ρ has an infinite path directed toward o, then rotor walk never completes its first excursion from o.

(ii) If rotor walk does not complete its first excursion, then it visits each vertex x at most $\deg(x)$ times by Lemma 10, so it must visit at least t/Δ_t distinct vertices.

(iii) If rotor walk is transient, then for some n it does not complete its nth excursion, so this follows from part (ii) taking C to be the total length of the first $n-1$ excursions. □

2.3.4 Uniform rotor walk on the comb

The 2-dimensional *comb* is the subgraph of the square lattice $\mathbb{Z}^2$ obtained by removing all of its horizontal edges except for those on the x-axis (Figure 2.4). Vertices on the x-axis have degree 4, and all other vertices have degree 2.

Recall that the **uniform rotor walk** starts with independent random initial rotors $\rho(v)$ with the uniform distribution on outgoing edges from v. The following result shows that the range of the uniform rotor walk on the comb is close to the diamond

$$D_n := \{(x,y) \in \mathbb{Z}^2 \, : \, |x| + |y| < n\}.$$

Theorem 9. Consider uniform rotor walk on the comb with any rotor mechanism. Let $n \geq 2$ and $t = \lfloor \frac{16}{3} n^3 \rfloor$. For any $a > 0$ there exist constants $c, C > 0$ such that

$$\mathbb{P}\{D_{n-\sqrt{cn\log n}} \subset R_t \subset D_{n+\sqrt{cn\log n}}\} > 1 - Cn^{-a}.$$

Since the bounding diamonds have area $2n^2(1 + o(1))$ while t has order n^3, it follows that the size of the range is of order $t^{2/3}$. More precisely, by the first Borel-Cantelli lemma,

$$\frac{\#R_t}{t^{2/3}} \to \left(\frac{3}{2}\right)^{2/3}$$

as $t \to \infty$, almost surely. See[34] for more details.

The proof of Theorem 9 is based on the observation that rotor walk on the

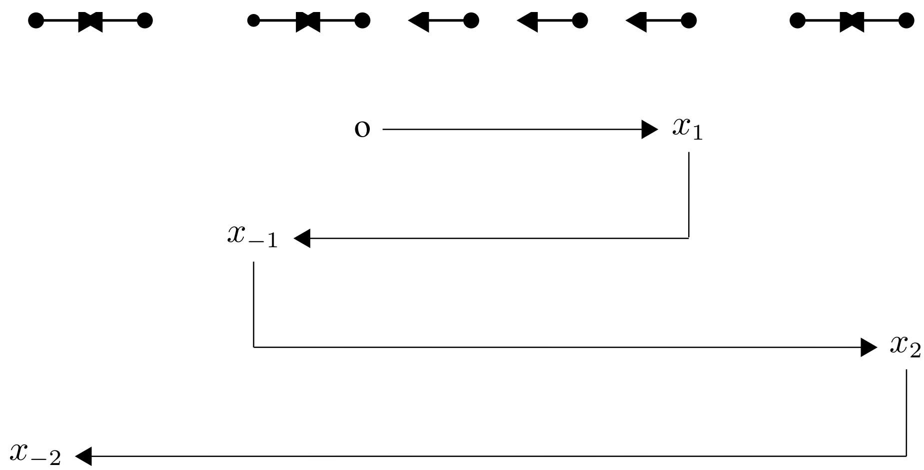


Figure 2.7: An initial rotor configuration on $\mathbb{Z}$ (top) and the corresponding rotor walk.

comb, viewed at the times when it is on the x-axis, is a rotor walk on $\mathbb{Z}$. If $0 < x_1 < x_2 < \dots$ are the positions of rotors on the positive x-axis that will send the walker left before right, and $0 > x_{-1} > x_{-2} > \dots$ are the positions on the negative x-axis that will send the walker right before left, then the x-coordinate of the rotor walk on the comb follows a zigzag path: right from 0 to x_1, then left to x_{-1}, right to x_2, left to x_{-2}, and so on (Figure 2.7).

Likewise, rotor walk on the comb, viewed at the times when it is on a fixed vertical line $x = k$, is also a rotor walk on $\mathbb{Z}$. Let $0 < y_{k,1} < y_{k,2} < \dots$ be the heights of the rotors on the line $x = k$ above the x-axis that initially send the walker down, and let $0 > y_{k,-1} > y_{k,-2} > \dots$ be the heights of the rotors on the line $x = k$ below the x-axis that initially send the walker up.

We only sketch the remainder of the proof; the full details are in[34]. For uniform initial rotors, the quantities x_i and $y_{k,i}$ are sums of independent geometric random variables of mean 2. We have $\mathbb{E}x_i = 2|i|$ and $\mathbb{E}y_{k,j} = 2|j|$. Standard concentration inequalities ensure that these quantities are close to their expectations, so that a rotor walk on the comb run for $n/2$ excursions visits each site $(x, 0) \in D_n$ about $(n - |x|)/2$ times, and hence visits each site $(x, y) \in D_n$ about $(n - |x| - |y|)/2$ times. Summing over $(x, y) \in D_n$ shows that the total time to complete these $n/2$ excursions is about $\frac{16}{3}n^3$. With high probability, every site in the smaller diamond $D_{n-\sqrt{cn \log n}}$ is visited at least once during these $n/2$ excursions, whereas no site outside the larger diamond $D_{n+\sqrt{cn \log n}}$ is visited.

2.3.5 Directed lattices and the mirror model

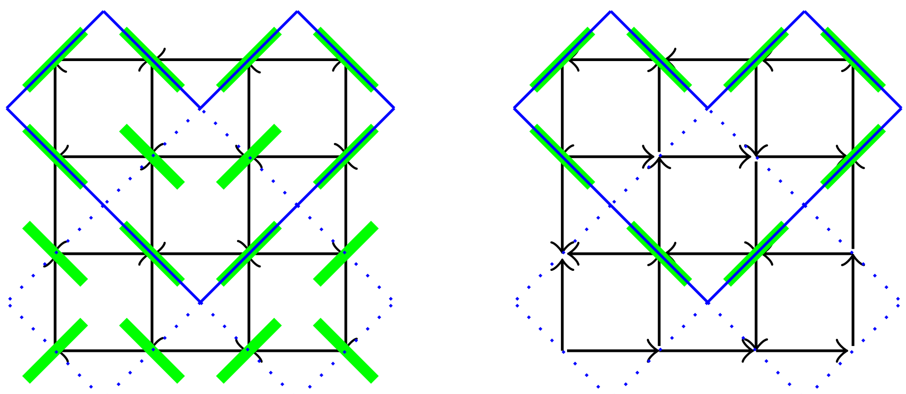
Figure 2.5 shows two different orientations of the square grid $\mathbb{Z}^2$: The *F- lattice* has outgoing vertical arrows (N and S) at even sites, and outgoing horizontal arrows (E and W) at odd sites. The *Manhattan lattice* has every even row pointing E, every odd row pointing W, every even column pointing S and every odd column pointing N. In these two lattices every vertex has outdegree 2, so there is a unique rotor mechanism on each lattice (namely, exits from a given vertex alternate between the two outgoing edges) and a rotor walk is completely specified by its starting point and the initial rotor configuration ρ.

In this section we relate the uniform rotor walk on these lattices to percolation and the Lorenz mirror model[43] §13.3. Consider the *half dual lattice* $\mathbb{L}$, a square grid whose vertices are the points $(x + \frac{1}{2}, y + \frac{1}{2})$ for $x, y \in \mathbb{Z}$ with $x + y$ even, and the usual lattice edges: $(x + \frac{1}{2}, y + \frac{1}{2}) - (x + \frac{1}{2}, y - \frac{1}{2})$, $(x + \frac{1}{2}, y + \frac{1}{2}) - (x - \frac{1}{2}, y + \frac{1}{2})$, $(x + \frac{1}{2}, y + \frac{1}{2}) - (x + \frac{3}{2}, y + \frac{1}{2})$, $(x + \frac{1}{2}, y + \frac{1}{2}) - (x + \frac{1}{2}, y + \frac{3}{2})$. We consider critical bond percolation on $\mathbb{L}$. Each possible lattice edge of $\mathbb{L}$ is either open or closed, independently with probability $\frac{1}{2}$.

Note that each vertex v of $\mathbb{Z}^2$ lies on a unique edge e_v of $\mathbb{L}$. We consider two different rules for placing two-sided mirrors at the vertices of $\mathbb{Z}^2$.

- F-lattice: Each vertex v has a mirror, which is oriented parallel to e_v if e_v is closed and perpendicular to e_v if e_v is open.

- Manhattan lattice: If e_v is closed then v has a mirror oriented parallel to e_v; otherwise v has no mirror.



(a) F-Lattice

(b) Manhattan lattice

Figure 2.8: Percolation on $\mathbb{L}$: dotted blue edges are open, solid blue edges are closed. Shown in green are the corresponding mirrors on the F-lattice (left) and Manhattan lattice.

Consider now the *first glance mirror walk*: Starting at the origin o, it travels along a uniform random outgoing edge $\rho(o)$. On its first visit to each vertex $v \neq \mathbb{Z}^2 - \{o\}$, the walker behaves like a light ray. If there is a mirror at v then the walker reflects by a right angle, and if there is no mirror then the walker continues straight. At this point v is assigned the rotor $\rho(v) = (v, w)$ where w

is the vertex of $\mathbb{Z}^2$ visited immediately after v. On all subsequent visits to v, the walker follows the usual rules of rotor walk.

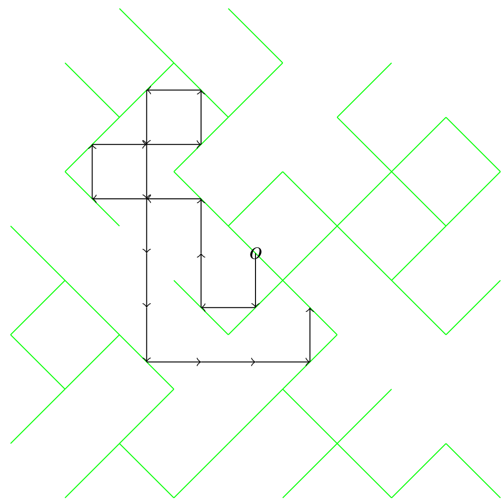


Figure 2.9: Mirror walk on the Manhattan lattice.

Lemma 13. With the mirror assignments described above, uniform rotor walk on the Manhattan lattice or the F-lattice has the same law as the first glance mirror walk.

Proof. The mirror placements are such that the first glance mirror walk must follow a directed edge of the corresponding lattice. The rotor $\rho(v)$ assigned by the first glance mirror walk when it first visits v is uniform on the outgoing edges from v; this remains true even if we condition on the past, because all previously assigned rotors are independent of the status of the edge e_v (open or closed), and changing the status of e_v changes $\rho(v)$. $\square$

Write $\beta_e = 1\{e \text{ is open}\}$. Given the random variables $\beta_e \in \{0, 1\}$ indexed by the edges of $\mathbb{L}$, we have described how to set up mirrors and run a rotor walk, using the mirrors to reveal the initial rotors as needed. The next lemma holds pointwise in β.

Lemma 14. If there is a cycle of closed edges in $\mathbb{L}$ surrounding o, then rotor walk started at o returns to o at least twice before visiting any vertex outside the cycle.

Proof. Denote by C the set of vertices v such that e_v lies on the cycle, and by A the set of vertices enclosed by the cycle. Let w be the first vertex not in $A \cup C$ visited by the rotor walk. Since the cycle surrounds o, the walker must arrive at w along an edge (v, w) where $v \in C$. Since e_v is closed, the walker reflects off the mirror e_v the first time it visits v, so only on the second visit to v does it use the outgoing edge (v, w). Moreover, the two incoming edges to v are on opposite sides of the mirror. Therefore by minimality of w, the walker must use the same incoming edge (u, v) twice before visiting w. The first edge to be used twice is incident to the origin by Lemma 10, so the walk must return to the origin twice before visiting w. □

Now we use a well-known theorem about critical bond percolation: there are infinitely many disjoint cycles of closed edges surrounding the origin. Together with Lemma 14 this completes the proof that the uniform rotor walk is recurrent both on the Manhattan lattice and the F-lattice.

To make a quantitative statement, consider the probability of finding a closed cycle within a given annulus. The following result is a consequence of the Russo-Seymour-Welsh estimate and FKG inequality; see[43] 11.72.

Theorem 10. Let $S_\ell = [-\ell, \ell] \times [-\ell, \ell]$. Then for all $\ell \geq 1$,

$$P(\textit{there exists a cycle of closed edges surrounding the origin in } S_{3\ell} - S_\ell) > p$$

for a constant p that does not depend on ℓ.

Let $u_t(o)$ be the number of visits to o by the first t steps of uniform rotor walk in the Manhattan or F-lattice.

Theorem 11. For any $a > 0$ there exists $c > 0$ such that

$$P(u_t(o) < c \log t) < t^{-a}.$$

Proof. By Lemma 14, the event $\{u_t(o) < k\}$ is contained in the event that at most $k/2$ of the annuli $S_{3^j} - S_{3^{j-1}}$ for $j = 1, \ldots, \frac{1}{10} \log t$ contain a cycle of closed edges surrounding the origin. Taking $k = c \log t$ for sufficiently small c, this event has probability at most t^{-a} by Theorem 10. □

Although we used the same technique to show that the uniform rotor walk on these two lattices is recurrent, experiments suggest that behavior of the two walks is rather different. The number of distinct sites visited in t steps appears to be of order $t^{2/3}$ on the Manhattan lattice but of order t for F-lattice. This difference is clearly visible in Figure 2.10.

2.4 Time for rotor walk to cover a finite Eulerian graph

Let $(X_t)_{t \geq 0}$ be a rotor walk on a finite connected Eulerian directed graph $G = (V, E)$ with diameter D. The *vertex cover time* is defined by

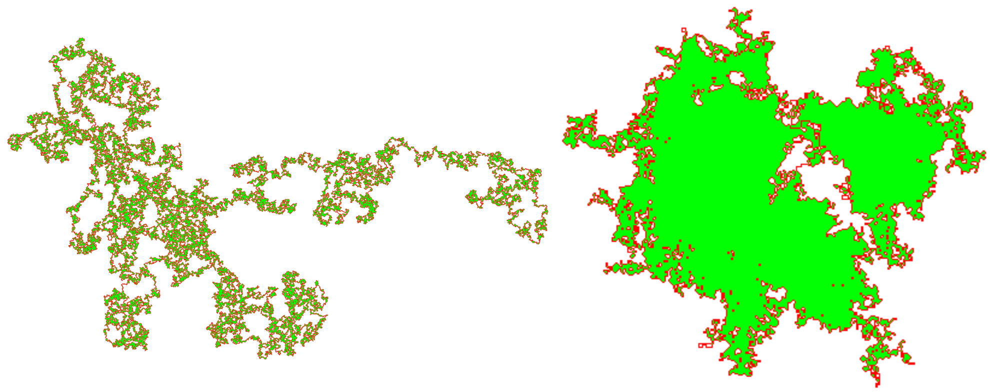


Figure 2.10: Set of sites visited by uniform rotor walk after 250000 steps on the F-lattice and the Manhattan lattice (right). Green represents at least two visits to the vertex and red one visit.

$$t_{\text{vertex}} = \min\{t \,:\, \{X_s\}_{s=1}^{t} = V\}.$$

The *edge cover time* is defined by

$$t_{\text{edge}} = \min\{t \,:\, \{(X_{s-1}, X_s)\}_{s=1}^{t} = E\}$$

where E is the set of directed edges. Yanovski, Wagner and Bruckstein[89] show $t_{\text{edge}} \leq 2D\#E$ for any Eulerian directed graph. The next result improves this bound slightly, replacing $2D$ by $D + 1$.

Theorem 12. For rotor walk on a finite Eulerian graph G of diameter D, with any rotor mechanism m and any initial rotor configuration ρ,

$$t_{\text{vertex}} \leq D\#E$$

and

$$t_{\text{edge}} \leq (D+1)\#E.$$

Proof. Consider the time $T(n)$ for rotor walk to complete n excursions from o. If G has diameter D then $A_D = V$ by Corollary 6.2, and $e_{D+1} \equiv \deg$ by Lemma 12(ii). It follows that $t_{\text{vertex}} \leq T(D)$ and $t_{\text{edge}} \leq T(D+1)$. By Lemma 10, each directed edge is used at most once per excursion so $T(n) \leq n\#E$ for all $n \geq 0$. □

Bampas et al.[8] prove a corresponding lower bound: on any finite undirected graph there exist a rotor mechanism m and initial rotor configuration ρ such that $t_{\text{vertex}} \geq \frac{1}{4}D\#E$.

2.4.1 Hitting times for random walk

The upper bounds for t_{vertex} and t_{edge} in Theorem 12 match (up to a constant factor) those found by Friedrich and Sauerwald[39] on an impressive variety of graphs: regular trees, stars, tori, hypercubes, complete graphs, lollipops and expanders. Using a theorem of Holroyd and Propp[48] relating rotor walk to the expected time $H(u,v)$ for *random* walk started at u to hit v, they infer that $t_{\text{vertex}} \leq K+1$ and $t_{\text{edge}} \leq 3K$, where

$$K := \max_{u,v \in V} H(u,v) + \frac{1}{2}\left(\#E + \sum_{(i,j)\in E} |H(i,v) - H(j,v) - 1|\right).$$

A curious consequence of the upper bound $t_{\text{vertex}} \leq K+1$ of[39] and the lower bound $\max_{m,\rho} t_{\text{vertex}}(m,\rho) \geq \frac{1}{4}D\#E$ of[8] is the following inequality.

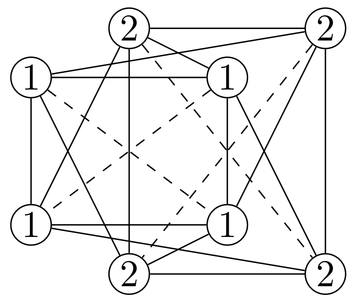


Figure 2.11: The thick cycle $G_{\ell,N}$ with $\ell = 4$ and $N = 2$. Long-range edges are dotted and short-range edges are solid.

Corollary 12.1. For any undirected graph G of diameter D we have

$$K \geq \frac{1}{4} D \# E - 1.$$

Is K always within a constant factor of $D\#E$? It turns out the answer is no. To construct a counterexample we will build a graph $G = G_{\ell,N}$ of small diameter which has so few long-range edges that random walk effectively does not feel them (Figure 2.11). Let $\ell, N \geq 2$ be integers and set $V = \{1, \ldots, \ell\} \times \{1, \ldots, N\}$ with edges $(x, y) \sim (x', y')$ if either $x' \equiv x \pm 1 \pmod{\ell}$ or $y' = y$. The diameter of G is 2: any two vertices (x, y) and (x', y') are linked by the path $(x, y) \sim (x + 1, y') \sim (x', y')$. Each vertex (x, y) has $2N$ short-range edges to $(x \pm 1, y')$ and $\ell - 3$ long-range edges to (x', y). It turns out that if ℓ is sufficiently large and N is much larger still ($N = \ell^5$), then $K > \frac{1}{10} \ell \# E$, showing that K can exceed $D\#E$ by an arbitrarily large factor. The details can be found in[34].

We conclude with a curious observation and a question. Corollary 12.1 is a fact purely about random walk on a graph. Can it be proved without resorting to rotor walk?

3

Diffusions

3.1 Frozen random walk

This chapter is based on papers[33] and[36].

The goal of this chapter is to understand the long term behavior of the mass evolution process which is a divisible version of the particle system "Frozen Random Walk". We define **Frozen-Boundary Diffusion** with parameter $\alpha \in (0,1)$ (or **FBD**-α) as follows. Informally it is a sequence μ_t of symmetric probability distributions on $\mathbb{Z}$. The sequence has the following recursive definition: given μ_t, the leftmost and rightmost $\frac{\alpha}{2}$ masses are constrained to not move, and the remaining $1 - \alpha$ mass diffuses according to one step of the discrete heat equation to yield μ_{t+1}. In other words, we split the mass at site x equally to its two neighbors. Formal descriptions appear later. We briefly remark that this process is similar to Stefan type problems, which have been studied for example in[42].

Now we also introduce the random counterpart of **FBD**-α. We define the frozen random walk process (Frozen Random Walk-$(n, 1/2)$) as follows: n particles start at the origin. At any discrete time the leftmost and rightmost $\lfloor n\frac{\alpha}{2}\rfloor$ particles are "frozen" and do not move. The remaining $n - 2\lfloor n\frac{\alpha}{2}\rfloor$ particles independently jump to the left and right uniformly, all at the same time. Letting $n \to \infty$ and fixing t, the mass distribution for the above random process converges to the t^{th} element, μ_t, in **FBD**-α. However, if t and n simultaneously go to ∞, one has to control the fluctuations to be able to prove any limiting statement. Figure 3.1 depicts the mass distribution μ_t and the frozen random walk process for $\alpha = \frac{1}{2}$.

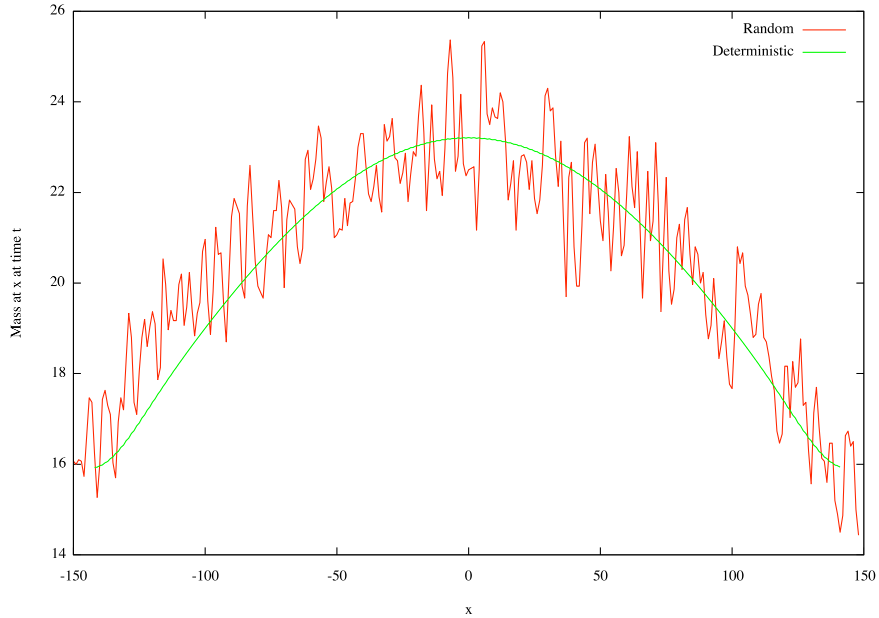


Figure 3.1: The free mass of **Frozen-Boundary Diffusion-**$\frac{1}{2}$ and **Frozen Random Walk-**$(10000, \frac{1}{2})$ averaged over 15 trials at $t = 25000$. For parity considerations, the values at x are the averages over x and $x+1$. We identify the limit of **FBD-**α in Theorem 13.

At every step t of **FBD-**α, we also keep track of the location of the boundary of the process, β_t, which we define as

$$\beta_t := \sup \left\{ x \in \mathbb{Z} : \mu_t \left([x, \infty) \right) \geq \frac{\alpha}{2} \right\}.$$

Our first result is

Lemma 15. For every $\alpha \in (0,1)$ there exist constants $a, b > 0$ such that

$$a\sqrt{t} < \beta_t < b\sqrt{t} \; , \forall \, t.$$

The lemma above suggests that a proper scaling of β_t is $\sqrt{t}$. Motivated by this behavior of the boundary β_t, one can ask the following natural questions:

Question 1. Does $\frac{\beta_t}{\sqrt{t}}$ converge?

Considering μ_t as a measure on $\mathbb{R}$, for $t = 0, 1, \dots$ define the Borel measure $\tilde{\mu}_t(\alpha) = \tilde{\mu}_t$ on $\mathbb{R}$ equipped with the Borel $\sigma-$algebra such that for any Borel set A,

$$\tilde{\mu}_t(A) = \mu_t(\{y\sqrt{t} : y \in A\}). \tag{3.1}$$

We can now ask

Question2. Does the sequence of probability measures $\tilde{\mu}_t$ have a weak limit?

Question 3. If $\tilde{\mu}_t$ has a weak limit, what is this limiting distribution?

We conjecture affirmative answers to **Q1** and **Q2**:

Conjecture 2. For every $\alpha \in (0, 1)$, there exists $\ell_\alpha > 0$ such that

$$\lim_{t\to\infty} \frac{\beta_t}{\sqrt{t}} = \ell_\alpha.$$

Conjecture 3. Fix $\alpha \in (0, 1)$. Then there exists a probability measure $\mu_\infty(\alpha)$ on $\mathbb{R}$ such that as $t \to \infty$,

$$\tilde{\mu}_t \overset{weak}{\Longrightarrow} \mu_\infty(\alpha),$$

where $\overset{weak}{\Longrightarrow}$ denotes weak convergence in the space of finite measures on $\mathbb{R}$.

That Conj. 3 implies Conj. 2 is the content of Lemma 17. We now state our main result which shows that Conj. 2 implies Conj. 3 and identifies the limiting distribution, thus answering **Q3.** To this end we need the following definition.

Definition 2. Let $\Phi(\cdot)$ be the standard Gaussian measure on $\mathbb{R}$. Also for any $q > 0$ denote by $\Phi_q(\cdot)$, the probability measure on $\mathbb{R}$ which is supported on $[-q, q]$ and whose density is the standard Gaussian density restricted on the interval $[-q, q]$ and properly normalized to have integral 1.

Theorem 13. Assuming that $\lim\limits_{t\to\infty} \dfrac{\beta_t}{\sqrt{t}}$ is a constant, the following is true:

$$\tilde{\mu}_t \overset{weak}{\Longrightarrow} \mu_\infty(\alpha),$$

where,

$$\mu_\infty(\alpha) = \frac{\alpha}{2}\delta(-q_\alpha) + (1-\alpha)\Phi_{q_\alpha} + \frac{\alpha}{2}\delta(q_\alpha),$$

and q_α is the unique positive number such that:

$$\frac{\alpha}{2}q_\alpha = \frac{(1-\alpha)e^{-q_\alpha^2/2}}{\sqrt{2\pi}\Phi([-q_\alpha, q_\alpha])}.$$

Remark 1. It is easy to show (see Lemma 17) that the above result implies that

$$\lim_{t\to\infty} \frac{\beta_t}{\sqrt{t}} = q_\alpha.$$

Thus observe that by the above result, just assuming that the boundary location properly scaled converges to a constant determines the value of the constant. This is a consequence of uniqueness of the root of a certain functional equation discussed in detail in Section 3.1.2.

3.1.1 Formal definitions

Let **FBD**-$\alpha := \{\mu_0, \mu_1, \ldots\}$: where for each $t = 0, 1, \ldots$, μ_t is a probability distribution on $\mathbb{Z}$. For brevity we suppress the dependence on α in the notation since there is no scope of confusion as α will remain fixed throughout any argument. We do the same for $\mu_\infty(\alpha)$, q_α and ℓ_α, replacing them by μ, q and ℓ. Thus

$$\mu := \mu_\infty(\alpha) = \frac{\alpha}{2}\delta(-q) + (1-\alpha)\Phi_q + \frac{\alpha}{2}\delta(q). \tag{3.2}$$

Let $\mu_0 \equiv \delta(0)$ be the delta function at 0. In the discrete setting, this function takes value 1 at 0 and 0 otherwise. By construction μ_t will be symmetric for all t. As described above, each μ_t contains a "constrained/frozen" part and a "free" part. Let the free mass and the frozen mass be denoted by the mass distributions ν_t and f_t respectively.

Recall the boundary of the process,

$$\beta_t = \sup\left\{x \in \mathbb{Z} : \mu_t\left([x, \infty)\right) \geq \frac{\alpha}{2}\right\}. \tag{3.3}$$

Then for all $y \geq 0$,

$$f_t(y) := \begin{cases} \mu_t(y) & y > \beta_t \\ \frac{\alpha}{2} - \sum_{z>\beta_t} \mu_t(z) & y = \beta_t \\ 0 & \text{otherwise.} \end{cases} \tag{3.4}$$

For $y < 0$ let $f_t(y) := f_t(-y)$. Thus f_t is the extreme $\alpha/2$ mass on both sides of the origin. Define the free mass to be $\nu_t := \mu_t - f_t$. With the above notation the heat diffusion is described by

$$\mu_{t+1}(x) = \frac{\nu_t(x-1) + \nu_t(x+1)}{2} + f_t(x). \tag{3.5}$$

Recall Lemma 15, which implies the diffusive nature of the boundary:

Lemma 1. For every $\alpha \in (0,1)$ there exist constants $a, b > 0$ such that

$$a\sqrt{t} < \beta_t < b\sqrt{t}\ , \forall\, t.$$

This result implies that in order to obtain any limiting statement about the measures μ_t, one has to scale space down by $\sqrt{t}$.

The proof of the lemma appears later. Let us first prove that the frozen mass f_t cannot be supported over many points.

Lemma 16. For all t, the frozen mass at time t, f_t, is supported on at most two points on each side of the origin, i.e., for all $y \in \mathbb{Z}$ such that $|y| \geq \beta_t + 2$, we have $f_t(y) = 0$.

Proof. The lemma follows by induction. Assume for all $k \leq t$, for all y such that $|y| \geq \beta_k + 2$, we have $\mu_k(y) = f_k(y) = 0$. The base case $t = 0$ is easy to check.

Now observe that by (3.5) and the above induction hypothesis,

$$\mu_{t+1}(y) = 0, \tag{3.6}$$

for all $|y| \geq \beta_t + 2$. Also notice that by (3.5) it easily follows that β_t is a non-decreasing function of t. Thus clearly for all y, with $|y| \geq \beta_{t+1} + 2 \geq \beta_t + 2$,

$$\mu_{t+1}(y) = 0.$$

Hence we are done by induction.

$\square$

We now return to the proof of the diffusive nature of the boundary of the process β_t.

Proof of Lemma 15. We consider the second moment of the mass distribution μ_t, which we denote as $M_2(t) := \sum_{x \in \mathbb{Z}} \mu_t(x) x^2$. This is at most $(\beta_t + 1)^2$ since μ_t is supported on $[-\beta_t - 1, \beta_t + 1]$ by Lemma 16. It is also at least $\alpha {\beta_t}^2$ since there exists mass α which is at a distance at least β_t from the origin. Now we observe how the second moment of the mass distribution evolves over time. Suppose a free mass m at x splits and moves to $x - 1$ and $x + 1$. Then the increase in the second moment is

$$\frac{m}{2}((x + 1)^2 + (x - 1)^2) - mx^2 = m.$$

Since at every time step exactly $1 - \alpha$ mass is moving, the net change in the second moment at every step is $1-\alpha$. So at time t the second moment is exactly

$$t(1-\alpha). \tag{3.7}$$

Hence $\alpha {\beta_t}^2 < t(1-\alpha) < (\beta_t + 1)^2$, and we are done. $\square$

We next prove Conjecture 2 (a stronger version of Lemma 15) assuming Conjecture 3.

Lemma 17. If Conjecture 3 holds, then so does Conjecture 2, i.e., for every $\alpha \in (0,1)$, there exists $\ell > 0$, such that

$$\lim_{t\to\infty} \frac{\beta_t}{\sqrt{t}} = \ell.$$

Proof. Fix $\alpha \in (0,1)$. From Lemma 15 we know that $\{\beta_t/\sqrt{t}\}$ is bounded. Hence, if $\beta_t/\sqrt{t}$ does not converge, there exists two subsequences $\{s_1, s_2, \ldots\}$ and $\{t_1, t_2, \ldots\}$ such that

$$\lim_{i\to\infty} \beta_{s_i}/\sqrt{s_i} = \ell_1 \text{ and } \lim_{j\to\infty} \beta_{t_j}/\sqrt{t_j} \to \ell_2,$$

for some $\ell_2, \ell_1 > 0$ such that $\ell_2 - \ell_1 := \delta > 0$. Recall $\mu_\infty(\alpha) := \mu$ from Conjecture 3. Now by hypothesis,

$$\lim_{i\to\infty} \tilde{\mu}_{s_i} \overset{weak}{\Longrightarrow} \mu, \quad \lim_{j\to\infty} \tilde{\mu}_{t_j} \overset{weak}{\Longrightarrow} \mu.$$

This yields a contradiction since the first relation implies μ assigns mass 0 to the interval $(l_2 - \frac{\delta}{2}, l_2 + \frac{\delta}{2})$ while the second one implies (by Lemma 16) that it assigns mass at least $\frac{\alpha}{2}$ to that interval. $\square$

3.1.2 Proof of Theorem 13

The proof follows by observing the moment evolutions of the mass distributions μ_t and using the moment method. The proof is split into several lemmas. Denote the k^{th} moment of μ_t as $M_k(t)$. We now make some simple observations which are consequences of the previously stated lemmas. Recall the free and frozen mass distributions ν_t and f_t. We denote the k^{th} moments of the measures ν_t (the free mass at time t), f_t (the frozen mass at time t), by $M_k^{\nu}(t)$ and $M_k^{f}(t)$ respectively. Also define $\tilde{f}_t$ and $\tilde{\nu}_t$ similarly to $\tilde{\mu}_t$ in (3.1). Assuming Conjecture 2, it follows from Lemma 16 that,

$$\tilde{f}_t \overset{weak}{\Longrightarrow} f \tag{3.8}$$

where $f := \frac{\alpha}{2}\delta(-\ell) + \frac{\alpha}{2}\delta(\ell)$, and $\ell = \ell_\alpha$ appears in the statement of Conjecture 2. This implies that

$$\frac{M_k^f(t)}{t^{k/2}} = \begin{cases} 0, & k \text{ odd} \\ \alpha\ell^k(1+o(1)), & k \text{ even} \end{cases} \tag{3.9}$$

where $o(1)$ goes to 0 as t goes to infinity.

The proof of Theorem 13 is in two steps: first we show that $\ell = q$ and then show that $\tilde{\nu}_t$ converges weakly to the part of μ which is absolutely continuous with respect to the Lebesgue measure. Clearly the above two results combined imply Theorem 13.

As mentioned this is done by observing the moment sequence $M_k(t)$. Now notice owing to symmetry of the measures μ_t for any t, $M_{2k+1}(t) = 0$ for all non-negative integers k.

Thus it suffices to consider $M_{2k}(t)$ for some non-negative integer k. We begin by observing that at any time t the change in the moment $M_{2k}(t+1) - M_{2k}(t)$, is caused by the movement of the free mass ν_t. The change caused by a mass m moving at a site x (already argued in the proof of Lemma 15 for $k = 1$) is

$$\frac{m((x+1)^{2k} + (x-1)^{2k})}{2} - mx^{2k} = m\left[\sum_{i=1}^{k} \binom{2k}{2k-2i} x^{2k-2i}\right]. \tag{3.10}$$

Now summing over x we get that,

$$M_{2k}(t+1) - M_{2k}(t) = \sum_{i=1}^{k} \binom{2k}{2k-2i} M^{\nu}_{2k-2i}(t). \tag{3.11}$$

Notice that the moments of the free mass distribution ν_t appear on the RHS since m in (3.10) was the free mass at a site x. Now using (3.11) we sum $M_{2k}(j+1) - M_{2k}(j)$ over $0 \leq j \leq t-1$ and normalize by t^k to get

$$\frac{M_{2k}(t)}{t^k} = \sum_{j=0}^{t-1}\left[\sum_{i=1}^{k} \binom{2k}{2k-2i} M^{\nu}_{2k-2i}(j)\frac{1}{t^k}\right]. \tag{3.12}$$

Recall that by Lemma 15, for any $k \geq 1$, $M^{\nu}_{2k-2}(j)$ is $O(j^{k-1})$. Moreover, the above equation allows us to make the following observation:

Claim. Assume (3.9) holds. Then for any $k \geq 1$, the existence of $\lim_{j\to\infty} \frac{M^{\nu}_{2k-2}(j)}{j^{k-1}}$ implies existence of $\lim_{j\to\infty} \frac{M^{\nu}_{2k}(j)}{j^{k}}$.

Proof of claim. Notice that by Lemma 15, $M^{\nu}_{2k-\ell}(j) = O(j^{k-2})$ for any $\ell \leq 4$.

Also let

$$\lim_{j\to\infty} \frac{M^{\nu}_{2k-2}(j)}{j^{k-1}} = M^{\nu}_{2k-2},$$

which exists by hypothesis.

Thus using (3.12) and the standard fact that

$$\lim_{t\to\infty} \sum_{j=0}^{t-1} \frac{j^{k-1}}{t^{k-1}} \frac{1}{t} = \int_0^1 x^{k-1} dx = \frac{1}{k}$$

we get by bounded convergence

$$\sum_{j=1}^{t-1} \left[\sum_{i=1}^{k} \binom{2k}{2k-2i} \frac{M^{\nu}_{2k-2i}(j)}{t^k} \right] = (2k-1)M^{\nu}_{2k-2} + o(1) + O\left(\frac{1}{t}\right).$$

Thus

$$\lim_{t\to\infty} \frac{M_{2k}(t)}{t^k} = (2k-1)M^{\nu}_{2k-2} \tag{3.13}$$

and since

$$M_{2k}(t) = M^{\nu}_{2k}(t) + M^{f}_{2k}(t), \tag{3.14}$$

we are done by (3.9).

□

Using the above claim, the fact that $\lim_{t\to\infty} \frac{M_k(t)}{t^{k/2}}$ and hence, $\lim_{t\to\infty} \frac{M^{\nu}_k(t)}{t^{k/2}}$ (by (3.14) and (3.9)) exists for all k, follows from the fact that $\frac{M_2(t)}{t} = (1-\alpha)$ (see (3.7)).

Let us call the limits M_k and M^{ν}_k respectively.

Thus we have

$$M_{2k} = M^{\nu}_{2k} + \alpha\ell^{2k} = (2k-1)M^{\nu}_{2k-2}, \tag{3.15}$$

where the first equality is by (3.14) and (3.9) and the second by (3.13). For $k = 1$ we get

$$\alpha\ell^2 + M^{\nu}_2 = 1 - \alpha.$$

Notice that this implies that for all k, M^{ν}_{2k} can be expressed in terms of a polynomial in ℓ of degree $2k$, which we denote as $P_k(\ell)$. Then, by (3.15) the polynomials P_k satisfy the following recurrence relation:

$$\begin{aligned} P_k(\ell) &= (2k-1)P_{k-1}(\ell) - \alpha\ell^{2k} \\ P_0 &= 1 - \alpha. \end{aligned} \tag{3.16}$$

By definition, we have

$$P_k(\ell) = M^{\nu}_{2k} = \lim_{t\to\infty} \frac{M^{\nu}_{2k}(t)}{t^k} = \lim_{t\to\infty} \frac{\sum\limits_{-\beta_t \le x \le \beta_t} x^{2k}\nu_t(x)}{t^k}. \tag{3.17}$$

Thus assuming Conj. 2 and the fact that $\sum\limits_{-\beta_t \le x \le \beta_t} \nu_t(x) = 1 - \alpha$ for all t, we get the following family of inequalities,

$$0 \le P_k(\ell) \le (1-\alpha)\ell^{2k} \;\; \forall\, k \ge 0. \tag{3.18}$$

We next show that the above inequalities are true only if $\ell = q$ where q appears in (3.2).

Lemma 18. The inequalities in (3.18) are satisfied by the unique number ℓ

such that

$$\frac{\alpha}{2}\ell = \frac{(1-\alpha)e^{-\ell^2/2}}{\sqrt{2\pi}\Phi([-\ell,\ell])}$$

where $\Phi(\cdot)$ is the standard Gaussian measure.

Thus the above implies that necessarily $\ell = q$ where q appears in (3.2). This was mentioned in Remark 1.

Proof. To prove this, first we write the inequalities in a different form so that the polynomials stabilize. To this goal, let us define

$$\tilde{P}_k = \frac{P_k}{(2k-1)!!}$$

where $(2k-1)!! = (2k-1)(2k-3)\ldots 1$. Then it follows from (3.16) that

$$\tilde{P}_k(\ell) = \tilde{P}_{k-1}(\ell) - \frac{\alpha}{(2k-1)!!}\ell^{2k}.$$

Hence

$$\tilde{P}_k(\ell) = \left(1 - \alpha - \sum_{i=1}^{k} \frac{\alpha\ell^{2i}}{(2i-1)!!}\right).$$

The inequalities in (3.18) translate to

$$0 \leq 1 - \alpha - \sum_{i=1}^{k} \frac{\alpha\ell^{2i}}{(2i-1)!!} \leq \frac{\ell^{2k}}{(2k-1)!!}. \tag{3.19}$$

Let us first identify the power series

$$g(x) = \sum_{i=1}^{\infty} \frac{x^{2i-1}}{(2i-1)!!}.$$

Clearly the power series converges absolutely for all values of x. It is also standard to show that one can interchange differentiation and the sum in the expression for $g(\cdot)$. Thus we have that $g(x)$ satisfies,

$$\frac{dg(x)}{dx} = 1 + xg(x).$$

Solving this differential equation using integrating factor $e^{-x^2/2}$ and the fact that $g(0) = 0$ we get

$$g(x) = e^{x^2/2} \int_0^x e^{-y^2/2} dy.$$

As $k \to \infty$, the upper bound in (3.19) converges to 0 for any value of ℓ. Also the expression in the middle converges to $1 - \alpha - \alpha \ell g(\ell)$. Thus taking the limit in (3.19) as $k \to \infty$ we get that $\ell > 0$ satisfies

$$\ell g(\ell) = \frac{1-\alpha}{\alpha}. \tag{3.20}$$

Clearly this is the same as the equation appearing in the statement of the lemma. Also notice that since $xg(x)$ is monotone on the positive real axis, by the uniqueness of the solution of (3.20) we get $\ell = q$ where q appears in (3.2). Hence we are done. $\square$

The value of ℓ that solves (3.20) when $\alpha = 1/2$ is approximately 0.878. Figure 3.2 shows the numerical convergence of $\beta_t/\sqrt{t}$ to q_α for various values of α.

Thus, assuming Conjecture 2, by Lemma 18, $\tilde{f}_t$ converges to f (as stated in (3.8)) which consists of two atoms of size $\frac{\alpha}{2}$ at q and $-q$. To conclude the proof of Theorem 13, we now show $\tilde{\nu}_t$ converges to the absolutely continuous part of

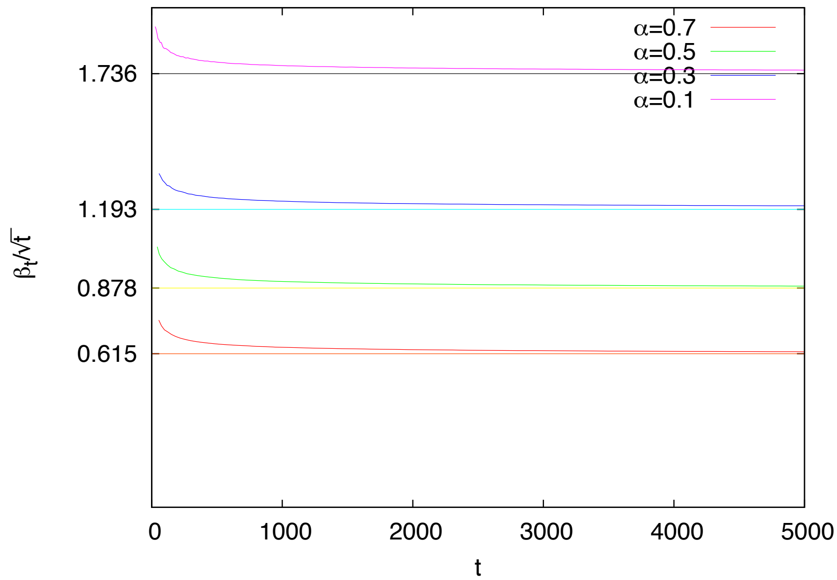


Figure 3.2: Convergence of $\beta_t/\sqrt{t}$ for various α. The horizontal lines denote the values q_α and the curves plot $\frac{\beta_t}{\sqrt{t}}$ as a function of time t.

μ (see (3.2)). Recall that by (3.17) and Lemma 18 the $2k^{th}$ moment of $\tilde{\nu}_t$ converges to $P_k(q)$. We will use the following well known result:

Lemma 19 (30.1,[14]). Let μ be a probability measure on the line having finite moments $\alpha_k = \int_{-\infty}^{\infty} x^k \mu(dx)$ of all orders. If the power series $\sum_k \alpha_k r^k/k!$ has a positive radius of convergence, then μ is the only probability measure with the moments $\alpha_1, \alpha_2, \ldots$.

Thus, to complete the proof of Theorem 13 we need to show the following:

Claim. The $2k^{th}$ moment of the measure $(1-\alpha)\Phi_q$ is $P_k(q)$ where q is the q_α appearing in Theorem 13.

To prove this claim, it suffices to show that the moments of $(1-\alpha)\Phi_q$ satisfy the recursion (3.16). Recall that $q = q_\alpha$. Let $C = C_\alpha := \frac{\sqrt{2\pi}\Phi([-q,q])}{1-\alpha}$. Using integration by parts we have:

$$\begin{aligned} \frac{\int_{-q}^{q} x^{2k} e^{-\frac{x^2}{2}} dx}{C} &= \frac{\int_{-q}^{q} x^{2k-1} x e^{-\frac{x^2}{2}} dx}{C} \\ &= -\frac{2q^{2k-1} e^{-\frac{q^2}{2}}}{C} + \frac{(2k-1)}{C} \int_{-q}^{q} x^{2k-2} e^{-\frac{x^2}{2}} dx. \end{aligned}$$

By the relation that q satisfies in the statement of Theorem 13, the first term on the RHS without the $-$ sign is αq^{2k}. Also, note that the second term is $(2k-1)$ times the $(2k-2)^{nd}$ moment of $(1-\alpha)\Phi_q$. Thus, the moments of $(1-\alpha)\Phi_q$ satisfy the same recursion as in (3.16).

Now from Example 30.1 in[14], we know that the absolute value of the k^{th} moment of the standard normal distribution is bounded by $k!$. Then, similarly, the absolute value of the k^{th} moment of our truncated Gaussian, Φ_q, is bounded by $c^k k!$ for a constant c. Then Lemma 19 implies that Φ_q is determined by its moments and quoting Theorem 30.2 in[14] we are done. □

3.1.3 Concluding Remarks

We conclude with a brief discussion about a possible approach towards proving Conjectures 2 and 3 and some experiments in higher dimensions.

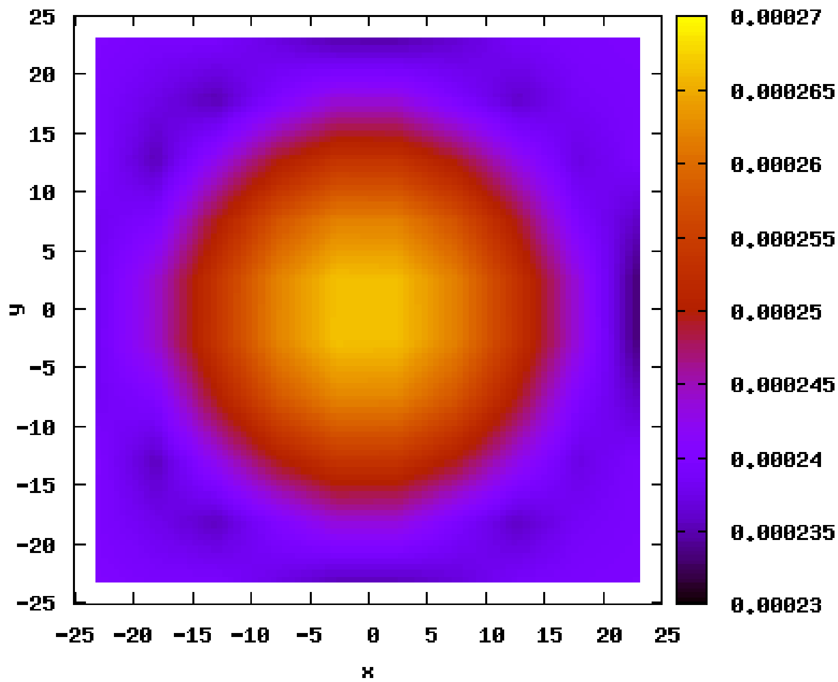


Figure 3.3: Heat map of the free mass distribution after 1000 steps in 2 dimensions for **FBD**-1/2.

The free part ν_t of the distribution μ_t could represent the distribution of a random walk in a growing interval. If the interval boundaries grow diffusively, the scaling limit of this random process will be a reflected Ornstein-Uhlenbeck process on this interval $[-q, q]$. We remark that the stationary measure for Ornstein-Uhlenbeck process reflected on the interval is known to be the same truncated Gaussian which appears in Theorem 13, see[62] (31). This connection could be useful in proving the conjectures.

We also note that similar results are expected in higher dimensions; in particular, the mass distribution should exhibit rotational symmetry.

3.2 Optimal controlled diffusion

Suppose that we have a unit mass at the origin of the d-dimensional lattice $\mathbb{Z}^d$ and we wish to move half of the mass to distance n. If the only moves we are allowed to make take a vertex and split the mass at the vertex equally among its neighbors, how many moves do we need to accomplish this goal? The one-dimensional case was solved by Paterson, Peres, Thorup, Winkler, and Zwick[74], who studied this question due to its connections with the maximum overhang problem[75,74]. The main result of this section solves this problem in $\mathbb{Z}^d$ for general d; the proof builds on the one-dimensional case, but requires new ideas. We also explore this question on several other graphs, such as the comb, regular trees, Galton-Watson trees, and more.

The problem also has a probabilistic interpretation. Suppose there is a particle at the origin of $\mathbb{Z}^d$, as well as a controller who cannot see the particle (but who knows that the particle is initially at the origin). The goal of the controller is to move the particle to distance n from the origin and it can give commands of the type "jump if you are at vertex v". The particle does not move unless the controller's command correctly identifies the particle's location, in which case the particle jumps to a neighboring vertex chosen uniformly at random. How many commands does the controller have to make in order for the particle to be at distance n with probability at least $1/2$?

3.2.1 Setting and main result

Definition 3 (Toppling moves)**.** Given a graph $G = (V, E)$ and a mass distribution μ on the vertex set V, a *toppling move* selects a vertex $v \in V$ with positive mass $\mu(v) > 0$ and topples (part of) the mass equally to its neighbors. We denote by T_v^m the toppling move that topples mass m at vertex v, resulting in the mass distribution $T_v^m \mu$.

Given a subset of the vertices $A \subset V$, mass $p > 0$, and an initial mass distribution μ_0, we define $N_p(G, A, \mu_0)$ to be the minimum number of toppling moves needed to move mass p outside of the set A, i.e., the minimum number of toppling moves needed to obtain a mass distribution μ such that $\sum_{v \notin A} \mu(v) \geq p$.

Our interest is in the case when the initial mass distribution is a unit mass δ_o at a given vertex o and A is the (open) ball of radius n around o, i.e., $A = B_n := \{u \in V : d_G(u, o) < n\}$, where d_G denotes graph distance in G. In other words, we wish to transport a mass of at least p to distance at least n away from o. Our results hold for p constant.

Our main result concerns the lattice $\mathbb{Z}^d$:

Theorem 14. Start with initial unit mass δ_o at the origin o of $\mathbb{Z}^d$, $d \geq 2$, and let $p \in (0, 1)$ be constant. The minimum number of toppling moves needed to transport mass p to distance at least n from the origin is

$$N_p\left(\mathbb{Z}^d, B_n, \delta_o\right) = \Theta\left(n^{d+2}\right),$$

where the implied constants depend only on d and p.

As mentioned previously, the one-dimensional case was studied and solved in[74], where the authors obtained the same result as in Theorem 14 for $d = 1$. We discuss the connection to the maximum overhang problem and related open problems in more detail at the end of the section (see Section 3.2.10).

3.2.1.1 Further results for other graphs

We start by giving a general upper bound on the number of toppling moves necessary to transport the mass from a vertex to outside a given set.

Theorem 15. Let $G = (V, E)$ be an infinite, connected, locally finite graph and let $\{X_t\}_{t \geq 0}$ be simple random walk on G with $X_0 = o$ for a vertex $o \in V$. Let $A \subset V$ be a set of vertices containing o and let T_A be the first exit time of the random walk from A. Start with initial unit mass δ_o at o. The minimum number of toppling moves needed to transport mass p to outside of the set A is

$$N_p \left(G, A, \delta_o\right) \leq \left(1 - p\right)^{-1} \mathrm{Vol}\left(A\right) \cdot \mathbb{E}_o \left[T_A\right], \tag{3.21}$$

where $\mathrm{Vol}\left(A\right) = \left|\{u \in A\}\right|$ denotes the volume of A, i.e., the number of vertices in A.

In Section 3.2.6 we give two proofs of this result: one using random walk on the graph to transport the mass and the other using a greedy algorithm. The two different arguments are useful because they can be extended in different ways, which, as we shall see, allows us to obtain sharper upper bounds in specific cases.

We now consider several specific graphs, starting with the comb graph $\mathbb{C}_2$, which is obtained from $\mathbb{Z}^2$ by removing all horizontal edges except those on the x axis.

Theorem 16. Start with initial unit mass δ_o at the origin o of the comb graph $\mathbb{C}_2$ and let $p \in (0,1)$ be constant. The minimum number of toppling moves needed to transport mass p to distance at least n from the origin is

$$N_p\left(\mathbb{C}_2, B_n, \delta_o\right) = \Theta\left(n^{7/2}\right),$$

where the implied constants depend only on p.

We also study various trees, starting with regular ones.

Theorem 17. Start with initial unit mass δ_ρ at the origin ρ of the d-ary tree $\mathbb{T}_d$, $d \geq 2$, and let $p \in (0,1)$ be constant. The minimum number of toppling moves needed to transport mass p to distance at least n from the origin is

$$N_p\left(\mathbb{T}_d, B_n, \delta_\rho\right) = \Theta\left(d^n\right),$$

where the implied constants depend only on d and p.

We prove a general result for graphs where random walk has positive speed ℓ and entropy h and which satisfy Shannon's theorem. This roughly states that $N_p\left(G, B_n, \delta_o\right) = \exp\left(n \cdot \frac{h}{\ell} \cdot (1 + o(1))\right)$; see Section 3.2.8 for a precise statement. This result can then be applied to specific examples, such as Galton-Watson trees and the product of two trees.

Theorem 18. Fix an offspring distribution with mean $m > 1$ and let GWT be a Galton-Watson tree obtained with this offspring distribution, on the event of nonextinction. Start with initial unit mass δ_ρ at the root ρ of GWT and let $p \in (0,1)$ be constant. The minimum number of toppling moves needed to transport mass p to distance at least n from the origin is almost surely

$$N_p\left(\mathrm{GWT}, B_n, \delta_\rho\right) = \exp\left(\mathbf{dim} \cdot n\ (1 + o(1))\right),$$

where **dim** is the dimension of harmonic measure and where the implied constants depend only on p and the offspring distribution.

When the offspring distribution is degenerate (i.e., every vertex has exactly m offspring and hence the tree is the m-ary tree $\mathbb{T}_m$), then Theorem 17 provides a sharper result than Theorem 18. However, when the offspring distribution is nondegenerate, then $\mathbf{dim} < \log m$ almost surely (see[64]) and hence the number of toppling moves necessary is exponentially smaller than the volume of B_n.

Theorem 19. Let $\mathbb{T}_d$ denote the $(d+1)$-regular tree. Start with initial unit mass δ_ρ at the origin ρ of the product of two regular trees, $\mathbb{T}_d \times \mathbb{T}_k$, and let $p \in (0,1)$ be constant. Assume that $d \geq k \geq 1$ and $d+k \geq 3$. The minimum number of toppling moves needed to transport mass p to distance at least n from the origin is

$$N_p \left(\mathbb{T}_d \times \mathbb{T}_k, B_n, \delta_\rho\right) = \theta \left(d, k\right)^{n(1+o(1))},$$

where $\theta\left(d, k\right) = d^{\frac{d-1}{d+k-2}} \cdot k^{\frac{k-1}{d+k-2}}$, and where the implied constants depend only on d, k, and p.

When $d > k \geq 2$, then the volume of a ball grows as $\mathrm{Vol}\left(B_n\right) = \Theta\left(d^n\right)$, whereas $\theta\left(d, k\right) < d$. Hence the number of toppling moves necessary to transport a constant mass to distance n from the root is exponentially smaller than the volume of the ball of radius n.

Finally, we consider graphs of bounded degree with exponential decay of the Green's function for simple random walk (see Definition 5).

Theorem 20. Let $G = (V, E)$ be an infinite, connected graph of bounded degree with exponential decay of the Green's function for simple random walk on G. Start with initial unit mass δ_o at a vertex $o \in V$ and let $p \in (0,1)$ be constant. The minimum number of toppling moves needed to transport mass p to distance at least n from o is

$$N_p\left(G, B_n, \delta_o\right) = \exp\left(\Theta\left(n\right)\right),$$

where the implied constants depend only on p, the maximum degree of G, and the exponent in the exponential bound on the Green's function.

See Section 3.2.9 where this result is restated more precisely as Theorem 28 and then proved, and where we illustrate this result with the example of the lamplighter graph.

3.2.2 Notation and preliminaries

Let $G = (V, E)$ be a graph and let $\mathcal{N}_v := \{y \in V : d_G(y, v) = 1\}$ denote the neighborhood of a vertex $v \in V$. All graphs we consider here are connected and locally finite (i.e., every vertex has finite degree). We also write $y \sim v$ for $y \in \mathcal{N}_v$. The discrete Laplacian Δ acting on functions $f : V \to \mathbb{R}$ is defined as

$$\Delta f\left(x\right) := \frac{1}{|\mathcal{N}_x|} \sum_{y \sim x} f\left(y\right) - f\left(x\right). \tag{3.22}$$

We can then write how a toppling move T_v^m acts on a mass distribution μ as

$$T_v^m \mu = \mu - m\delta_v + \frac{m}{|\mathcal{N}_v|} \sum_{y \sim v} \delta_y = \mu + m\Delta\delta_v. \tag{3.23}$$

We recall the well-known fact that if G is a regular graph and f and g are two functions from V to $\mathbb{R}$, with at least one of them having finite support, then

$$\sum_{x \in V} f\left(x\right) \Delta g\left(x\right) = \sum_{x \in V} \Delta f\left(x\right) g\left(x\right), \tag{3.24}$$

an equality which we refer to as summation by parts.

We also define the second moment of a mass distribution μ on $\mathbb{Z}^d$ as

$$M_2[\mu] = \sum_{v \in \mathbb{Z}^d} \mu(v) \cdot \|v\|_2^2 . \tag{3.25}$$

3.2.3 Upper bound on $\mathbb{Z}^d$ and preliminaries for the lower bound

We start with an upper bound on $N_p\left(\mathbb{Z}^d, B_n, \delta_o\right)$, stated as Theorem 21 below, which can be obtained by a greedy algorithm. We then introduce preliminaries for a lower bound argument which uses an appropriately defined potential. As we shall see, applying this argument directly leads to a lower bound of the correct order only in the case of $d = 1$. Additional ideas are required to obtain a tight lower bound for $d \geq 2$, which are then presented in Section 3.2.4.

3.2.3.1 A greedy upper bound on $\mathbb{Z}^d$

We use a greedy algorithm to provide an upper bound on the number of toppling moves needed to transport mass p to distance n from the origin in $\mathbb{Z}^d$.

Theorem 21. Start with initial unit mass δ_o at the origin o of $\mathbb{Z}^d$, $d \geq 1$. The minimum number of toppling moves needed to transport mass p to distance at least n from the origin satisfies

$$N_p\left(\mathbb{Z}^d, B_n, \delta_o\right) < \frac{2^d}{(1-p) \times d!} n^{d+2}. \tag{3.26}$$

Proof. Consider the following greedy algorithm for choosing toppling moves: until the mass outside of B_n is at least p, choose $v \in B_n$ with the largest mass in

B_n (break ties arbitrarily) and topple the full mass at v. Let $\mu_0 \equiv \delta_o, \mu_1, \mu_2, \ldots$ denote the resulting mass distributions, let v_i denote the vertex that was toppled to get from μ_{i-1} to μ_i, and let m_i denote the mass that was toppled at this step. By (3.23) we can then write

$$\mu_i = \mu_{i-1} + m_i \Delta \delta_{v_i}. \tag{3.27}$$

Furthermore, let t denote the number of moves necessary for this greedy algorithm to transport mass p to distance at least n from the origin, i.e., $t = \min\{i \geq 1 : \mu_i(B_n) \leq 1 - p\}$.

We first compute how the second moment of the mass distribution changes after each toppling move. By (3.27) we can write

$$M_2[\mu_i] - M_2[\mu_{i-1}] = \sum_{x \in \mathbb{Z}^d} \mu_i(x) \|x\|_2^2 - \sum_{x \in \mathbb{Z}^d} \mu_{i-1}(x) \|x\|_2^2 = m_i \sum_{x \in \mathbb{Z}^d} \Delta \delta_{v_i}(x) \cdot \|x\|_2^2.$$

Now using summation by parts (see (3.24)) and the fact that $\Delta \|x\|_2^2 = 1$ for every $x \in \mathbb{Z}^d$, we get that

$$\sum_{x \in \mathbb{Z}^d} \Delta \delta_{v_i}(x) \cdot \|x\|_2^2 = \sum_{x \in \mathbb{Z}^d} \delta_{v_i}(x) \cdot \Delta \|x\|_2^2 = \sum_{x \in \mathbb{Z}^d} \delta_{v_i}(x) = 1.$$

Putting the previous two displays together we thus obtain that

$$M_2[\mu_i] - M_2[\mu_{i-1}] = m_i.$$

The greedy choice implies that for every $i \leq t$ we must have that

$$m_i \geq \frac{\mu_{i-1}(B_n)}{|B_n|} > \frac{1-p}{|B_n|}.$$

This gives us the following lower bound on the second moment of μ_t:

$$M_2 [\mu_t] = \sum_{i=1}^{t} (M_2 [\mu_i] - M_2 [\mu_{i-1}]) = \sum_{i=1}^{t} m_i > t \times \frac{(1-p)}{|B_n|}. \tag{3.28}$$

On the other hand, all vertices with positive mass at time t have (graph) distance at most n from the origin, and hence $\|v\|_2^2 \leq n^2$ for every $v \in \mathbb{Z}^d$ such that $\mu_t(v) > 0$, which implies that $M_2 [\mu_t] \leq n^2$. Combining this with (3.28) we obtain that $t < |B_n| \times n^2 / (1-p)$. The claim in (3.26) then follows from the estimate $|B_n| \leq \left(2^d/d!\right) n^d$ on the size of the (open) ball of radius n. □

Remark 2. The greedy algorithm described in the proof above requires a tie-breaking rule, which breaks the symmetries of $\mathbb{Z}^d$. It is also natural to consider a greedy algorithm that keeps the symmetries of $\mathbb{Z}^d$. The same proof as above shows that this also transports mass p to distance at least n in at most $O\left(n^{d+2}\right)$ toppling moves.

3.2.3.2 Energy of measure and potential kernel

To obtain a lower bound it is natural to combine the second moment estimates with estimates for an appropriately defined potential function. We consider here a quantity called the *energy* of the measure. This subsection contains the necessary definitions, together with properties of the Green's function for random walk on $\mathbb{Z}^d$, which are required for subsequent estimates.

Definition 4. The *energy* of a measure μ on $\mathbb{Z}^d$ is defined as

$$\mathcal{E}_a [\mu] = \sum_{x,y \in \mathbb{Z}^d} a(x-y) \mu(x) \mu(y),$$

where a is the *potential kernel* function.

The energy of measure is a classical quantity; for more details regarding the physical context in which it arises, see, for example,[26]. We will use the energy of the measure μ with the potential kernel function defined using the Green's function for random walk on $\mathbb{Z}^d$, which we introduce next.

Definition 5 (Green's function)**.** For a random walk $\{X_k\}_{k\geq 0}$ on a graph $G = (V, E)$, the Green's function $g : V \times V \to [0, \infty]$ is defined as

$$g(x,y) := \mathbb{E}_x\left[\#\left\{k \geq 0 : X_k = y\right\}\right] = \sum_{k=0}^{\infty} \mathbb{P}_x\left(X_k = y\right) = \sum_{k=0}^{\infty} p^k\left(x,y\right),$$

where $\mathbb{P}_x$ and $\mathbb{E}_x$ denote probabilities and expectations given that $X_0 = x$, and $p^k(\cdot,\cdot)$ denotes the k-step transition probabilities. That is, $g(x,y)$ is the expected number of visits to y by the random walk started at x.

Since $\mathbb{Z}^d$ is translation invariant, we have that $g(x,y) = g(o, y-x)$ for simple random walk on $\mathbb{Z}^d$, where o denotes the origin of $\mathbb{Z}^d$. It is thus natural to define $g(x) := g(o,x)$ as the Green's function in $\mathbb{Z}^d$. Note that $g(x) = g(-x)$ by symmetry. For $d \geq 3$, simple random walk is transient in $\mathbb{Z}^d$ and hence $g(x)$ is finite for every $x \in \mathbb{Z}^d$. Since simple random walk is recurrent in $\mathbb{Z}$ and $\mathbb{Z}^2$, we have $g(x) = \infty$ for every $x \in \mathbb{Z}$ and $x \in \mathbb{Z}^2$. Thus we define instead

$$g_n(x) := \mathbb{E}_o\left[\#\left\{k \in \{0,1,\ldots,n\} : X_k = x\right\}\right], \tag{3.29}$$

the expected number of visits to x until time n by simple random walk started at o. With these notions we are ready to define the potential kernel function we will use in $\mathbb{Z}^d$.

Definition 6 (Potential kernel function for $\mathbb{Z}^d$)**.** For $d \geq 1$, define the potential kernel function $a : \mathbb{Z}^d \to \mathbb{R}$ as

$$a(x) := \lim_{n \to \infty} \{g_n(o) - g_n(x)\},$$

where g_n is defined as in (3.29).

This definition ensures that $a(x)$ is finite for $d = 1$ and $d = 2$ as well: for $d = 1$ we have that $a(x) = |x|$, and for $d = 2$ see, e.g.,[56] Theorem 1.6.1. For $d \geq 3$ we simply have that $a(x) = g(o) - g(x)$ and we can then write the energy of a probability measure μ with this potential kernel function as

$$\mathcal{E}_a[\mu] = g(o) - \mathcal{E}_g[\mu],$$

where

$$\mathcal{E}_g[\mu] := \sum_{x,y \in \mathbb{Z}^d} g(x,y)\,\mu(x)\,\mu(y).$$

By conditioning on the first step of the random walk one can check that the discrete Laplacians of the functions g and a satisfy $\Delta g(x) = \Delta a(x) = 0$ for $x \neq o$, while at the origin o we have that $\Delta g(o) = -1$ and $\Delta a(o) = 1$.

We will use the following estimates for the asymptotics of the Green's function on $\mathbb{Z}^d$ far from the origin; more precise estimates are known[56,40], but are not required for our purposes. First, when $d = 2$ then there exists an absolute constant C_2 such that

$$\left|a(x) - \frac{2}{\pi}\ln\|x\|_2 - \kappa\right| \leq \frac{C_2}{\|x\|_2^2} \tag{3.30}$$

for all $x \neq o$, where κ is an explicit constant whose value is not relevant for our purposes[40]. Second, for every $d \geq 3$ there exists an absolute constant C_d such that

$$\left|g(x) - a_d\|x\|_2^{2-d}\right| \leq \frac{C_d}{\|x\|_2^{d-1}} \tag{3.31}$$

for all $x \neq o$, where $a_d = (d/2)\Gamma(d/2-1)\pi^{-d/2} = \frac{2}{(d-2)\omega_d}$, where ω_d is the volume of the L_2 unit ball in $\mathbb{R}^d$ (see[56] Theorem 1.5.4).

3.2.3.3 Comparing the energy with the second moment

To obtain a lower bound on $N_p\left(\mathbb{Z}^d, B_n, \delta_o\right)$ we need to compare the second moment of a mass distribution with its energy, as defined in the previous subsection. This comparison is done in the following lemma.

Lemma 20. Let $\mu_0, \mu_1, \ldots, \mu_t$ be a sequence of mass distributions on $\mathbb{Z}^d$ resulting from toppling moves and let a be the potential kernel function defined in Definition 6. Then we have that

$$t\left(\mathcal{E}_a[\mu_t] - \mathcal{E}_a[\mu_0]\right) \geq \left(M_2[\mu_t] - M_2[\mu_0]\right)^2. \tag{3.32}$$

Proof. For $i \in [t]$ let v_i denote the vertex that was toppled to get from μ_{i-1} to μ_i, and let m_i denote the mass that was toppled at this step. From Section 3.2.3.1 we know that $M_2[\mu_i] - M_2[\mu_{i-1}] = m_i$ for each $i \in [t]$. Turning to the energy of the measure, we first recall from (3.27) that $\mu_i = \mu_{i-1} + m_i\Delta\delta_{v_i}$ for every $i \in [t]$.

We can use this to write how the energy changes after each toppling move as follows:

$$\mathcal{E}_a[\mu_i] - \mathcal{E}_a[\mu_{i-1}] = \sum_{x,y\in\mathbb{Z}^d} a(x-y)[\mu_i(x)\mu_i(y) - \mu_{i-1}(x)\mu_{i-1}(y)]$$
$$= \sum_{x,y\in\mathbb{Z}^d} a(x-y)[\{\mu_{i-1}(x) + m_i\Delta\delta_{v_i}(x)\}\{\mu_{i-1}(y) + m_i\Delta\delta_{v_i}(y)\} - \mu_{i-1}(x)\mu_{i-1}(y)]$$
$$= m_i\sum_{x,y\in\mathbb{Z}^d} a(x-y)\Delta\delta_{v_i}(x)\mu_{i-1}(y) + m_i\sum_{x,y\in\mathbb{Z}^d} a(x-y)\Delta\delta_{v_i}(y)\mu_{i-1}(x)$$
$$+m_i^2\sum_{x,y\in\mathbb{Z}^d} a(x-y)\Delta\delta_{v_i}(x)\Delta\delta_{v_i}(y).$$

We compute each term in this sum separately. Recall that $\Delta a(x) = \delta_0(x)$ and hence for every $y \in \mathbb{Z}^d$, $(\Delta a(\cdot - y))(x) = \delta_y(x)$. Using summation by parts we have for every fixed $y \in \mathbb{Z}^d$ that

$$\sum_{x\in\mathbb{Z}^d} a(x-y)\cdot\Delta\delta_{v_i}(x) = \sum_{x\in\mathbb{Z}^d}(\Delta a(\cdot - y))(x)\cdot\delta_{v_i}(x) = \delta_{v_i}(y).$$

For the first term above we thus have:

$$\sum_{x,y\in\mathbb{Z}^d} a(x-y)\Delta\delta_{v_i}(x)\mu_{i-1}(y) = \sum_{y\in\mathbb{Z}^d}\delta_{v_i}(y)\mu_{i-1}(y) = \mu_{i-1}(v_i).$$

Since $a(x-y) = a(y-x)$ we have that the second term in the sum above is equal to the first one. Finally we can compute the third term similarly:

$$\sum_{x,y\in\mathbb{Z}^d} a(x-y)\Delta\delta_{v_i}(x)\Delta\delta_{v_i}(y) = \sum_{y\in\mathbb{Z}^d}\delta_{v_i}(y)\Delta\delta_{v_i}(y) = \Delta\delta_{v_i}(v_i) = -1.$$

Putting together the previous two displays with the sum above, we can conclude that

$$\mathcal{E}_a[\mu_i] - \mathcal{E}_a[\mu_{i-1}] = 2m_i\mu_{i-1}(v_i) - m_i^2 \geq m_i^2,$$

where the last step follows because $m_i \leq \mu_{i-1}(v_i)$.

The claimed inequality (3.32) now follows by the Cauchy-Schwarz inequality:

$$(M_2[\mu_t] - M_2[\mu_0])^2 = \left(\sum_{i=1}^{t} m_i\right)^2 \leq t \sum_{i=1}^{t} m_i^2 \leq t\left(\mathcal{E}_a[\mu_t] - \mathcal{E}_a[\mu_0]\right). \qquad \square$$

3.2.3.4 An initial lower bound argument

A lower bound of the correct order in dimension $d = 1$ now follows (see also[74] where this argument first appeared). Suppose that a sequence of t toppling moves are applied to obtain mass distributions $\mu_0 \equiv \delta_o, \mu_1, \mu_2, \ldots, \mu_t$ that satisfy $\mu_i = T_{v_i}^{m_i}\mu_{i-1}$ for every $i \in [t]$ and $\mu_t(B_n) \leq 1 - p$. We may assume that $|v_i| \leq n - 1$ for every $i \leq t$; any other toppling move can be removed from the sequence to obtain a shorter sequence that still moves mass p to distance n from the origin.

Now recall that $a(x) = |x|$ when $d = 1$. Since $\mu_t(\{v : |v| > n\}) = 0$, we have that $\mathcal{E}_a[\mu_t] \leq 2n$. On the other hand, since $\mu_t(\{v : |v| \geq n\}) \geq p$, we have that $M_2[\mu_t] \geq pn^2$. By Lemma 20 we thus have that $t \times 2n \geq (pn^2)^2$, implying that

$$N_p(\mathbb{Z}, B_n, \delta_o) \geq \frac{p^2}{2} n^3,$$

which matches the upper bound of Theorem 21 up to constant factors in p.

However, the same argument for $d \geq 2$ (using the estimates for the Green's function from (3.30) and (3.31); we leave the details to the reader) only provides the following estimates: there exists a constant C depending only on d and p such that

$$N_p\left(\mathbb{Z}^2, B_n, \delta_o\right) \geq \frac{Cn^4}{\log(n)},$$

and

$$N_p\left(\mathbb{Z}^d, B_n, \delta_o\right) \geq Cn^4,$$

for $d \geq 3$. Therefore, to obtain a tight lower bound in dimensions $d \geq 2$, a new idea is needed. The idea, presented in the following section, is to perform an initial smoothing of the mass distribution.

3.2.4 Smoothing and the lower bound on $\mathbb{Z}^d$

The previous section provides the basis for the proof of Theorem 14, but applying the arguments directly leads to a suboptimal lower bound, as described in Section 3.2.3.4. The remedy is to perform an initial smoothing of the mass distribution. In this section we first describe the smoothing operation in general in Section 3.2.5, followed by describing the specifics of smoothing in $\mathbb{Z}^d$ in Section 3.2.5.1. We conclude with the proof of the lower bound on $\mathbb{Z}^d$ in Section 3.2.5.2.

3.2.5 Smoothing of distributions

For the proofs of the lower bounds on most families of graphs investigated in this section we use a certain smoothing of the mass distribution. That is, we first perform some toppling moves to obtain a mass distribution $\widetilde{\mu}$ that is "smooth" in the sense that it is approximately uniform over a subset of the ball B_n. In this subsection we show that it is valid to use smoothing for lower bound arguments, since the minimum number of toppling moves necessary to transport mass p outside of a set A cannot increase by smoothing. What then remains to be estimated (for each family of graphs separately) is the minimum number of toppling moves necessary to transport mass p to distance n started from the smooth distribution $\widetilde{\mu}$.

Lemma 21 (Smoothing weakly reduces the minimum number of toppling moves)**.** Start with mass distribution μ on a graph G, and let $A \subseteq V(G)$. Suppose that toppling mass m at vertex $v \in A$ is a valid toppling move. We then have that

$$N_p(G, A, T_v^m \mu) \leq N_p(G, A, \mu). \tag{3.33}$$

Proof. We prove the statement by induction on $t := N_p(G, A, \mu)$. For the base case of $t = 0$, if $N_p(G, A, \mu) = 0$, then $\mu(A) \leq 1 - p$. Since $v \in A$, no mass can enter A from outside of A in the toppling move, so $T_v^m \mu(A) \leq 1 - p$ and hence $N_p(G, A, T_v^m \mu) = 0$.

For the induction step, let $t = N_p(G, A, \mu)$ and let $\mu \equiv \mu_0, \mu_1, \ldots, \mu_t$ be a series of mass distributions such that μ_i is obtained from μ_{i-1} by a toppling move at vertex v_i with mass m_i being toppled, i.e., $\mu_i = T_{v_i}^{m_i} \mu_{i-1}$, and such that $\mu_t(A) \leq 1 - p$. Due to the optimality of the sequence of toppling moves we have that $N_p(G, A, \mu_1) = t - 1$.

Consider first the case that $v \neq v_1$. In this case the toppling moves T_v^m and $T_{v_1}^{m_1}$ commute, i.e., $T_{v_1}^{m_1} T_v^m \mu = T_v^m T_{v_1}^{m_1} \mu$. Hence

$$N_p(G,A,T_v^m\mu) \le N_p\left(G,A,T_{v_1}^{m_1}T_v^m\mu\right)+1 = N_p(G,A,T_v^m\mu_1)+1 \le N_p(G,A,\mu_1)+1 = t,$$

where the second inequality is due to the induction hypothesis.

Now consider the case that $v = v_1$. If $m_1 > m$, then

$$N_p(G,A,T_v^m\mu) \le N_p\left(G,A,T_{v_1}^{m_1-m}T_v^m\mu\right) + 1 = N_p(G,A,\mu_1) + 1 = t.$$

If $m \ge m_1$, then

$$N_p(G,A,T_v^m\mu) = N_p\left(G,A,T_v^{m-m_1}\mu_1\right) \le N_p(G,A,\mu_1) = t-1,$$

where the inequality is again due to the induction hypothesis. $\square$

Iterating this lemma we immediately obtain the following corollary.

Corollary 21.1. Let $\mu_0, \mu_1, \ldots, \mu_t$ be a sequence of mass distributions on a graph G such that for every $i \in [t]$, the mass distribution μ_i is obtained from μ_{i-1} by applying a toppling move at vertex $v_i \in V(G)$, toppling a mass m_i, i.e., $\mu_i = T_{v_i}^{m_i}\mu_{i-1}$. Let $A \subseteq V(G)$ and assume that $v_i \in A$ for every $i \in [t]$. Then we have that

$$N_p(G,A,\mu_t) \le N_p(G,A,\mu_0).$$

Another corollary of the lemma above is that we can assume without loss of generality that at every move we topple *all* the mass at a given vertex. Given a graph $G = (V,E)$, a subset of the vertices $A \subset V$, mass $p > 0$, and an initial mass distribution μ_0, we define $N_p^{\text{full}}(G,A,\mu_0)$ to be the minimum number of toppling moves needed to move mass p outside of the set A, where at every toppling move we have to topple all the mass at a given vertex.

Corollary 21.2. We have that $N_p(G, A, \mu_0) = N_p^{\text{full}}(G, A, \mu_0)$.

Proof. Since allowing only full topplings is more restrictive than allowing partial topplings, we have that $N_p(G, A, \mu_0) \leq N_p^{\text{full}}(G, A, \mu_0)$. We prove the other inequality, i.e., that $N_p(G, A, \mu_0) \geq N_p^{\text{full}}(G, A, \mu_0)$, by induction on $t := N_p(G, A, \mu_0)$. For the base of $t = 0$, if $N_p(G, A, \mu_0) = 0$, then $\mu_0(A) \leq 1 - p$, and hence $N_p^{\text{full}}(G, A, \mu_0) = 0$.

For the induction step, let $t = N_p(G, A, \mu_0)$ and let $\mu_0, \mu_1, \ldots, \mu_t$ be a series of mass distributions such that μ_i is obtained from μ_{i-1} by a toppling move at vertex v_i with mass m_i being toppled, i.e., $\mu_i = T_{v_i}^{m_i}\mu_{i-1}$, and such that $\mu_t(A) \leq 1 - p$. Due to the optimality of the sequence of toppling moves we have that $N_p(G, A, \mu_1) = t - 1$. Define the mass distribution $\mu_1' = T_{v_1}^{\mu_0(v_1)}\mu_0$, which corresponds to toppling all the original mass at v_1, and note that $\mu_1' = T_{v_1}^{\mu_0(v_1)-m_1}\mu_1$. By Lemma 21 we have that $N_p(G, A, \mu_1') \leq N_p(G, A, \mu_1)$ and by the induction hypothesis we have that $N_p(G, A, \mu_1') = N_p^{\text{full}}(G, A, \mu_1')$. Therefore we obtain that

$$N_p^{\text{full}}(G, A, \mu_0) \leq N_p^{\text{full}}(G, A, \mu_1') + 1 = N_p(G, A, \mu_1') + 1 \leq t. \qquad \square$$

3.2.5.1 Smoothing in $\mathbb{Z}^d$

For the initial smoothing in $\mathbb{Z}^d$ we leverage connections between our controlled diffusion setting and the *divisible sandpile model*, and use results by Levine and

Peres[60] on this model. In the divisible sandpile each site $x \in \mathbb{Z}^d$ starts with mass $\nu_0(x) \in \mathbb{R}_{\geq 0}$. A site x is *full* if its mass is at least 1. A *divisible sandpile move* at x, denoted by $\mathcal{D}_x$, consists of no action if x is not full, and consists of keeping mass 1 at x and splitting any excess mass equally among its neighbors if x is full.

Recall from (3.23) that the mass distribution after a toppling move can be written as $T_v^m \mu = \mu + m\Delta\delta_v$. Similarly, for a mass distribution μ and a site $x \in \mathbb{Z}^d$, the mass distribution after a divisible sandpile move at x can be written as

$$\mathcal{D}_x \mu = \mu + \max\{\mu(x) - 1, 0\} \Delta\delta_x. \tag{3.34}$$

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

Proposition 2 (Levine and Peres[60])**.** *Let $x_1, x_2, \cdots \in \mathbb{Z}^d$ be a sequence with the property that for any $x \in \mathbb{Z}^d$ there are infinitely many terms $x_k = x$. Let ν_0 denote the initial mass distribution and assume that ν_0 has finite support. Let*

$$\begin{aligned} u_k(x) = &\ \textit{total mass emitted by } x \textit{ after divisible sandpile moves } x_1, \ldots, x_k; \\ \nu_k(x) = &\ \textit{amount of mass present at } x \textit{ after divisible sandpile moves } x_1, \ldots, x_k. \end{aligned}$$

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 limit ν represents the final mass distribution and sites $x \in \mathbb{Z}^d$ with $\nu(x) = 1$ are called *fully occupied*. We are interested primarily in the case when the initial mass distribution is a point mass at the origin: $\nu_o = m\delta_o$ for some $m > 0$. The natural question then is to identify the shape of the resulting domain D_m of fully occupied sites. The following result states that D_m is very close to a Euclidean ball. Since here the notation B_n is reserved for the L_1 ball (and the graph distance ball more generally), we denote by $B_r^{(2)} = \left\{x \in \mathbb{Z}^d : \|x\|_2 < r\right\}$ the (open) L_2 ball around the origin.

Theorem 22 (Levine and Peres[60])**.** 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 and c' depending only on d such that

$$B_{r-c}^{(2)} \subset D_m \subset B_{r+c'}^{(2)},$$

where $r = (m/\omega_d)^{1/d}$ and ω_d is the volume of the L_2 unit ball in $\mathbb{R}^d$.

We note that the sequence of divisible sandpile moves started from a pile of mass m at the origin could potentially be infinite. However, there exists a finite K such that $\nu_K(x) \leq 1 + \epsilon$ for every $x \in \mathbb{Z}^d$ and for some small $\epsilon > 0$. This is useful for proving the following corollary of the theorem above.

Corollary 22.1. For every $c \in (0,1)$ there exists a finite sequence of toppling moves that takes the mass distribution δ_o on $\mathbb{Z}^d$ to a mass distribution μ on $\mathbb{Z}^d$ for which the following two properties hold:

$$\forall x \in B_{cn}^{(2)} : \quad \mu(x) \leq \frac{2}{\mathrm{Vol}\left(B_{cn}^{(2)}\right)}, \tag{3.35}$$

$$\forall x \notin B_{cn}^{(2)} : \quad \mu(x) = 0, \tag{3.36}$$

where $\mathrm{Vol}\left(B_{cn}^{(2)}\right) = \left|\left\{x \in \mathbb{Z}^d : \|x\|_2 < cn\right\}\right|$ denotes the volume of the ball $B_{cn}^{(2)}$.

Proof. The result follows from Theorem 22 by scaling the masses by m, for both the mass distributions and the divisible sandpile moves. □

3.2.5.2 A lower bound on $\mathbb{Z}^d$

We are now ready to prove Theorem 14. By performing an initial smoothing as detailed in Section 3.2.5.1, we are able to obtain a lower bound that matches the upper bound of Theorem 21 up to constant factors.

Theorem 23. Start with initial unit mass δ_o at the origin o of $\mathbb{Z}^d$, $d \geq 2$. There exists a constant C depending only on d and p such that the minimum number of toppling moves needed to transport mass p to distance at least n from the origin satisfies

$$N_p\left(\mathbb{Z}^d, B_n, \delta_o\right) \geq C n^{d+2}. \tag{3.37}$$

Proof. The first step is to smooth the distribution δ_o. Let $c := \sqrt{p/(2d)}$. By Corollary 22.1 there exists a finite sequence of toppling moves taking δ_o to a mass distribution μ satisfying (3.35) and (3.36). By Corollary 21.1 we have that $N_p\left(\mathbb{Z}^d, B_n, \delta_o\right) \geq N_p\left(\mathbb{Z}^d, B_n, \mu\right)$, so it suffices to bound $N_p\left(\mathbb{Z}^d, B_n, \mu\right)$ from below.

Suppose that starting from μ a sequence of t toppling moves are applied to obtain mass distributions $\mu_0 \equiv \mu, \mu_1, \mu_2, \ldots, \mu_t$ that satisfy $\mu_t\left(B_n\right) \leq 1 - p$. Let v_i denote the vertex that was toppled to get from μ_{i-1} to μ_i, and let m_i denote the mass that was toppled at this step. We may assume that $\|v_i\|_1 \leq n-1$ for every $i \leq t$, since any other toppling move can be removed from the sequence to obtain a shorter sequence that still moves mass p to distance n from the origin.

By Lemma 20 we have that

$$t \geq \frac{\left(M_2\left[\mu_t\right] - M_2\left[\mu_0\right]\right)^2}{\mathcal{E}_a\left[\mu_t\right] - \mathcal{E}_a\left[\mu_0\right]} \tag{3.38}$$

and in the following we bound the numerator and the denominator separately, starting with the numerator.

Since $\mu_t\left(\left\{x \in \mathbb{Z}^d : \|x\|_1 \geq n\right\}\right) \geq p$ and $\|x\|_2 \geq \|x\|_1 / \sqrt{d}$, we have that $M_2\left[\mu_t\right] \geq \frac{pn^2}{d}$. On the other hand, the support of μ_0 is contained within $B_{cn}^{(2)}$ and so $M_2\left[\mu_0\right] \leq c^2 n^2 = \frac{pn^2}{2d}$. Putting these two estimates together we obtain that

$$\left(M_2\left[\mu_t\right] - M_2\left[\mu_0\right]\right)^2 \geq \frac{p^2}{4d^2} n^4. \tag{3.39}$$

From (3.38) and (3.39) we have that in order to show (3.37), what remains is to show that

$$\mathcal{E}_a\left[\mu_t\right] - \mathcal{E}_a\left[\mu_0\right] \leq C' n^{2-d} \tag{3.40}$$

for some constant C' depending only on d and p. At this point the proof slightly differs for $d = 2$ and $d \geq 3$. We start with the case of $d \geq 3$.

Recall from Section 3.2.3.2 that when $d \geq 3$ then

$$\mathcal{E}_a\left[\mu_t\right] - \mathcal{E}_a\left[\mu_0\right] = \mathcal{E}_g\left[\mu_0\right] - \mathcal{E}_g\left[\mu_t\right] \leq \mathcal{E}_g\left[\mu_0\right].$$

We estimate this latter quantity by dividing $\mathbb{Z}^d \times \mathbb{Z}^d$ into shells

$$E_k := \left\{ (x,y) \in \mathbb{Z}^d \times \mathbb{Z}^d : \frac{2n}{2^k} < \|x-y\|_2 \leq \frac{2n}{2^{k-1}} \right\}$$

and estimating the sum on each shell separately. Since the support of μ_0 is contained in $B^{(2)}_{cn}$, we can write

$$\mathcal{E}_g[\mu_0] = \sum_{x \in B^{(2)}_{cn}} g(o)\,\mu_0(x)^2 + \sum_{k=1}^{K} \sum_{(x,y)\in E_k} g(x-y)\,\mu_0(x)\,\mu_0(y), \tag{3.41}$$

where $K = \lceil \log_2(2n) \rceil$. Using (3.35) we have that the first term in (3.41) can be bounded as follows:

$$\sum_{x \in B^{(2)}_{cn}} g(o)\,\mu_0(x)^2 \leq \sum_{x \in B^{(2)}_{cn}} \frac{4g(o)}{\mathrm{Vol}\left(B^{(2)}_{cn}\right)^2} = \frac{4g(o)}{\mathrm{Vol}\left(B^{(2)}_{cn}\right)} = O\left(n^{-d}\right), \tag{3.42}$$

where in the last estimate we used that $\mathrm{Vol}\left(B^{(2)}_{cn}\right) = \Theta\left(n^d\right)$. Now if $x \neq y$ then we have from (3.31) that

$$g(x-y) \leq a_d \|x-y\|_2^{2-d} + C_d \|x-y\|_2^{1-d}$$

and so if $(x,y) \in E_k$ then

$$g(x-y) \leq a_d \left(\frac{2n}{2^{k-1}}\right)^{2-d} + C_d \left(\frac{2n}{2^{k-1}}\right)^{1-d} = O\left(n^{2-d} \times 2^{dk} \times \left(2^{-2k} + n^{-1} \times 2^{-k}\right)\right). \tag{3.43}$$

Now to bound the mass of a shell first note that for any $x \in \mathbb{Z}^d$ we have that

$$\sum_{y:(x,y)\in E_k} \mu_0(y) \leq \text{Vol}\left(B^{(2)}_{\frac{2n}{2^{k-1}}}\right) \times \frac{2}{\text{Vol}\left(B^{(2)}_{cn}\right)} = O\left(2^{-dk}\right),$$

where we used again that $\text{Vol}\left(B^{(2)}_r\right) = \Theta\left(r^d\right)$. This then implies that

$$\sum_{(x,y)\in E_k} \mu_0(x)\,\mu_0(y) = O\left(2^{-dk}\right). \tag{3.44}$$

Putting together (3.43) and (3.44) we obtain that

$$\sum_{(x,y)\in E_k} g(x-y)\,\mu_0(x)\,\mu_0(y) = O\left(n^{2-d} \times \left(2^{-2k} + n^{-1} \times 2^{-k}\right)\right).$$

Summing this over k we get that

$$\sum_{k=1}^{K} \sum_{(x,y)\in E_k} g(x-y)\,\mu_0(x)\,\mu_0(y) = O\left(n^{2-d}\right),$$

which, together with (3.42), shows that $\mathcal{E}_g[\mu_0] = O\left(n^{2-d}\right)$. This concludes the proof of (3.40) for $d \geq 3$.

The case of $d = 2$ is similar, but the Green's function behaves differently, and we cannot neglect the energy of the mass distribution μ_t as we did for $d \geq 3$. We first bound $\mathcal{E}_a[\mu_t]$ from above. Recall that $a(o) = 0$ and that for every $x \neq o$ we have the estimate $a(x) \leq \frac{2}{\pi}\ln\|x\|_2 + \kappa + C_2\|x\|_2^{-2}$ (see (3.30)). We know that every x in the support of μ_t satisfies $\|x\|_1 \leq n$ and hence also $\|x\|_2 \leq n$. Thus by the triangle inequality if both x and y are in the support of μ_t then $\|x-y\|_2 \leq 2n$. Therefore

$$\mathcal{E}_a[\mu_t] = \sum_{x,y\in\mathbb{Z}^d : x\neq y} a(x-y)\,\mu_t(x)\,\mu_t(y)$$

$$\leq \left(\frac{2}{\pi}\ln\left(2n\right) + \kappa + C_2\right) \sum_{x,y\in\mathbb{Z}^d : x\neq y} \mu_t\left(x\right)\mu_t\left(y\right)$$
$$= \frac{2}{\pi}\ln\left(n\right) + O\left(1\right). \tag{3.45}$$

Next we bound from below the energy $\mathcal{E}_a\left[\mu_0\right]$. Noting again that $a\left(o\right) = 0$, we can write $\mathcal{E}_a\left[\mu_0\right]$ similarly to (3.41):

$$\mathcal{E}_a\left[\mu_0\right] = \sum_{k=1}^{K} \sum_{(x,y)\in E_k} a\left(x-y\right)\mu_0\left(x\right)\mu_0\left(y\right). \tag{3.46}$$

For $x \neq o$ we have the estimate $a\left(x\right) \geq \frac{2}{\pi}\ln\left\|x\right\|_2 - C_2\left\|x\right\|_2^{-2}$ (see (3.30) and note that $\kappa > 0$), and thus if $(x,y) \in E_k$ then

$$a\left(x-y\right) \geq \frac{2}{\pi}\ln\left(2n\right) - \frac{2}{\pi}\ln\left(2\right)\times k - C_2 \times \frac{2^{2k}}{4n^2}. \tag{3.47}$$

Plugging the estimate (3.47) into (3.46) we get three terms which we can each estimate separately. First, observing that

$$\sum_{k=1}^{K} \sum_{(x,y)\in E_k} \mu_0\left(x\right)\mu_0\left(y\right) = 1 - \sum_{x\in\mathbb{Z}^d} \mu_0\left(x\right)^2 \geq 1 - \frac{4}{\mathrm{Vol}\left(B_{cn}^{(2)}\right)} = 1 - O\left(n^{-2}\right),$$

we get that

$$\sum_{k=1}^{K} \sum_{(x,y)\in E_k} \frac{2}{\pi}\ln\left(2n\right)\mu_0\left(x\right)\mu_0\left(y\right) = \frac{2}{\pi}\ln\left(n\right) - O\left(1\right). \tag{3.48}$$

For the second term in (3.47) we use (3.44) to obtain that

$$\sum_{k=1}^{K} \sum_{(x,y)\in E_k} \left(\frac{2}{\pi}\ln(2)\times k\right)\mu_0(x)\,\mu_0(y) = \sum_{k=1}^{K} O\left(k\times 2^{-2k}\right) = O(1). \tag{3.49}$$

For the third term in (3.47) we again use (3.44), together with the fact that $K = \lceil \log_2(2n) \rceil$, to get that

$$\sum_{k=1}^{K} \sum_{(x,y)\in E_k} \left(C_2 \times \frac{2^{2k}}{4n^2}\right)\mu_0(x)\,\mu_0(y) = \frac{1}{n^2}\sum_{k=1}^{K} O(1) = O\left(\frac{\log(n)}{n^2}\right). \tag{3.50}$$

Putting together (3.48), (3.49), and (3.50) with (3.46) and (3.47) we obtain that

$$\mathcal{E}_a[\mu_0] = \frac{2}{\pi}\ln(n) - O(1). \tag{3.51}$$

Finally, putting together (3.45) and (3.51) we obtain (3.40) for $d = 2$. $\square$

3.2.6 A general upper bound

In this section we provide two proofs of Theorem 15.

Proof of Theorem 15 using random walk. We write $|A| := \mathrm{Vol}(A)$ to abbreviate notation. Let $x_1, x_2, \ldots, x_{|A|}$ denote the vertices of A in some specific order. We define a sequence of toppling moves that proceeds in rounds by repeatedly cycling through the vertices of A in this specified order and at each move toppling all of the mass that was at the given vertex at the beginning of the round. That is, letting $\mu_0 := \delta_o$, we let $\mu_1 := T_{x_1}^{\mu_0(x_1)}\mu_0$, then $\mu_2 := T_{x_2}^{\mu_0(x_2)}\mu_1$, and so on. In general, for a positive integer i, let i^* be the unique integer in $\{1, 2, \ldots, |A|\}$ such that $i - i^*$ is divisible by $|A|$. We then have that

$$\mu_i := T_{x_{i^*}}^{m_i} \mu_{i-1}, \qquad \text{with} \qquad m_i = \mu_{i-i^*}\left(x_{i^*}\right). \tag{3.52}$$

We call each group of $|A|$ toppling moves a *round* of the toppling process.

Let $\{Z_t\}_{t\geq 0}$ denote the random walk on G that is killed when it exits A, i.e., $Z_t = X_{t\wedge T_A}$, with initial condition $Z_0 = o$. Observe that all the toppling moves of a given round can be executed in parallel, since the mass that is toppled at each vertex only depends on the mass distribution at the beginning of the round. Since all of the mass that is present in A at the beginning of the round is toppled, each round of the toppling process defined in (3.52) perfectly simulates a step of the killed random walk $\{Z_t\}_{t\geq 0}$. That is, for every nonnegative integer t, the measure $\mu_{t|A|}$ agrees with the distribution of Z_t.

Let

$$M := \inf\left\{i \geq 0 : \mu_{i|A|}\left(A\right) \leq 1 - p\right\}$$

denote the first time that the distribution of the killed random walk has mass at least p outside of the set A. By the definition of the exit time T_A we have that

$$\mathbb{E}_o\left[T_A\right] = \sum_{k=1}^{\infty} \mathbb{P}_o\left(T_A \geq k\right) = \sum_{k=1}^{\infty} \mathbb{P}_o\left(Z_{k-1} \in A\right) = \sum_{k=1}^{\infty} \mu_{(k-1)|A|}\left(A\right). \tag{3.53}$$

Now by the definition of M we have that for every $m < M$, the measure $\mu_{m|A|}$ satisfies $\mu_{m|A|}\left(A\right) > 1 - p$. Therefore keeping only the first M terms in the sum in (3.53) we obtain the bound

$$\mathbb{E}_o\left[T_A\right] \geq \sum_{k=1}^{M} \mu_{(k-1)|A|}\left(A\right) > M\left(1-p\right).$$

By the definition of M this immediately implies that

$$N_p\left(G, A, \delta_o\right) \leq M \times |A| < \left(1-p\right)^{-1} \mathbb{E}_o\left[T_A\right] \times |A|. \qquad \square$$

Theorem 15 can also be proven using a greedy algorithm, similarly to the proof of the greedy upper bound on $\mathbb{Z}^d$ presented in Section 3.2.3.1. The only part of that proof that was specific to $\mathbb{Z}^d$ was the use of the second moment of the mass distribution. In particular, the key property of the second moment that we used was that $\Delta \left\|x\right\|_2^2 = 1$ for every $x \in \mathbb{Z}^d$. For a general graph $G = (V, E)$ and a subset of the vertices $A \subset V$, the expected first exit time from A starting from a given vertex is a function whose discrete Laplacian is constant on A. This is because by conditioning on the first step of the random walk we have that

$$\mathbb{E}_x\left[T_A\right] = 1 + \frac{1}{|\mathcal{N}_x|} \sum_{y \sim x} \mathbb{E}_y\left[T_A\right] \tag{3.54}$$

for every $x \in A$.

Proof of Theorem 15 using a greedy algorithm. Consider the following greedy algorithm for choosing toppling moves: until the mass outside of A is at least p, choose $v \in A$ with the largest mass in A (break ties arbitrarily) and topple the full mass at v. Let $\mu_0 \equiv \delta_o, \mu_1, \mu_2, \ldots$ denote the resulting mass distributions, let v_i denote the vertex that was toppled to get from μ_{i-1} to μ_i, and let

m_i denote the mass that was toppled at this step. By (3.23) we can then write $\mu_i = \mu_{i-1} + m_i \Delta \delta_{v_i}$. Furthermore, let t denote the number of moves necessary for this greedy algorithm to transport mass p to distance at least n from the origin, i.e., $t = \min\{i \geq 1 : \mu_i(A) \leq 1 - p\}$.

For $x \in V$, let $h(x) := -\mathbb{E}_x[T_A]$. We have $h(x) = 0$ for every $x \notin A$, and, by (3.54), we have that $\Delta h(x) = 1$ for every $x \in A$. For a mass distribution μ define $\widetilde{M}[\mu] := \sum_{x \in V} \mu(x) h(x)$. We have that $\widetilde{M}[\mu_0] = -\mathbb{E}_o[T_A]$ and $\widetilde{M}[\mu_t] \leq 0$. We first compute how $\widetilde{M}$ changes after each toppling move:

$$\widetilde{M}[\mu_i] - \widetilde{M}[\mu_{i-1}] = \sum_{x \in V} \mu_i(x) h(x) - \sum_{x \in V} \mu_{i-1}(x) h(x) = m_i \sum_{x \in V} \Delta \delta_{v_i}(x) \cdot h(x).$$

Now by definition we have that

$$\begin{aligned}
\sum_{x \in V} \Delta \delta_{v_i}(x) \cdot h(x) &= \sum_{x \in V} \left(-\delta_{v_i}(x) + \frac{1}{|\mathcal{N}_{v_i}|} \sum_{y \sim v_i} \delta_y(x) \right) h(x) \\
&= -h(v_i) + \frac{1}{|\mathcal{N}_{v_i}|} \sum_{y \sim v_i} h(y) = \Delta h(v_i) = 1,
\end{aligned}$$

where the last equality follows from the fact that $\Delta h(x) = 1$ for every $x \in A$. Putting the previous two displays together we obtain that

$$\widetilde{M}[\mu_i] - \widetilde{M}[\mu_{i-1}] = m_i$$

for every $i \leq t$. Note that the greedy choice implies that $m_i > (1-p)/|A|$ for every $i \leq t$. Therefore we obtain that

$$\mathbb{E}_o\left[T_A\right] \geq \widetilde{M}\left[\mu_t\right] - \widetilde{M}\left[\mu_0\right] = \sum_{i=1}^{t}\left(\widetilde{M}\left[\mu_i\right] - \widetilde{M}\left[\mu_{i-1}\right]\right) = \sum_{i=1}^{t} m_i > t \times \frac{1-p}{|A|},$$

and the result follows by rearranging this inequality. □

3.2.7 Controlled diffusion on the comb

The general upper bound given by Theorem 15 applied directly to the comb $\mathbb{C}_2$ gives a bound of

$$N_p\left(\mathbb{C}_2, B_n, \delta_o\right) \leq C\left(1-p\right)^{-1} n^4 \tag{3.55}$$

for some constant C, since $\mathrm{Vol}\left(B_n\right) = \Theta\left(n^2\right)$ and $\mathbb{E}_o\left[T_{B_n}\right] = \Theta\left(n^2\right)$. However, this bound is not tight. Recall that the general upper bound that gives (3.55) is proven by simulating a random walk within B_n. The key observation that improves (3.55) to a tight bound is that one can restrict the random walk on $\mathbb{C}_2$ to the rectangle $R_{C,n} := [-C\sqrt{n}, C\sqrt{n}] \times [-n, n]$ for large enough C. This is because with probability close to 1, the random walk will exit B_n before it exits the rectangle $R_{C,n}$. Since $\mathrm{Vol}\left(R_{C,n}\right) = \Theta\left(n^{3/2}\right)$, this gives the improved upper bound of $O\left(n^{7/2}\right)$.

To obtain a matching lower bound, we first smooth the mass distribution by simulating the random walk killed when it exits the rectangle $[-C\sqrt{n}, C\sqrt{n}] \times (-n/2, n/2)$. The resulting mass distribution μ has almost all of its mass on the "ends of the teeth", i.e., on the set

$$S := \left\{\left(i, \pm\frac{n}{2}\right) : |i| \leq C\sqrt{n}\right\}.$$

Moreover, most of the mass is roughly uniformly spread on S, in the sense that $\mu(x) = O(1/\sqrt{n})$ for every $x \in S$ (after potentially throwing away a tiny constant mass). So in order to move a constant mass p to distance n from the origin o, we need to move a constant fraction of the mass present at $\Omega(\sqrt{n})$ points in S. Since each "tooth" of the comb is locally a line, this requires $\Omega(n^3)$ toppling moves along each tooth (by Theorem 14 for $d = 1$, proven in[74]), resulting in $\Omega\left(n^{7/2}\right)$ toppling moves in total.

The rest of this section makes the two preceding paragraphs precise and proves Theorem 16. Let $\{X_t\}_{t\geq 0}$ denote random walk on $\mathbb{C}_2$ started at the origin, i.e., with $X_0 = o$. We write $R \equiv R_{C,n}$ when the implied parameters are clear from the context. Let T_R denote the first exit time of the random walk $\{X_t\}_{t\geq 0}$ from R and let $Z_t := X_{t\wedge T_R}$ denote the random walk killed when it exits R. We write $X_t = \left(X_t^{(1)}, X_t^{(2)}\right)$ and $Z_t = \left(Z_t^{(1)}, Z_t^{(2)}\right)$ for the coordinates of X_t and Z_t.

The following lemma says that by making C large enough, one can make the probability that the random walk exits $R_{C,n}$ along one of the "teeth" arbitrarily close to 1.

Lemma 22. For every $\epsilon > 0$ there exists $C = C(\epsilon) < \infty$ such that

$$\mathbb{P}_o\left(X^{(2)}_{T_{R_{C,n}}} = 0\right) \leq \epsilon.$$

Proof. It suffices to show that if we run the random walk on $\mathbb{C}_2$ for cn^2 steps, where c is large enough, then the probability that the random walk has not yet reached $\pm n$ in the second coordinate is small and the probability that the random walk has reached $\pm C\sqrt{n}$ in the first coordinate is also small. More precisely, the statement follows from the following two inequalities:

$$\mathbb{P}_o\left(\left|X_t^{(2)}\right| \leq n \text{ for every } t \leq cn^2\right) \leq \epsilon/2, \tag{3.56}$$

$$\mathbb{P}_o\left(\left|X_t^{(1)}\right| \geq C\sqrt{n} \text{ for some } t \leq cn^2\right) \leq \epsilon/2. \tag{3.57}$$

Note that $\left\{X_t^{(2)}\right\}_{t\geq 0}$ is Markovian: when away from 0 it behaves like simple symmetric random walk on $\mathbb{Z}$ and at 0 it becomes lazy, i.e., it stays put with probability 1/2, and it jumps to ± 1 with probability 1/4 each. Therefore (3.56) follows from classical random walk estimates (for instance, it follows from the central limit theorem, see, e.g.,[84] Theorem 2.9), provided $c = c(\epsilon)$ is large enough.

Now fix c such that (3.56) holds. Again by classical estimates (see, e.g.,[84] Theorem 9.11) there exists a constant c' such that

$$\#\left\{t : t \leq cn^2, X_t^{(2)} = 0\right\} \leq c'n$$

with probability at least $1-\epsilon/4$. Note that $\left\{X_t^{(1)}\right\}_{t\geq 0}$ only moves at times when $X_t^{(2)} = 0$, and when it does, it moves according a lazy random walk, staying in put with probability 1/2. Let $\{Y_t\}_{t\geq 0}$ denote such a lazy random walk. By classical estimates we have that

$$\mathbb{P}_o\left(|Y_t| \geq C\sqrt{n} \text{ for some } t \leq c'n\right) \leq \epsilon/4$$

provided that C is large enough. Putting everything together gives us (3.57).

$\square$

With this lemma in hand we are now ready to prove Theorem 16. We start with the upper bound and we again give two proofs, one using random walk and one using a greedy algorithm.

Proof of the upper bound of Theorem 16 using random walk. Fix $\epsilon \in (0, 1-p)$, let $C = C(\epsilon)$ be the constant given by Lemma 22, and let $R := R_{C,n}$. Just like in the proof of Theorem 15, we define a sequence of toppling moves $\mu_0 := \delta_o, \mu_1, \mu_2, \ldots$ that simulate the killed random walk $\{Z_t\}_{t\geq 0}$, i.e., for every non-negative integer t, the distribution $\mu_{t|R|}$ agrees with the distribution of Z_t.

Let

$$M := \inf\left\{i \geq 0 : \mu_{i|R|}(R) \leq 1-p-\epsilon\right\}$$

denote the first time that the distribution of the killed random walk has mass at least $p+\epsilon$ outside of the rectangle R. By Lemma 22 we have that

$$\mu_{M|R|}\left(\left\{\left(-C\sqrt{n}-1,0\right)\right\} \cup \left\{\left(C\sqrt{n}+1,0\right)\right\}\right) \leq \mathbb{P}_o\left(X_{T_R}^{(2)} = 0\right) \leq \epsilon,$$

i.e., there is mass at most ϵ that is not at the "ends of the teeth" of R. Since every other vertex in the support of $\mu_{M|R|}$ that is outside of R has distance at least n from the origin, it follows that $\mu_{M|R|}(B_n) \leq 1-p$, which implies that

$$N_p(\mathbb{C}_2, B_n, \delta_o) \leq M|R|.$$

Just like in the proof of Theorem 15, one can show that

$$M < (1-p-\epsilon)^{-1}\,\mathbb{E}_o[T_R].$$

The upper bound now follows by putting together the previous two displays and using the facts that $|R| = \Theta\left(n^{3/2}\right)$ and $\mathbb{E}_o\left[T_R\right] = \Theta\left(n^2\right)$. $\square$

Proof of the upper bound of Theorem 16 using a greedy algorithm. Fix $\epsilon \in (0, 1-p)$, let $C = C(\epsilon)$ be the constant given by Lemma 22, and let $R := R_{C,n}$. Consider the following greedy algorithm for choosing toppling moves: until the mass outside of R is at least $p+\epsilon$, choose $v \in R$ with the largest mass in R (break ties arbitrarily) and topple the full mass at v. Let $\mu_0 \equiv \delta_o, \mu_1, \mu_2, \ldots$ denote the resulting mass distributions, let v_i denote the vertex that was toppled to get from μ_{i-1} to μ_i, and let m_i denote the mass that was toppled at this step. Furthermore, let t denote the number of moves necessary for this greedy algorithm to transport mass $p+\epsilon$ outside of R, i.e., $t = \min\left\{i \geq 1 : \mu_i(R) \leq 1-p-\epsilon\right\}$.

Just as in the proof of Theorem 21 we can compute how the second moment of the mass distribution changes after each toppling move and we obtain that

$$M_2\left[\mu_i\right] - M_2\left[\mu_{i-1}\right] = m_i.$$

The greedy choice implies that for every $i \leq t$ we must have that

$$m_i \geq \frac{\mu_{i-1}(R)}{|R|} > \frac{1-p-\epsilon}{|R|}.$$

This gives us the following lower bound on the second moment of μ_t:

$$M_2\left[\mu_t\right] = \sum_{i=1}^{t}\left(M_2\left[\mu_i\right] - M_2\left[\mu_{i-1}\right]\right) = \sum_{i=1}^{t} m_i > t \times \frac{1-p-\epsilon}{|R|}.$$

On the other hand, there exists a constant $C' < \infty$ such that $\|v\|_2^2 \leq C'n^2$ for every $v \in \mathbb{C}_2$ such that $\mu_t(v) > 0$, which implies that $M_2\left[\mu_t\right] \leq C'n^2$. Combining this with the display above we obtain that $t < C'n^2 \times |R| / (1-p-\epsilon)$. Since $|R| = \Theta\left(n^{3/2}\right)$ we thus have that $t = O\left(n^{7/2}\right)$.

What remains to show is that the mass distribution μ_t has mass at least p at distance at least n from the origin, i.e., that $\mu_t(B_n) \leq 1-p$. Note that there are only two vertices in the vertex boundary of R that are at distance less than n from the origin: $(-C\sqrt{n}-1,0)$ and $(C\sqrt{n}+1,0)$. Thus we have that

$$\mu_t(B_n) \leq \mu_t(R) + \mu_t\left(\left(-C\sqrt{n}-1,0\right)\right) + \mu_t\left(\left(C\sqrt{n}+1,0\right)\right),$$

and since $\mu_t(R) \leq 1-p-\epsilon$, what remains to show is that

$$\mu_t\left(\left(-C\sqrt{n}-1,0\right)\right) + \mu_t\left(\left(C\sqrt{n}+1,0\right)\right) \leq \epsilon. \tag{3.58}$$

For $x \in \mathbb{C}_2$ let $h(x) := \mathbb{P}_x\left(X^{(2)}_{T_R} = 0\right)$. By Lemma 22 we have that $h(o) \leq \epsilon$, and hence $\sum_{x\in\mathbb{C}_2} h(x)\mu_0(x) \leq \epsilon$. The function h is harmonic in R, which implies that $\sum_{x\in\mathbb{C}_2} h(x)\mu_i(x) = \sum_{x\in\mathbb{C}_2} h(x)\mu_{i-1}(x)$ for every $i \geq 1$, and hence $\sum_{x\in\mathbb{C}_2} h(x)\mu_t(x) \leq \epsilon$. The inequality 3.58 then immediately follows from the fact that $h((-C\sqrt{n}-1,0)) = h((C\sqrt{n}+1,0)) = 1$. $\square$

Proof of the lower bound of Theorem 16. Given $p \in (0,1)$, let $\epsilon := p/4$. In the following we fix $c = c(\epsilon)$ and $C = C(\epsilon)$ to be large enough constants; we shall see soon the specific criterion for choosing these constants.

We start by smoothing the initial mass distribution appropriately. Define the rectangle $R' \equiv R'_{C,n} := [-C\sqrt{n}, C\sqrt{n}] \times (-n/2, n/2)$ and let $Z'_t := X_{t\wedge T_{R'}}$ denote the random walk killed when it exits R'. Starting with the initial mass distribution δ_o, we apply a sequence of $cn^2 \times \mathrm{Vol}(R')$ toppling moves that simulate cn^2

steps of the killed random walk $\{Z_t'\}_{t\geq 0}$, to arrive at a new mass distribution μ. In the same way as in the proof of Lemma 22, we can argue that most of the mass of the resulting measure μ is on the "ends of the teeth", i.e., it is on the set

$$S := \left\{ \left(i, \pm \frac{n}{2} \right) : |i| \leq C\sqrt{n} \right\}.$$

More precisely, if c and C are chosen appropriately, then $\mu(S) \geq 1 - \epsilon$. Furthermore, most of the mass is roughly uniformly spread on S. Specifically, we claim that there exists a constant K such that we can write the mass measure μ restricted to S as the sum of two mass measures, $\mu|_S = \mu_1 + \mu_2$, such that

$$\mu_1(x) \leq \frac{K}{\sqrt{n}}, \quad \forall x \in S, \qquad \text{and} \qquad \mu_2(S) \leq \epsilon. \tag{3.59}$$

Before proving (3.59), we show how to conclude the proof assuming that (3.59) holds. First of all, from Corollary 21.1 we have that $N_p(\mathbb{C}_2, B_n, \delta_o) \geq N_p(\mathbb{C}_2, B_n, \mu)$, so it suffices to bound from below this latter quantity. Now suppose that a sequence of toppling moves takes μ to a mass distribution μ' satisfying $\mu'(B_n) \leq 1 - p$, and for $x \in S$ let $\nu(x) \in [0, \mu(x)]$ denote the amount of mass that was originally (under μ) at x, but through the toppling moves was transported outside of B_n. We can write $\nu(x) = \nu_1(x) + \nu_2(x)$ in accordance with how we have $\mu(x) = \mu_1(x) + \mu_2(x)$. Since $\mu(S) \geq 1 - \epsilon$ and $\mu'(B_n) \leq 1 - p$, we must have that

$$\sum_{x \in S} \nu(x) \geq p - \epsilon. \tag{3.60}$$

Since $\nu_2(S) \leq \mu_2(S) \leq \epsilon$, we must then have that

$$\sum_{x \in S} \nu_1(x) \geq p - 2\epsilon. \tag{3.61}$$

Let $S_{\mathrm{lg}} := \{x \in S : \nu_1(x) \geq \epsilon/(5C\sqrt{n})\}$ and $S_{\mathrm{sm}} := S \setminus S_{\mathrm{lg}}$, and break the sum in (3.61) into two parts accordingly. Using that $|S| = 4C\sqrt{n} + 2 \leq 5C\sqrt{n}$, we have that $\sum_{x \in S_{\mathrm{sm}}} \nu_1(x) \leq \epsilon$, and so

$$\sum_{x \in S_{\mathrm{lg}}} \nu_1(x) \geq p - 3\epsilon = p/4.$$

On the other hand, (3.59) implies that

$$\sum_{x \in S_{\mathrm{lg}}} \nu_1(x) \leq |S_{\mathrm{lg}}| \times \frac{K}{\sqrt{n}}$$

and so we must have that $|S_{\mathrm{lg}}| \geq \frac{p}{4K}\sqrt{n}$. Notice that for every $x \in S_{\mathrm{lg}}$ we have that $\nu_1(x)/\mu_1(x) \geq \epsilon/(5CK)$, i.e., a constant fraction of the mass at x (under μ_1) is transported outside of B_n. In order to transport mass from $x = (x_1, x_2) \in S$ to outside of B_n, the mass necessarily has to go through either $(x_1, x_2 + n/4)$ or $(x_1, x_2 - n/4)$. Since the graph between these two points is a line of length $\Omega(n)$, we know from Theorem 14 for $d = 1$ (proven in[74]) that $\Omega(n^3)$ toppling moves are necessary to do this. Since this holds for every $x \in S_{\mathrm{lg}}$, we see that $\Omega\left(n^{7/2}\right)$ toppling moves are necessary altogether.

Finally, we turn back to proving (3.59). First, note that there exists $\delta = \delta(\epsilon)$ such that with probability at least $1 - \epsilon/2$, the killed random walk $\{Z'_t\}_{t \geq 0}$ has not exited the rectangle R' by time δn^2 (this follows by classical estimates for simple random walk, see, e.g.,[84]). On this event, which we shall denote by

$\mathcal{A}$, the killed random walk $\{Z'_t\}_{t=0}^{\delta n^2}$ and the simple random walk $\{X_t\}_{t=0}^{\delta n^2}$ agree. Now let $N_t^{Z'}$ denote the number of visits to the x axis of the killed random walk until time t, i.e.,

$$N_t^{Z'} := \# \left\{ k \in \{0, 1, \dots, t\} : Z'^{(2)}_k = 0 \right\},$$

and similarly define N_t^X for the simple random walk. Under the event $\mathcal{A}$, we have that

$$N_{\delta n^2}^{Z'} = N_{\delta n^2}^X.$$

By classical estimates on the local time at 0 (see, e.g.,[84] Theorem 9.11), there exists $\gamma = \gamma(\epsilon)$ such that with probability at least $1 - \epsilon/2$, we have that

$$N_{\delta n^2}^X \geq \gamma n. \tag{3.62}$$

Denote by $\mathcal{B}$ the event that the inequality in (3.62) holds and note that $\mathbb{P}(\mathcal{A} \cap \mathcal{B}) \geq 1 - \epsilon$. In the following we assume that the event $\mathcal{A} \cap \mathcal{B}$ holds; whatever happens on the event $(\mathcal{A} \cap \mathcal{B})^c$ we put into the mass measure μ_2, which hence has mass at most ϵ.

Under the event $\mathcal{A} \cap \mathcal{B}$ we have that $N := N_{cn^2}^{Z'} \geq N_{\delta n^2}^{Z'} = N_{\delta n^2}^X \geq \gamma n$. Let $\{Y_t\}_{t \geq 0}$ denote a lazy random walk on $\mathbb{Z}$ that stays put with probability $1/2$, and otherwise does a step according to simple random walk, just like in the proof of Lemma 22. Conditioned on N, we have that $Z'^{(1)}_{cn^2}$ has the same distribution as Y_N. For fixed N, the local limit theorem says that there exists K' such that

$$\sup_{\ell \in \mathbb{Z}} \mathbb{P}\left(Y_N = \ell\right) \leq \frac{K'}{\sqrt{N}}.$$

Hence there exists K such that

$$\sup_{\ell \in \mathbb{Z}} \mathbb{P}\left({Z'}^{(1)}_{cn^2} = \ell \,\middle|\, \mathcal{A} \cap \mathcal{B} \right) \leq \frac{K}{\sqrt{n}},$$

which implies the claim. □

3.2.8 Graphs where random walk has positive speed

In this section we study graphs on which simple random walk has positive speed. As a warm-up, we study d-ary trees in Section 3.2.8.1, followed by general results in Section 3.2.8.2. We then apply the general results to two examples: Galton-Watson trees (Section 3.2.8.3) and product of trees (Section 3.2.8.4). The main observation for these latter results is that in these cases one can a priori specify an exponentially small subset of the vertices of the ball of radius n with the property that the random walk on the graph started from the center of the ball does not exit this subset with probability close to 1. Thus simple random walk can be simulated approximately by performing toppling moves only on this exponentially small subset of B_n, leading to much better bounds than the general upper bound of Theorem 15.

3.2.8.1 d-ary trees

The general upper bound of Theorem 15 applied directly to the d-ary tree $\mathbb{T}_d$ gives

$$N_p(\mathbb{T}_d, B_n, \delta_\rho) < C(1-p)^{-1} \cdot n \cdot d^n$$

for some constant $C < \infty$, since $\mathrm{Vol}(B_n) = \Theta(d^n)$ and $\mathbb{E}_\rho[T_{B_n}] = \Theta(n)$. However, this bound is not tight, as Theorem 17 states that $N_p(\mathbb{T}_d, B_n, \delta_\rho) = \Theta(d^n)$. This example is interesting because the factor coming from the exit time of the random walk is completely absent from $N_p(\mathbb{T}_d, B_n, \delta_\rho)$. The proof requires a more careful analysis of the greedy algorithm.

In the rest of this subsection we prove Theorem 17, starting with the lower bound. We define the level of a vertex $v \in \mathbb{T}_d$ to be its distance from the root: $\ell(v) := d_{\mathbb{T}_d}(v, \rho)$.

Proof of the lower bound in Theorem 17. We begin by smoothing the initial mass distribution in such a way that most of the mass is on the vertices at level $n-1$, where it is uniformly spread. More precisely, for any $\epsilon > 0$ it is possible to obtain, via a finite sequence of toppling moves, a mass distribution μ such that $\mu(v) \in \left((1-\epsilon)\, d^{-(n-1)}, d^{-(n-1)}\right)$ for every vertex v at level $n-1$. By Corollary 21.1 we have that $N_p(\mathbb{T}_d, B_n, \delta_\rho) \geq N_p(\mathbb{T}_d, B_n, \mu)$, so it suffices to bound from below this latter quantity.

Fix $\epsilon \in (0, p)$. In order to transport mass at least p to level n starting from μ, it is necessary to transport mass at least $p - \epsilon$ to level n from the vertices at level $n-1$. However, each vertex at level $n-1$ has mass at most $d^{-(n-1)}$. Hence mass from at least $(p-\epsilon)d^{n-1}$ vertices at level $n-1$ needs to transported to level n, and this requires at least $(p-\epsilon)d^{n-1}$ toppling moves. Hence $N_p(\mathbb{T}_d, B_n, \mu) \geq (p-\epsilon)d^{n-1}$. □

The greedy algorithm provides an upper bound of the correct order. In order to analyze it we study the average level of a mass distribution μ, defined as

$$M_1[\mu] := \sum_{v\in\mathbb{T}_d} \mu(v)\,\ell(v).$$

We will make use of the following lemma, which states that if the average level is not too large, then there must be a reasonably large mass at some vertex.

Lemma 23. If μ is a mass distribution on $\mathbb{T}_d$ such that $M_1[\mu] \leq \ell$, then there exists $v \in \mathbb{T}_d$ such that $\ell(v) \leq \ell$ and $\mu(v) \geq d^{-(\ell+1)}/4$.

Proof. We prove the statement by contradiction. Suppose that $\mu(v) < d^{-(\ell+1)}/4$ for every $v \in \mathbb{T}_d$ such that $\ell(v) \leq \ell$; our goal is to show that then $M_1[\mu] > \ell$. To bound $M_1[\mu]$ from below, we can first bound $\ell(v)$ by $\ell+1$ for every v such that $\ell(v) \geq \ell+1$ to obtain that

$$\begin{aligned} M_1[\mu] &\geq \sum_{v:\ell(v)\leq\ell} \mu(v)\,\ell(v) + (\ell+1)\left(1 - \sum_{v:\ell(v)\leq\ell}\mu(v)\right) \\ &= \ell+1-\sum_{v:\ell(v)\leq\ell}\mu(v)\,(\ell+1-\ell(v)). \end{aligned}$$

Using the assumption that $\mu(v) < d^{-(\ell+1)}/4$ for every $v \in \mathbb{T}_d$ such that $\ell(v) \leq \ell$, we thus have that

$$M_1[\mu] \geq \ell+1-\frac{1}{4}d^{-(\ell+1)}\sum_{v:\ell(v)\leq\ell}(\ell+1-\ell(v)).$$

Finally, we have that

$$\sum_{v:\ell(v)\leq\ell} (\ell+1-\ell(v)) = \sum_{k=0}^{\ell} (\ell+1-k)\, d^k = \frac{1}{d-1}\left[d\cdot\frac{d^{\ell+1}-1}{d-1} - (\ell+1)\right] \leq 2d^{\ell+1},$$

and so $M_1[\mu] \geq \ell + 1/2$. □

Proof of the upper bound in Theorem 17. Consider the following greedy algorithm for choosing toppling moves: until the mass outside of B_n is at least p, choose $v \in B_n$ with the largest mass in B_n (break ties arbitrarily) and topple the full mass at v. Let $\mu_0 \equiv \delta_\rho, \mu_1, \mu_2, \ldots$ denote the resulting mass distributions, let v_i denote the vertex that was toppled to get from μ_{i-1} to μ_i, and let m_i denote the mass that was toppled at this step. Let t denote the number of moves necessary for this greedy algorithm to transport mass p to distance at least n from the root, i.e., $t = \min\{i \geq 0 : \mu_i(B_n) \leq 1-p\}$. Finally, for every $\ell \in \mathbb{N}$, let t_ℓ denote the number of moves necessary for this greedy algorithm to make the average level of the mass distribution at least ℓ, i.e., $t_\ell := \min\{i \geq 0 : M_1[\mu_i] \geq \ell\}$.

We first consider how the average level of the mass distribution changes with each toppling move. If $v_i = \rho$, then all the mass goes to the first level and hence we have that $M_1[\mu_i] - M_1[\mu_{i-1}] = m_i$. If $v_i \neq \rho$, then a $1/(d+1)$ fraction of the mass goes one level lower, while the rest of the mass goes one level higher, so $M_1[\mu_i] - M_1[\mu_{i-1}] = \frac{d-1}{d+1} m_i$. In every case we have that

$$M_1[\mu_i] - M_1[\mu_{i-1}] \geq \frac{d-1}{d+1} m_i.$$

Now fix $\ell < n$. By Lemma 23, for every $i < t_\ell$ we have that $m_i \geq d^{-(\ell+1)}/4$.

This implies that

$$M_1\left[\mu_{t_\ell - 1}\right] - M_1\left[\mu_{t_{\ell-1}}\right] \geq \left(t_\ell - 1 - t_{\ell-1}\right) \times \frac{d-1}{d+1} \times \frac{1}{4d^{\ell+1}}.$$

On the other hand, by the definition of t_ℓ we have that

$$M_1\left[\mu_{t_\ell - 1}\right] - M_1\left[\mu_{t_{\ell-1}}\right] < \ell - (\ell - 1) = 1.$$

Putting the previous displays together we obtain that

$$t_\ell - t_{\ell-1} = O\left(d^\ell\right) \tag{3.63}$$

for every $\ell < n$, where the implied constant depends only on d. Summing (3.63) over ℓ from 1 to $n-1$ we obtain that

$$t_{n-1} = O\left(d^n\right).$$

Thus what remains is to show that $t - t_{n-1} = O\left(d^n\right)$. Recall that for every $i < t$ we have that $\mu_i\left(B_n\right) > 1 - p$. Since $\mathrm{Vol}\left(B_n\right) < d^n$, there must exist $v \in B_n$ such that $\mu_i\left(v\right) > \left(1-p\right)/d^n$. Hence for every $i \in \left(t_{n-1}, t\right]$ we have that $m_i > \left(1-p\right)/d^n$. Thus

$$M_1\left[\mu_t\right] - M_1\left[\mu_{t_{n-1}}\right] > \left(t - t_{n-1}\right) \frac{d-1}{d+1} \left(1-p\right)/d^n.$$

On the other hand, since the support of μ_t is contained in B_{n+1}, we have that $M_1\left[\mu_t\right] \leq n$, so

$$M_1\left[\mu_t\right] - M_1\left[\mu_{t_{n-1}}\right] \leq n - (n-1) = 1.$$

Putting the previous two displays together we obtain that $t - t_{n-1} < \left(1-p\right)^{-1} \frac{d+1}{d-1} d^n$.

$\square$

3.2.8.2 A general bound for graphs where random walk has positive speed and entropy

In this subsection we present a general result for graphs where simple random walk has positive speed and entropy. Let $G = (V, E)$ be an infinite, connected, locally finite graph with $o \in V$ a specified vertex, and let $\{X_t\}_{t \geq 0}$ denote simple random walk on G started from o, i.e., with $X_0 = o$. We denote by $p_t(\cdot, \cdot)$ the t step probability transition kernel. We start by introducing the basic notions of speed and entropy for random walk.

Definition 7. The (asymptotic) *speed* of the random walk $\{X_t\}_{t \geq 0}$ on G is defined as

$$\boldsymbol{\ell} := \lim_{t \to \infty} \frac{d(X_0, X_t)}{t}.$$

Note that the triangle inequality implies subadditivity, that is, $d(X_0, X_{s+t}) \leq d(X_0, X_s) + d(X_s, X_{s+t})$, and hence the speed of the random walk exists almost surely by Kingman's subadditive ergodic theorem (see, e.g.,[65] Theorem 14.44).

Recall that the entropy of a discrete random variable X taking values in $\mathcal{X}$ is defined as

$$H(X) = -\sum_{x \in \mathcal{X}} \mathbb{P}(X = x) \log \mathbb{P}(X = x),$$

where here we use log to denote the natural logarithm.

Definition 8. The *asymptotic entropy*, also known as the *Avez entropy*, of the random walk $\{X_t\}_{t \geq 0}$ on G is defined as

$$\boldsymbol{h} := \lim_{t \to \infty} \frac{H(X_t)}{t},$$

provided that this limit exists.

When G is transitive, the sequence $\{H(X_t)\}_{t \geq 0}$ is subadditive, and hence the Avez entropy exists by Fekete's lemma (see, e.g.,[65] Section 14.1).

We recall two results concerning the asymptotic speed and the Avez entropy of the random walk. First, the positivity of these two quantities are related, as stated in the following theorem.

Theorem 24. [[52],[65] Theorem 14.1] Let G be a Cayley graph. Then the random walk has positive asymptotic speed, i.e., $\boldsymbol{\ell} > 0$, if and only if the Avez entropy of the random walk is positive, i.e., $\boldsymbol{h} > 0$.

The following result is known as Shannon's theorem for random walks.

Theorem 25. [[52] Theorem 2.1,[65] Theorem 14.10] Assume the setup described in the first paragraph of Section 3.2.8.2 and in addition assume that G is a transitive graph. Then we have that

$$\lim_{t \to \infty} \frac{1}{t} \log p_t(o, X_t) = -\boldsymbol{h}$$

almost surely.

In the main result of this subsection, we provide sharp bounds in the exponent for the number of toppling moves necessary to transport mass p to distance n for graphs where simple random walk has positive asymptotic speed, positive Avez entropy, and which satisfy Shannon's theorem.

Theorem 26. Let $G = (V, E)$ be an infinite, connected, locally finite graph with $o \in V$ a specified vertex, and let $\{X_t\}_{t\geq 0}$ denote simple random walk on G started from o, i.e., with $X_0 = o$. Assume that the following three conditions hold:

1. Simple random walk on G has positive asymptotic speed, i.e., $\boldsymbol{\ell} > 0$.

2. Simple random walk on G has positive Avez entropy, i.e., $\boldsymbol{h} > 0$.

3. We have that

$$\lim_{t\to\infty} \frac{1}{t} \log p_t (o, X_t) = -\boldsymbol{h} \tag{3.64}$$

   almost surely.

Then the minimum number of toppling moves needed to transport mass p to distance at least n from o is

$$N_p (G, B_n, \delta_o) = \exp \left(n \times \frac{\boldsymbol{h}}{\boldsymbol{\ell}} (1 + o (1)) \right). \tag{3.65}$$

Proof. To prove the upper bound, we define a sequence of toppling moves that simulates the random walk, killed when it exits B_n, until time $t^* = (1 + \epsilon) n/\boldsymbol{\ell}$, by which time most of the mass is outside of B_n. However, in order to get an upper bound of the correct order, we only do the toppling moves at the subset of sites that the random walk typically visits. The rest of the proof makes this precise.

Fix $\epsilon > 0$ and let $t^* = (1+\epsilon)\, n/\boldsymbol{\ell}$. We first define the set of vertices on which we simulate the random walk. Let

$$r_n := \max\{r : |B_r| \leq n\}$$

and note that $\lim_{n\to\infty} r_n = \infty$ due to the assumptions on G. Define also

$$V_{t,n} := \left\{ x \in B_n : \frac{1}{t} \log p_t(o,x) \in (-\boldsymbol{h}(1+\epsilon), -\boldsymbol{h}(1-\epsilon)) \right\}, \tag{3.66}$$

and note that $|V_{t,n}| \leq \exp(t\boldsymbol{h}(1+\epsilon))$ for every t, since $p_t(o,x) \geq \exp(-t\boldsymbol{h}(1+\epsilon))$ for every $x \in V_{t,n}$. Now define

$$U_n := B_{r_n} \cup \bigcup_{t=r_n}^{t^*} V_{t,n}$$

and let $Z_t := X_{t \wedge T_{U_n}}$ denote the random walk started at o and killed when it exits U_n. We can simulate the killed random walk $\{Z_t\}_{t=0}^{t^*}$ using $t^* |U_n|$ toppling moves. We shall show that

$$\mathbb{P}_o(Z_{t^*} \notin B_n) \geq p \tag{3.67}$$

if n is large enough, which thus implies that

$$N_p(G, B_n, \delta_o) \leq t^* |U_n| \leq t^* (n + t^* \exp(t^* \boldsymbol{h}(1+\epsilon)))$$

if n is large enough. Since this holds for every $\epsilon > 0$, we get the desired upper bound stated in (3.65).

So what remains is to show (3.67). There are two ways that Z_{t^*} can be in the ball B_n: either it is in the set U_n, or the random walk exited U_n before exiting the ball B_n, and thus we have that

$$\mathbb{P}_o\left(Z_{t^*} \in B_n\right) = \mathbb{P}_o\left(Z_{t^*} \in U_n\right) + \mathbb{P}_o\left(Z_{t^*} \in B_n \setminus U_n\right). \tag{3.68}$$

The first scenario is unlikely due to Assumption 1. Specifically, if the killed random walk has not exited U_n, then its distance from $X_0 = o$ is less than n, so we have that

$$\mathbb{P}_o\left(Z_{t^*} \in U_n\right) \leq \mathbb{P}_o\left(d\left(X_0, X_{t^*}\right) < n\right) = \mathbb{P}_o\left(\tfrac{1}{t^*} d\left(X_0, X_{t^*}\right) < \boldsymbol{\ell}/\left(1+\epsilon\right)\right).$$

Assumption 1 implies that this latter probability goes to 0, since $t^* \to \infty$ as $n \to \infty$. In particular, if n is large enough then we have that $\mathbb{P}_o\left(Z_{t^*} \in U_n\right) \leq (1-p)/2$. The second probability on the right hand side of (3.68) is small due to Assumption 3. First note that the random walk satisfies $Z_t \in U_n$ for all $t < r_n$ due to the construction of U_n. Now if the random walk exited U_n before exiting B_n, then by the definition of U_n there must exist a time $t \in \{r_n, r_n+1, \ldots, t^*\}$ such that $X_t \in B_n \setminus V_{t,n}$. This implies that

$$\mathbb{P}_o\left(Z_{t^*} \in B_n \setminus U_n\right) \leq \mathbb{P}_o\left(\exists\, t \geq r_n : \tfrac{1}{t} \log p_t\left(X_0, X_t\right) \notin \left(-\boldsymbol{h}\left(1+\epsilon\right), -\boldsymbol{h}\left(1-\epsilon\right)\right)\right).$$

Assumption 3 implies that this latter probability converges to 0 as $r_n \to \infty$. Since $r_n \to \infty$ as $n \to \infty$, we have in particular that $\mathbb{P}_o\left(Z_{t^*} \in B_n \setminus U_n\right) \leq (1-p)/2$ if n is large enough. This concludes the proof of (3.67).

To prove the lower bound stated in (3.65), we again start by smoothing the initial mass distribution, by simulating simple random walk on G until time $t^{**} := (1-\epsilon)\, n/\boldsymbol{\ell}$. As we shall see, the mass distribution is then approximately uniformly distributed on a subset of B_n of size approximately $\exp(t^{**}\boldsymbol{h})$. In order to transport a constant mass outside of B_n, it is then necessary to topple the mass at a constant fraction of the vertices in this subset, which leads to the desired lower bound. The rest of the proof makes this precise.

Fix $\epsilon > 0$ and let $t^{**} := (1-\epsilon)\, n/\boldsymbol{\ell}$. The choice of t^{**} is due to the fact that, by Assumption 1, with probability close to 1, simple random walk on G does not exit the ball B_n until time t^{**}. Let $Z'_t := X_{t \wedge T_{B_n}}$ denote the simple random walk on G killed when it exits B_n. Starting with the initial mass distribution δ_o, we apply a sequence of $t^{**} \times \mathrm{Vol}(B_n)$ toppling moves that simulate t^{**} steps of the killed random walk $\{Z'_t\}_{t=0}^{t^{**}}$, to arrive at a new mass distribution μ. By Corollary 21.1 we have that $N_p(G, B_n, \delta_o) \geq N_p(G, B_n, \mu)$, so it suffices to bound from below this latter quantity. Recall the definition of $V_{t,n}$ from (3.66). By the definition of t^{**} and Assumptions 1 and 3, it follows that

$$\mu(V_{t^{**},n}) \geq 1 - \frac{p}{2}$$

if n is large enough. Therefore, in order to transport mass p outside of B_n starting from the mass distribution μ, it is necessary to transport mass at least $p/2$ from vertices in $V_{t^{**},n}$. However, $\mu(x) \leq \exp(-t^{**}\boldsymbol{h}(1-\epsilon))$ for every $x \in V_{t^{**},n}$, so at least

$$\frac{p}{2} \times \exp\left(t^{**}\boldsymbol{h}\left(1-\epsilon\right)\right) = \frac{p}{2} \times \exp\left(n \times \frac{\boldsymbol{h}}{\boldsymbol{\ell}}\left(1-\epsilon\right)^2\right)$$

vertices in $V_{t^{**},n}$ need to be toppled at least once. Since this holds for any $\epsilon > 0$, the result follows. ☐

3.2.8.3 Galton-Watson trees

The behavior of simple random walk on Galton-Watson trees was studied in great detail by Lyons, Pemantle, and Peres[64]. Using their results, combined with the general results of Section 3.2.8.2, we can prove Theorem 18.

Specifically, Lyons, Pemantle, and Peres[64] showed that the three conditions of Theorem 26 hold for almost every Galton-Watson tree. Furthermore, they also show that the ratio of the asymptotic entropy and speed is equal to the Hausdorff dimension of harmonic measure on the boundary of a Galton-Watson tree. Here we state the basic results necessary to conclude Theorem 26, and refer to[64] for much more detailed results, including formulas for the asymptotic speed and entropy as a function of the offspring distribution of the Galton-Watson branching process. We state this result for nondegenerate offspring distributions, as degenerate offspring distributions (giving rise to m-ary trees) are treated more carefully in Section 3.2.8.1.

Theorem 27. [[64] Theorem 1.1, Theorem 3.2, Theorem 9.7] Fix a nondegenerate offspring distribution with mean $m > 1$ and let GWT be a Galton-Watson tree obtained with this offspring distribution, on the event of nonextinction. Let $\{X_t\}_{t\geq 0}$ denote simple random walk on GWT started from the root ρ, i.e., with

$X_0 = \rho$, and let $p_t(\cdot,\cdot)$ denote the t step probability transition kernel. For almost every Galton-Watson tree GWT the following statements hold. The asymptotic speed $\boldsymbol{\ell}$ and Avez entropy $\boldsymbol{h}$ of the random walk exist and are positive almost surely. Moreover, we have that

$$\frac{\boldsymbol{\ell}}{\boldsymbol{h}} = \mathbf{dim}$$

almost surely, where **dim** is the dimension of harmonic measure, which is almost surely a constant less than $\log m$. Furthermore, we have that

$$\lim_{t\to\infty} \frac{1}{t} \log p_t(o, X_t) = -\boldsymbol{h}$$

almost surely.

Proof of Theorem 18. Theorem 27 shows that the three conditions of Theorem 26 hold for almost every Galton-Watson tree. Hence Theorem 18 follows from Theorem 26. □

3.2.8.4 Product of trees

In this subsection we apply the general result derived in Section 3.2.8.2 to obtain tight bounds for the specific case of the product of trees. As we shall see, the key observation is that random walk typically does not visit the entire ball B_n on the product of trees, due to its different speeds on the edges belonging to different trees.

Let $\mathbb{T}_d$ denote the $(d+1)$-regular tree.* We define the Cartesian product $\mathbb{T}_d \times$

*In Section 3.2.8.1, $\mathbb{T}_d$ denotes the d-ary tree, which differs from the $(d+1)$-regular tree in

$\mathbb{T}_k$ to have vertex set $V(\mathbb{T}_d \times \mathbb{T}_k) = V(\mathbb{T}_d) \times V(\mathbb{T}_k)$ and edge set defined as follows:

$$(u,v) \sim (u',v') \Longleftrightarrow \begin{cases} u \sim u' \text{ and } v = v', & \text{or} \\ u = u' \text{ and } v \sim v'. \end{cases}$$

Note that $\mathbb{T}_d \times \mathbb{T}_k$ is a $(d+k+2)$-regular graph. Note also that $\mathbb{T}_1$ is isomorphic to $\mathbb{Z}$, and so $T_1 \times T_1$ is isomorphic to $\mathbb{Z}^2$; this graph is covered by Theorem 14, and hence we may assume that $d + k \geq 3$.

Proof of Theorem 19. We prove this result by appealing to the general result of Theorem 26. Therefore we need to check that the three assumptions of Theorem 26 hold and we also need to compute the asymptotic speed $\boldsymbol{\ell}$ and the Avez entropy $\boldsymbol{h}$ for simple random walk on $\mathbb{T}_d \times \mathbb{T}_k$.

Let $\{X_t\}_{t\geq 0}$ denote simple random walk on $\mathbb{T}_d \times \mathbb{T}_k$ with $X_0 = \rho$. We start by computing the speed of random walk. Recall that the speed of random walk on the $(d+1)$-regular tree $\mathbb{T}_d$ is $\frac{d-1}{d+1}$. Moreover, the probability of random walk on $\mathbb{T}_d \times \mathbb{T}_k$ making a step in the first coordinate (corresponding to $\mathbb{T}_d$) is $\frac{d+1}{d+k+2}$. Hence the speed of random walk $\{X_t\}_{t\geq 0}$ is the convex combination of the speeds of random walk on the regular trees $\mathbb{T}_d$ and $\mathbb{T}_k$:

$$\boldsymbol{\ell} = \frac{d+1}{d+k+2} \times \frac{d-1}{d+1} + \frac{k+1}{d+k+2} \times \frac{k-1}{k+1} = \frac{d+k-2}{d+k+2}. \tag{3.69}$$

Since $d + k \geq 3$, the speed is positive: $\boldsymbol{\ell} > 0$.

Since $\mathbb{T}_d \times \mathbb{T}_k$ is a transitive graph, we know from Theorem 25 that (3.64)

that the root ρ has degree d instead of $d+1$. This difference is not important for the questions we consider, so we allow ourselves this abuse of notation.

holds. Thus what remains is to compute the Avez entropy of $\{X_t\}_{t\geq 0}$ and to show that it is positive. We start by computing the Avez entropy of random walk on $\mathbb{T}_d$. Let $\{Y_t\}_{t\geq 0}$ denote simple random walk on $\mathbb{T}_d$ started from the root, i.e., with $Y_0 = \rho$, and let $|Y_t|$ denote the distance of Y_t from the root ρ. By the chain rule of conditional entropy we have that

$$H(Y_t) = H(|Y_t|) + H(Y_t \,|\, |Y_t|).$$

Since $|Y_t|$ takes values in $\{0, 1, \ldots, t\}$, we have that $H(|Y_t|) \in [0, \log(t+1)]$. For $i \in [t]$, conditioned on $|Y_t| = i$, the random variable Y_t is uniformly distributed among all $(d+1)\, d^{i-1}$ vertices at distance i from the root. Hence, using the fact that the asymptotic speed of $\{Y_t\}_{t\geq 0}$ is $\frac{d-1}{d+1}$, we have that

$$\begin{aligned} H(Y_t \,|\, |Y_t|) &= \sum_{i=1}^{t} \mathbb{P}(|Y_t| = i) \times \log\left((d+1)\, d^{i-1}\right) = \log(1 + 1/d) \times \mathbb{P}(|Y_t| \neq 0) + \log(d) \times \mathbb{E}[|Y_t|] \\ &= \log(d) \times \frac{d-1}{d+1} \times t\,(1 + o(1)). \end{aligned}$$

We conclude that the Avez entropy of $\{Y_t\}_{t\geq 0}$ is

$$\boldsymbol{h}_Y = \log(d) \times \frac{d-1}{d+1}.$$

Now let $\{Z_t\}_{t\geq 0}$ denote simple random walk on $\mathbb{T}_k$ started from the root, i.e., with $Z_0 = \rho$, and let $\{Y_t\}_{t\geq 0}$ and $\{Z_t\}_{t\geq 0}$ be independent. Furthermore, independently of everything else, let $\{W_i\}_{i\geq 1}$ be i.i.d. Bernoulli random variables with expectation $\frac{d+1}{d+k+2}$, and let $S_t := \sum_{i=1}^{t} W_i$. Then, by construction, $\{(Y_{S_t}, Z_{t-S_t})\}_{t\geq 0}$ has the same distribution as $\{X_t\}_{t\geq 0}$. We can again use the chain rule of conditional entropy, this time conditioning on S_t, to get that

$$H(X_t) = H(S_t) + H((Y_{S_t}, Z_{t-S_t}) \mid S_t).$$

Since S_t takes values in $\{0, 1, \ldots, t\}$, we have that $H(S_t) \in [0, \log(t+1)]$. Conditioning on S_t, the random variables Y_{S_t} and Z_{t-S_t} are independent, and hence $H((Y_{S_t}, Z_{t-S_t}) \mid S_t) = H(Y_{S_t} \mid S_t) + H(Z_{t-S_t} \mid S_t)$. Therefore, using the computation from above of the entropy of random walk on a regular tree, together with the fact that $S_t = \frac{d+1}{d+k+2} t (1 + o(1))$ with high probability, we obtain that the Avez entropy of $\{X_t\}_{t \geq 0}$ is

$$\begin{aligned}
\boldsymbol{h}_X &= \frac{d+1}{d+k+2} \boldsymbol{h}_Y + \frac{k+1}{d+k+2} \boldsymbol{h}_Z \\
&= \frac{d-1}{d+k+2} \log(d) + \frac{k-1}{d+k+2} \log(k).
\end{aligned}$$

Since at least one of d and k is greater than 1, the Avez entropy $\boldsymbol{h}_X$ is positive. Plugging in the values of $\boldsymbol{\ell}$ and $\boldsymbol{h}$ into the conclusion of Theorem 26, we obtain the desired result. $\square$

3.2.9 Graphs with bounded degree and exponential decay of the Green's function

In this section we study graphs of bounded degree with exponential decay of the Green's function, showing that the minimum number of toppling moves necessary to transport a constant mass to distance at least n is exponential in n.

Let $G = (V, E)$ be an infinite and connected graph with bounded degree. Recall the definition of the Green's function g for simple random walk on G from Definition 5. We say that the Green's function has *exponential decay* if there exist positive and finite constants a and a' depending only on G such that

$$g(x,y) \leq \exp(-a \times d(x,y) + a') \tag{3.70}$$

for every $x, y \in V$, where d denotes graph distance. Note that the Green's function cannot decay faster than exponentially as a function of the distance.

If simple random walk on G has positive speed and positive entropy, then the Green's function has exponential decay (see[10,15]). However, the reverse implication does not hold, and hence the method described in Section 3.2.8.2 to bound the minimum number of toppling moves $N_p(G, B_n, \delta_o)$ does not work in general. As an example, we shall investigate the lamplighter graph with base graph $\mathbb{Z}$, for which it has been shown that the speed and entropy of simple random walk are both zero (see[52] Proposition 6.2).

We restate Theorem 20 more precisely before proving it.

Theorem 28. Let $G = (V, E)$ be an infinite and connected graph such that every vertex has degree at most D and the Green's function g for simple random walk on G satisfies (3.70). Start with initial unit mass δ_o at a vertex $o \in V$ and let $p \in (0, 1)$ be constant. The minimum number of toppling moves needed to transport mass p to distance at least n from o is

$$N_p(G, B_n, \delta_o) = \exp(\Theta(n)),$$

where the implied constants depend only on p, D, a, and a'.

Proof. For the upper bound we use the general bound given by Theorem 15. Since G has bounded degree, the volume of a ball grows at most exponentially: $\mathrm{Vol}\left(B_n\right) \leq \sum_{i=0}^{n-1} D^i \leq D^n$. Furthermore, the exit time of random walk from a ball can also be bounded, e.g., in the following crude way. The exit time T_{B_n} is equal to the number of visits to vertices in B_n before the random walk exits B_n, and hence can be bounded by the total number of visits to vertices in B_n. Thus we obtain the following crude bound: $\mathbb{E}_o\left[T_{B_n}\right] \leq \sum_{x \in B_n} g\left(o,x\right) \leq e^{a'}\mathrm{Vol}\left(B_n\right)$. Hence using Theorem 15 we have that

$$N_p\left(G, B_n, \delta_o\right) \leq \left(1-p\right)^{-1} e^{a'} D^{2n}.$$

For the lower bound we again perform smoothing of the initial mass distribution. Let $\left\{X_t\right\}_{t \geq 0}$ denote simple random walk on G with $X_0 = o$, and let $Z_t := X_{t \wedge T_{B_{n-1}}}$ denote the random walk killed when it exits the ball B_{n-1}. Let t^* be such that

$$\mathbb{P}_o\left(Z_{t^*} \in B_{n-1}\right) \leq p/2. \tag{3.71}$$

Starting with the initial mass distribution δ_o, we apply a sequence of $t^* \times \mathrm{Vol}\left(B_n\right)$ toppling moves that simulate t^* steps of the killed random walk $\left\{Z_t\right\}_{t \geq 0}$, to arrive at a new mass distribution μ. By Corollary 21.1 we have that $N_p\left(G, B_n, \delta_o\right) \geq N_p\left(G, B_n, \mu\right)$, so it suffices to bound from below this latter quantity.

Denote the boundary of B_{n-1} by $\partial B_{n-1} := \left\{x \in V : d\left(o,x\right) = n-1\right\}$. For every $x \in \partial B_{n-1}$ we can bound the mass at x using the Green's function:

$$\mu(x) = \mathbb{P}_o(Z_{t^*} = x) \le \mathbb{P}_o\left(X_{T_{B_{n-1}}} = x\right) \le \sum_{k=0}^{\infty} \mathbb{P}_o(X_k = x) = g(o,x) \le \exp(-an + a + a'),$$

where in the last inequality we used 3.70. Now (3.71) implies that $\mu(\partial B_{n-1}) \ge 1 - p/2$, and so in order to transport mass at least p to outside of B_n starting from μ, it is necessary to transport mass at least $p/2$ from the vertices in ∂B_{n-1}. However, the display above shows that every $x \in \partial B_{n-1}$ has mass at most $\exp(-an + a + a')$, so this requires at least $(p/2) \times \exp(an - a - a')$ toppling moves. Hence

$$N_p(G, B_n, \mu) \ge \frac{p}{2\exp(a + a')} \times e^{an}. \qquad \square$$

3.2.9.1 The lamplighter graph

We illustrate the results above with the lamplighter graph, which is an example of a graph with bounded degree and exponential decay of the Green's function.

Definition 9. The *lamplighter group* is the wreath product $\mathbb{Z}_2 \wr \mathbb{Z}$. The elements of the group are pairs of the form (η, y), where $\eta : \mathbb{Z} \to \mathbb{Z}_2$ and $y \in \mathbb{Z}$. The group operation is

$$(\eta_1, y_1)(\eta_2, y_2) := (\eta, y_1 + y_2),$$

where $\eta(x) = \eta_1(x) + \eta_2(x - y_1) \mod 2$.

The reason for the name is that we may think of a lamp being present at each vertex of $\mathbb{Z}$, with a lamplighter walking on $\mathbb{Z}$ and turning lights on and off. A group element (η, y) corresponds to the on/off configuration of the lamps η and the position of the lamplighter y. Multiplying with the group elements $(\mathbf{0}, 1)$ and $(\mathbf{0}, -1)$ corresponds to the lamplighter moving to the right or to the left, and multiplying with $(\mathbf{1}_0, 0)$ corresponds to flipping the light at the position of the lamplighter. Consider the random walk on the lamplighter group associated with the measure $\nu * \mu * \nu$, where μ is a simple random walk step by the lamplighter, and ν is a measure causing the lamplighter to randomize the current lamp. That is, $\mu(\mathbf{0}, \pm 1) = 1/2$ and $\nu(\mathbf{1}_0, 0) = \nu(\mathbf{0}, 0) = 1/2$. In words, each step of the random walk corresponds to a "randomize-move-randomize" triple. We call the graph corresponding to this random walk the lamplighter graph and denote it by $\mathcal{G}$. The transition probabilities for this random walk have been well studied, which allow us to conclude the following result.

Theorem 29. Let o denote the identity element of the lamplighter group $\mathbb{Z}_2 \wr \mathbb{Z}$, start with initial unit mass δ_o at o, and let $p \in (0,1)$ be constant. The minimum number of toppling moves needed to transport mass p to distance at least n from o is

$$N_p(\mathcal{G}, B_n, \delta_o) = \exp(\Theta(n)).$$

Proof. In order to apply Theorem 28 we need to check that the two conditions of the theorem hold. First, $\mathcal{G}$ is 8-regular, so the first condition holds. The fact that the Green's function decays exponentially follows directly from[83] Theorems 1 and 2. $\square$

3.2.10 Open problems

- **Connections to maximum overhang problems.** Paterson et al.[74] studied the controlled diffusion problem on $\mathbb{Z}$ due to its connections with the maximum overhang problem in one dimension: how far can a stack of n identical blocks be made to hang over the edge of a table?

  The answer was widely believed to be of order $\log(n)$, by considering harmonic stacks in which n unit length blocks are placed one on top of the other, with the i^{th} block from the top extending by $1/(2i)$ beyond the block below it. This construction has an overhang of $\sum_{i=1}^{n} 1/(2i) \sim \frac{1}{2}\ln(n)$.

  However, Paterson and Zwick showed that this belief is false, by constructing an example with overhang on the order of $n^{1/3}$[75]. Subsequently, Paterson et al. showed that this is best possible up to a constant factor[74]. The authors proved this result by connecting the overhang problem to the controlled diffusion problem on $\mathbb{Z}$.

  This connection naturally leads to the following question: are the results presented in this section relevant for maximum overhang problems in higher dimensions?

- **Effectiveness of the greedy algorithm.** Under what circumstances is the greedy algorithm (approximately) optimal?

- **Small mass asymptotics.** What is the dependence of $N_p(G, B_n, o)$ on p as $p \to 0$?

4

Stochastic block model

This chapter is based on paper[37].

The stochastic block model is a widely studied model of community detection in random graphs, introduced by[46]. A simple description of the model is as follows: we start with n vertices, divided into two or more communities, then add edges independently at random, with probabilities depending on which communities the endpoints belong to. The algorithmic task is then to infer the communities from the graph structure.

A different class of models of random computational problems with planted solutions is that of planted satisfiability problems: we start with an assignment

σ to n Boolean variables and then choose clauses independently at random that are satisfied by σ. The task is to recover σ given the random formula. A closely related problem is that of recovering the planted assignment in[41]'s one-way function, see Section 4.0.4.

A priori, the stochastic block model and planted satisfiability may seem only tangentially related. Nevertheless, two observations reveal a strong connection:

1. Planted satisfiability can be viewed as a k-uniform hypergraph stochastic block model, with the set of $2n$ Boolean literals partitioned into two communities of true and false literals under the planted assignment, and clauses represented as hyperedges.

2. [31] gave a general algorithm for a unified model of planted satisfiability problems which reduces a random formula with a planted assignment to a bipartite stochastic block model with planted partitions in each of the two parts.

The bipartite stochastic block model in[31] has the distinctive feature that the two sides of the bipartition are extremely unbalanced; in reducing from a planted k-satisfiability problem on n variables, one side is of size $\Theta(n)$ while the other can be as large as $\Theta(n^{k-1})$.

We study this bipartite block model in detail, first locating a sharp threshold for detection and then studying the performance of spectral algorithms.

Our main contributions are the following:

1. When the ratio of the sizes of the two parts diverge, we locate a sharp

threshold below which detection is impossible and above which an efficient algorithm succeeds (Theorems 30 and 31). The proof of impossibility follows that of[69] in the stochastic block model, with the change that we couple the graph to a broadcast model on a two-type Poisson Galton-Watson tree. The algorithm we propose involves a reduction to the stochastic block model and the algorithms of[66,68].

2. We next consider spectral algorithms and show that computing the singular value decomposition (SVD) of the biadjacency matrix M of the model can succeed in recovering the planted partition even when the norm of the 'signal', $\|\mathbb{E}M\|$, is much smaller than the norm of the 'noise', $\|M - \mathbb{E}M\|$ (Theorem 32).

3. We show that at a sparser density, the SVD fails due to a localization phenomenon in the singular vectors: almost all of the weight of the top singular vectors is concentrated on a vanishing fraction of coordinates (Theorem 33).

4. We propose a modification of the SVD algorithm, Diagonal Deletion SVD, that succeeds at a sparser density still, far below the failure of the SVD (Theorem 32).

5. We apply the first algorithm to planted hypergraph partition and planted satisfiability problems to find the best known general bounds on the density at which the planted partition or assignment can be recovered efficiently (Theorem 34).

4.0.1 The model and main results

4.0.1.0.1 The bipartite stochastic block model

Fix parameters $\delta \in [0, 2]$, $n_1 \leq n_2$, and $p \in [0, 1/2]$. Then we define the bipartite stochastic block model as follows:

- Take two vertex sets V_1, V_2, with $|V_1| = n_1$, $|V_2| = n_2$.
- Assign labels '+' and '-' independently with probability 1/2 to each vertex in V_1 and V_2. Let $\sigma \in \{\pm 1\}^{n_1}$ denote the labels of the vertices in V_1 and $\tau \in \{\pm 1\}^{n_2}$ denote the labels of V_2.
- Add edges independently at random between V_1 and V_2 as follows: for $u \in V_1, v \in V_2$ with $\sigma(u) = \tau(v)$, add the edge (u, v) with probability δp; for $\sigma(u) \neq \tau(v)$, add (u, v) with probability $(2 - \delta)p$.

Algorithmic task: Determine the labels of the vertices given the bipartite graph, and do so with an efficient algorithm at the smallest possible edge density p.

4.0.1.0.2 Preliminaries and assumptions

In the application to planted satisfiability, it suffices to recover σ, the partition of the smaller vertex set, V_1, and so we focus on that task here; we will accomplish that task even when the number of edges is much smaller than the size of V_2. For a planted k-SAT problem or k-uniform hypergraph partitioning problem on n variables or vertices, the reduction gives vertex sets of size $n_1 = \Theta(n), n_2 = \Theta(n^{k-1})$, and so the relevant cases are extremely unbalanced.

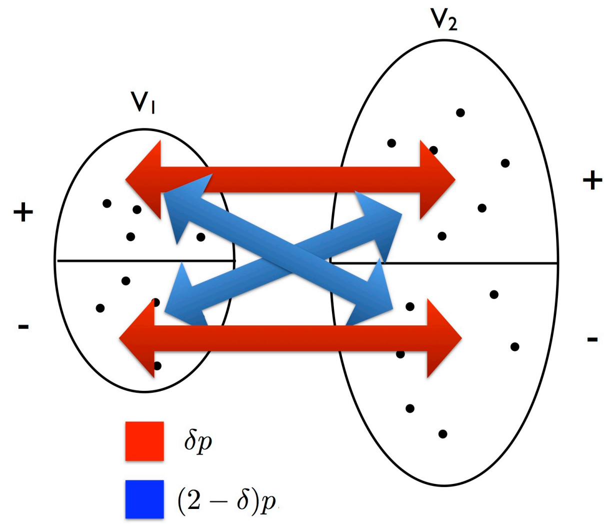


Figure 4.1: Bipartite stochastic block model on V_1 and V_2. Red edges are added with probability δp and blue edges are added with probability $(2-\delta)p$.

We will say that an algorithm *detects* the partition if for some fixed $\epsilon > 0$, independent of n_1, whp it returns an ϵ-correlated partition, i.e. a partition that agrees with σ on a $(1/2+\epsilon)$-fraction of vertices in V_1 (again, up to the sign of σ).

We will say an algorithm *recovers* the partition of V_1 if whp the algorithm returns a partition that agrees with σ on $1-o(1)$ fraction of vertices in V_1. Note that agreement is up to sign as σ and $-\sigma$ give the same partition.

4.0.1.1 Optimal algorithms for detection

On the basis of heuristic analysis of the belief propagation algorithm,[24] made the striking conjecture that in the two part stochastic block model, with interior edge probability a/n, crossing edge probability b/n, there is a *sharp threshold* for detection: for $(a-b)^2 > 2(a+b)$ detection can be achieved with an efficient algorithm, while for $(a-b)^2 \leq 2(a+b)$, detection is impossible for any algorithm. This conjecture was proved by[69,68] and[66].

Our first result is an analogous sharp threshold for detection in the bipartite stochastic block model at $p = (\delta-1)^{-2}(n_1 n_2)^{-1/2}$, with an algorithm based on a reduction to the SBM, and a lower bound based on a connection with the non-reconstruction of a broadcast process on a tree associated to a two-type Galton Watson branching process (analogous to the proof for the SBM[69] which used a single-type Galton Watson process).

Algorithm: SBM Reduction.

1. Construct a graph G' on the vertex set V_1 by joining u and w if they are both connected to the same vertex $v \in V_2$ and v has degree exactly 2.
2. Randomly sparsify the graph (as detailed in Section 4.0.6).
3. Apply an optimal algorithm for detection in the SBM from[66,68,17] to partition V_1.

Theorem 30. Let $\delta \in [0, 2] \setminus \{1\}$ be fixed and $n_2 = \omega(n_1)$. Then there is a polynomial-time algorithm that detects the partition $V_1 = A_1 \cup B_1$ whp if

$$p > \frac{1+\epsilon}{(\delta-1)^2\sqrt{n_1 n_2}}$$

for any fixed $\epsilon > 0$.

Theorem 31. On the other hand, if $n_2 \geq n_1$ and

$$p \leq \frac{1}{(\delta-1)^2\sqrt{n_1 n_2}},$$

then no algorithm can detect the partition whp.

Note that for $p \leq \frac{1}{\sqrt{n_1 n_2}}$ it is clear that detection is impossible: whp there is no giant component in the graph. The content of Theorem 31 is finding the sharp dependence on δ.

4.0.2 Spectral algorithms

One common approach to graph partitioning is spectral: compute eigenvectors or singular vectors of an appropriate matrix and round the vector(s) to partition the vertex set. In our setting, we can take the $n_1 \times n_2$ rectangular biadjacency matrix M, with rows and columns indexed by the vertices of V_1 and V_2 respectively, with a 1 in the entry (u, v) if the edge (u, v) is present, and a 0 otherwise. The matrix M has independent entries that are 1 with probability δp or $(2-\delta)p$ depending on the label of u and v and 0 otherwise.

Algorithm: Singular Value Decomposition.

1. Compute the left singular vector of M corresponding to the second largest singular value.
2. Round the singular vector to a vector $z \in \{\pm 1\}^{n_1}$ by taking the sign of each entry.

A typical analysis of spectral algorithms requires that the second largest eigenvalue or singular value of the expectation matrix $\mathbb{E}M$ is much larger than the spectral norm of the noise matrix, $(M - \mathbb{E}M)$. But here we have $\|M - \mathbb{E}M\| = \tilde{\Theta}(\sqrt{n_2 p})$, which is in fact much larger than $\lambda_2(\mathbb{E}M) = \Theta(p\sqrt{n_1 n_2})$ when $p = o(n_1^{-1})$. Does this doom the spectral approach at lower densities?

Question 1. For what values of $p = p(n_1, n_2)$ is the singular value decomposition (SVD) of M correlated with the vector σ indicating the partition of V_1?

In particular, this question was asked by[31]. We show that there are two thresholds, both well below $p = n_1^{-1}$: at $p = \tilde{\Omega}(n_1^{-2/3} n_2^{-1/3})$ the second singular vector of M is correlated with the partition of V_1, but below this density, it is uncorrelated with the partition, and in fact localized. Nevertheless, we give a simple spectral algorithm based on modifications of M that matches the bound $p = \tilde{O}((n_1 n_2)^{-1/2})$ achieved with subsampling by[31]. In the case of very unbalanced sizes, in particular in the applications noted above, these thresholds can differ by a polynomial factor in n_1.

Algorithm: Diagonal Deletion SVD.

1. Let $B = MM^T - \text{diag}(MM^T)$ (set the diagonal entries of MM^T to 0).
2. Compute the second eigenvector of B.
3. Round the eigenvector to a vector $z \in \{\pm 1\}^{n_1}$ by taking the sign of each entry.

Our results locate two different thresholds for spectral algorithms for the bipartite block model: while the usual SVD is only effective with $p = \tilde{\Omega}(n_1^{-2/3} n_2^{-1/3})$, the modified diagonal deletion algorithm is effective already at $p = \tilde{\Omega}(n_1^{-1/2} n_2^{-1/2})$, which is optimal up to logarithmic factors. In particular, when $n_1 = n, n_2 = n^{k-1}$ for some $k \geq 3$, as in the application above, these thresholds are separated by a polynomial factor in n.

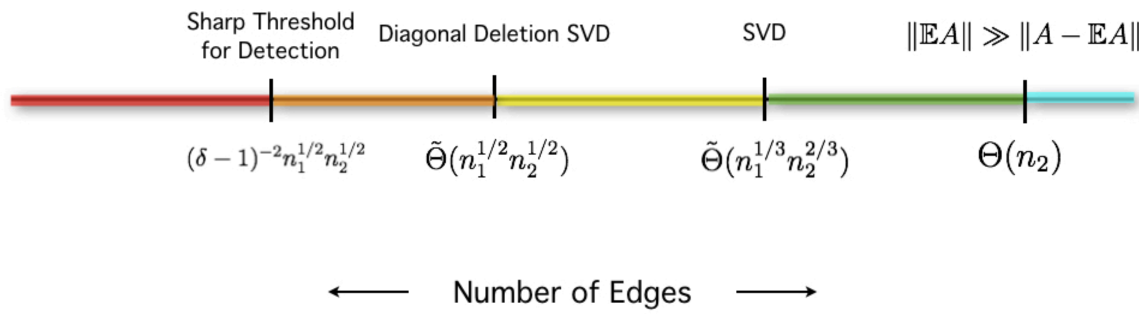


Figure 4.2: Main theorems illustrated.

First we give positive results for recovery using the two spectral algorithms.

Theorem 32. Let $n_2 \geq n_1 \log^4 n_1$, with $n_1 \to \infty$. Let $\delta \in [0, 2] \setminus \{1\}$ be fixed with respect to n_1, n_2. Then there exists a universal constant $C > 0$ so that

1. If $p = C(n_1 n_2)^{-1/2} \log n_1$, then whp the diagonal deletion SVD algorithm recovers the partition $V_1 = A_1 \cup B_1$.

2. If $p = C n_1^{-2/3} n_2^{-1/3} \log n_1$, then whp the unmodified SVD algorithm recovers the partition.

Next we show that below the recovery threshold for the SVD, the top left singular vectors are in fact *localized*: they have nearly all of their mass on a vanishingly small fraction of coordinates.

Theorem 33.

Let $n_2 \geq n_1 \log^4 n_1$. For any constant $c > 0$, let $p = cn_1^{-2/3}n_2^{-1/3}$, $t \leq n_1^{1/3}$, and $r = n_1/\log n_1$. Let $\overline{\sigma} = \sigma/\sqrt{n_1}$, and $v_1, v_2, \dots v_t$ be the top t left unit-norm singular vectors of M.

Then, whp, there exists a set $S \subset \{1, \dots, n_1\}$ of coordinates, $|S| \leq r$, so that for all $1 \leq i \leq t$, there exists a unit vector u_i supported on S so that

$$\|v_i - u_i\| = o(1).$$

That is, each of the first t singular vectors has nearly all of its weight on the coordinates in S. In particular, this implies that for all $1 \leq i \leq t$, v_i is asymptotically uncorrelated with the planted partition:

$$|\overline{\sigma} \cdot v_i| = o(1).$$

One point of interest in Theorem 33 is that in this case of a random biadjacency matrix of unbalanced dimension, the localization and delocalization of the singular vectors can be understood and analyzed in a simple manner, in contrast to the more delicate phenomenon for random square adjacency matrices.

Our techniques use bounds on the norms of random matrices and eigenvector perturbation theorems, applied to carefully chosen decompositions of the matrices of interest. In particular, our proof technique suggested the Diagonal Deletion SVD, which proved much more effective than the usual SVD algorithm on these unbalanced bipartite block models, and has the advantage over more sophisticated approaches of being extremely simple to describe and implement. We believe it may prove effective in many other settings.

Under what conditions might we expect the Diagonal Deletion SVD outperform the usual SVD? The SVD is a central algorithm in statistics, machine learning, and computer science, and so any general improvement would be useful. The bipartite block model addressed here has two distinctive characteristics: the dimensions of the matrix M are extremely unbalanced, and the entries are very sparse Bernoulli random variables, a distribution whose fourth moment is much larger than the square of its second moment. These two facts together lead to the phenomenon of multiple spectral thresholds and the outperformance of the SVD by the Diagonal Deletion SVD. Under both of these conditions we expect dramatic improvement by using diagonal deletion, while under one or the other condition, we expect mild improvement. We expect diagonal deletion will be effective in the more general setting of recovering a low-rank matrix in the presence of random noise, beyond our setting of adjacency matrices of graphs.

4.0.3 Planted k-SAT and hypergraph partitioning

[31] reduce three planted problems to solving the bipartite block model: planted hypergraph partitioning, planted random k-SAT, and Goldreich's planted CSP. We describe the reduction here and calculate the density at which our algorithm can detect the planted solution by solving the resulting bipartite block model.

We state the general model in terms of hypergraph partitioning first.

4.0.3.0.1 Planted hypergraph partitioning

Fix a function $Q : \{\pm 1\}^k \to [0, 1]$ so that $\sum_{x \in \{\pm 1\}^k} Q(x) = 1$. Fix parameters n and $p \in (0, 1)$ so that

$\max_x Q(x)2^k p \leq 1$. Then we define the planted k-uniform hypergraph partitioning model as follows:

- Take a vertex set V of size n.
- Assign labels '+' and '-' independently with probability 1/2 to each vertex in V. Let $\sigma \in \{\pm 1\}^n$ denote the labels of the vertices.
- Add (ordered) k-uniform hyperedges independently at random according to the distribution

  $$\Pr(e) = 2^k p \cdot Q(\sigma(e))$$

  where $\sigma(e)$ is the evaluation of σ on the vertices in e.

Algorithmic task: Determine the labels of the vertices given the hypergraph, and do so with an efficient algorithm at the smallest possible edge density p.

Usually Q will be symmetric in the sense that $Q(x)$ depends only on the number of $+1$'s in the vector x, and in this case we can view hyperedges as unordered. We assume that Q is not identically 2^{-k} as this distribution would simply be uniform and the planted partition would not be evident.

Planted k-satisfiability is defined similarly: we fix an assignment σ to n Boolean variables which induces a partition of the set of $2n$ literals (Boolean variables and their negations) into true and false literals. Then we add k-clauses independently at random, with probability proportional to the evaluation of Q on the k literals of the clause.

Planting distributions for the above problems are classified by their *distribution complexity*, $r = \min_{S \neq \emptyset}\{|S| : \hat{Q}(S) \neq 0\}$, where $\hat{Q}(S)$ is the discrete Fourier coefficient of Q corresponding to the subset $S \subseteq [k]$. This is an integer between 1 and k, where k is the uniformity of the hyperedges or clauses.

A consequence of Theorem 30 is the following:

Theorem 34. There is an efficient algorithm to detect the planted partition in the random k-uniform hypergraph partitioning problem, with planting function Q, when

$$p > (1+\epsilon) \min_{S \subseteq [k]} \frac{1}{2^{2k}\hat{Q}(S)^2 n^{k-|S|/2}}$$

for any fixed $\epsilon > 0$. Similarly, in the planted k-satisfiability model with planting function Q, there is an efficient algorithm to detect the planted assignment when

$$p > (1+\epsilon) \min_{S \subseteq [k]} \frac{1}{2^{2k}\hat{Q}(S)^2 (2n)^{k-|S|/2}}.$$

In both cases, if the distribution complexity of Q is at least 3, we can achieve full recovery at the given density.

Proof. Suppose Q has distribution complexity r. Fix a set $S \subseteq [k]$ with $\hat{Q}(S) \neq 0$, and $|S| = r$. The first step of the reduction of[31] transforms each k-uniform hyperedge into an r-uniform hyperedge by selecting the vertices indicated by the set S. Then a bipartite block model is constructed on vertex sets V_1, V_2, with V_1 the set of all vertices in the hypergraph (or literals in the formula), and V_2 the set of all $(r-1)$-tuples of vertices or literals. An edge is added by taking each r-uniform edge and splitting it randomly into sets of size 1 and $r-1$ and joining the associated vertices in V_1 and V_2. The parameters in our model are $n_1 = n$ and $n_2 \sim n^{r-1}$ (considering ordered $(r-1)$-tuples of vertices or literals).

These edges appear with probabilities that depend on the parity of the number of vertices on one side of the original partition in the joined sets, exactly the bipartite block model addressed in this section; the parameter δ in the model is given by $\delta = 1 + 2^k \hat{Q}(S)$ (see Lemma 1 of[31]). Theorems 30 then states that detection in the resulting block model exhibits a sharp threshold at edge density p^* , with $p^* = \frac{1}{2^{2k}\hat{Q}(S)^2 n^{k-r/2}}$. The difference in bounds in Theorem 34 is due to the two models having n vertices and $2n$ literals respectively.

To go from an ϵ-correlated partition to full recovery, if $r \geq 3$, we can appeal to Theorem 2 of[16] and achieve full recovery using only a linear number of additional hyperedges or clauses, which is lower order than the $\Theta(n^{r/2})$ used by our algorithm. □

Note that Theorem 31 says that no further improvement can be gained by analyzing this particular reduction to a bipartite stochastic block model.

There is some evidence that up to constant factors in the clause or hyperedge density, there may be no better efficient algorithms[72,30], unless the constraints induce a consistent system of linear equations. But in the spirit of[24], we can ask if there is in fact a sharp threshold for detection of planted solutions in these models. In one special case, such sharp thresholds have been conjectured:[54] have conjectured threshold densities based on fixed points of belief propagation equations. The planted k-SAT distributions covered, however, are only those with distribution complexity $r = 2$: those that are known to be solvable with a linear number of clauses. We ask if there are sharp thresholds for detection in the general case, and in particular for those distributions with distribution complexity $r \geq 3$ that cannot be solved by Gaussian elimination. In particular, in the case of the parity distribution we conjecture that there is a sharp threshold for detection.

Conjecture 4. Partition a set of n vertices at random into sets A, B. Add k-uniform hyperedges independently at random with probability δp if the number of vertices in the edge from A is even and $(2-\delta)p$ if the number of vertices from A is odd. Then for any $\delta \in (0, 2)$ there is a constant c_δ so that $p = c_\delta n^{-k/2}$ is a sharp threshold for detection of the planted partition by an efficient algorithm. That is, if $p > (1+\epsilon)c_\delta n^{-k/2}$, then there is a polynomial-time algorithm that detects the partition whp, and if $p \leq c_\delta n^{-k/2}$ then no polynomial-time algorithm can detect the partition whp.

This is a generalization to hypergraphs of the SBM conjecture of[24]; the $k = 2$ parity distribution is that of the stochastic block model. We do not venture a guess as to the precise constant c_δ, but even a heuristic as to what the constant might be would be very interesting.

4.0.4 Relation to Goldreich's generator

[41]'s pseudorandom generator or one-way function can be viewed as a variant of planted satisfiability. Fix an assignment σ to n Boolean variables, and fix a predicate $P : \{\pm 1\}^k \to \{0, 1\}$. Now choose m k-tuples of variables *uniformly*

at random, and label the k-tuple with the evaluation of P on the tuple with the Boolean values given by σ. In essence this generates a uniformly random k-uniform hypergraph with labels that depend on the planted assignment and the fixed predicate P. The task is to recover σ given this labeled hypergraph. The algorithm we describe above will work in this setting by simply discarding all hyperedges labeled 0 and working with the remaining hypergraph.

4.0.5 Related work

The stochastic block model has been a source of considerable recent interest. There are many algorithmic approaches to the problem, including algorithms based on maximum-likelihood methods[86], belief propagation[24], spectral methods[67], modularity maximization[13], and combinatorial methods[18],[28],[50],[21].[20] gave the first algorithm to detect partitions in the sparse, constant average degree regime.[24] conjectured the precise achievable constant and subsequent algorithms[66,68,17,2] achieved this bound. Sharp thresholds for full recovery (as opposed to detection) have been found by[70,1,44].

[16] used ideas for reconstructing assignments to random 3-SAT formulas in the planted 3-SAT model to show that Goldreich's construction of a one-way function in[41] is not secure when the predicate correlates with either one or two of its inputs. For more on Goldreich's PRG from a cryptographic perspective see the survey of[5].

[31] gave an algorithm to recover the partition of V_1 in the bipartite stochastic block model to solve instances of planted random k-SAT and planted hyper-

graph partitioning using subsampled power iteration.

A key part of our analysis relies on looking at an auxiliary graph on V_1 with edges between vertices which share a common neighbor; this is known as the one-mode projection of a bipartite graph:[90] give an approach to recommendation systems using a weighted version of the one-mode projection. One-mode projections are implicitly used in studying collaboration networks, for example in[71]'s analysis of scientific collaboration networks.[55] defined a general model of bipartite block models, and propose a community detection algorithm that does not use one-mode projection.

The behavior of the singular vectors of a low rank rectangular matrix plus a noise matrix was studied by[9]. The setting there is different: the ratio between n_1 and n_2 converges, and the entries of the noise matrix are mean 0 variance 1.

[19] and[45] both consider the case of recovering a planted submatrix with elevated mean in a random rectangular Gaussian matrix.

4.0.5.0.1 Notation All asymptotics are as $n_1 \to \infty$, so for example, 'E occurs whp' means $\lim_{n_1 \to \infty} \Pr(E) = 1$. We write $f(n_1) = \tilde{O}(g(n_1))$ and $f(n_1) = \tilde{\Omega}(g(n_1))$ if there exist constants C, c so that $f(n_1) \leq C \log^c(n_1) \cdot g(n_1)$ and $f(n_1) \geq g(n_1)/(C \log^c(n_1))$ respectively. For a vector, $\|v\|$ denotes the l_2 norm. For a matrix, $\|A\|$ denotes the spectral norm, i.e. the largest singular value (or largest eigenvalue in absolute value for a square matrix). For ease of reading, C will always denote an absolute constant, but the value may change during the course of the proofs.

4.0.6 Theorem 30: detection

In this section we sketch the proof of Theorem 30, giving an optimal algorithm for detection in the bipartite stochastic block model when $n_2 = \omega(n_1)$. The main idea of the proof is that almost all of the information in the bipartite block model is in the subgraph induced by V_1 and the vertices of degree two in V_2. From this induced subgraph of the bipartite graph we form a graph G' on V_1 by replacing each path of length two from V_1 to V_2 back to V_1 with a single edge between the two endpoints in V_1. We then apply an algorithm from[66,68], or[17] to detect the partition.

Fix $\epsilon > 0$. Given an instance G of the bipartite block model with

$$p = (1+\epsilon)(\delta-1)^{-2}(n_1 n_2)^{-1/2},$$

we reduce to a graph G' on V_1 as follows:

- Sort V_2 according to degrees and remove any vertices (with their accompanying edges) which are not of degree 2.
- We now have a union of 2-edge paths from vertices in V_1 to vertices in V_2 and back to vertices in V_1. Create a multi-set of edges $\mathcal{E}$ on V_1 by replacing each 2-path $u - v - w$ by the edge (u, w).
- Choose N from the distribution Poisson$((1+\epsilon)(\delta-1)^{-4}n_1/2)$.
- If $N > |\mathcal{E}|$, then stop and output 'failure'. Otherwise, select N edges uniformly at random from $\mathcal{E}$ to form the graph G' on V_1, replacing any edge of multiplicity greater than one with a single edge.

- Apply an SBM algorithm to G' to partition V_1.

From the construction above, conditioned on σ the distribution of G' is that of the stochastic block model on V_1 with partition σ: each edge interior to the partition is present with probability a/n_1, each crossing edge with probability b/n_1, and all edges are independent.

For σ such that $\beta_1 = o(n^{-1/3})$, we have

$$a = \frac{(1+\epsilon)(2-2\delta+\delta^2)}{(\delta-1)^4}(1+o(1))$$
$$b = \frac{(1+\epsilon)(2\delta-\delta^2)}{(\delta-1)^4}(1+o(1))$$

For these values of a and b the condition for detection in the SBM, $(a-b)^2 \geq (1+\epsilon)2(a+b)$ is satisfied and so whp the algorithms from[66,68,17] will find a partition that agrees with σ on $1/2+\epsilon'$ fraction of vertices.

4.0.7 Theorem 31: impossibility

The proof of impossibility below the threshold $(a-b)^2 = 2(a+b)$ in[69] proceeds by showing that the $\log n$ depth neighborhood of a vertex ρ, along with the accompanying labels, can be coupled to a binary symmetric broadcast model on a Poisson Galton-Watson tree. In this model, it was shown by[29] that reconstruction, recovering the label of the root given the labels at depth R of the tree, is impossible as $R \to \infty$, for the corresponding parameter values (the critical case was shown by[76]).

In the binary symmetric broadcast model, the root of a tree is labeled with a uniformly random label $+1$ or -1, and then each child takes its parent's label with probability $1-\eta$ and the opposite label with probability η, independently over all of the parent's children. The process continues in each successive generation of the tree.

The criteria for non-reconstruction can be stated as $(1-2\eta)^2 B \leq 1$, where B is the branching number of the tree T. The branching number is $B = p_c(T)^{-1}$, where p_c is the critical probability for bond percolation on T (see[63] for more on the branching number).

Assume first that $n_2 \sim cn_1$ for some constant c, and that $p = d/n_1$. Then there is a natural multitype Poisson branching process that we can associate to the bipartite block model: nodes of type 1, corresponding to vertices in V_1, have a Poisson(cd) number of children of type 2; nodes of type 2, corresponding to vertices in V_2, have a Poisson(d) number of children of type 1. The branching number of this distribution on trees is $\sqrt{c} \cdot d$, an easy calculation by reducing to a one-type Galton Watson process by combining two generations into one. Transferring the block model labeling to the branching process gives $\eta = \delta/2$, and so the threshold for reconstruction is given by

$$(\delta - 1)^2 \sqrt{c} d \leq 1$$

or in other words,

$$p \leq \frac{1}{(\delta-1)^2 \sqrt{n_1 n_2}}$$

exactly the threshold in Theorem 31. In fact, in this case the proof from[69] can be carried out in essentially the exact same way in our setting.

Now take $n_2 = \omega(n_1)$. A complication arises: the distribution of the number of neighbors of a node of type 1 does not converge (its mean is $n_2 p \to \infty$), and the distribution of the number of neighbors of a node of type 2 converges to a delta mass at 0. But this can be fixed by ignoring the vertices in V_2 of degree 0 and 1. Now we explore from a vertex $\rho \in V_1$, but discard any vertices from V_2 that do not have a second neighbor. We denote by $\hat{G}$ the subgraph of G induced by V_1 and the vertices of V_2 of degree at least 2. Let T be the branching process associated to this modified graph: nodes of type 1 have Poisson(d^2) neighbors of type 2, and nodes of type 2 have exactly 1 neighbor of type 1, where here $p = d/\sqrt{n_1 n_2}$. The branching number of this process is d, and the reconstruction threshold is $(\delta - 1)^2 d \leq 1$, again giving the threshold $p \leq \frac{1}{(\delta-1)^2\sqrt{n_1 n_2}}$, as required.

As in[69], the proof of impossibility shows the stronger statement that conditioned on the label of a fixed vertex $w \in V_1$ and the graph G, the variance of the label of another fixed vertex ρ tends to 1 as $n_1 \to \infty$. The proof of this fact has two main ingredients: showing that the depth R neighborhood of a vertex ρ in the bipartite block model (with vertices of degree 0 and 1 in V_2 removed) can be coupled with the branching process described above, and showing that conditioned on the labels on the boundary of the neighborhood, the label of ρ is asymptotically independent of the rest of the graph and the labels outside of the neighborhood.

4.0.8 Theorem 32: Recovery

We follow a similar framework in proving both parts of Theorem 32. Recalling M to be the adjacency matrix, let $B = MM^T - \text{diag}(MM^T)$ and $D_V = \text{diag}(MM^T)$.

A simple computation shows that the second eigenvector of $\mathbb{E}B$ is the vector σ that we wish to recover; we will consider the different perturbations of $\mathbb{E}B$ that arise with the three spectral algorithms and show that at the respective thresholds, the second eigenvector of the resulting matrix is close to σ. To analyze the diagonal deletion SVD, we must show that the second eigenvector of B is highly correlated with σ (the addition of a constant multiple of the identity matrix does not change the eigenvectors). The main step is to bound the spectral norm $\|B - \mathbb{E}B\|$. Since the entries of B are not independent, we will decompose B into a sequence of matrices based on subgraphs induced by vertices of a given degree in V_2. This (Lemma 24) is the most technical part of the work.

To analyze the unmodified SVD, we write $MM^T = \mathbb{E}B + (B - \mathbb{E}B) + \mathbb{E}D_V + (D_V - \mathbb{E}D_V)$. The left singular vectors of M are the eigenvectors of MM^T. $\mathbb{E}B$ has σ as its second eigenvector and $\mathbb{E}D_V$ is a multiples of the identity matrix and so adding it does not change the eigenvectors. As above we bound $\|B - \mathbb{E}B\|$ and what remains is showing that the difference of the matrix D_V with its expectation has small spectral norms at the respective thresholds; this involves simple bounds on the fluctuations of independent random variables.

We will assume that σ and τ assign $+1$ and -1 labels to an equal number of vertices; this allows for a clearer presentation, but is not necessary to the argument. We will treat σ and τ as unknown but fixed, and so expectations and probabilities will all be conditioned on the labelings.

The main technical lemma is the following:

Lemma 24. Define B, D_V as above. Assume n_1, n_2, and p are as in Theorem 32. Then there exists an absolute constant C so that

1. $\mathbb{E}B = \lambda_1 J/n_1 + \lambda_2 \sigma\sigma^T/n_1$, with $\lambda_1 = n_1 n_2 p^2$ and $\lambda_2 = (\delta-1)^2 n_1 n_2 p^2$, where J is the all ones $n_1 \times n_1$ matrix.

2. For $p \geq n_1^{-1/2} n_2^{-1/2} \log n_1$, $\|B - \mathbb{E}B\| \leq C n_1^{1/2} n_2^{1/2} p$ whp.

3. $\mathbb{E}D_V$ is a multiple of the identity matrix.

4. For $p \geq n_1^{-2/3} n_2^{-1/3} \log n_1$, $\|D_V - \mathbb{E}D_V\| \leq C\sqrt{n_2 p \log n_1}$ whp.

We also will use the following lemma from[58] to round a unit vector with high correlation with σ to a ± 1 vector that denotes a partition:

Lemma 25 ([58])**.** For any $x \in \{-1, +1\}^n$ and $y \in \mathbb{R}^n$ with $\|y\| = 1$ we have

$$d(x, \operatorname{sign}(y)) \leq n \left\| \frac{x}{\sqrt{n}} - y \right\|^2,$$

where d represents the Hamming distance.

The next lemma is a classic eigenvector perturbation theorem. Denote by $P_A(S)$ the orthogonal projection onto the subspace spanned by the eigenvectors of A corresponding to those of its eigenvalues that lie in S.

Lemma 26 ([23])**.** Let A be an $n \times n$ symmetric matrix with $|\lambda_1| \geq |\lambda_2| \geq \dots,$ with $|\lambda_k| - |\lambda_{k+1}| \geq 2\delta$. Let B be a symmetric matrix with $\|B\| < \delta$. Let A_k and $(A+B)_k$ be the spaces spanned by the top k eigenvectors of the respective matrices. Then

$$\sin(A_k, (A+B)_k) = \|P_{A_k} - P_{(A+B)_k}\| \leq \frac{\|B\|}{\delta}$$

In particular, If $|\lambda_1| - |\lambda_2| \geq 2\delta$, $|\lambda_2| - |\lambda_3| \geq 2\delta$, $\|B\| < \delta$, and $e_2(A)$, $e_2(A+B)$ are the second (unit) eigenvectors of A and $A+B$, respectively, satisfying $e_2(A) \cdot e_2(A+B) \geq 0$, then $\|e_2(A) - e_2(A+B)\| \leq \frac{4\|B\|}{\delta}$.

Now using Lemmas 24, 25, and 26 we prove parts 1 and 2 of Theorem 32.

4.0.8.0.1 **Diagonal deletion SVD** Let $p \geq n_1^{-1/2} n_2^{-1/2} \log n_1$. Part 1 of Lemma 24 shows that if we had access to the second eigenvector of $\mathbb{E}B$, we would recover σ exactly. (The addition of a multiple of the identity matrix does not change the eigenvectors). Instead we have access to $B = \mathbb{E}B + (B - \mathbb{E}B)$, a noisy version of the matrix we want. We use a matrix perturbation inequality to show that the top eigenvectors of the noisy version are not too far from the original eigenvectors.

Let y_1 and y_2 be the top two eigenvectors of B, and $\hat{B}$ be the space spanned by y_1 and y_2, and $(\mathbb{E}B)_2$ the space spanned by the top two eigenvectors of $\mathbb{E}B$. Then Lemma 26 gives

$$\sin((\mathbb{E}B)_2, \hat{B}) \leq \frac{C\|B - \mathbb{E}B\|}{\lambda_2} \leq C \frac{n_1^{1/2} n_2^{1/2} p}{(\delta - 1)^2 n_1 n_2 p^2} = O\left(\frac{1}{\log n_1}\right)$$

where the inequality holds whp by Lemma 24. Assuming $\delta \in (0, 2)$, we use the particular case of Lemma 26 to show that $\|y_2 - \sigma/\sqrt{n_1}\| = O(\log^{-1} n_1)$. We round y_2 by signs to get z, and then apply Lemma 25 to show that whp the algorithm recovers $1 - o(1)$ fraction of the coordinates of σ. (If $\delta = 0$ or 2, then instead of taking the second eigenvector, we take the component of $\hat{B}$ perpendicular to the all ones vector and get the same result).

4.0.8.0.2 **The SVD** Let $p \geq n_1^{-2/3} n_2^{-1/3} \log n_1$. Let y_1 and y_2 be the top two left singular vectors of M, and M_2 be the space spanned by y_1 and y_2. y_1 and y_2 are the top two eigenvectors of $MM^T = B + D_V$. Again Lemma 26 gives that whp,

$$\sin((\mathbb{E}B)_2, M_2) \leq C \frac{\|B - \mathbb{E}B\| + \|D_V - \mathbb{E}D_V\|}{\lambda_2} \leq \frac{C_1 n_1^{1/2} n_2^{1/2} p + C_2 \sqrt{n_2 p \log n_1}}{(\delta - 1)^2 n_1 n_2 p^2} = O\left(\frac{1}{\log n_1}\right).$$

This gives $\|y_2 - \sigma/\sqrt{n_1}\| = O(\log^{-1} n_1)$, and shows that the SVD algorithm recovers σ whp. Note that in this case $\|D_V - \mathbb{E}D_V\| \gg \|B - \mathbb{E}B\|$. It is these fluctuations on the diagonal that explain the poor performance of the SVD and its need for a higher edge density for success.

4.0.9 Theorem 33: Failure of the vanilla SVD

Here we again use a matrix perturbation lemma, but in the opposite way: we will show that the 'noise matrix' $(D_V - \mathbb{E}D_V)$ has a large spectral norm (and an eigenvalue gap), and thus adding the 'signal matrix' approximately preserves the space spanned by the top eigenvalues. This shows that the top t eigenvectors of $B + D_V$ have almost all their weight on a small number of coordinates and is enough to conclude that they cannot be close to the planted vector σ.

The perturbation lemma we use is a generalization of the Davis-Kahan theorem found in[11].

Lemma 27 ([11])**.** Let A and B be $n \times n$ symmetric matrices with the eigenvalues of A ordered $\lambda_1 \geq \lambda_2 \geq \dots \lambda_n$. Suppose $r > k$, $\lambda_k - \lambda_r > 2\delta$, and $\|B\| \leq \delta$. Let A_r denote the subspace spanned by the first r eigenvectors of A and likewise for $(A+B)_k$. Then

$$\|P_{A_r^\perp} P_{(A+B)_k}\| \leq \frac{\|B\|}{\delta}.$$

In particular, if v_k is the k^{th} unit eigenvector of $(A+B)$, then there is some unit vector $u \in A_r$ so that

$$\|u - v_k\| \leq \frac{4\|B\|}{\delta}.$$

We also need to analyze the degrees of the vertices in V_1. The following lemma gives some basic information about the degree sequence:

Lemma 28. Let $d_1, \dots d_{n_1}$ be the sequence of degrees of vertices in V_1. Then there exist constants c_1, c_2, c_3 so that

1. The d_i's are independent and identically distributed, with distribution $d_i \sim \mathrm{Bin}(n_2/2, \delta p) + \mathrm{Bin}(n_2/2, (2-\delta)p)$.

2. $\mathbb{E} d_i = n_2 p$.

3. Whp, $\max_i d_i \le n_2 p + c_1\sqrt{n_2 p \log n_1}$.

4. Whp, $\left|\{i : d_i \ge n_2 p + c_2\sqrt{n_2 p \log n_1}\}\right| \ge n_1^{1/3}$.

5. Whp, $\left|\{i : d_i \ge n_2 p + c_3\sqrt{n_2 p \log\log n_1}\}\right| \le n_1/\log n_1$.

Now we can prove Theorem 33. Let $p = cn_1^{-2/3}n_2^{-1/3}$. The left singular vectors of M are the eigenvectors of $B + D_V$. Recall that D_V is a diagonal matrix with the ith entry the degree of the ith vertex of V_1. $\mathbb{E}D_V$ is therefore a multiple of the identity matrix, and so subtracting $\mathbb{E}D_V$ from $B + D_V$ does not change its eigenvectors. The standard basis vectors form an orthonormal set of eigenvectors of $D_V - \mathbb{E}D_V$.

For the constants c_2, c_3 in Lemma 28, let $\eta_1 = c_2\sqrt{n_2 p \log n_1}$ and $\eta_2 = c_3\sqrt{n_2 p \log\log n_1}$. Order the eigenvalues of $D_V - \mathbb{E}D_V$ as $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_n$ and let r be the smallest integer such that $\lambda_r < \eta_2$. Then we have $\lambda_i - \lambda_r \ge c\sqrt{n_2 p \log n_1}$ for all $1 \le i \le t$. From Lemma 28, $r \le n_1/\log n_1$.

We now bound

$$\|B\| \le \|\mathbb{E}B\| + \|B - \mathbb{E}B\| \le n_1 n_2 p^2 + Cn_1^{1/2}n_2^{1/2}p.$$

Now Lemma 27 says that if v_i is the ith eigenvector of $D_V - \mathbb{E}D_V + B$, then there is a vector u in the span of the first r eigenvectors of $D_V - \mathbb{E}D_V$ so that

$$\|v_i - u\| \le C\frac{n_1 n_2 p^2 + n_1^{1/2}n_2^{1/2}p}{\sqrt{n_2 p \log n_1}} = O\left(\frac{1}{\sqrt{\log n_1}}\right).$$

The span of the first r eigenvectors of $D_V - \mathbb{E}D_V$ is supported on only r coordinates, so u is far from $\overline{\sigma} = \sigma/\sqrt{n_1}$:

$$\|u - \overline{\sigma}\| \geq \sqrt{2 - 2\sqrt{r/n_1}} = \sqrt{2} - O(1/\sqrt{\log n_1}).$$

By the triangle inequality, v_i must also be far from $\overline{\sigma}$: $|v_i \cdot \overline{\sigma}| = O(1/\sqrt{\log n_1})$. This proves Theorem 33.

References

[1] Abbe, E., Bandeira, A. S., & Hall, G. (2016). Exact recovery in the stochastic block model. *Information Theory, IEEE Transactions on*, 62(1), 471–487.

[2] Abbe, E. & Sandon, C. (2015). Detection in the stochastic block model with multiple clusters: proof of the achievability conjectures, acyclic bp, and the information-computation gap. *arXiv preprint arXiv:1512.09080*.

[3] Angel, O. & Holroyd, A. E. (2011a). Recurrent rotor-router configurations. `arXiv:1101.2484`.

[4] Angel, O. & Holroyd, A. E. (2011b). Rotor walks on general trees. *SIAM J. Discrete Math.*, 25, 423–446. `arXiv:1009.4802`.

[5] Applebaum, B. (2013). Cryptographic hardness of random local functions–survey. In *Theory of Cryptography* (pp. 599–599). Springer.

[6] Bak, P. & Sneppen, K. (1993). Punctuated equilibrium and criticality in a simple model of evolution. *Phys. Rev. Lett.*, 71, 4083–4086.

[7] Bak, P., Tang, C., & Wiesenfeld, K. (1987). Self-organized criticality: An explanation of the $1/f$ noise. *Phys. Rev. Lett.*, 59, 381–384.

[8] Bampas, E., Gasieniec, L., Hanusse, N., Ilcinkas, D., Klasing, R., & Kosowski, A. (2009). Euler tour lock-in problem in the rotor-router model. In *Distributed Computing* (pp. 423–435). Springer.

[9] Benaych-Georges, F. & Nadakuditi, R. R. (2012). The singular values and vectors of low rank perturbations of large rectangular random matrices. *Journal of Multivariate Analysis*, 111, 120–135.

[10] Benjamini, I. & Peres, Y. (1994). Tree-indexed random walks on groups and first passage percolation. *Probability Theory and Related Fields*, 98(1), 91–112.

[11] Bhatia, R. (1997). *Matrix analysis*, volume 169. Springer Science & Business Media.

[12] Bhatt, S. N., Even, S., Greenberg, D. S., & Tayar, R. (2002). Traversing directed Eulerian mazes. *J. Graph Algorithms Appl.*, 6(2), 157–173.

[13] Bickel, P. J. & Chen, A. (2009). A nonparametric view of network models and newmanÐgirvan and other modularities. *Proceedings of the National Academy of Sciences*, 106(50), 21068–21073.

[14] Billingsley, P. (1995). *Probability and Measure*. New York, NY: Wiley, 3rd edition.

[15] Blachère, S., Haissinsky, P., & Mathieu, P. (2008). Asymptotic entropy and Green speed for random walks on countable groups. *Ann. Probab.*, 36(3), 1134–1152.

[16] Bogdanov, A. & Qiao, Y. (2009). On the security of goldreich's one-way function. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques* (pp. 392–405). Springer.

[17] Bordenave, C., Lelarge, M., & Massoulié, L. (2015). Non-backtracking spectrum of random graphs: community detection and non-regular ramanujan graphs. In *Foundations of Computer Science (FOCS), 2015 IEEE 56th Annual Symposium on* (pp. 1347–1357).: IEEE.

[18] Bui, T., Chaudhuri, S., Leighton, F., & Sipser, M. (1987). Graph bisection algorithms with good average case behavior. *Combinatorica*, 7(2), 171–191.

[19] Butucea, C., Ingster, Y. I., & Suslina, I. A. (2015). Sharp variable selection of a sparse submatrix in a high-dimensional noisy matrix. *ESAIM: Probability and Statistics*, 19, 115–134.

[20] Coja-Oghlan, A. (2010). Graph partitioning via adaptive spectral techniques. *Combinatorics, Probability & Computing*, 19(2), 227.

[21] Condon, A. & Karp, R. M. (2001). Algorithms for graph partitioning on the planted partition model. *Random Struct. Algorithms*, 18(2), 116–140.

[22] Cooper, J. N. & Spencer, J. (2006). Simulating a random walk with constant error. *Combinatorics, Probability and Computing*, 15(06), 815–822. arXiv:math/0402323.

[23] Davis, C. & Kahan, W. M. (1970). The rotation of eigenvectors by a perturbation. iii. *SIAM Journal on Numerical Analysis*, 7(1), 1–46.

[24] Decelle, A., Krzakala, F., Moore, C., & Zdeborová, L. (2011). Asymptotic analysis of the stochastic block model for modular networks and its algorithmic applications. *Physical Review E*, 84(6), 066106.

[25] Doerr, B., Friedrich, T., & Sauerwald, T. (2008). Quasirandom rumor spreading. In *Proc. 19th SODA* (pp. 773–781). San Francisco, CA, USA: SIAM.

[26] Doob, J. L. (1984). *Classical Potential Theory and Its Probabilistic Counterpart*. Grundlehren der mathematischen Wissenschaften. New York, Berlin, Heidelberg: Springer-Verlag.

[27] Drossel, B. & Schwabl, F. (1992). Self-organized critical forest-fire model. *Phys. Rev. Lett.*, 69, 1629–1632.

[28] Dyer, M. & Frieze, A. (1989). The solution of some random np-hard problems in polynomial expected time. *Journal of Algorithms*, 10(4), 451 – 489.

[29] Evans, W., Kenyon, C., Peres, Y., & Schulman, L. J. (2000). Broadcasting on trees and the ising model. *Annals of Applied Probability*, (pp. 410–433).

[30] Feldman, V., Perkins, W., & Vempala, S. (2015a). On the complexity of random satisfiability problems with planted solutions. In *STOC 2015: 47th Annual Symposium on the Theory of Computing*.

[31] Feldman, V., Perkins, W., & Vempala, S. (2015b). Subsampled power iteration: a new algorithm for block models and planted csp's. In *NIPS*.

[32] Florescu, L., Ganguly, S., Levine, L., & Peres, Y. Escape rates for rotor walks in $\mathbb{Z}^d$. *SIAM J. Discrete Math*, 28, 323–334.

[33] Florescu, L., Ganguly, S., Peres, Y., & Spencer, J. (2015). Heat diffusion with frozen boundary. *Journal of Statistical Physics*, 161, 521–531.

[34] Florescu, L., Levine, L., & Peres, Y. The range of a rotor walk (long version). arXiv:1408.5533.

[35] Florescu, L., Levine, L., & Peres, Y. (2016). The range of a rotor walk. *American Mathematical Monthly*.

[36] Florescu, L., Peres, Y., & Racz, M. Optimal control for diffusions on graphs.

[37] Florescu, L. & Perkins, W. (2016). Spectral thresholds in the bipartite stochastic block model. In *29th Annual Conference on Learning Theory* (pp. 943–959).

[38] Friedrich, T., Gairing, M., & Sauerwald, T. (2012). Quasirandom load balancing. *SIAM Journal on Computing*, 41(4), 747–771.

[39] Friedrich, T. & Sauerwald, T. (2010). The cover time of deterministic random walks. *The Electronic Journal of Combinatorics*, 17, R167. arXiv:1006.3430.

[40] Fukai, Y. & Uchiyama, K. (1996). Potential kernel for two-dimensional random walk. *The Annals of Probability*, 24(4), 1979–1992.

[41] Goldreich, O. (2000). Candidate one-way functions based on expander graphs. *IACR Cryptology ePrint Archive*, 2000, 63.

[42] Gravner, J. & Quastel, J. (2000). Internal dla and the stefan problem. *Ann. Probab.*, 28(4), 1528–1562.

[43] Grimmett, G. (1999). *Percolation*. Springer, second edition.

[44] Hajek, B., Wu, Y., & Xu, J. (2015a). Achieving exact cluster recovery threshold via semidefinite programming. In *Information Theory (ISIT), 2015 IEEE International Symposium on* (pp. 1442–1446).: IEEE.

[45] Hajek, B., Wu, Y., & Xu, J. (2015b). Submatrix localization via message passing. *arXiv preprint arXiv:1510.09219*.

[46] Holland, P. W., Laskey, K. B., & Leinhardt, S. (1983). Stochastic blockmodels: First steps. *Social Networks*, 5(2), 109–137.

[47] Holroyd, A. E., Levine, L., Mészáros, K., Peres, Y., Propp, J., & Wilson, D. B. (2008). Chip-firing and rotor-routing on directed graphs. In *In and out of equilibrium. 2*, volume 60 of *Progr. Probab.* (pp. 331–364). Basel: Birkhäuser. arXiv:0801.3306.

[48] Holroyd, A. E. & Propp, J. G. (2010). Rotor walks and Markov chains. *Algorithmic Probability and Combinatorics*, 520, 105–126. arXiv:0904.4507.

[49] Huss, W. & Sava, E. (2011). Rotor-router aggregation on the comb. *Electronic Journal of Combinatorics*, 18, P224. arXiv:1103.4797.

[50] Jerrum, M. & Sorkin, G. B. (1998). The metropolis algorithm for graph bisection. *Discrete Applied Mathematics*, 82(1â€“3), 155 – 175.

[51] Kager, W. & Levine, L. (2010). Rotor-router aggregation on the layered square lattice. *The Electronic Journal of Combinatorics*, 17(1), R152. arXiv:1003.4017.

[52] Kaimanovich, V. & Vershik, A. (1983). Random walks on discrete groups: boundary and entropy. *The Annals of Probability*, 11(3), 457–490.

[53] Kapri, R. & Dhar, D. (2009). Asymptotic shape of the region visited by an Eulerian walker. *Phys. Rev. E*, 80(5). arXiv:0906.5506.

[54] Krzakala, F., Mézard, M., & Zdeborová, L. (2014). Reweighted belief propagation and quiet planting for random k-sat. *Journal on Satisfiability, Boolean Modeling and Computation*, 8, 149–171.

[55] Larremore, D. B., Clauset, A., & Jacobs, A. Z. (2014). Efficiently inferring community structure in bipartite networks. *Physical Review E*, 90(1), 012805.

[56] Lawler, G. F. (1991). *Intersections of Random Walks*. Springer Science & Business Media.

[57] Lawler, G. F. (1996). *Intersections of random walks*. Birkhäuser Boston.

[58] Lelarge, M., Massoulié, L., & Xu, J. (2013). Reconstruction in the labeled stochastic block model. In *Information Theory Workshop (ITW), 2013 IEEE* (pp. 1–5).: IEEE.

[59] Levine, L. & Peres, Y. (2009a). Strong spherical asymptotics for rotor-router aggregation and the divisible sandpile. *Potential Analysis*, 30, 1–27. arXiv:0704.0688.

[60] Levine, L. & Peres, Y. (2009b). Strong Spherical Asymptotics for Rotor-Router Aggregation and the Divisible Sandpile. *Potential Analysis*, 30, 1–27.

[61] Li, G., Reis, S. D. S., Moreira, A. A., Havlin, S., Stanley, H. E., & Andrade, J. S. (2010). Towards design principles for optimal transport networks. *Phys. Rev. Lett.*, 104, 018701.

[62] Linetsky, V. (2005). On the transition densities for reflected diffusions. *Advances in Applied Probability*, 37, 435–460.

[63] Lyons, R. (1990). Random walks and percolation on trees. *The annals of Probability*, (pp. 931–958).

[64] Lyons, R., Pemantle, R., & Peres, Y. (1995). Ergodic theory on Galton-Watson trees: Speed of random walk and dimension of harmonic measure. *Ergodic Theory and Dynamical Systems*, 15, 593–619.

[65] Lyons, R. & Peres, Y. (2016). *Probability on Trees and Networks.* New York: Cambridge University Press. Available at http://pages.iu.edu/~rdlyons/.

[66] Massoulié, L. (2014). Community detection thresholds and the weak ramanujan property. In *STOC 2014: 46th Annual Symposium on the Theory of Computing* (pp. 1–10).

[67] McSherry, F. (2001). Spectral partitioning of random graphs. In *Foundations of Computer Science, 2001. Proceedings. 42nd IEEE Symposium on* (pp. 529–537).: IEEE.

[68] Mossel, E., Neeman, J., & Sly, A. (2013). A proof of the block model threshold conjecture. *arXiv preprint arXiv:1311.4115.*

[69] Mossel, E., Neeman, J., & Sly, A. (2014). Reconstruction and estimation in the planted partition model. *Probability Theory and Related Fields*, (pp. 1–31).

[70] Mossel, E., Neeman, J., & Sly, A. (2015). Consistency thresholds for the planted bisection model. In *Proceedings of the Forty-Seventh Annual ACM on Symposium on Theory of Computing* (pp. 69–75).: ACM.

[71] Newman, M. E. (2001). Scientific collaboration networks. i. network construction and fundamental results. *Physical review E*, 64(1), 016131.

[72] O'Donnell, R. & Witmer, D. (2014). Goldreich's prg: Evidence for near-optimal polynomial stretch. In *CCC*.

[73] Pach, J. & Tardos, G. (2013). The range of a random walk on a comb. *The Electronic Journal of Combinatorics*, 20(3), P59. arXiv:1309.6360.

[74] Paterson, M., Peres, Y., Thorup, M., Winkler, P., & Zwick, U. (2009). Maximum overhang. *American Mathematical Monthly*, 116, 763–787.

[75] Paterson, M. & Zwick, U. (2006). Overhang. In *Proceedings of the 17th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)* (pp. 231–240).: Society for Industrial and Applied Mathematics.

[76] Pemantle, R. & Peres, Y. (2010). The critical ising model on trees, concave recursions and nonlinear capacity. *The Annals of Probability*, 38(1), 184–206.

[77] Povolotsky, A. M., Priezzhev, V. B., & Shcherbakov, R. R. (1998). Dynamics of Eulerian walkers. *Physical Review E*, 58(5), 5449. arXiv:cond-mat/9802070.

[78] Priezzhev, V., Dhar, D., Dhar, A., & Krishnamurthy, S. (1996a). Eulerian walkers as a model of self-organised criticality. *Phys. Rev. Lett.*, 77, 5079–5082. arXiv:cond-mat/9611019.

[79] Priezzhev, V. B., Dhar, D., Dhar, A., & Krishnamurthy, S. (1996b). Eulerian walkers as a model of self-organised criticality. *Phys. Rev. Lett.*, 77, 5079–5082. arXiv:cond-mat/9611019.

[80] Priezzhev, V. B., Dhar, D., Dhar, A., & Krishnamurthy, S. (1996). Eulerian Walkers as a Model of Self-Organized Criticality. *Physical Review Letters*, 77, 5079–5082.

[81] Propp, J. (2003). Random walk and random aggregation, derandomized. http://research.microsoft.com/apps/video/default.aspx?id=104906.

[82] Reddy A, T. R. (2010). A recurrent rotor-router configuration in $\mathbb{Z}^3$. arXiv:1005.3962.

[83] Revelle, D. (2003). Heat kernel asymptotics on the lamplighter group. *Elect. Comm. in Probab.*, (pp. 142–154).

[84] Révész, P. (2005). *Random Walk in Random and Non-Random Environments*. World Scientific.

[85] Schramm, O. (2003). Personal communication.

[86] Snijders, T. A. B. & Nowicki, K. (1997). Estimation and Prediction for Stochastic Blockmodels for Graphs with Latent Block Structure. *Journal of Classification*, 14(1), 75–100.

[87] Sornette, A. & Sornette, D. (1989). Self-organized criticality and earthquakes. *EPL (Europhysics Letters)*, 9(3), 197.

[88] Wagner, I. A., Lindenbaum, M., & Bruckstein, A. M. (1996). Smell as a computational resource – a lesson we can learn from the ant. In *4th Israel*

Symposium on Theory of Computing and Systems (ISTCS '96) (pp. 219–230).

[89] Yanovski, V., Wagner, I. A., & Bruckstein, A. M. (2003). A distributed ant algorithm for efficiently patrolling a network. *Algorithmica*, 37(3), 165–186.

[90] Zhou, T., Ren, J., Medo, M., & Zhang, Y.-C. (2007). Bipartite network projection and personal recommendation. *Physical Review E*, 76(4), 046115.