OPTIMAL CONTROL FOR DIFFUSIONS ON GRAPHS[*]

LAURA FLORESCU[†], YUVAL PERES[‡], AND MIKLÓS Z. RÁCZ[‡]

Abstract. Starting from a unit mass on a vertex of a graph, we investigate the minimum number of *controlled diffusion* steps needed to transport a constant mass p outside of the ball of radius n. In a step of a controlled diffusion process we may select any vertex with positive mass and topple its mass equally to its neighbors. Our initial motivation comes from the maximum overhang question in one dimension, but the more general case arises from optimal mass transport problems. On $\mathbb{Z}^d$ we show that $\Theta(n^{d+2})$ steps are necessary and sufficient to transport the mass. We also give sharp bounds on the comb graph and d-ary trees. Furthermore, we consider graphs where a simple random walk has positive speed and entropy and which satisfy Shannon's theorem, and show that the minimum number of controlled diffusion steps is $\exp\left(n \cdot h/\ell(1 + o(1))\right)$, where h is the Avez asymptotic entropy and ℓ is the speed of a random walk. As examples, we give precise results on Galton–Watson trees and the product of trees $\mathbb{T}_d \times \mathbb{T}_k$.

Key words. optimal transport, optimal control, combinatorial optimization, controlled diffusion, potential theory

AMS subject classifications. 05C81, 90C27, 60J45

DOI. 10.1137/17M1125923

1. Introduction. 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 et al. [10], who studied this question due to its connections with the maximum overhang problem [11, 10]. The main result of this paper 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 a 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 $1/2$?

1.1. Setting and main result.

DEFINITION 1.1 (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*

[†]Department of Mathematics, New York University, New York, NY 10003 (florescu@cims.nyu.edu).

[‡]Microsoft Research, Redmond, WA 98052 (peres@microsoft.com, miracz@microsoft.com).
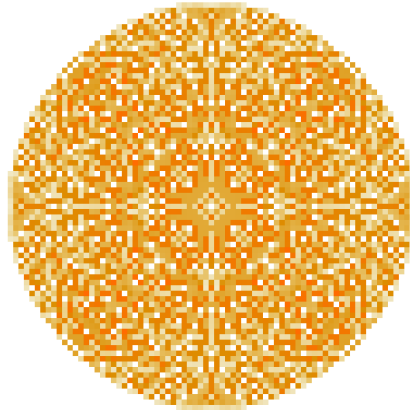
FIG. 1. *Starting with a unit mass at the origin of $\mathbb{Z}^2$, consider the process that at each step takes the vertex with largest mass and topples its mass equally to its neighbors. In case of a tie, topple all vertices with largest mass simultaneously. The figure shows the mass distribution after 10^5 such greedy steps, with darker squares representing larger masses.*

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 a distance at least n away from o. Our results hold for p constant.

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

THEOREM 1.2. *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 a 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 [10], where the authors obtained the same result as in Theorem 1.2 for $d = 1$. We discuss the connection to the maximum overhang problem and related open problems in more detail at the end of the paper (see section 8). Figure 1 illustrates a greedy algorithm on $\mathbb{Z}^2$, which indeed transports mass p to a distance at least n in $O\left(n^4\right)$ toppling moves.

1.2. 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 1.3. *Let $G = (V, E)$ be an infinite, connected, locally finite graph and let $\{X_t\}_{t \geq 0}$ be a simple random walk on G with $X_0 = o$ for a vertex $o \in V$. Let $A \subset V$*
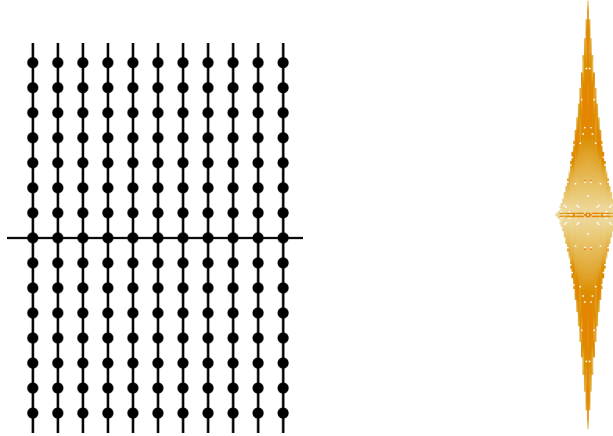
FIG. 2. *The comb graph $\mathbb{C}_2$ is on the left. On the right is the mass distribution on the comb after 10^6 steps of a greedy algorithm (starting from a unit mass at the origin), with darker squares representing larger masses. The algorithm uses the same symmetric tie-breaking rule as described in the caption of Figure 1. The 10^6 greedy steps resulted in 3439472 toppling moves.*

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

$$(1.1) \qquad N_p\left(G, A, \delta_o\right) \le (1-p)^{-1} \operatorname{Vol}\left(A\right) \cdot \mathbb{E}_o\left[T_A\right],$$

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

In section 4 we give two proofs of this result: one using a 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; see Figure 2 for an illustration.

THEOREM 1.4. *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 a 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.

Figure 2 also illustrates a greedy algorithm which achieves the upper bound in Theorem 1.4.

We also study various trees, starting with regular ones.[1]

---

[1]In the following we let $\mathbb{T}_d$ denote both the $(d+1)$-regular tree, where every vertex has degree $d+1$, and the d-ary tree, which differs from the $(d+1)$-regular tree in 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. Whether $\mathbb{T}_d$ denotes the d-ary tree or the $(d+1)$-regular tree will always be clear from the context (for instance, this is explicitly stated in Theorems 1.5 and 1.7).

THEOREM 1.5. *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 a distance at least n from the origin is*

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

where the implied constants depend only on d and p.

We prove a general result for graphs where the random walk has positive speed ℓ and entropy h and which satisfy Shannon's theorem. This roughly states that the minimum number of toppling moves satisfies $N_p(G, B_n, \delta_o) = \exp\left(n \cdot \frac{h}{\ell} \cdot (1 + o(1))\right)$; see section 6 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 1.6. *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 a distance at least n from the origin is almost surely*

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

where $\mathbf{dim}$ 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 1.5 provides a sharper result than Theorem 1.6. However, when the offspring distribution is nondegenerate, then $\mathbf{dim} < \log m$ almost surely (see [8]) and hence the number of toppling moves necessary is exponentially smaller than the volume of B_n.

THEOREM 1.7. *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 a distance at least n from the origin is*

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

where $\theta(d, k) = d^{\frac{d-1}{d+k-2}} \cdot k^{\frac{k-1}{d+k-2}}$, 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 $\mathrm{Vol}(B_n) = \Theta(d^n)$, whereas $\theta(d, k) < d$. Hence the number of toppling moves necessary to transport a constant mass to a 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 a simple random walk (see Definition 2.4).

THEOREM 1.8. *Let $G = (V, E)$ be an infinite, connected graph of bounded degree with exponential decay of the Green's function for a 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 a 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, the maximum degree of G, and the exponent in the exponential bound on the Green's function.

See section 7 where this result is restated more precisely as Theorem 7.1 and then proved, and where we illustrate this result with the example of the lamplighter graph.

1.3. 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 in this paper 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

$$(1.2) \qquad \Delta f(x) := \frac{1}{|\mathcal{N}_x|} \sum_{y \sim x} f(y) - f(x).$$

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

$$(1.3) \qquad T_v^m \mu = \mu - m\delta_v + \frac{m}{|\mathcal{N}_v|} \sum_{y \sim v} \delta_y = \mu + m\Delta\delta_v.$$

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

$$(1.4) \qquad \sum_{x \in V} f(x) \Delta g(x) = \sum_{x \in V} \Delta f(x) g(x),$$

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

$$(1.5) \qquad M_2[\mu] = \sum_{v \in \mathbb{Z}^d} \mu(v) \cdot \|v\|_2^2.$$

2. Upper bound on $\mathbb{Z}^d$ and preliminaries for the lower bound. We start with an upper bound on $N_p(\mathbb{Z}^d, B_n, \delta_o)$, stated as Theorem 2.1 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.

2.1. 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 a distance n from the origin in $\mathbb{Z}^d$.

THEOREM 2.1. *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 a distance at least n from the origin satisfies*

$$(2.1) \qquad N_p\left(\mathbb{Z}^d, B_n, \delta_o\right) < \frac{2^d}{(1-p) \times d!} n^{d+2}.$$

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 (1.3) we can then write

$$(2.2) \qquad \mu_i = \mu_{i-1} + m_i \Delta\delta_{v_i}.$$

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

We first compute how the second moment of the mass distribution changes after each toppling move. By (2.2) 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 (1.4)) 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:

$$(2.3) \qquad M_2[\mu_t] = \sum_{i=1}^{t} (M_2[\mu_i] - M_2[\mu_{i-1}]) = \sum_{i=1}^{t} m_i > t \times \frac{(1-p)}{|B_n|}.$$

On the other hand, all vertices with positive mass at time t have a (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 (2.3) we obtain that $t < |B_n| \times n^2/(1-p)$. The claim in (2.1) then follows from the estimate $|B_n| \leq (2^d/d!) n^d$ on the size of the (open) ball of radius n. $\square$

Remark 2.2. 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 $\mathbb{Z}^d$, such as the one described and illustrated in Figure 1. The same proof as above shows that this also transports mass p to a distance at least n in at most $O(n^{d+2})$ toppling moves.

2.2. 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 a random walk on $\mathbb{Z}^d$, which are required for subsequent estimates.

DEFINITION 2.3. *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, [3]. We will use the energy of the measure μ with the potential kernel function defined using the Green's function for a random walk on $\mathbb{Z}^d$, which we introduce next.

DEFINITION 2.4 (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*

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

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 a 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 $d \geq 3$, a simple random walk is transient in $\mathbb{Z}^d$ and hence $g(x)$ is finite for every $x \in \mathbb{Z}^d$. Since a 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

$$(2.4) \qquad g_n(x) := \mathbb{E}_o\left[\#\{k \in \{0, 1, \ldots, n\} : X_k = x\}\right],$$

the expected number of visits to x until time n by a 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 2.5 (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 (2.4).

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., [6, 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 [6, 4], but are not required for our purposes. First, when $d = 2$ then there exists an absolute constant C_2 such that

$$(2.5) \qquad \left|a(x) - \frac{2}{\pi}\ln\|x\|_2 - \kappa\right| \leq \frac{C_2}{\|x\|_2^2}$$

for all $x \neq o$, where κ is an explicit constant whose value is not relevant for our purposes [4]. Second, for every $d \geq 3$ there exists an absolute constant C_d such that

$$(2.6) \qquad \left| g\left(x\right) - a_d \left\| x \right\|_2^{2-d} \right| \leq \frac{C_d}{\left\| x \right\|_2^{d-1}}$$

for all $x \neq o$, where $a_d = (d/2)\Gamma\left(d/2 - 1\right)\pi^{-d/2} = \frac{2}{(d-2)\omega_d}$, where ω_d is the volume of the L_2 unit ball in $\mathbb{R}^d$ (see [6, Theorem 1.5.4]).

2.3. 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 2.6. *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 2.5. Then we have that*

$$(2.7) \qquad t\left(\mathcal{E}_a\left[\mu_t\right] - \mathcal{E}_a\left[\mu_0\right]\right) \geq \left(M_2\left[\mu_t\right] - M_2\left[\mu_0\right]\right)^2.$$

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 2.1 we know that $M_2\left[\mu_i\right] - M_2\left[\mu_{i-1}\right] = m_i$ for each $i \in [t]$. Turning to the energy of the measure, we first recall from (2.2) 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:

$$\mathcal{E}_a\left[\mu_i\right] - \mathcal{E}_a\left[\mu_{i-1}\right] = \sum_{x,y \in \mathbb{Z}^d} a\left(x - y\right)\left[\mu_i\left(x\right)\mu_i\left(y\right) - \mu_{i-1}\left(x\right)\mu_{i-1}\left(y\right)\right]$$

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

$$- \sum_{x,y \in \mathbb{Z}^d} a\left(x - y\right)\mu_{i-1}\left(x\right)\mu_{i-1}\left(y\right)$$

$$(2.8) \qquad = m_i \sum_{x,y \in \mathbb{Z}^d} a\left(x - y\right)\Delta \delta_{v_i}\left(x\right)\mu_{i-1}\left(y\right) + m_i \sum_{x,y \in \mathbb{Z}^d} a\left(x - y\right)\Delta \delta_{v_i}\left(y\right)\mu_{i-1}\left(x\right)$$

$$(2.9) \qquad + m_i^2 \sum_{x,y \in \mathbb{Z}^d} a\left(x - y\right)\Delta \delta_{v_i}\left(x\right)\Delta \delta_{v_i}\left(y\right).$$

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

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

For the first term in (2.8) we thus have

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

Since $a\left(x - y\right) = a\left(y - x\right)$ we have that the second term in (2.8) is equal to the first one. Finally we can compute the third term similarly:

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

Putting together the previous two displays with (2.8) and (2.9), we can conclude that

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

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

The claimed inequality (2.7) now follows by the Cauchy–Schwarz inequality:

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

2.4. An initial lower bound argument. A lower bound of the correct order in dimension $d = 1$ now follows (see also [10] 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 $\mu_t\left(B_n\right) \leq 1-p$. 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 a distance n from the origin.

Now recall that $a\left(x\right) = |x|$ when $d = 1$. Since $\mu_t\left(\{v : |v| > n\}\right) = 0$, we have that $\mathcal{E}_a\left[\mu_t\right] \leq 2n$. On the other hand, since $\mu_t\left(\{v : |v| \geq n\}\right) \geq p$, we have that $M_2\left[\mu_t\right] \geq pn^2$. By Lemma 2.6 we thus have that $t \times 2n \geq \left(pn^2\right)^2$, implying that

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

which matches the upper bound of Theorem 2.1 up to constant factors in p.

However, the same argument for $d \geq 2$ (using the estimates for the Green's function from (2.5) and (2.6); 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

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

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. Smoothing and the lower bound on $\mathbb{Z}^d$. The previous section provides the basis for the proof of Theorem 1.2, but applying the arguments directly leads to a suboptimal lower bound, as described in section 2.4. 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 3.1, followed by describing the specifics of smoothing in $\mathbb{Z}^d$ in section 3.2. We conclude with the proof of the lower bound on $\mathbb{Z}^d$ in section 3.3.

3.1. Smoothing of distributions. For the proofs of the lower bounds on most families of graphs investigated in this paper 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 a distance n started from the smooth distribution $\widetilde{\mu}$.

LEMMA 3.1 (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*

$$(3.1) \qquad N_p(G, A, T_v^m \mu) \leq N_p(G, A, \mu).$$

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 $v \neq v_1$. Observe that

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

since one way to move mass p outside of A starting from the mass distribution $T_v^m \mu$ is to first topple mass m_1 at vertex v_1 (this is possible since $m_1 \leq \mu(v_1) \leq (T_v^m \mu)(v_1)$), resulting in the mass distribution $T_{v_1}^{m_1} T_v^m \mu$, and then to move mass p outside of A in $N_p\left(G, A, T_{v_1}^{m_1} T_v^m \mu\right)$ additional toppling moves. Now note that the toppling moves T_v^m and $T_{v_1}^{m_1}$ commute, that is, $T_{v_1}^{m_1} T_v^m \mu = T_v^m T_{v_1}^{m_1} \mu$. Thus we have that $N_p\left(G, A, T_{v_1}^{m_1} T_v^m \mu\right) = N_p(G, A, T_v^m \mu_1)$. Finally, by the induction hypothesis we have that $N_p(G, A, T_v^m \mu_1) \leq N_p(G, A, \mu_1) = t - 1$. Putting everything together we thus obtain that

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

and so indeed the desired inequality holds.

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

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

If $m \geq m_1$, then

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

where the inequality is again due to the induction hypothesis. □

Iterating this lemma we immediately obtain the following corollary.

COROLLARY 3.2. *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*

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

Another corollary of the lemma above is that we can assume without loss of generality that at every move of an optimal toppling sequence 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 3.3. *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 case 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 $\mu_t(A) \leq 1 - p$. We must have that $v_i \in A$ for every $i \in [t]$, since one can always remove topplings of a vertex outside of A to obtain a shorter toppling sequence that still moves mass p outside of A. 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 3.1 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. 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 [7] 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 (1.3) 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

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

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

PROPOSITION 3.4 (Levine and Peres [7]). *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*

$u_k(x) =$ *total mass emitted by* x *after divisible sandpile moves* $x_1, \ldots, x_k$;

$\nu_k(x) =$ *amount of mass present at* x *after divisible sandpile moves* $x_1, \ldots, x_k$.

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

The 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 in this paper 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 3.5 (see [7, Levine and Peres]). *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 $\nu_K(x) \leq 1 + \varepsilon$ for every $x \in \mathbb{Z}^d$ and for some small $\varepsilon > 0$. This is useful for proving the following corollary of the theorem above.

COROLLARY 3.6. *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:*

$$(3.3) \qquad \forall x \in B_{cn}^{(2)} : \quad \mu(x) \leq \frac{2}{\mathrm{Vol}\left(B_{cn}^{(2)}\right)},$$

$$(3.4) \qquad \forall x \notin B_{cn}^{(2)} : \quad \mu(x) = 0,$$

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 3.5 by scaling the masses by m, for both the mass distributions and the divisible sandpile moves. $\quad\square$

3.3. A lower bound on $\mathbb{Z}^d$. We are now ready to prove Theorem 1.2. By performing an initial smoothing as detailed in section 3.2, we are able to obtain a lower bound that matches the upper bound of Theorem 2.1 up to constant factors.

THEOREM 3.7. *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 a distance at least n from the origin satisfies*

$$(3.5) \qquad N_p\left(\mathbb{Z}^d, B_n, \delta_o\right) \geq Cn^{d+2}.$$

Proof. The first step is to smooth the distribution δ_o. Let $c := \sqrt{p/(2d)}$. By Corollary 3.6 there exists a finite sequence of toppling moves taking δ_o to a mass distribution μ satisfying (3.3) and (3.4). By Corollary 3.2 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 $\mu_t(B_n) \leq 1 - p$. 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 $\|v_i\|_1 \le n-1$ for every $i \le t$, since any other toppling move can be removed from the sequence to obtain a shorter sequence that still moves mass p to a distance n from the origin.

By Lemma 2.6 we have that

$$(3.6) \qquad t \ge \frac{\left(M_2\left[\mu_t\right] - M_2\left[\mu_0\right]\right)^2}{\mathcal{E}_a\left[\mu_t\right] - \mathcal{E}_a\left[\mu_0\right]}$$

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

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

$$(3.7) \qquad \left(M_2\left[\mu_t\right] - M_2\left[\mu_0\right]\right)^2 \ge \frac{p^2}{4d^2}n^4.$$

From (3.6) and (3.7) we have that in order to show (3.5), what remains is to show that

$$(3.8) \qquad \mathcal{E}_a\left[\mu_t\right] - \mathcal{E}_a\left[\mu_0\right] \le C'n^{2-d}$$

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

Recall from section 2.2 that when $d \ge 3$ then

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

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

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

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

$$(3.9) \qquad \mathcal{E}_g\left[\mu_0\right] = \sum_{x \in B_{cn}^{(2)}} g(o)\,\mu_0(x)^2 + \sum_{k=1}^{K} \sum_{(x,y) \in E_k} g(x-y)\,\mu_0(x)\,\mu_0(y),$$

where $K = \lceil \log_2(2n) \rceil$. Using (3.3) we have that the first term in (3.9) can be bounded as follows:

$$(3.10) \qquad \sum_{x \in B_{cn}^{(2)}} g(o)\,\mu_0(x)^2 \le \sum_{x \in B_{cn}^{(2)}} \frac{4g(o)}{\mathrm{Vol}\left(B_{cn}^{(2)}\right)^2} = \frac{4g(o)}{\mathrm{Vol}\left(B_{cn}^{(2)}\right)} = O\left(n^{-d}\right),$$

where in the last estimate we used that $\mathrm{Vol}(B_{cn}^{(2)}) = \Theta\left(n^d\right)$. Now if $x \ne y$ then we have from (2.6) that

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

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

$$g(x - y) \leq a_d \left(\frac{2n}{2^{k-1}}\right)^{2-d} + C_d \left(\frac{2n}{2^{k-1}}\right)^{1-d}$$

(3.11)
$$= O\left(n^{2-d} \times 2^{dk} \times \left(2^{-2k} + n^{-1} \times 2^{-k}\right)\right).$$

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

(3.12)
$$\sum_{(x,y)\in E_k} \mu_0(x)\mu_0(y) = O\left(2^{-dk}\right).$$

Putting together (3.11) and (3.12) 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 (3.10), shows that $\mathcal{E}_g[\mu_0] = O\left(n^{2-d}\right)$. This concludes the proof of (3.8) 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 (2.5)). 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

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

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

(3.13)
$$= \frac{2}{\pi} \ln(n) + O(1).$$

Next we bound from below the energy $\mathcal{E}_a[\mu_0]$. Noting again that $a(o) = 0$, we can write $\mathcal{E}_a[\mu_0]$ similarly to (3.9):

(3.14)
$$\mathcal{E}_a[\mu_0] = \sum_{k=1}^{K} \sum_{(x,y)\in E_k} a(x - y)\mu_0(x)\mu_0(y).$$

For $x \neq o$ we have the estimate $a(x) \geq \frac{2}{\pi} \ln \|x\|_2 - C_2 \|x\|_2^{-2}$ (see (2.5) and note that $\kappa > 0$) and, thus, if $(x, y) \in E_k$, then

$$(3.15) \qquad a(x - y) \geq \frac{2}{\pi} \ln(2n) - \frac{2}{\pi} \ln(2) \times k - C_2 \times \frac{2^{2k}}{4n^2}.$$

Plugging the estimate (3.15) into (3.14) we get three terms which we can each estimate separately. First, observing that

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

we get that

$$(3.16) \qquad \sum_{k=1}^{K} \sum_{(x,y) \in E_k} \frac{2}{\pi} \ln(2n) \mu_0(x) \mu_0(y) = \frac{2}{\pi} \ln(n) - O(1).$$

For the second term in (3.15) we use (3.12) to obtain that

$$(3.17) \qquad \sum_{k=1}^{K} \sum_{(x,y) \in E_k} \left(\frac{2}{\pi} \ln(2) \times k\right) \mu_0(x) \mu_0(y) = \sum_{k=1}^{K} O\left(k \times 2^{-2k}\right) = O(1).$$

For the third term in (3.15) we again use (3.12), together with the fact that $K = \lceil \log_2(2n) \rceil$, to get that

$$(3.18) \qquad \sum_{k=1}^{K} \sum_{(x,y) \in E_k} \left(C_2 \times \frac{2^{2k}}{4n^2}\right) \mu_0(x) \mu_0(y) = \frac{1}{n^2} \sum_{k=1}^{K} O(1) = O\left(\frac{\log(n)}{n^2}\right).$$

Putting together (3.16), (3.17), and (3.18) with (3.14) and (3.15) we obtain that

$$(3.19) \qquad \mathcal{E}_a[\mu_0] = \frac{2}{\pi} \ln(n) - O(1).$$

Finally, putting together (3.13) and (3.19) we obtain (3.8) for $d = 2$. $\qquad \square$

4. A general upper bound. In this section we provide two proofs of Theorem 1.3.

Proof of Theorem 1.3 *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

$$(4.1) \qquad \mu_i := T_{x_{i^*}}^{m_i} \mu_{i-1} \qquad \text{with} \qquad m_i = \mu_{i-i^*}(x_{i^*}).$$

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 (4.1) 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

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

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

$$(4.2) \qquad \mathbb{E}_o[T_A] = \sum_{k=1}^{\infty} \mathbb{P}_o(T_A \geq k) = \sum_{k=1}^{\infty} \mathbb{P}_o(Z_{k-1} \in A) = \sum_{k=1}^{\infty} \mu_{(k-1)|A|}(A).$$

Now by the definition of M we have that for every $m < M$, the measure $\mu_{m|A|}$ satisfies $\mu_{m|A|}(A) > 1 - p$. Therefore keeping only the first M terms in the sum in (4.2) we obtain the bound

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

By the definition of M this immediately implies that

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

Theorem 1.3 can also be proven using a greedy algorithm, similarly to the proof of the greedy upper bound on $\mathbb{Z}^d$ presented in section 2.1. 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 $\Delta \|x\|_2^2 = 1$ 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

$$(4.3) \qquad \mathbb{E}_x[T_A] = 1 + \frac{1}{|\mathcal{N}_x|} \sum_{y \sim x} \mathbb{E}_y[T_A]$$

for every $x \in A$.

Proof of Theorem 1.3 *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 $\mu_0 \equiv \delta_o, \mu_1, \mu_2, \dots$ 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 (1.3) 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 a 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 (4.3), 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

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

$$= -h(v_i) + \frac{1}{|\mathcal{N}_{v_i}|} \sum_{y \sim v_i} h(y) = \Delta h(v_i) = 1,$$

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

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

and the result follows by rearranging this inequality. $\square$

5. Controlled diffusion on the comb. The general upper bound given by Theorem 1.3 applied directly to the comb $\mathbb{C}_2$ gives a bound of

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

for some constant C, since $\mathrm{Vol}(B_n) = \Theta(n^2)$ and $\mathbb{E}_o[T_{B_n}] = \Theta(n^2)$. However, this bound is not tight. Recall that the general upper bound that gives (5.1) is proven by simulating a random walk within B_n. The key observation that improves (5.1) 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 $\mathrm{Vol}(R_{C,n}) = \Theta(n^{3/2})$, this gives the improved upper bound of $O(n^{7/2})$.

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 a 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 path, this requires $\Omega(n^3)$ toppling moves along each tooth (by Theorem 1.2 for $d = 1$, proven in [10]), resulting in $\Omega(n^{7/2})$ toppling moves in total.

The rest of this section makes the two preceding paragraphs precise and proves Theorem 1.4. Let $\{X_t\}_{t \geq 0}$ denote a 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 $X_t = (X_t^{(1)}, X_t^{(2)})$ and $Z_t = (Z_t^{(1)}, Z_t^{(2)})$ 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 5.1. *For every $\varepsilon > 0$ there exists $C = C(\varepsilon) < \infty$ such that*

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

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:

$$(5.2) \qquad\qquad \mathbb{P}_o\left(\left|X^{(2)}_t\right| \leq n \text{ for every } t \leq cn^2\right) \leq \varepsilon/2,$$

$$(5.3) \qquad\qquad \mathbb{P}_o\left(\left|X^{(1)}_t\right| \geq C\sqrt{n} \text{ for some } t \leq cn^2\right) \leq \varepsilon/2.$$

Note that $\{X^{(2)}_t\}_{t\geq 0}$ is Markovian: when away from 0 it behaves like a 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 (5.2) follows from classical random walk estimates (for instance, it follows from the central limit theorem; see, e.g., [13, Theorem 2.9]), provided $c = c(\varepsilon)$ is large enough.

Now fix c such that (5.2) holds. Again by classical estimates (see, e.g., [13, Theorem 9.11]) there exists a constant c' such that

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

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

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

provided that C is large enough. Putting everything together gives us (5.3). □

With this lemma in hand we are now ready to prove Theorem 1.4. We start with the upper bound and we again give two proofs, one using a random walk and one using a greedy algorithm.

Proof of the upper bound of Theorem 1.4 using random walk. Fix $\varepsilon \in (0, 1-p)$, let $C = C(\varepsilon)$ be the constant given by Lemma 5.1, and let $R := R_{C,n}$. Just like in the proof of Theorem 1.3, 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 nonnegative integer t, the distribution $\mu_{t|R|}$ agrees with the distribution of Z_t.

Let

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

denote the first time that the distribution of the killed random walk has mass at least $p + \varepsilon$ outside of the rectangle R. By Lemma 5.1 we have that

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

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 a 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 1.3, one can show that

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

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

Proof of the upper bound of Theorem 1.4 *using a greedy algorithm.* First, fix $\varepsilon \in (0, 1 - p)$, let $C = C(\varepsilon)$ be the constant given by Lemma 5.1, and let $R := R_{C,n}$. Consider the following greedy algorithm for choosing toppling moves: until the mass outside of R is at least $p + \varepsilon$, 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 $p + \varepsilon$ outside of R, i.e., $t = \min \{i \geq 1 : \mu_i(R) \leq 1 - p - \varepsilon\}$.

Just as in the proof of Theorem 2.1 we can compute how the second moment of the mass distribution changes after each toppling move and we 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

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

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

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

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 $M_2 [\mu_t] \leq C' n^2$. Combining this with the display above we obtain that $t < C' n^2 \times |R| / (1 - p - \varepsilon)$. Since $|R| = \Theta(n^{3/2})$ we thus have that $t = O(n^{7/2})$.

What remains to show is that the mass distribution μ_t has mass at least p at a 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

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

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

$$(5.4) \qquad \mu_t((-C\sqrt{n} - 1, 0)) + \mu_t((C\sqrt{n} + 1, 0)) \leq \varepsilon .$$

For $x \in \mathbb{C}_2$ let $h(x) := \mathbb{P}_x(X_{T_R}^{(2)} = 0)$. By Lemma 5.1 we have that $h(o) \leq \varepsilon$ and, hence, $\sum_{x \in \mathbb{C}_2} h(x) \mu_0(x) \leq \varepsilon$. 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, $\sum_{x \in \mathbb{C}_2} h(x) \mu_t(x) \leq \varepsilon$. The inequality (5.4) then immediately follows from the fact that $h((-C\sqrt{n}-1,0)) = h((C\sqrt{n}+1,0)) = 1$. $\qquad\square$

Proof of the lower bound of Theorem 1.4. Given $p \in (0,1)$, let $\varepsilon := p/4$. In the following we fix $c = c(\varepsilon)$ and $C = C(\varepsilon)$ 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 5.1, 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 $\mu(S) \geq 1-\varepsilon$. 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

$$(5.5) \qquad \mu_1(x) \leq \frac{K}{\sqrt{n}} \quad \forall x \in S, \qquad \text{and} \qquad \mu_2(S) \leq \varepsilon.$$

Before proving (5.5), we show how to conclude the proof assuming that (5.5) holds. First of all, from Corollary 3.2 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.[2] 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 $\mu(S) \geq 1-\varepsilon$ and $\mu'(B_n) \leq 1-p$, we must have that

$$(5.6) \qquad \sum_{x \in S} \nu(x) \geq p - \varepsilon.$$

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

$$(5.7) \qquad \sum_{x \in S} \nu_1(x) \geq p - 2\varepsilon.$$

---

[2]One way to define ν unambiguously is as follows. First, associate the color blue with all vertices not in S and with each vertex in S associate its own unique and nonblue color. Next, fix an ordering π of the vertices in S. Now suppose that we obtained μ' from μ via a sequence of toppling moves: $\mu' = \prod_{i=1}^t T_{v_i}^{m_i} \mu$. When we topple mass m at vertex v, it is possible that the mass at v has multiple "colors" and not all of this mass is toppled. We may then fix an order of priority between the colors, given by the permutation π among colors associated with S, and giving the color blue lowest priority. In this way we first topple all mass of the highest priority color, we then topple the mass of the second highest priority color, and so on, until we exhaust the mass m that we want to topple at vertex v. This procedure does not at all affect the sequence of mass distributions, its only purpose is to allow the tracking of how mass moves from some vertices to others. Finally, we can define $\nu(x)$ as the mass of the color associated with x that is outside of B_n under the mass distribution μ'. The mass measures ν_1 and ν_2 can be defined analogously by introducing two subcolors for each color.

Let $S_{\text{lg}} := \{x \in S : \nu_1(x) \geq \varepsilon/(5C\sqrt{n})\}$ and $S_{\text{sm}} := S \setminus S_{\text{lg}}$, and break the sum in (5.7) into two parts accordingly. Using that $|S| = 4C\sqrt{n} + 2 \leq 5C\sqrt{n}$, we have that $\sum_{x \in S_{\text{sm}}} \nu_1(x) \leq \varepsilon$, and so

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

On the other hand, (5.5) implies that

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

and so we must have that $|S_{\text{lg}}| \geq \frac{p}{4K}\sqrt{n}$. Notice that for every $x \in S_{\text{lg}}$ we have that $\nu_1(x)/\mu_1(x) \geq \varepsilon/(5CK)$, 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 path of length $\Omega(n)$, we know from Theorem 1.2 for $d = 1$ (proven in [10]) that $\Omega(n^3)$ toppling moves are necessary to do this. Since this holds for every $x \in S_{\text{lg}}$, we see that $\Omega(n^{7/2})$ toppling moves are necessary altogether.

Finally, we turn back to proving (5.5). First, note that there exists $\delta = \delta(\varepsilon)$ such that with probability at least $1 - \varepsilon/2$, 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 a simple random walk; see, e.g., [13]). 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.,

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

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., [13, Theorem 9.11]), there exists $\gamma = \gamma(\varepsilon)$ such that with probability at least $1 - \varepsilon/2$, we have that

(5.8) $$N_{\delta n^2}^X \geq \gamma n.$$

Denote by $\mathcal{B}$ the event that the inequality in (5.8) holds and note that $\mathbb{P}(\mathcal{A} \cap \mathcal{B}) \geq 1 - \varepsilon$. 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 5.1. 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

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

Hence there exists K such that

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

which implies the claim. □

6. Graphs where a random walk has positive speed. In this section we study graphs on which a simple random walk has positive speed. As a warmup, we study d-ary trees in section 6.1, followed by general results in section 6.2. We then apply the general results to two examples: Galton–Watson trees (section 6.3) and product of trees (section 6.4). 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 a 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 1.3.

6.1. d-ary trees. The general upper bound of Theorem 1.3 applied directly to the d-ary tree $\mathbb{T}_d$ gives

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

for some constant $C < \infty$, since $\mathrm{Vol}\left(B_n\right) = \Theta\left(d^n\right)$ and $\mathbb{E}_\rho\left[T_{B_n}\right] = \Theta\left(n\right)$. However, this bound is not tight, as Theorem 1.5 states that $N_p\left(\mathbb{T}_d, B_n, \delta_\rho\right) = \Theta\left(d^n\right)$. This example is interesting because the factor coming from the exit time of the random walk is completely absent from $N_p\left(\mathbb{T}_d, B_n, \delta_\rho\right)$. The proof requires a more careful analysis of the greedy algorithm.

In the rest of this subsection we prove Theorem 1.5, starting with the lower bound. We define the level of a vertex $v \in \mathbb{T}_d$ to be its distance from the root: $\ell\left(v\right) := d_{\mathbb{T}_d}\left(v, \rho\right)$.

Proof of the lower bound in Theorem 1.5. First, we smooth 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 $\varepsilon > 0$ it is possible to obtain, via a finite sequence of toppling moves, a mass distribution μ such that $\mu\left(v\right) \in \left(\left(1 - \varepsilon\right) d^{-(n-1)}, d^{-(n-1)}\right)$ for every vertex v at level $n - 1$.

One way to see this is by the following explicit construction. Define a "sweep through level k" as the following sequence of toppling moves: first, topple the root, then topple all vertices at level 1, then topple all vertices at level 2, and so on, concluding by toppling all vertices at level k. Note that a sweep through level k is a total of $\left(d^{k+1} - 1\right) / \left(d - 1\right)$ toppling moves. We now argue that $M\left(\varepsilon, d\right) := \left\lceil \frac{1-\varepsilon}{\varepsilon}\left(\frac{d+1}{d}\right)^{n-2}\right\rceil$ sweeps through level $n - 2$ result in the desired mass distribution. First, notice that throughout this sequence of toppling moves, whenever all vertices are toppled on a given level, the mass on each level is distributed uniformly among vertices on that level. Hence we only have to understand the mass distribution on the levels. Observe that, starting from a unit mass at the root, after the first sweep through level $n - 2$ we have mass $(d/(d+1))^{n-2}$ at level $n - 1$; the rest of the mass is spread over levels 0 through $n - 2$. During each subsequent sweep through level $n - 2$ we have that at least a $(d/(d+1))^{n-2}$ fraction of mass that was not at level $n - 1$ before the sweep is at level $n - 1$ after the sweep. That is, as long as there is mass at least ε not at level $n - 1$, each subsequent sweep through level $n - 2$ adds mass at

least $\varepsilon \left(d/\left(d+1\right) \right)^{n-2}$ to level $n-1$. Hence after $M\left(\varepsilon, d\right)$ sweeps through level $n-2$ there will be mass at least $1-\varepsilon$ at level $n-1$. By the discussion above this mass is uniformly spread among vertices at level $n-1$ and so the resulting mass distribution μ indeed satisfies $\mu\left(v\right) \in \left(\left(1-\varepsilon\right) d^{-(n-1)}, d^{-(n-1)}\right)$ for every vertex v at level $n-1$.

By Corollary 3.2 we have that $N_p\left(\mathbb{T}_d, B_n, \delta_\rho\right) \geq N_p\left(\mathbb{T}_d, B_n, \mu\right)$, so it suffices to bound from below this latter quantity.

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

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

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

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 6.1. *If μ is a mass distribution on $\mathbb{T}_d$ such that $M_1\left[\mu\right] \leq \ell$, then there exists $v \in \mathbb{T}_d$ such that $\ell\left(v\right) \leq \ell$ and $\mu\left(v\right) \geq d^{-(\ell+1)}/2$.*

Proof. We prove the statement by contradiction. Suppose that $\mu\left(v\right) < d^{-(\ell+1)}/2$ for every $v \in \mathbb{T}_d$ such that $\ell\left(v\right) \leq \ell$; our goal is to show that then $M_1\left[\mu\right] > \ell$. To bound $M_1\left[\mu\right]$ from below, we can first bound $\ell\left(v\right)$ by $\ell+1$ for every v such that $\ell\left(v\right) \geq \ell+1$ to obtain that

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

Using the assumption that $\mu\left(v\right) < d^{-(\ell+1)}/2$ for every $v \in \mathbb{T}_d$ such that $\ell\left(v\right) \leq \ell$, we thus have that

$$M_1\left[\mu\right] > \ell+1 - \frac{1}{2}d^{-(\ell+1)} \sum_{v:\ell(v)\leq \ell} \left(\ell+1-\ell\left(v\right)\right).$$

Finally, we have that

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

and so $M_1\left[\mu\right] > \ell$. □

Proof of the upper bound in Theorem 1.5. We 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 6.1, for every $i < t_\ell$ we have that $m_i \geq d^{-(\ell+1)}/2$. This implies that

$$M_1[\mu_{t_\ell - 1}] - M_1[\mu_{t_{\ell-1}}] \geq (t_\ell - 1 - t_{\ell-1}) \times \frac{d-1}{d+1} \times \frac{1}{2d^{\ell+1}}.$$

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

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

Putting the previous displays together we obtain that

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

for every $\ell < n$, where the implied constant depends only on d. Summing (6.1) 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 $t - t_{n-1} = O(d^n)$. Recall that for every $i < t$ we have that $\mu_i(B_n) > 1 - p$. Since $\text{Vol}(B_n) < d^n$, there must exist $v \in B_n$ such that $\mu_i(v) > (1 - p)/d^n$. Hence for every $i \in (t_{n-1}, t]$ we have that $m_i > (1 - p)/d^n$. Thus

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

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

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

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

6.2. A general bound for graphs where a random walk has positive speed and entropy. In this subsection we present a general result for graphs where a 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 a 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 a random walk.

DEFINITION 6.2. *The (asymptotic)* speed *of the random walk* $\{X_t\}_{t\geq 0}$ *on G is defined as*

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

Note that the triangle inequality implies subadditivity, that is, $d\left(X_0, X_{s+t}\right) \leq d\left(X_0, X_s\right) + d\left(X_s, X_{s+t}\right)$ and, hence, the speed of the random walk exists almost surely by Kingman's subadditive ergodic theorem (see, e.g., [9, Theorem 14.44]).

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

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

where in this paper we use log to denote the natural logarithm.

DEFINITION 6.3. *The* asymptotic entropy, *also known as the* Avez entropy, *of the random walk* $\{X_t\}_{t\geq 0}$ *on G is defined as*

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

provided that this limit exists.

When G is transitive, the sequence $\{H\left(X_t\right)\}_{t\geq 0}$ is subadditive and, hence, the Avez entropy exists by Fekete's lemma (see, e.g., [9, section 14.1]).

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

THEOREM 6.4 (see [5], [9, 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 6.5 (see [5, Theorem 2.1], [9, Theorem 14.10]). *Assume the setup described in the first paragraph of section* 6.2 *and, in addition, assume that G is a transitive graph. Then we have that*

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

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 a distance n for graphs where a simple random walk has positive asymptotic speed and positive Avez entropy, and which satisfy Shannon's theorem.

THEOREM 6.6. *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 a simple random walk on G started from o, i.e., with $X_0 = o$. Assume that the following three conditions hold:*
1. *A simple random walk on G has positive asymptotic speed, i.e., $\boldsymbol{\ell} > 0$.*
2. *A simple random walk on G has positive Avez entropy, i.e., $\boldsymbol{h} > 0$.*
3. *We have that*

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

almost surely.

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

$$(6.3) \qquad N_p\left(G, B_n, \delta_o\right) = \exp\left(n \times \frac{\boldsymbol{h}}{\boldsymbol{\ell}}\left(1 + o\left(1\right)\right)\right).$$

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

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

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

$$(6.4) \qquad V_{t,n} := \left\{x \in B_n : \frac{1}{t}\log p_t\left(o, x\right) \in \left(-\boldsymbol{h}\left(1+\varepsilon\right), -\boldsymbol{h}\left(1-\varepsilon\right)\right)\right\},$$

and note that $|V_{t,n}| \leq \exp\left(t\boldsymbol{h}\left(1+\varepsilon\right)\right)$ for every t, since $p_t\left(o, x\right) \geq \exp\left(-t\boldsymbol{h}\left(1+\varepsilon\right)\right)$ 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

$$(6.5) \qquad \mathbb{P}_o\left(Z_{t^*} \notin B_n\right) \geq p$$

if n is large enough, which thus implies that

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

if n is large enough. Since this holds for every $\varepsilon > 0$, we get the desired upper bound stated in (6.3).

So what remains is to show (6.5). 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

$$(6.6) \qquad \mathbb{P}_o\left(Z_{t^*} \in B_n\right) = \mathbb{P}_o\left(Z_{t^*} \in U_n\right) + \mathbb{P}_o\left(Z_{t^*} \in B_n \setminus U_n\right).$$

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

$$\mathbb{P}_o\left(Z_{t^*} \in U_n\right) \leq \mathbb{P}_o\left(d\left(X_0, X_{t^*}\right) < n\right) = \mathbb{P}_o\left(\tfrac{1}{t^*}d\left(X_0, X_{t^*}\right) < \boldsymbol{\ell}/\left(1+\varepsilon\right)\right).$$

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 (6.6) 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

$$\mathbb{P}_o\left(Z_{t^*} \in B_n \setminus U_n\right) \leq \mathbb{P}_o\left(\exists t \geq r_n : \tfrac{1}{t} \log p_t\left(X_0, X_t\right) \notin \left(-\boldsymbol{h}\left(1+\varepsilon\right), -\boldsymbol{h}\left(1-\varepsilon\right)\right)\right).$$

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 (6.5).

To prove the lower bound stated in (6.3), we again start by smoothing the initial mass distribution, by simulating a simple random walk on G until time $t^{**} := (1-\varepsilon)\, n/\boldsymbol{\ell}$. 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 $\varepsilon > 0$ and let $t^{**} := (1-\varepsilon)\, n/\boldsymbol{\ell}$. The choice of t^{**} is due to the fact that, by Assumption 1, with probability close to 1, a 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 $t^{**} \times \operatorname{Vol}(B_n)$ 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 3.2 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. Recall the definition of $V_{t,n}$ from (6.4). By the definition of t^{**} and Assumptions 1 and 3, it follows that

$$\mu\left(V_{t^{**},n}\right) \geq 1 - \frac{p}{2}$$

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, $\mu(x) \leq \exp\left(-t^{**}\boldsymbol{h}\left(1-\varepsilon\right)\right)$ for every $x \in V_{t^{**},n}$, so at least

$$\frac{p}{2} \times \exp\left(t^{**}\boldsymbol{h}\left(1-\varepsilon\right)\right) = \frac{p}{2} \times \exp\left(n \times \frac{\boldsymbol{h}}{\boldsymbol{\ell}}\left(1-\varepsilon\right)^2\right)$$

vertices in $V_{t^{**},n}$ need to be toppled at least once. Since this holds for any $\varepsilon > 0$, the result follows. $\square$

6.3. Galton–Watson trees. The behavior of a simple random walk on Galton–Watson trees was studied in great detail by Lyons, Pemantle, and Peres [8]. Using their results, combined with the general results of section 6.2, we can prove Theorem 1.6.

Specifically, Lyons, Pemantle, and Peres [8] showed that the three conditions of Theorem 6.6 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 6.6, and refer to [8] 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 6.1.

THEOREM 6.7 (see [8, 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 a 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 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 $\mathbf{dim}$ 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 1.6. Theorem 6.7 shows that the three conditions of Theorem 6.6 hold for almost every Galton–Watson tree. Hence Theorem 1.6 follows from Theorem 6.6.   □

6.4. Product of trees. In this subsection we apply the general result derived in section 6.2 to obtain tight bounds for the specific case of the product of trees. As we shall see, the key observation is that a 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 $\mathbb{T}_d \times \mathbb{T}_k$ 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:

$$(u, v) \sim (u', v') \iff \begin{cases} u \sim u' \text{ and } v = v', \quad \text{or} \\ u = u' \text{ and } v \sim v'. \end{cases}$$

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 1.2, and hence we may assume that $d + k \geq 3$.

Proof of Theorem 1.7. We prove this result by appealing to the general result of Theorem 6.6. Therefore we need to check that the three assumption's of Theorem 6.6 hold and we also need to compute the asymptotic speed $\boldsymbol{\ell}$ and the Avez entropy $\boldsymbol{h}$ for a simple random walk on $\mathbb{T}_d \times \mathbb{T}_k$.

Let $\{X_t\}_{t\geq 0}$ denote a simple random walk on $\mathbb{T}_d \times \mathbb{T}_k$ with $X_0 = \rho$. We start by computing the speed of a random walk. Recall that the speed of a random walk on the $(d+1)$-regular tree $\mathbb{T}_d$ is $\frac{d-1}{d+1}$. Moreover, the probability of a 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 a random walk $\{X_t\}_{t\geq 0}$ is the convex combination of the speeds of a random walk on the regular trees $\mathbb{T}_d$ and $\mathbb{T}_k$:

$$(6.7) \qquad \boldsymbol{\ell} = \frac{d+1}{d+k+2} \times \frac{d-1}{d+1} + \frac{k+1}{d+k+2} \times \frac{k-1}{k+1} = \frac{d+k-2}{d+k+2}.$$

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 6.5 that (6.2) 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 a random walk on $\mathbb{T}_d$. Let

$\{Y_t\}_{t\geq 0}$ denote a 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 a 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 a 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 independent and identically distributed 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

$$H(X_t) = H(S_t) + H((Y_{S_t}, Z_{t-S_t}) \,|\, S_t).$$

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, $H((Y_{S_t}, Z_{t-S_t}) \,|\, S_t) = H(Y_{S_t} \,|\, S_t) + H(Z_{t-S_t} \,|\, S_t)$. Therefore, using the computation from above of the entropy of a 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 6.6, we obtain the desired result. □

7. 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 a 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 a simple random walk on G from Definition 2.4. 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

$$(7.1) \qquad g(x, y) \leq \exp(-a \times d(x, y) + a')$$

for every $x, y \in V$, where d denotes the graph distance. Note that the Green's function cannot decay faster than exponentially as a function of the distance.

If a simple random walk on G has positive speed and positive entropy, then the Green's function has exponential decay (see [1, 2]). However, the reverse implication does not hold, and hence the method described in section 6.2 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 a simple random walk are both zero (see [5, Proposition 6.2]).

We restate Theorem 1.8 more precisely before proving it.

THEOREM 7.1. *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 a simple random walk on G satisfies (7.1). 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 1.3. Since G has bounded degree, the volume of a ball grows at most exponentially: $\mathrm{Vol}(B_n) \leq \sum_{i=0}^{n-1} D^i \leq D^n$. Furthermore, the exit time of a 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: $\mathbb{E}_o[T_{B_n}] \leq \sum_{x \in B_n} g(o, x) \leq e^{a'} \mathrm{Vol}(B_n)$. Hence using Theorem 1.3 we have that

$$N_p(G, B_n, \delta_o) \leq (1 - p)^{-1} e^{a'} D^{2n}.$$

For the lower bound we again perform smoothing of the initial mass distribution. Let $\{X_t\}_{t \geq 0}$ denote a 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

$$(7.2) \qquad \mathbb{P}_o(Z_{t^*} \in B_{n-1}) \leq p/2.$$

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 $\{Z_t\}_{t \geq 0}$, to arrive at a new mass distribution μ. By Corollary 3.2 we have that $N_p(G, B_n, \overline{\delta}_o) \geq N_p(G, B_n, \mu)$, so it suffices to bound from below this latter quantity.

Denote the boundary of B_{n-1} by $\partial B_{n-1} := \{x \in V : d(o, x) = n - 1\}$. For every $x \in \partial B_{n-1}$ we can bound the mass at x using the Green's function:

$$\mu(x) = \mathbb{P}_o(Z_{t^*} = x) \leq \mathbb{P}_o\left(X_{T_{B_{n-1}}} = x\right)$$

$$\leq \sum_{k=0}^{\infty} \mathbb{P}_o(X_k = x) = g(o, x) \leq \exp(-an + a + a'),$$

where in the last inequality we used (7.1). Now (7.2) implies that $\mu\left(\partial B_{n-1}\right) \geq 1-p/2$, and so in order to transport mass at least p to the 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 $\exp\left(-an+a+a'\right)$, so this requires at least $(p/2) \times \exp\left(an-a-a'\right)$ toppling moves. Hence

$$N_p\left(G, B_n, \mu\right) \geq \frac{p}{2\exp\left(a+a'\right)} \times e^{an}. \qquad \square$$

7.1. 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 7.2. *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*

$$\left(\eta_1, y_1\right)\left(\eta_2, y_2\right) := \left(\eta, y_1+y_2\right),$$

where $\eta\left(x\right) = \eta_1\left(x\right) + \eta_2\left(x-y_1\right) \mod 2.$

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 by 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, $\mu\left(\mathbf{0}, \pm 1\right) = 1/2$ and $\nu\left(\mathbf{1}_0, 0\right) = \nu\left(\mathbf{0}, 0\right) = 1/2$. 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 7.3. *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 a distance at least* n *from* o *is*

$$N_p\left(\mathcal{G}, B_n, \delta_o\right) = \exp\left(\Theta\left(n\right)\right).$$

Proof. In order to apply Theorem 7.1 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 [12, Thms. 1 and 2]. $\square$

8. Open problems.
  • *Connections to maximum overhang problems.* Paterson et al. [10] studied the controlled diffusion problem on $\mathbb{Z}$ due to its connections to 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 ith block from the top extending by $1/(2i)$ beyond the block below it. This construction has an overhang of $\sum_{i=1}^{n} 1/(2i) \sim \frac{1}{2} \ln\left(n\right)$.

However, Paterson and Zwick showed that this belief is false, by constructing an example with overhang on the order of $n^{1/3}$ [11]. Subsequently, Paterson et al. showed that this is best possible up to a constant factor [10]. 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 paper 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 $p \to 0$?

REFERENCES

[1] I. Benjamini and Y. Peres, *Tree-indexed random walks on groups and first passage percolation*, Probab. Theory Related Fields, 98 (1994), pp. 91–112, https://doi.org/10.1007/BF01311350.

[2] S. Blachère, P. Haissinsky, and P. Mathieu, *Asymptotic entropy and Green speed for random walks on countable groups*, Ann. Probab., 36 (2008), pp. 1134–1152, https://doi.org/10.1214/07-AOP356.

[3] J. L. Doob, *Classical Potential Theory and Its Probabilistic Counterpart*, Grundlehren Math. Wiss. 262, Springer, New York, 1984.

[4] Y. Fukai and K. Uchiyama, *Potential kernel for two-dimensional random walk*, Ann. Probab., 24 (1996), pp. 1979–1992.

[5] V. Kaimanovich and A. Vershik, *Random walks on discrete groups: Boundary and entropy*, Ann. Probab., 11 (1983), pp. 457–490.

[6] G. F. Lawler, *Intersections of Random Walks*, Springer, Berlin, 1991.

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

[8] R. Lyons, R. Pemantle, and Y. Peres, *Ergodic theory on Galton-Watson trees: Speed of random walk and dimension of harmonic measure*, Ergodic Theory Dynam. Systems, 15 (1995), pp. 593–619.

[9] R. Lyons and Y. Peres, *Probability on Trees and Networks*, Cambridge University Press, New York, 2016; available at http://pages.iu.edu/~rdlyons/.

[10] M. Paterson, Y. Peres, M. Thorup, P. Winkler, and U. Zwick, *Maximum overhang*, Amer. Math. Monthly, 116 (2009), pp. 763–787.

[11] M. Paterson and U. Zwick, *Overhang*, in Proceedings of the 17th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), SIAM, Philadelphia, 2006, pp. 231–240.

[12] D. Revelle, *Heat kernel asymptotics on the lamplighter group*, Electron. Commun. Probab., 8 (2003), pp. 142–154.

[13] P. Révész, *Random Walk in Random and Non-Random Environments*, World Scientific, Hackensack, NJ, 2005.