Deterministic Random Walks on Regular Trees

Joshua Cooper,[1] Benjamin Doerr,[2] Tobias Friedrich,[2] Joel Spencer[3]

[1]*Department of Mathematics, LeConte College, University of South Carolina,
   Columbia, South Carolina 29208; e-mail: cooper@math.sc.edu*

[2]*Department 1: Algorithms and Complexity, Max-Planck-Institut für Informatik,
   Saarbrücken 66123, Germany; e-mail: firstname.lastname@mpg-inf.mpg.de*

[3]*Courant Institute of Mathematical Sciences, New York University, New York,
   New York 10012-1185; e-mail: spencer@cims.nyu.edu*

ABSTRACT: Jim Propp's rotor–router model is a deterministic analog of a random walk on a graph. Instead of distributing chips randomly, each vertex serves its neighbors in a fixed order.

Cooper and Spencer (Comb Probab Comput 15 (2006) 815–822) show a remarkable similarity of both models. If an (almost) arbitrary population of chips is placed on the vertices of a grid $\mathbb{Z}^d$ and does a simultaneous walk in the Propp model, then at all times and on each vertex, the number of chips on this vertex deviates from the expected number the random walk would have gotten there by at most a constant. This constant is independent of the starting configuration and the order in which each vertex serves its neighbors.

This result raises the question if all graphs do have this property. With quite some effort, we are now able to answer this question negatively. For the graph being an infinite k-ary tree ($k \geq 3$), we show that for any deviation D there is an initial configuration of chips such that after running the Propp model for a certain time there is a vertex with at least D more chips than expected in the random walk model. However, to achieve a deviation of D it is necessary that at least $\exp(\Omega(D^2))$ vertices contribute by being occupied by a number of chips not divisible by k at a certain time. © 2010 Wiley Periodicals, Inc.   Random Struct. Alg., 37, 353–366, 2010

Keywords: quasirandomness; discrepancy; rotor-router model

1. INTRODUCTION

The rotor–router model is a simple deterministic process first introduced by Priezzhev et al. [9] and later popularized by Jim Propp. It can be viewed as an attempt to derandomize

---

random walks on graphs. So far, the "Propp machine" has been studied primarily on infinite grids $\mathbb{Z}^d$. There, each vertex $x \in \mathbb{Z}^d$ is equipped with a "rotor" together with a cyclic permutation (called a "rotor sequence") of the $2d$ cardinal directions of $\mathbb{Z}^d$. While a chip (particle, coin, ...) performing a random walk leaves a vertex in a random direction, in the Propp model it always goes in the direction the rotor is pointing. After a chip is sent, the rotor is rotated according to the fixed rotor sequence. This rule ensures that chips are distributed highly evenly among the neighbors of a vertex.

The Propp machine has attracted considerable attention recently. It has been shown that it closely resembles a random walk in several respects. The first results were due to Levine and Peres [8] (and later Landau and Levine [7]) who compared random walk and Propp machine in an aggregating model called Internal Diffusion-Limited Aggregation (IDLA).

Cooper and Spencer [1] compared both models in terms of the single vertex discrepancy. Apart from a technicality, they place arbitrary numbers of chips on the vertices. Then they run the Propp machine on this initial configuration for a certain number of rounds. A round consists of each chip (in arbitrary order) performing one move as directed by the Propp machine. For the resulting chip arrangement, they compare the number of chips at each vertex with the expected number of chips that the random walk would have put there, starting from the same initial configuration and running for the same time. Cooper and Spencer showed that for all grids $\mathbb{Z}^d$, these differences can be bounded by a constant c_d independent of the initial setup (in particular, the total number of chips) and the run-time. For the case $d = 1$, that is, the graph being the infinite path, the optimal constant c_1 is ~ 2.29 [2], for the two-dimensional grid it is $c_2 \approx 7.87$ [3].

This raises the question whether the Propp machine on all graphs simulates the random walk that well. In this work, we show that the infinite k-regular tree behaves different. We prove that here arbitrarily large discrepancies can result from suitable initial configurations. However, to obtain a discrepancy of D, at least $\exp(\Omega(D^2))$ vertices have to participate by being occupied by a number of chips not divisible by k at some time.

While the work cited above and ours in this paper primarily aims at understanding random walks and their deterministic counterparts from a foundations perspective, we would like to mention that meanwhile the rotor–router mechanism also led to improvements in applications. An example is the quasirandom analog of the randomized rumor spreading protocol to broadcast information in networks [4].

2. PRELIMINARIES

To bound the single vertex discrepancy between the Propp machine and a random walk on the k-regular tree, we first introduce several requisite definitions and notational conventions.

Let $G = (V, E)$ be the infinite k-regular tree, also known as the Cayley tree and the Bethe lattice. We fix an arbitrary node to be its origin **0**. $|\mathbf{x}|$ denotes the distance between the origin and vertex **x**.

We first describe the Propp machine in detail. Each vertex **x** is equipped with a rotor pointing to one of its neighbors. If **A** is the "direction" the rotor is pointing to, then we denote this neighbor by $\mathbf{x} + \mathbf{A}$. This notation is in analogy to the one used in the grid case. Though we keep the additive notation, here the directions would rather be the generators of the nonabelian group generated by the expression $\langle \{d_i\}_{i=1}^k \mid d_i^2 = 1 \rangle$, whose Cayley graph is an infinite k-regular tree. Denote by DIR $:= \{d_1, \ldots, d_k\}$ the set of generators ("directions").

Each vertex also holds a rule describing how the rotor is moved. This rule is encoded by a cyclic permutation of the directions called rotor sequence. Again, if the rotor of vertex $\mathbf{x}$ is pointing in direction $\mathbf{A}$, then we denote by $\text{NEXT}(\mathbf{A}) := \text{NEXT}(\mathbf{A}, \mathbf{x})$ the new direction after one move. Since there will be no danger of confusion, we will always omit the explicit mention of the vertex. Nevertheless, we should stress that we do not require all vertices to have the same rotor sequence.

Chips move according to the rotors. That is, if at a certain time t there is a single chip on vertex $\mathbf{x}$ (with rotor pointing to $\mathbf{A}$), then at time $t + 1$ this chips is on vertex $\mathbf{x} + \mathbf{A}$, the neighbor the rotor pointed to, and the rotor is updated to point in direction $\text{NEXT}(\mathbf{A})$. If there are more than one chip on a vertex, they all move in one round, but applying the rotor principle one after the other (in an arbitrary order—we do not care about individual chips but only about the number of chips present on each vertex).

With these rules, the Propp machine describes a process that is fully determined by the initial setting of the chips and the positions of the rotors. For all $\mathbf{x} \in V$ and $t \in \mathbb{N}_0$, we denote by $f(\mathbf{x}, t)$ the number of chips on vertex $\mathbf{x}$ and $\text{ARR}(\mathbf{x}, t)$ the direction of the rotor associated with $\mathbf{x}$ after t steps of the Propp machine. Hence, $f(\cdot, 0)$ and $\text{ARR}(\cdot, 0)$ describe the initial configuration, which determines all other values of f and ARR.

For completeness, let us briefly review the random walk model. Here, we do not have rotors. Instead, in each round, each chip independently and uniformly at random chooses a neighbor of his current position and moves to there. This process is described by the initial setting of the chips together with all random decisions made.

However, since we shall be only interested in the expected number of chips on the vertices, we may instead regard the following linear machine [1]. Here, in each time step every vertex splits its pile of chips evenly. When a pile of ℓ chips splits evenly on some vertex, ℓ/k chips go to each neighbor. By the harmonic property of random walks, the (possibly non-integral) number of chips at vertex $\mathbf{x}$ at time t is exactly the expected number of chips in the random walk model.

Note that the linear machine again is a deterministic process fully described by the initial numbers of chips on each vertex. Given such an initial configuration, we denote by $E(\mathbf{x}, t)$ the (fractional) number of chips at time t on vertex $\mathbf{x}$. Note that, again, $E(\cdot, 0)$ is just this initial configuration. Note further that $E(\mathbf{x}, t) = \frac{1}{k} \sum_{\mathbf{A} \in \text{DIR}} E(\mathbf{x} + \mathbf{A}, t - 1)$ for all $t \in \mathbb{N}$ and $\mathbf{x} \in V$.

To compare the Propp machine with the linear machine, we start both with identical settings of the chips. Hence, we have $f(\mathbf{x}, 0) = E(\mathbf{x}, 0)$ for all $\mathbf{x} \in V$. A configuration describes the current state of machine we regard. A configuration of the Propp machine assigns to each vertex $\mathbf{x} \in V$ its current (integral) number of chips and the current direction of the rotor. For the linear machine, a configuration is a simple mapping $V \to \mathbb{Q}_{\geq 0}$, describing the (fractional) number of chips on each vertex.

As pointed out in the introduction, there is one limitation without which neither the results of [1–3] nor our results hold. Note that since G is a bipartite graph, chips that start on even vertices never mix with those starting on odd vertices. It looks as if we are playing two noninteracting games at once. However, this is not true. Chips in different bipartition classes may affect each other through the rotors. We, therefore, require the initial configuration to have chips only on one class of the bipartition. Without loss of generality, we consider only even initial configurations, i. e., chip configurations supported on vertices at an even distance from the origin.

To analyze the behavior of the linear machine, we need the following notation. By $H(x, t)$ we denote the probability that a chip from a vertex with distance x to the origin arrives at the origin after t random steps ("at time t") in a simple random walk. Then,

$$H(x,t) = k^{-t}n(x,t) \tag{1}$$

with $n(x,t)$ counting the number of paths of length t between two vertices at distance x on the infinite k-regular tree. It is easy to verify the following properties of $n(x,t)$:

$$n(0,0) = 1,$$
$$n(x,0) = 0 \quad \text{for all } x \geq 1,$$
$$n(0,t) = kn(1,t-1) \quad \text{for all } t \geq 1,$$
$$n(x,t) = n(x-1,t-1) + (k-1)n(x+1,t-1) \quad \text{for all } x, t \geq 1. \tag{2}$$

Finally, we write $\mathbf{x} \sim t$ to mean that $|\mathbf{x}| \equiv t \pmod{2}$.

3. MOD-K-FORCING THEOREM

For a deterministic process like the Propp machine, it is obvious that the initial configuration (that is, the location of each chip and the direction of each rotor), determines all subsequent configurations. The following theorem shows a partial converse, namely that (roughly speaking) we may prescribe the number of chips modulo k on all vertices at all times by finding an appropriate initial configuration. An analogous result for the one-dimensional Propp machine has been shown in [2].

Theorem 1 (Mod-k-forcing Theorem). *For all initial directions of the rotors and any $\pi: V \times \mathbb{N}_0 \rightarrow \{0, 1, \ldots, (k-1)\}$ with $\pi(\mathbf{x}, t) = 0$ for all $\mathbf{x} \nsim t$, there is an initial even configuration $f(\mathbf{x}, 0)$ that results in subsequent configurations satisfying $f(\mathbf{x}, t) \equiv \pi(\mathbf{x}, t) \pmod{k}$ for all $\mathbf{x}$ and $t \geq 0$.*

Proof. Fix arbitrary initial directions for the rotors. Let us call f, or more precisely, an initial configuration $(f(\mathbf{x}, 0))_{\mathbf{x} \in V}$ fine for a subset $S \subseteq V \times \mathbb{N}_0$,

- if it yields $f(\mathbf{x}, t) \equiv \pi(\mathbf{x}, t) \pmod{k}$ for all $(\mathbf{x}, t) \in S$, and
- if it is even, that is, for all $\mathbf{x} \in V$ we have $f(\mathbf{x}, 0) = 0$ if $\mathbf{x} \nsim 0$.

Consider first the initial configuration defined by $f^{(0)}(\mathbf{x}, 0) := \pi(\mathbf{x}, 0)$ for all $\mathbf{x} \in V$. Clearly, $f^{(0)}$ is fine for $V \times \{0\}$.

Assume now that there is a $T \geq 0$ and an initial configuration $f^{(T)}$ which is fine for $V \times \{0, \ldots, T\}$. In order to obtain an $f^{(T+1)}$ that is fine for $V \times \{0, \ldots, T+1\}$, we shall modify $f^{(T)}$ by defining $f^{(T+1)}(\mathbf{x}, 0) := f^{(T)}(\mathbf{x}, 0) + \varepsilon(\mathbf{x}) k^{T+1}$ for appropriately chosen $\varepsilon(\mathbf{x}) \in \{0, 1, \ldots, (k-1)\}$ with $\varepsilon(\mathbf{x}) = 0$ if $\mathbf{x} \nsim 0$.

Observe that a pile of k^{T+1} chips splits evenly $T+1$ times. Therefore, for all choices of the $\varepsilon(\mathbf{x}), f^{(T+1)}(\mathbf{x}, t) \equiv f^{(T)}(\mathbf{x}, t) \pmod{k}$ for all $(\mathbf{x}, t)$ with $t \leq T$. This implies that $f^{(T+1)}$ is fine for $V \times \{0, \ldots, T\}$.

We shall define ε inductively such that $f^{(T+1)}$ is also fine for $V \times \{T+1\}$. Let $\varepsilon^{(0)}(\mathbf{x}) := 0$ for all $\mathbf{x} \in V$. Assume that for some $\theta \geq 0$, $\varepsilon^{(\theta)}$ is such that the initial configuration $f^{(\theta)}(\mathbf{x}, 0) := f^{(T)}(\mathbf{x}, 0) + \varepsilon^{(\theta)}(\mathbf{x}) k^{T+1}$ is fine for all $(\mathbf{x}, T+1)$ with $|\mathbf{x}| < \theta$. Trivially, this is fulfilled for $\theta = 0$. For each $\mathbf{x}$ with $|\mathbf{x}| = \theta$, we choose some $\mathbf{y}(\mathbf{x})$ with $|\mathbf{y}(\mathbf{x})| = T + 1 + \theta$ such that the distance between $\mathbf{x}$ and $\mathbf{y}(\mathbf{x})$ is $T + 1$. Note that this implies that the mapping $\mathbf{x} \mapsto \mathbf{y}(\mathbf{x})$ is injective. We define $\varepsilon^{(\theta+1)}$ as follows. For all $\mathbf{x}$ with $|\mathbf{x}| = \theta$, let

$\varepsilon^{\langle\theta+1\rangle}(\mathbf{y}(\mathbf{x})) := (\pi(\mathbf{x}, T+1) - f^{\langle\theta\rangle}(\mathbf{x}, T+1)) \bmod k$. For all $\mathbf{z} \in V \setminus \{\mathbf{y}(\mathbf{x}): |\mathbf{x}| = \theta\}$, we set $\varepsilon^{\langle\theta+1\rangle}(\mathbf{z}) := \varepsilon^{\langle\theta\rangle}(\mathbf{z})$. We first argue that $f^{\langle\theta+1\rangle}$ is an even initial configuration. As $f^{\langle\theta\rangle}$ is an even initial configuration, it suffices to show that for all $\mathbf{x}$ with $|\mathbf{x}| = \theta$ and $\mathbf{y}(\mathbf{x}) \not\sim 0$, $\varepsilon^{\langle\theta+1\rangle}(\mathbf{y}(\mathbf{x})) = 0$. However, if $|\mathbf{y}(\mathbf{x})|$ is odd, then $|\mathbf{x}|$ and $T+1$ have different parity. Hence, $\pi(\mathbf{x}, T+1) = 0$ and $f^{\langle\theta\rangle}(\mathbf{x}, T+1) = 0$, and by construction, $\varepsilon^{\langle\theta+1\rangle}(\mathbf{y}(\mathbf{x})) = 0$. Moreover, for all $\mathbf{x}$ with $|\mathbf{x}| = \theta$,

$$f^{\langle\theta+1\rangle}(\mathbf{x}, T+1) = f^{\langle\theta\rangle}(\mathbf{x}, T+1) + \varepsilon^{\langle\theta+1\rangle}(\mathbf{y}(\mathbf{x})) \equiv \pi(\mathbf{x}, T+1) \ (\bmod k)$$

and for all $\mathbf{x}$ with $|\mathbf{x}| < \theta$,

$$f^{\langle\theta+1\rangle}(\mathbf{x}, T+1) = f^{\langle\theta\rangle}(\mathbf{x}, T+1) \equiv \pi(\mathbf{x}, T+1) \ (\bmod k).$$

Therefore, $f^{\langle\theta+1\rangle}$ is fine for all $(\mathbf{x}, T+1)$ with $|\mathbf{x}| < \theta+1$.

For all vertices $\mathbf{x}$ and all $\theta \geq \max(0, |\mathbf{x}| - T)$, $\varepsilon^{\langle\theta\rangle}(\mathbf{x}) = \varepsilon^{\langle\theta+1\rangle}(\mathbf{x})$. Hence, $\varepsilon(\mathbf{x}) := \lim_{\theta\to\infty} \varepsilon^{\langle\theta\rangle}(\mathbf{x})$ and $f^{(T+1)}(\mathbf{x}, 0) := f^{(T)}(\mathbf{x}, 0) + \varepsilon(\mathbf{x}) k^{T+1}$ are well-defined for all $\mathbf{x} \in V$. Also, $f^{(T+1)}$ is fine for $V \times \{0, \ldots, T+1\}$.

We now note that for all vertices $\mathbf{x} \in V$ and for all $T \geq |\mathbf{x}|$, we have that $f^{(T)}(\mathbf{x}, 0) = f^{(T+1)}(\mathbf{x}, 0)$. Hence again, $f(\mathbf{x}, 0) := \lim_{T\to\infty} f^{(T)}(\mathbf{x}, 0)$ is well-defined for all $\mathbf{x} \in V$. Therefore, f is fine for $V \times \mathbb{N}_0$, which finishes the proof. ∎

4. THE BASIC METHOD

In this section, we lay the foundations for our analysis of the maximal possible single-vertex discrepancy. Closely following the arguments of [1], we will see that it is possible to determine the contribution of a vertex to the discrepancy at another one independent from all other vertices.

For the moment, in addition to the notations given in Section 2, we also use the following mixed notation. By $E(\mathbf{x}, t_1, t_2)$ we denote the (possibly fractional) number of chips at location x after first performing t_1 steps with the Propp machine and then $t_2 - t_1$ steps with the linear machine.

We are interested in bounding the discrepancies $|f(\mathbf{x}, t) - E(\mathbf{x}, t)|$ for all vertices $\mathbf{x}$ and all times t. It suffices to consider the vertex $\mathbf{x} = \mathbf{0}$. From

$$E(\mathbf{0}, 0, t) = E(\mathbf{0}, t),$$
$$E(\mathbf{0}, t, t) = f(\mathbf{0}, t),$$

we obtain

$$f(\mathbf{0}, t) - E(\mathbf{0}, t) = \sum_{s=0}^{t-1} (E(\mathbf{0}, s+1, t) - E(\mathbf{0}, s, t)).$$

Let $|\mathbf{x}|$ denote the distance of a vertex $\mathbf{x}$ to $\mathbf{0}$. Now $E(\mathbf{0}, s+1, t) - E(\mathbf{0}, s, t) = \sum_{\mathbf{x}\in V} \sum_{\ell=1}^{f(\mathbf{x},s)} (H(|\mathbf{x} + \text{NEXT}^{\ell-1}(\text{ARR}(\mathbf{x}, s))|, t - s - 1) - H(|\mathbf{x}|, t - s))$ motivates the definition of the influence of a Propp move (compared to a random walk move) from vertex $\mathbf{x}$ in direction $A \in \{-1, +1\}$ on the discrepancy of $\mathbf{0}$ (t time steps later) by

$$\text{INF}(x, A, t) := H(x + A, t - 1) - H(x, t).$$

In order to ultimately reduce all ARRs involved to the initial arrow settings $\text{ARR}(\cdot, 0)$, we define $s_i(\mathbf{x}) := \min\{u \geq 0 \mid i < \sum_{t=0}^{u} f(\mathbf{x}, t)\}$ for all $i \in \mathbb{N}_0$. Hence at time $s_i(\mathbf{x})$ the location $\mathbf{x}$ is occupied by its ith chip (where, to be consistent with [2], we start counting with the 0th chip).

Consider the discrepancy at $\mathbf{0}$ at time T. Then the above yields

$$f(\mathbf{0}, T) - E(\mathbf{0}, T) = \sum_{\mathbf{x} \in V} \sum_{\substack{i \geq 0, \\ s_i(\mathbf{x}) < T}} \text{INF}(|\mathbf{x}|, \text{NEXT}^i(\text{ARR}(\mathbf{x}, 0)), T - s_i(\mathbf{x})). \qquad (3)$$

Since the inner sum of Eq. (3) will occur frequently in the remainder, let us define the contribution of a vertex $\mathbf{x}$ to be

$$\text{CON}(\mathbf{x}) := \sum_{\substack{i \geq 0, \\ s_i(\mathbf{x}) < T}} \text{INF}\big(|\mathbf{x}|, \text{NEXT}^i(\text{ARR}(\mathbf{x}, 0)), T - s_i(\mathbf{x})\big),$$

where we both suppress the initial configuration leading to the $s_i(\cdot)$ as well as the run-time T.

We summarize the discussion so far in the following theorem. It shows that it suffices to examine each vertex $\mathbf{x}$ separately.

Theorem 2. *The discrepancy between Propp machine and linear machine after T time steps is the sum of the contributions* $\text{CON}(\mathbf{x})$ *of all vertices $\mathbf{x}$, i. e.,*

$$f(\mathbf{0}, T) - E(\mathbf{0}, T) = \sum_{\mathbf{x} \in V} \text{CON}(\mathbf{x}).$$

5. DIVERGENCE OF THE MODELS

In this section, we analyze a specific initial configuration and show that the Propp machine may deviate from the linear machine by an arbitrarily large number of chips.

We choose the initial directions of the rotors to point inwards in the direction of the origin. For a fixed time T at which we aim to maximize the discrepancy $f(\mathbf{0}, T) - E(\mathbf{0}, T)$ we examine a configuration in which all vertices $\mathbf{x}$ with $0 < |\mathbf{x}| \leq T/\lambda$ and $\lambda := \frac{k}{k-2}$ are occupied by a number of chips not divisible by k only once. We assume that at time $T - t_{|\mathbf{x}|}$ with $t_{|\mathbf{x}|} := \llbracket \lambda |\mathbf{x}| \rrbracket$[1] an "odd chip" is sent in the direction of the origin. That is, the number of chips at $\mathbf{x}$ at time $t_{|\mathbf{x}|}$ is $f(\mathbf{x}, t_{|\mathbf{x}|}) \equiv 1 \pmod{k}$ while $f(\mathbf{x}, t) \equiv 0 \pmod{k}$ for all times $t \neq t_{|\mathbf{x}|}$. For the given initial direction of the rotors such a configuration exists by Theorem 1. We will prove the following theorem.

Theorem 3. *For any initial direction of the rotors and any $T > 0$, there is an even initial configuration such that the single vertex discrepancy between the Propp machine and linear machine after T time steps is $\Omega(\sqrt{k\,T})$.*

---

[1]With $\llbracket \lambda |\mathbf{x}| \rrbracket$ we denote the smallest integer t with $t \geq \lambda |\mathbf{x}|$ and $\mathbf{x} \sim T - t$.

By Theorem 2, the discrepancy at the origin at time T of the above described initial configuration is

$$f(\mathbf{0}, T) - E(\mathbf{0}, T) = \sum_{\substack{\mathbf{x} \in V, \\ |\mathbf{x}| \leq T/\lambda}} (H(|\mathbf{x}| - 1, t_{|\mathbf{x}|} - 1) - H(|\mathbf{x}|, t_{|\mathbf{x}|}))$$

$$= \sum_{x=1}^{\lfloor T/\lambda \rfloor} k(k-1)^{x-1}(H(x-1, t_x - 1) - H(x, t_x))$$

$$= \sum_{x=1}^{\lfloor T/\lambda \rfloor} k(k-1)^{x-1} \left(\frac{n(x-1, t_x - 1)}{k^{(t_x - 1)}} - \frac{n(x, t_x)}{k^{t_x}} \right)$$

$$= \sum_{x=1}^{\lfloor T/\lambda \rfloor} \frac{(k-1)^{x-1}}{k^{t_x - 1}}(k\, n(x-1, t_x - 1) - n(x, t_x))$$

$$= \sum_{x=1}^{\lfloor T/\lambda \rfloor} \frac{(k-1)^{x-1}}{k^{t_x - 1}}\, i(x, t_x) \tag{4}$$

with

$$i(x, t) := k\, n(x-1, t-1) - n(x, t)$$

for all $x, t \geq 1$. Let us define $i(x, t) = 0$ otherwise. For $x, t \geq 2$ we get

$$i(x, t) = kn(x-1, t-1) - n(x, t)$$
$$= kn(x-2, t-2) + k(k-1)n(x, t-2) - n(x-1, t-1) - (k-1)n(x+1, t-1)$$
$$= i(x-1, t-1) + (k-1)i(x+1, t-1).$$

It remains to examine the cases $(x, t) \in (\mathbb{N} \times \mathbb{N}) \setminus (\mathbb{N}_{\geq 2} \times \mathbb{N}_{\geq 2})$. Right from the definition, we get $i(1, 1) = k - 1$. For $x \geq 2$ and $t = 1$ we have

$$i(x, 1) = kn(x-1, 0) - n(x, 1)$$
$$= 0$$
$$= i(x-1, 0) + (k-1)i(x+1, 0).$$

Also, for $x = 1$ and $t \geq 2$ we have

$$i(1, t) = kn(0, t-1) - n(1, t)$$
$$= (k-1)\,(n(0, t-1) - n(2, t-1))$$
$$= (k-1)\,(kn(1, t-2) - n(2, t-1))$$
$$= i(0, t-1) + (k-1)\,i(2, t-1).$$

Summarizing the above, we see that $i(x, t)$ can be defined recursively as follows.

$i(x, 0) = 0 \quad$ for all $x \geq 0$,

$i(0, t) = 0 \quad$ for all $t \geq 0$,

$i(1, 1) = k - 1$,

$i(x, t) = i(x-1, t-1) + (k-1)\,i(x+1, t-1) \quad$ for $(x, t) \in \mathbb{N}_{\geq 1}^2 \setminus \{(1, 1)\}$.

This recursive view of $i(x, t)$ reveals another interpretation of these quantities. Apart from a factor $(k - 1)^{(\frac{t-x}{2}+1)}$, $i(x, t)$ counts the number of lattice paths from the origin $(0, 0)$ to (x, t) of steps $(+1, +1)$ and $(-1, +1)$, which do not cross the line $x = 0$. This can be described by the well-known Ballot numbers. The classical description is as follows. Suppose A and B are candidates for president. Let A receive a total number of a votes and B one of b votes. Let $a \geq b$. Now consider the progress of counting the votes, one after the other. Then the probability that throughout the counting B never has more votes than A is $(a - b + 1)/(a + 1)$ [6]. This implies that for given positive integers a, b with $a > b$, the number of lattice paths starting at the origin and consisting of a upsteps $(+1, +1)$ and b downsteps $(+1, -1)$ such that no step ends on the x-axis is $\frac{a-b}{a+b}\binom{a+b}{a}$. We are interested in the number of lattice paths starting at $(0, x)$ and consisting of $(t - x)/2$ upsteps $(+1, +1)$ and $(t + x)/2$ downsteps $(+1, -1)$ such that no step ends below the x-axis. Therefore,

$$i(x, t) = (k - 1)^{\left(\frac{t-x}{2}+1\right)} \frac{x}{t}\binom{t}{\frac{t+x}{2}} \tag{5}$$

for $x, t > 0$. Note that this can be continued to all $(x, t) \in (\mathbb{R}_+)^2$.

Here is a more intuitive interpretation why the Ballot numbers come into play. We want the effect of having one chip at distance x and time $t := t_x$ make its first move toward the origin. So, we can reduce to comparing two chips, one at distance $x - 1$ and the other at distance $x + 1$ at time $t - 1$.

We consider the two chips on a common random data set. The data set is a string s of length $t - 1$ from $\{I, O\}$ with probabilities $\frac{1}{k}$, $\frac{k-1}{k}$ for I (in, toward the origin), O (out, away from the origin), respectively. The string s determines where the chip goes (in terms of how far away it is from the origin, which is what concerns us): with O the distance x goes up by one and with I it goes down by one unless the current distance is zero in which case it goes up by one.

Say a string s has the property MINUS if for some initial segment the I's exceed the O's by precisely x. In a "fictitious walk," allowing distance x to go negative, the chip starting at x would reach -1. But consider just before this happens for the first time. The two chips are at $0, 2$, respectively, and an I is received. Now the two chips are both at 1. Hence, whatever comes after, they end up at the same point. Thus the two chips have the same probability of MINUS and ending at zero. The times that the near chip ends at zero but the far chip does not are then precisely the strings s which do not have MINUS and for which the I's exceed the O's by precisely $x - 1$. And as the far chip is never closer than the near chip, it never happens that the far chip ends at zero but the near chip does not. Thus the difference in the probabilities is given by the Ballot problem as calculated above.

Equations (4) and (5) now give

$$f(\mathbf{0}, T) - E(\mathbf{0}, T) = \sum_{x=1}^{\lfloor T/\lambda \rfloor} \frac{(k-1)^{x-1}}{k^{(t_x-1)}} (k-1)^{\left(\frac{t_x-x}{2}+1\right)} \frac{x}{t_x}\binom{t_x}{\frac{t_x+x}{2}}$$

$$= \sum_{x=1}^{\lfloor T/\lambda \rfloor} \frac{(k-1)^{\left(\frac{t_x+x}{2}\right)}}{k^{(t_x-1)}} \frac{x}{t_x}\binom{t_x}{\frac{t_x+x}{2}}$$

$$> \sum_{x=1}^{\lfloor T/\lambda \rfloor} \frac{(k-1)^{\frac{\lambda+1}{2}x}}{2\lambda\, k^{\lambda x}}\binom{\llbracket \lambda x\rrbracket}{\frac{\llbracket \lambda x\rrbracket+x}{2}}. \tag{6}$$

It remains to bound the binomial coefficient. From the Stirling formula we know that

$$\frac{5}{2} \sqrt{n} \left(\frac{n}{e}\right)^n < n! < \sqrt{\frac{15}{2}} \sqrt{n} \left(\frac{n}{e}\right)^n \tag{7}$$

for $n > 0$ and get

$$\frac{n^{n+1/2}}{3 \, (n-k)^{n-k+1/2} \, k^{k+1/2}} < \binom{n}{k} < \frac{n^{n+1/2}}{2 \, (n-k)^{n-k+1/2} \, k^{k+1/2}}. \tag{8}$$

for $n, k > 0$. Note that Eq. (8) also holds for the generalized binomial coefficient which is defined for nonintegral n, k by using the Gamma function Γ. The following lemma shows that inequality (6) also holds without the ceiling function in the binomial coefficient.

Lemma 4. $\binom{2x}{x+y}$ *is monotonic nondecreasing in x for* $0 \leq y \leq x$.

Proof. With Ψ denoting the logarithmic derivative of the gamma function, i. e., the Digamma function, we get

$$\frac{\partial \binom{2x}{x+y}}{\partial x} = (2 \, \Psi(2x+1) - \Psi(x+y+1) - \Psi(x-y+1)) \binom{2x}{x+y}.$$

The lemma follows by the fact that above binomial coefficient is positive for $x, y \geq 0$ and that $\Psi(x)$ is monotonic nondecreasing for $x > 0$ ([5, Theorem 7]). ∎

By the definitions of λ and t_x we now get with Lemma 4 and Eq. (8)

$$\binom{\llbracket \lambda x \rrbracket}{(\llbracket \lambda x \rrbracket + x)/2} \geq \binom{\lambda x}{\frac{\lambda+1}{2} x} = \binom{\frac{k}{k-2} x}{\frac{1}{k-2} x} > \frac{k^{\frac{k}{k-2} x + \frac{1}{2}} (k-2)^{\frac{1}{2}}}{3 \, (k-1)^{\frac{k-1}{k-2} x + \frac{1}{2}} \sqrt{x}} \tag{9}$$

for $x > 0, k > 2$.

Using this we obtain for all $k > 2$

$$f(\mathbf{0}, T) - E(\mathbf{0}, T) > \sum_{x=1}^{\lfloor T/\lambda \rfloor} \frac{(k-2)^{\frac{3}{2}}}{6 \, k^{\frac{1}{2}} (k-1)^{\frac{1}{2}} \sqrt{x}} = \Omega(\sqrt{k \, T}),$$

which proves Theorem 3. Using the same arguments as the following two sections one can also prove that for all even initial configurations the single vertex discrepancy after T time steps is at most $O(\sqrt{kT})$.

6. CONVERGENCE OF THE MODELS

The previous section showed that for very special configurations the single vertex discrepancy can be unbounded. In this section, we show that, on the other hand, many configurations have a bounded discrepancy. We will prove the following theorem.

Theorem 5. *If $f(\mathbf{x}, t) \equiv 0 \pmod{k}$ for all $\mathbf{x}$ and t such that $(1-\varepsilon)\lambda|\mathbf{x}| < T-t < (1+\varepsilon)\lambda|\mathbf{x}|$ with $\lambda := \frac{k}{k-2}$, then the discrepancy between Propp machine and linear machine at time T and vertex 0 is bounded by a constant depending only on k and $\varepsilon > 0$.*

Our aim is to bound Eq. (3). To that end, we further examine INF. From its definition and Eqs. (2) and (5) we know for $x, t > 0$

$$\text{INF}(x, -1, t) = \frac{i(x, t)}{k^t},$$

$$\text{INF}(x, +1, t) = \frac{kn(x + 1, t - 1) - n(x, t)}{k^t}$$

$$= \frac{\frac{1}{k-1}n(x, t) - \frac{k}{k-1}n(x + 1, t - 1)}{k^t}$$

$$= \frac{-i(x, t)}{(k-1)\,k^t}.$$

This also shows

$$\text{INF}(x, -1, t) + (k - 1)\,\text{INF}(x, 1, t) = 0. \tag{10}$$

Therefore, the absolute value of the influence $|\text{INF}(x, A, t)|$ of sending one chip towards $\mathbf{0}$ is $(k-1)$ times larger than sending one chip in the opposite direction.

Note that

$$\text{CON}(\mathbf{x}) = \sum_{\substack{i \geq 0, \\ s_i(\mathbf{x}) < T}} A^{(i)} \frac{i(|\mathbf{x}|, t_i)}{k^{t_i}}$$

with

$$t_i := T - s_i(\mathbf{x})$$

$$A^{(i)} := \begin{cases} \dfrac{-1}{k - 1} & \text{for NEXT}^i(\text{ARR}(\mathbf{x}, 0)) = +1 \\ 1 & \text{for NEXT}^i(\text{ARR}(\mathbf{x}, 0)) = -1. \end{cases}$$

To bound this alternating sum, we use the following elementary fact.

Lemma 6. *Let $f : X \to \mathbb{R}$ be non-negative and monotone nondecreasing with $X \subseteq \mathbb{R}$. Let $A^{(0)}, \ldots, A^{(n)} \in \mathbb{R}$ and $t_0, \ldots, t_n \in X$ such that $t_0 \leq \cdots \leq t_n$ and $|\sum_{i=a}^{b} A^{(i)}| \leq 1$ for all $0 \leq a \leq b \leq n$. Then*

$$\left| \sum_{i=0}^{n} A^{(i)} f(t_i) \right| \leq \max_{x \in X} f(x).$$

Let $X \subseteq \mathbb{R}$. We call a mapping $f : X \to \mathbb{R}$ unimodal, if there is a $t_1 \in X$ such that $f|_{x \leq t_1}$ as well as $f|_{x \geq t_1}$ are monotone. The following lemma shows that $\text{INF}(x, A, t)$ is unimodal.

Lemma 7. *The function $i(x, t)/k^t$ is unimodal in t for all x and k, and it is maximized over all $t \in \mathbb{N}$ at some $t_{\max}(x) = \lambda x + c_x$ with $|c_x| \leq 15$.*

Proof. Using $\binom{n+2}{k+1} = \frac{n^2+3n+2}{(k+1)(n-k+1)}\binom{n}{k}$ we get

$$\frac{i(x,t+2)}{k^{t+2}} - \frac{i(x,t)}{k^t} = \frac{(k-1)^{\frac{t-x}{2}} x \, p_x(t)}{(t+x+2)(t-x+2)k^{t+2}t}\binom{t}{\frac{t+x}{2}}$$

with $p_x(t) := -(k^3 - 5k^2 + 8k - 4)t^2 - (4k^3 - 8k^2 + 8k - 4)t + (k^3 - k^2)(x^2 - 4)$. Hence, the above difference is non-negative if

$$t \geq (\sqrt{8k^3 - 4k^2 - 8k + 4 + k^2(k-2)^2 x^2} - 2k^2 + 2k - 2)/(k-2)^2 \qquad (11)$$

and non-positive otherwise. Thus, we have unimodality with $i(x,t)/k^t$ taking its maximum when t is the smallest integer having the same parity as x and satisfying Eq. (11). Using $k \geq 3$, the claim follows. ∎

Armed with Lemmas 6 and 7, we can now prove Theorem 5.

Proof of Theorem 5. By Lemmas 6 and 7 we know

$$\mathrm{CON}(\mathbf{x}) \leq \frac{i(|\mathbf{x}|, \lfloor(1-\varepsilon)\lambda|\mathbf{x}|\rfloor)}{k^{\lfloor(1-\varepsilon)\lambda|\mathbf{x}|\rfloor}} + \frac{i(|\mathbf{x}|, \lceil(1+\varepsilon)\lambda|\mathbf{x}|\rceil)}{k^{\lceil(1+\varepsilon)\lambda|\mathbf{x}|\rceil}}.$$

By Theorem 2 it therefore remains to show that

$$f(\mathbf{0},T) - E(\mathbf{0},T) \leq \sum_{x>0}\left(\frac{(k-1)^{x-1}}{k^{\lfloor(1-\varepsilon)\lambda x\rfloor-1}} i(x, \lfloor(1-\varepsilon)\lambda x\rfloor) + \frac{(k-1)^{x-1}}{k^{\lceil(1+\varepsilon)\lambda x\rceil-1}} i(x, \lceil(1+\varepsilon)\lambda x\rceil)\right)$$

is bounded. We now show that the second summand is bounded, the first one can be handled analogously. For this, we choose ε' such that $(1+\varepsilon')\lambda x = \lceil(1+\varepsilon)\lambda x\rceil$. By Eq. (5),

$$\frac{(k-1)^{x-1}}{k^{(1+\varepsilon')\lambda x-1}} i(x, (1+\varepsilon')\lambda x) = \frac{(k-1)^{((1+\varepsilon')\lambda+1)x/2} x}{k^{(1+\varepsilon')\lambda x-1}((1+\varepsilon')\lambda x)}\binom{(1+\varepsilon')\lambda x}{((1+\varepsilon')\lambda+1)x/2}$$

With Eq. (8) we get,

$$\frac{(k-1)^{x-1}}{k^{(1+\varepsilon')\lambda x-1}} i(x, (1+\varepsilon')\lambda x) < \sqrt{\frac{k\,(k-2)^3}{(1+\varepsilon')(k\varepsilon'+2)(2k-2+k\varepsilon')}}\, p(\varepsilon')^{\frac{x}{k-2}}$$

with

$$p(\varepsilon') := \frac{2k-2+k\varepsilon'}{(k\varepsilon'+2)(k-1)}\left(\frac{k-1}{(k\varepsilon'+2)(2k-2+k\varepsilon')}\right)^{k\varepsilon'/2}\left(\frac{(k-1)(2+2\varepsilon')^{\varepsilon'+1}}{2k-2+k\varepsilon'}\right)^k$$

As $0 < p(\varepsilon') < 1$ for all $\varepsilon' > 0$ and $k \geq 3$, we have shown that $\sum_{x>0}\frac{(k-1)^{x-1}}{k^{(1+\varepsilon')\lambda x-1}} i(x, (1+\varepsilon')\lambda x)$ is bounded above by a constant depending on $\varepsilon' > 0$. The same arguments demonstrate that $\sum_{x>0}\frac{(k-1)^{x-1}}{k^{\lfloor(1-\varepsilon)\lambda x\rfloor-1}} i(x, \lfloor(1-\varepsilon)\lambda x\rfloor)$ can also be bounded above by a constant depending on $\varepsilon > 0$, which finishes the proof. ∎

7. NUMBER OF CHIPS

In the previous sections, we have shown that the discrepancy on a single vertex can be arbitrarily high provided we use a sufficient amount of time. However, a closer look at the proof also reveals that, assuming k to be constant, to obtain a discrepancy of D, at least $\exp(\Omega(D^2))$ vertices have to contribute (by holding a number of chips that is not a multiple of k at a suitable time). We now discuss that also $\exp(\Omega(D^2))$ chips are necessary to obtain a discrepancy of D.

Theorem 8. *Let $k \geq 3$ be a constant. Then all single vertex discrepancies arising from an even initial configuration with κ chips are bounded by $O(\sqrt{\log \kappa})$.*

Proof. Let us examine the discrepancy at the origin after T time steps, starting with an arbitrary even initial configuration which uses only κ chips. We bound the contribution of each sphere $S_x := \{\mathbf{x} \in V \mid |\mathbf{x}| = x\}, x \in \mathbb{N}_0$, separately.

Lemma 7 shows that $\mathrm{INF}(x, -1, \cdot)$ is unimodal. By Lemma 6, this gives

$$\mathrm{CON}(S_x) := \sum_{\mathbf{x} \in S_x} \mathrm{CON}(\mathbf{x})$$

$$\leq \sum_{\mathbf{x} \in S_x} \max_t \mathrm{INF}(|\mathbf{x}|, -1, t)$$

$$= k(k-1)^{x-1} \max_t \frac{i(x,t)}{k^t} \tag{12}$$

By Eq. (5) and Lemma 7, which showed that $i(x,t)/k^t$ is maximized at $t = \lambda x + c_x$ with $|c_x| \leq 15$, we obtain

$$\max_t \frac{i(x,t)}{k^t} = \frac{i(x, \lambda x + c_x)}{k^{\lambda x + c}}$$

$$= x \, k^{-\lambda x - c_x} (k-1)^{\frac{(\lambda-1)x + c_x}{2} + 1} (\lambda x + c_x)^{-1} \binom{\lambda x + c_x}{\frac{(\lambda+1)x + c_x}{2}}.$$

Since both $|c_x|$ and $\lambda = k/(k-2)$ are bounded by absolute constants, we have

$$\binom{\lambda x + c_x}{\frac{(\lambda+1)x + c_x}{2}} \Big/ \binom{\lambda x}{\frac{(\lambda+1)x}{2}}$$

$$= \frac{(\lambda x + c_x) \cdot \cdots \cdot (\lambda x + 1)}{\left[\left(\frac{\lambda+1}{2}x + \frac{c_x}{2}\right) \cdot \cdots \cdot \left(\frac{\lambda+1}{2}x + 1\right)\right] \cdot \left[\left(\frac{\lambda-1}{2}x + \frac{c_x}{2}\right) \cdot \cdots \cdot \left(\frac{\lambda-1}{2}x + 1\right)\right]} = O(1).$$

Now Eq. (8) and using that k is seen as constant yields

$$\max_t \frac{i(x,t)}{k^t} = O\left(\frac{(k-1)^{\frac{(\lambda-1)x}{2}}}{k^{\lambda x}} \binom{\lambda x}{\frac{(\lambda+1)x}{2}}\right)$$

$$= O\left(\frac{(k-1)^{\frac{1}{k-2}x}}{k^{\frac{k}{k-2}x}} \frac{k^{\frac{k}{k-2}x + \frac{1}{2}} (k-2)^{\frac{1}{2}}}{(k-1)^{\frac{k-1}{k-2}x + \frac{1}{2}} \sqrt{x}}\right)$$

$$= O(x^{-1/2}(k-1)^{-x}). \tag{13}$$

From Eqs. (12) and (13) we conclude

$$\text{CON}(S_x) = O(x^{-1/2}). \tag{14}$$

For large x, there are not enough chips such that each vertex on S_x contribute fully to the discrepancy. We use this to obtain a second bound for $\text{CON}(S_x)$. Theorem 5 implies that the sum of all contributions $\text{CON}(S_x)$ for all time steps t with $T - t > \frac{1}{2}\lambda x$ or $T - t < \frac{3}{2}\lambda x$ is bounded by a constant C_k. Hence, it suffices to examine times t with $\frac{1}{2}\lambda x \leq T - t \leq \frac{3}{2}\lambda x$. In every time step, each chip in S_x can contribute only at most $O(x^{-1/2}k^{-x})$ by Eq. (13). As there are κ chips and λx possible time slots, the total contribution of S_x in the time interval $\frac{1}{2}\lambda x < T - t < \frac{3}{2}\lambda x$ is

$$O(\kappa x^{1/2} k^{-x}). \tag{15}$$

With Theorem 2, the two bounds of Eqs. (14) and (15) now yield

$$f(\mathbf{0}, T) - E(\mathbf{0}, T) \leq \sum_{x>0} \text{CON}(S_x)$$

$$= \sum_{1 \leq x \leq \log_k(\kappa)} \text{CON}(S_x) + \sum_{x > \log_k(\kappa)} \text{CON}(S_x)$$

$$\leq \sum_{1 \leq x \leq \log_k(\kappa)} O(x^{-1/2}) + C_k + \sum_{x > \log_k(\kappa)} O(\kappa\, x^{1/2}\, k^{-x})$$

$$= O(\sqrt{\log \kappa}) + O(\sqrt{\log \kappa}),$$

still assuming k to be a constant.   ∎

8. CONCLUSION

In this paper, we showed that k-ary trees ($k \geq 3$) do not admit a constant bound on the single vertex discrepancies. Nevertheless, also on these trees, the Propp machine is a very good simulation of the random walk. For simultaneous walks of κ chips, the discrepancies are bounded by $O(\sqrt{\log \kappa})$.

With this work showing that infinite regular trees do not have discrepancies bounded by a constant, but previous work of Cooper and Spencer [1] showing this property for higher-dimensional grids, the natural open problem arising from this work is to give more insight to the question of which graphs display the one or the other behavior, or ideally, to give a characterization of those graphs having constant bounds for the single vertex discrepancies.

ACKNOWLEDGMENT

REFERENCES

[1] J. Cooper and J. Spencer, Simulating a random walk with constant error, Comb Probab Comput 15 (2006), 815–822.

[2] J. Cooper, B. Doerr, J. Spencer, and G. Tardos, Deterministic random walks on the integers, Eur J Comb 28 (2007), 2072–2090.

[3] B. Doerr and T. Friedrich, Deterministic random walks on the two-dimensional grid, Comb Probab Comput 18 (2009), 123–144.

[4] B. Doerr, T. Friedrich, and T. Sauerwald, Quasirandom rumor spreading, S.-H. Teng (editor), Proceedings of the Nineteenth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2008 (San Francisco, CA, 2008), SIAM, Philadelphia, PA, 2008, pp. 773–781.

[5] S. S. Dragomir, R. P. Agarwal, and N. S. Barnett, Inequalities for beta and gamma functions via some classical and new integral inequalities, J Inequal Appl 5 (2000), 103–165.

[6] P. Hilton and J. Pedersen, Catalan numbers, their generalization, and their uses, Math Intelligencer 13 (1991), 64–75.

[7] I. Landau and L. Levine, The rotor–router model on regular trees, J Comb Theory Ser A 116 (2009), 421–433.

[8] L. Levine and Y. Peres, The rotor–router shape is spherical, Math Intelligencer 27 (2005), 9–11.

[9] V. B. Priezzhev, D. Dhar, A. Dhar, and S. Krishnamurthy, Eulerian walkers as a model of self-organized criticality, Phys Rev Lett 77 (1996), 5079–5082.