Asymptotic Height Distribution in High-Dimensional Sandpiles

Antal A. Járai · Minwei Sun

Received: 21 May 2019 / Revised: 26 October 2019 / Published online: 14 November 2019
© The Author(s) 2019

Abstract
We give an asymptotic formula for the single-site height distribution of Abelian sandpiles on $\mathbb{Z}^d$ as $d \to \infty$, in terms of Poisson(1) probabilities. We provide error estimates.

Keywords Abelian sandpile · Uniform spanning forest · Wilson’s method · Loop-erased random walk

Mathematics Subject Classification (2010) Primary 60K35 · Secondary 82C20

1 Introduction

We consider the Abelian sandpile model on the nearest neighbour lattice $\mathbb{Z}^d$; see Sect. 1.1 for definitions and background. Let $\mathbf{P}$ denote the weak limit of the stationary distributions $\mathbf{P}_L$ in finite boxes $[-L, L]^d \cap \mathbb{Z}^d$. Let $\eta$ denote a sample configuration from the measure $\mathbf{P}$. Let $p_d(i) = \mathbf{P}[\eta(0) = i]$, $i = 0, \ldots, 2d - 1$, denote the height probabilities at the origin in $d$ dimensions. The following theorem is our main result that states the asymptotic form of these probabilities as $d \to \infty$.

Theorem 1.1 (i) For $0 \leq i \leq d^{1/2}$, we have

$$p_d(i) = \sum_{j=0}^{i} e^{-1} \frac{1}{j!} \frac{i}{2d-j} + O\left(\frac{i}{d^2}\right) = \frac{1}{2d} \sum_{j=0}^{i} e^{-1} \frac{1}{j!} + O\left(\frac{i}{d^2}\right). \quad (1.1)$$
(ii) If \( d^{1/2} < i \leq 2d - 1 \), we have

\[
p_d(i) = p_d(d^{1/2}) + O(d^{-3/2}).
\]

In particular, \( p_d(i) \sim (2d)^{-1} \), if \( i, d \to \infty \).

The appearance of the Poisson(1) distribution in the above formula is closely related to the result of Aldous [1] that the degree distribution of the origin in the uniform spanning forest in \( \mathbb{Z}^d \) tends to 1 plus a Poisson(1) random variable as \( d \to \infty \). Indeed our proof of (1.1) is achieved by showing that in the uniform spanning forest of \( \mathbb{Z}^d \), the number of neighbours \( w \) of the origin \( o \), such that the unique path from \( w \) to infinity passes through \( o \) is asymptotically the same as the degree of \( o \) minus 1, that is, Poisson(1).

In [11], we compared the formula (1.1) to numerical simulations in \( d = 32 \) on a finite box with \( L = 128 \), and there is excellent agreement with the asymptotics already for these values.

Other graphs where information on the height distribution is available are as follows. Dhar and Majumdar [7] studied the Abelian sandpile model on the Bethe lattice, and the exact expressions for various distribution functions including the height distribution at a vertex were obtained using combinatorial methods. For the single-site height distribution they obtained (see [7, Eqn. (8.2)])

\[
p_{\text{Bethe},d}(i) = \frac{1}{(d^2 - 1)d^d} \sum_{j=0}^{i} \binom{d+1}{j} (d-1)^{d-j+1}.
\]

If one lets the degree \( d \to \infty \) in this formula, one obtains the form in the right hand side of (1.1) for any fixed \( i \) (with \( 2d \) replaced by \( d \)).

Exact expressions for the distribution of height probabilities were derived by Papoyan and Shcherbakov [20] on the Husimi lattice of triangles with an arbitrary coordination number \( q \). However, on \( d \)-dimensional cubic lattices of \( d \geq 2 \), exact results for the height probability are only known for \( d = 2 \); see [13,14,18,21,22].

1.1 Definitions and Background

Sandpiles are a lattice model of self-organized criticality, introduced by Bak, Tang and Wiesenfeld [3] and have been studied in both physics and mathematics. See the surveys [6,9,10,15,23]. Although the model can easily be defined on an arbitrary finite connected graph, in this paper we will restrict to subsets of \( \mathbb{Z}^d \).

Let \( V_L = [-L, L]^d \cap \mathbb{Z}^d \) be a box of radius \( L \), where \( L \geq 1 \). For simplicity, we suppress the \( d \)-dependence in our notation. We let \( G_L = (V_L \cup \{s\}, E_L) \) denote the graph obtained from \( \mathbb{Z}^d \) by identifying all vertices in \( \mathbb{Z}^d \setminus V_L \) that becomes \( s \), and removing loop-edges at \( s \). We call \( s \) the sink. A sandpile \( \eta \) is a collection of indistinguishable particles on \( V_L \), specified by a map \( \eta : V_L \to \{0, 1, 2, \ldots \} \).

We say that \( \eta \) is stable at \( x \in V_L \), if \( \eta(x) < 2d \). We say that \( \eta \) is stable, if \( \eta(x) < 2d \), for all \( x \in V_L \). If \( \eta \) is unstable (i.e. \( \eta(x) \geq 2d \) for some \( x \in V_L \)), \( x \) is allowed to topple
which means that $x$ passes one particle along each edge to its neighbours. When the vertex $x$ topples, the particles are re-distributed as follows:

$$\eta(x) \to \eta(x) - 2d;$$
$$\eta(y) \to \eta(y) + 1, \quad y \in V_L, y \sim x.$$ 

Particles arriving at $s$ are lost, so we do not keep track of them. Toppling a vertex may generate further unstable vertices. Given a sandpile $\xi$ on $V_L$, we define its stabilization $\xi^o \in \Omega_L := \{\text{all stable sandpiles on } V_L\} = \{0, 1, \ldots, 2d - 1\}^{V_L}$ by carrying out all possible topplings, in any order, until a stable sandpile is reached. It was shown by Dhar [5] that the map $\xi \mapsto \xi^o$ is well-defined, that is, the order of topplings does not matter.

We now define the sandpile Markov chain. The state space is the set of stable sandpiles $\Omega_L$. Fix a positive probability distribution $p$ on $V_L$, i.e. $\sum_{x \in V_L} p(x) = 1$ and $p(x) > 0$ for all $x \in V_L$. Given the current state $\eta \in \Omega_L$, choose a random vertex $X \in V$ according to $p$, add one particle at $X$ and stabilize. The one-step transition of the Markov chain moves from $\eta$ to $(\eta + 1_X)^o$. Considering the sandpile Markov chain on $G_L$, there is only one recurrent class [5]. We denote the set of recurrent sandpiles by $R_L$. It is known [5] that the invariant distribution $P_L$ of the Markov chain is uniformly distributed on $R_L$.

Majumdar and Dhar [19] gave a bijection between $R_L$ and spanning trees of $G_L$. This maps the uniform measure $P_L$ on $R_L$ to the uniform spanning tree measure $\text{UST}_L$. A variant of this bijection was introduced by Priezzhev [22] and is described in more generality in [8,12]. The latter bijection enjoys the following property that we will exploit in this paper. Orient the spanning tree towards $s$, and let $\pi_L(x)$ denote the oriented path from a vertex $x$ to $s$. Let

$$W_L = \{x \in V_L : o \in \pi_L(x)\}.$$ 

Then, we have that

conditional on $\deg_{W_L}(o) = i$, the height $\eta(o)$ is uniformly distributed over the values $i, i + 1, \ldots, 2d - 1$. \hfill (1.2) 

This has the following consequence for the height probabilities. Let $q^L(i) = \text{UST}_L[\deg_{W_L}(o) = i], i = 0, \ldots, 2d - 1$. Then,

$$p^L(i) := P_L[\eta(o) = i] = \sum_{j=0}^{i} \frac{q^L(j)}{2d - j}.$$ 

The measures $P_L$ have a weak limit $P = \lim_{L \to \infty} P_L$ [2], and hence, $p(i) = \lim_{L \to \infty} p^L(i)$ exist, $i = 0, \ldots, 2d - 1$. Although the $q^L(i)$ depend on the non-local
variable $W_L$, one also has that $q(i) = \lim_{L \to \infty} q^L(i)$ exist, $i = 0, \ldots, 2d - 1$; see [12]. In fact, $q(i)$ is given by the following natural analogue of its finite volume definition. Consider the uniform spanning forest measure $USF$ on $\mathbb{Z}^d$, defined as the weak limit of $UST_L$; see [16, Chapter 10]. Let $\pi(x)$ denote the unique infinite self-avoiding path in the spanning forest starting at $x$, and let

$$W = \{x \in \mathbb{Z}^d : o \in \pi(x)\}.$$ 

Then, $q(i) = \text{USF}[\deg_W(o) = i]$, $i = 0, \ldots, 2d - 1$.

Therefore, we have

$$p(i) := \mathbb{P}[\eta(o) = i] = \sum_{j=0}^{i} \frac{q(j)}{2d - j}. \quad (1.3)$$

### 1.2 Wilson’s Method

Given a finite path $\gamma = [s_0, s_1, \ldots, s_k]$ in $\mathbb{Z}^d$, we erase loops from $\gamma$ chronologically, as they are created. We trace $\gamma$ until the first time $t$, if any, when $s_t \in \{s_0, s_1, \ldots, s_{t-1}\}$, i.e. there is a loop. We suppose $s_t = s_i$, for some $i \in \{0, 1, \ldots, t - 1\}$ and remove the loop $[s_i, s_{i+1}, \ldots, s_t = s_i]$. Then, we continue tracing $\gamma$ and follow the same procedure to remove loops until there are no more loops to remove. This gives the loop-erasure $\pi = LE(\gamma)$ of $\gamma$, which is a self-avoiding path [17]. If $\gamma$ is generated from a random walk process, the loop-erasure of $\gamma$ is called the loop-erased random walk (LERW).

When $d \geq 3$, the USF on $\mathbb{Z}^d$ can be sampled via Wilson’s method rooted at infinity [4], [16, Section 10], that is described as follows. Let $s_1, s_2, \ldots$ be an arbitrary enumeration of the vertices and let $T_0$ be the empty forest with no vertices. We start a simple random walk $\gamma_n$ at $s_n$ and $\gamma_n$ stops when $T_{n-1}$ is hit, otherwise we let it run indefinitely. $LE(\gamma_n)$ is attached to $T_{n-1}$, and the resulting forest is denoted by $T_n$. We continue the same procedure until all the vertices are visited. The above gives a random sequence of forests $T_1 \subset T_2 \subset \ldots$, where $T = \cup_n T_n$ is a spanning forest of $\mathbb{Z}^d$. The extension of Wilson’s theorem [24] to transient infinite graphs proved in [4] implies that $T$ is distributed as the USF.

### 2 Proof of the Main Theorem

Let $(S^x_n)_{n \geq 0}$ be a simple random walk started at $x$ (independent between x’s on $\mathbb{Z}^d$) and let $\pi(x)$ be the path in the USF from $x$ to infinity. We introduce the events:

$$E_i = \{\{w \sim o : \pi(w) \text{ passes through } o\} = i\}, \quad i = 0, \ldots, 2d - 1;$$

$$E_i(x_1, x_2, \ldots, x_i) = \{w \sim o : \pi(w) \text{ passes through } o\} = \{x_1, x_2, \ldots, x_i\}.$$
Then, recall that
\[ q_d(i) = P[\text{deg}_W(o) = i] = P[E_i] = \sum_{x_1, \ldots, x_i \sim o}^{\text{distinct}} P[E_i(x_1, \ldots, x_i)]. \quad (2.1) \]

### 2.1 Preliminary

**Lemma 2.1** We have \( P[S_\infty^n = o \text{ for some } n \geq 2] = O(1/d) \) and \( P[S_\infty^n = o \text{ for some } n \geq 4] = O(1/d^2), \) as \( d \to \infty. \)

**Proof** Let \( \hat{D}(k) = \frac{1}{d} \sum_{j=1}^{d} \cos(k_j), \ k \in [-\pi, \pi]^d \) be the Fourier transform in \( d \) dimensions of the one-step distribution of RW. Lemma A.3 in [17] states that for all non-negative integers \( n \) and all \( d \geq 1, \) we have
\[
\| \hat{D}^n \|_1 = (2\pi)^{-d} \int_{[-\pi, \pi]^d} |\hat{D}(k)| d^d k = \left( \frac{\pi d}{4n} \right)^{d/2}.
\]

Based on above, we have
\[
P[S_\infty^n = o \text{ for some } n \geq 4] \leq \frac{1}{(2\pi)^d} \sum_{n=4}^{\infty} \int \hat{D}^n(k) dk \leq \frac{1}{(2\pi)^d} \sum_{n=4}^{d-1} \int \hat{D}^n(k) dk + \sum_{n=d}^{\infty} \left( \frac{\pi d}{4n} \right)^{d/2}. \quad (2.2)
\]

Since \( \int \hat{D}^n(k) dk \) and \( \int \hat{D}^6(k) dk \) state the probability that \( S^n_\infty \) returns to \( o \) in 4 and 6 steps each, by counting the number of ways to return, they are bounded by dimension-independent multiples of \( 1/d^2 \) and \( 1/d^3, \) respectively. We have \( \int \hat{D}^n(k) dk = 0 \) with odd \( n, \) and for \( 6 < n \leq d - 1 \) and \( n \) even, we have \( \int \hat{D}^n(k) dk \leq \int \hat{D}^6(k) dk. \) Hence,
\[
\frac{1}{(2\pi)^d} \int \hat{D}^n(k) dk = O\left( \frac{1}{d^3} \right), \quad 6 \leq n \leq d - 1.
\]

The last sum in (2.2) can be bounded as:
\[
\left( \frac{\pi d}{4} \right)^{d/2} \sum_{n=d}^{\infty} n^{-d/2} \leq \left( \frac{\pi d}{4} \right)^{d/2} \int_{d-1}^{\infty} x^{-d/2} dx = \left( \frac{\pi d}{4} \right)^{d/2} (d - 1)^{1 - \frac{d}{2}} \frac{d}{d/2 - 1}
\]
\[
= \left( \frac{d - 1}{d/2 - 1} \right) \left( \frac{d}{d - 1} \right)^{\frac{d}{2}} \left( \frac{\pi}{4} \right)^{\frac{d}{2}} \leq C e^{-cd},
\]
since we can take \( d > 4 \) and \( \frac{\pi}{4} < 1. \)
Hence, we have the required results

\[ P[S_n^o = o \text{ for some } n \geq 4] \leq \int \hat{D}^4(k) dk + d \int \hat{D}^6(k) dk + Ce^{-cd} \]

\[ = O\left(\frac{1}{d^2}\right) + d \times O\left(\frac{1}{d^3}\right) = O\left(\frac{1}{d^2}\right), \]

\[ P[S_n^o = o \text{ for some } n \geq 2] \leq \left(\frac{1}{2d}\right) + P[S_n^o = o \text{ for some } n \geq 4] = O\left(\frac{1}{d}\right). \]

\[ \square \]

### 2.2 Lower Bounds

Let us fix the vertices \(x_1, \ldots, x_i \sim o\). Let

\[ A_0 = \left\{ S_1^o \notin \{x_1, \ldots, x_i\}, S_n^o \notin N \text{ for } n \geq 2 \right\}, \]

where \(N = \{y \in \mathbb{Z}^d : |y| \leq 1\}\).

**Lemma 2.2** We have \(P[A_0] \geq 1 - O(i/d)\).

**Proof**

\[ P[A_0] = P[S_1^o \neq x_1, \ldots, x_i]P[S_n^o \notin N \text{ for } n \geq 2|S_1^o \neq x_1, \ldots, x_i]. \]

We have \(P[S_1^o \neq x_1, \ldots, x_i] = 1 - O(i/d)\) and the probability for the remaining steps is at least \(1 - O(1/d)\), shown as follows. The probabilities \(P[S_2^o \neq o|S_1^o \neq x_1, \ldots, x_i] \) and \(P[S_3^o \notin N|S_2^o \neq o, S_1^o \neq x_1, \ldots, x_i]\) are both equal to \(1 - O(1/d)\). Considering the s.r.w starting at the position \(S_3^o\), it hits at most three neighbours of \(o\) in two further steps, the remaining neighbours will need at least 4 steps to hit, so, by Lemma 2.1, we have

\[ \sum_{\text{at most 3 neighbours } x_j} \sum_{k \geq 1} P_{2k}(S_3^o, x_j) \leq O\left(\frac{1}{d}\right), \]

\[ \sum_{\text{the remaining neighbours } x_j} \sum_{k \geq 2} P_{2k}(S_3^o, x_j) \leq O(d)O\left(\frac{1}{d^2}\right) = O\left(\frac{1}{d}\right), \]

since \(P_{2k}(x, y) \leq P_{2k}(o, o)\) for all \(x, y\). Therefore, combining above results together, we get \(P[S_n^o \notin N \text{ for } n \geq 2|S_1^o \neq x_1, \ldots, x_i] \geq 1 - O(1/d)\) as required. \[ \square \]

Let us label the neighbours of \(o\) different from \(x_1, \ldots, x_i\) as \(x_{i+1}, \ldots, x_{2d}\), in any order. On the event \(A_0\), the first step of \(\pi(o)\) is to a neighbour of \(o\) in \(\{x_{i+1}, \ldots, x_{2d}\}\) and we could assume \(x_{2d}\) to be the first step of \(\pi(o)\). Then, \(\pi(o)\) does not visit other vertices in \(N \setminus \{o\}\). Define \(A_j = \{S_1^{x_j} = o\}\) for \(j = 1, 2, \ldots, i\) and then \(P[A_j] = 1/2d\).
Using Wilson’s algorithm, consider random walks first started at $o, x_1, \ldots, x_i$ and then started at $x_{i+1}, \ldots, x_{2d-1}$. We obtain the following:

\[
\mathbb{P}[E_i(x_1, \ldots, x_i)] \geq \mathbb{P}[A_0] \times \prod_{j=1}^{i} \mathbb{P}[A_j] \times \mathbb{P}[E_i(x_1, \ldots, x_i)|A_0 \cap A_1 \cap \cdots \cap A_i] \\
\geq \left(1 - O\left(\frac{i}{d}\right)\right) \left(\frac{1}{2d}\right)^i \mathbb{P}[E_i(x_1, \ldots, x_i)|A_0 \cap A_1 \cap \cdots \cap A_i].
\]

(2.3)

Define $B_k = \{S_n^{x_k} \neq o, S_n^{x_k} \notin \{x_1, \ldots, x_i\} \text{ for } n \geq 2\}$ for $k = i + 1, \ldots, 2d-1$.

**Lemma 2.3** $\mathbb{P}[B_k] \geq 1 - 1/2d - O(i/d^2)$, where $i + 1 \leq k \leq 2d - 1$.

**Proof** We have $\mathbb{P}[S_1^{x_k} \neq o] = 1 - 1/2d$. If the first step is not to $o$, the first step could be in one of the $e_1, \ldots, e_i$ directions, say $e_j$, with probability $i/2d$. Then, the probability to hit $x_j$ is $1/2d + O(1/d^2)$. Hence, the probability that $S_n^{x_k}$ hits $\{x_1, \ldots, x_i\}$ is $O(i/d^2)$.

**Lemma 2.4** $q_d(i) \geq e^{-1 \frac{1}{7!}} \left(1 + O\left(\frac{i^2}{d}\right)\right)$.

**Proof** By (2.3), we have

\[
\mathbb{P}[E_i(x_1, \ldots, x_i)] \geq \left(1 - O\left(\frac{i}{d}\right)\right) \left(\frac{1}{2d}\right)^i \left(1 - \frac{1}{2d} + O\left(\frac{i}{d^2}\right)\right)^{2d-1-i}.
\]

Then, by (2.1),

\[
q_d(i) \geq \binom{2d}{i} \left(1 - O\left(\frac{i}{d}\right)\right) \left(\frac{1}{2d}\right)^i \left(1 - \frac{1}{2d} + O\left(\frac{i}{d^2}\right)\right)^{2d-1-i} \\
= \frac{2d(2d-1)\ldots(2d-i+1)}{i!(2d)^i} \left(1 - O\left(\frac{i}{d}\right)\right) \left(1 - \frac{1}{2d} + O\left(\frac{i}{d^2}\right)\right)^{2d} \left(1 + O\left(\frac{i}{d}\right)\right),
\]

where

\[
\left(1 - \frac{1}{2d} + O\left(\frac{i}{d^2}\right)\right)^{2d} = \exp\left(2d \times \log\left(1 - \frac{1}{2d} + O\left(\frac{i}{d^2}\right)\right)\right) \\
= \exp\left(2d\left(-\frac{1}{2d} + O\left(\frac{i}{d^2}\right)\right)\right) \\
= \exp\left(-1 + O\left(\frac{i}{d}\right)\right) = e^{-1}\left(1 + O\left(\frac{i}{d}\right)\right).
\]

\begin{center}
\textcopyright Springer
\end{center}
and
\[
\frac{2d(2d-1)\ldots(2d-i+1)}{(2d)^i} = 1 \left( 1 - \frac{1}{2d} \right) \left( 1 - \frac{2}{2d} \right) \ldots \left( 1 - \frac{i}{2d} + \frac{1}{2d} \right) = \left( 1 + O\left( \frac{i^2}{d} \right) \right).
\]

Then, the result follows
\[
q_d(i) \geq e^{-1} \left( 1 + O\left( \frac{i}{d} \right) \right) \left( 1 + O\left( \frac{i^2}{d} \right) \right) \left( 1 + O\left( \frac{i}{d} \right) \right) \left( 1 - O\left( \frac{i}{d} \right) \right) = e^{-1} \left( 1 + O\left( \frac{i^2}{d} \right) \right) .
\]

\[\square\]

The above lemma gives a lower bound for \( q_d \), and we now prove an upper bound.

### 2.3 Upper Bounds

Recall that \( \pi(o) \) denotes the unique infinite self-avoiding path in the spanning forest starting at \( o \) and let \( \hat{A}_o = \{ \pi(o) \text{ visits only one neighbour of } o \} \).

**Lemma 2.5** \( P[\pi(o) \text{ visits more than one neighbour of } o] = P[\hat{A}_o^c] = O(1/d) \).

**Proof** The first step of \( \pi(o) \) must visit a neighbour of \( o \), denoted by \( w \), then
\[
P[\pi(w) \text{ visits } o \text{ but not at the first step}]
\]
\[
= P[\text{The second step of } \pi(o) \text{ visits } x \neq 2w, \text{ the third step visits } w' \sim o, w' \neq w] + O\left( \frac{1}{d^2} \right)
\]
\[
= \left( \frac{1}{2d} \right) \left( \frac{2d-1}{2d} \right) + O\left( \frac{1}{d^2} \right) = O\left( \frac{1}{d} \right).
\]

\[\square\]

Let \( \hat{A}_{\text{all}} = \{ \forall w \sim o: \text{either } \pi(w) \text{ does not visit } o \text{ or } \pi(w) \text{ visits } o \text{ at the first step} \} \).

**Lemma 2.6** \( P[\exists w \sim o : \pi(w) \text{ visits } o \text{ but not at the first step}] = P[\hat{A}_{\text{all}}^c] = O(1/d) \).

**Proof** For a given \( w, w \sim o \), use Wilson’s algorithm with a walk started at \( w \). Consider that if \( S_1^w \neq o \), or \( S_1^w = o \), but \( S_w^w \) returns to \( w \) subsequently and then this loop starting from \( w \) in \( S_w^w \) is erased, \( \pi(w) \) does not visit \( o \) at the first step. Hence, we have the inequality:

\[
P[\pi(w) \text{ visits } o \text{ but not at the first step}]
\]
\[
\leq P[S_w^w \text{ visits } o \text{ but not at the first step}] + P[S_1^w = o, S_n^w = w \text{ for some } n \geq 2].
\]

(2.4)
We bound the two terms as follows. For the first term, let us append a step from \( o \) to \( w \) at the beginning of the walk and analyse it as if the walk started at \( o \). Since \( S_1^o \in N \setminus \{o\} \), by symmetry, we may assume \( S_1^o = w \). Then, if \( S_2^o \neq o \), \( S^o \) will need at least 2 more steps to return to \( o \).

For the second term in the right hand side of (2.4), we first note that we have

\[
P[S_1^w = o, S_2^w = w] = 1/(2d)^2.
\]

If \( S^w \) does not return to \( w \) in the first two steps, \( S^w \) will need at least 4 steps to return to \( w \). Then, we have that the right hand side of (2.4) is

\[
\leq P[S^o \text{ returns to } o \text{ in at least } 4 \text{ steps}] + \frac{1}{(2d)^2} + P[S^w \text{ returns to } w \text{ in at least } 4 \text{ steps}]
\]

\[
= 2 \times P[S^o \text{ returns to } o \text{ in at least } 4 \text{ steps}] + O\left(\frac{1}{d^2}\right).
\]

Therefore, by Lemma 2.1, we have the required result

\[
P[\exists w \sim o : \pi(w) \text{ visits } o \text{ but not at the first step }]
\]

\[
= 2d \times P[\pi(w) \text{ visits } o \text{ but not at the first step for a fixed } w \sim o] = O\left(\frac{1}{d}\right).
\]

Due to Lemmas 2.5 and 2.6, we have

\[
g_d(i) \leq O\left(\frac{1}{d}\right) + P[\tilde{A}_o \cap \tilde{A}_{\text{all}} \cap E_i]
\]

\[
= O\left(\frac{1}{d}\right) + \sum_{x_1, \ldots, x_i \sim o}^{\text{distinct}} P[\tilde{A}_o \cap \tilde{A}_{\text{all}} \cap E_i(x_1, \ldots, x_i)].
\]

Here,

\[
\tilde{A}_o \cap \tilde{A}_{\text{all}} \cap E_i(x_1, \ldots, x_i)
\]

\[
\subset \tilde{A}_o \cap \tilde{A}_{\text{all}} \cap \{\text{the first step of } \pi(x_j) \text{ is to } o, j = 1, \ldots, i\} \cap F_i(x_1, \ldots, x_i),
\]

(2.5)

where

\[
F_i(x_1, \ldots, x_i) = \{\pi(x_j) \text{ does not go through } o, j = i + 1, \ldots, 2d\}.
\]

The right hand side of (2.5) is contained in the event

\[
\tilde{A}_o \cap \{\pi(o) \text{ does not visit } x_1, \ldots, x_i\} \cap \tilde{A}_{\text{rest}} \cap \bigcap_{1 \leq j \leq i} H_j \cap F_i(x_1, \ldots, x_i),
\]
where
\[ \tilde{A}_{\text{rest}} = \{ \pi(x_j) \text{ goes through at most one} \} \]
\[ x_{j'}, j = i + 1, \ldots, 2d, i + 1 \leq j' \leq 2d, j' \neq j \]
and \( H_j = \{ \text{the first step of } \pi(x_j) \text{ is } o \} \) for \( j = 1, \ldots, i \).

We denote \( A_o \cap \{ \pi(o) \text{ does not visit } x_1, \ldots, x_i \} \) by \( \tilde{A}_{o,x_1,\ldots,x_i} \). Then,
\[
P\left[ \tilde{A}_{o,x_1,\ldots,x_i} \cap \tilde{A}_{\text{rest}} \cap \bigcap_{1 \leq j \leq i} H_j \cap F_i(x_1, \ldots, x_i) \right]
\]
\[ = P[\tilde{A}_{o,x_1,\ldots,x_i}] \prod_{j=1}^{i} P[H_j \bigcap \bigcap_{1 \leq j' < j} H_{j'} \cap \tilde{A}_{o,x_1,\ldots,x_i}]
\times P[F_i(x_1, \ldots, x_i) \cap \tilde{A}_{\text{rest}} \tilde{A}_{o,x_1,\ldots,x_i} \cap \bigcap_{1 \leq j \leq i} H_j].
\]

Therefore, we have
\[
q_d(i) \leq O\left( \frac{1}{d} \right) + \sum_{x_1,\ldots,x_i \sim o} \left( \prod_{j=1}^{i} P[H_j \bigcap \bigcap_{1 \leq j' < j} H_{j'} \cap \tilde{A}_{o,x_1,\ldots,x_i}] \right)
\times P[F_i(x_1, \ldots, x_i) \cap \tilde{A}_{\text{rest}} \tilde{A}_{o,x_1,\ldots,x_i} \cap \bigcap_{1 \leq j \leq i} H_j]. \tag{2.6}
\]

**Lemma 2.7** \( P[H_j | \tilde{A}_{o,x_1,\ldots,x_i} \cap \bigcap_{1 \leq j' < j} H_{j'}] = 1/2d + O(1/d^2) \), where \( j = 1, \ldots, i \).

**Proof** Given that \( \pi(o) \) visits only one neighbour of \( o \) which is not in \( \{x_1, \ldots, x_i\} \) and the first steps of \( \pi(x_1), \ldots, \pi(x_{j-1}) \) are all to \( o \), the probability that \( H_j \) happens is \( P[S_i^{x_j} = o] = 1/2d \) with the error term of \( O(1/d^2) \) due to the loop-erasure. \( \square \)

**Lemma 2.8**
\[
P[F_i(x_1, \ldots, x_i) \cap \tilde{A}_{\text{rest}} \tilde{A}_{o,x_1,\ldots,x_i} \cap \bigcap_{1 \leq j \leq i} H_j]
\]
\[ \leq E\left[ \left( 1 - \frac{1}{2d} + O\left( \frac{1}{d^2} \right) \right)^{2d-i-1-N} \mathbf{1}_{\tilde{A}_{\text{rest}}} \right]. \tag{2.7}
\]

where \( N = \{ i + 1 \leq j \leq 2d - 1 : \exists i + 1 \leq j' < j \text{ s.t. } \pi(x_j) \text{ goes through } x_j \} \).

**Proof** Consider Wilson’s algorithm with random walks started at the remaining neighbours \( x_{i+1}, \ldots, x_{2d} \). Assume \( x_{2d} \) to be the neighbour of \( o \) that \( \pi(o) \) goes through. The probability that \( \pi(x_k) \) does not go through \( o \) is \( 1 - 1/2d + O(1/d^2) \) for \( k \in \{ i + 1, \ldots, 2d - 1 \} \).

If \( \pi(x_k) \) visits \( x_{k'} \), where \( k < k' \leq 2d - 1 \), the probability that \( \pi(x_{k'}) \) does not go through \( o \) is \( 1 \) instead of \( 1 - 1/2d + O(1/d^2) \), since the LERW from \( x_{k'} \) stops immediately and \( \pi(x_{k'}) \subset \pi(x_k) \), which does not go through \( o \). \( \square \)
Lemma 2.9  On the event $\bar{A}_{\text{rest}}$, $N \leq B$, where $B \sim \text{Binom}(2d - i - 1, p)$, $p = 1/2d + O(1/d^2)$.

Proof  Since we have $(2d - i - 1)$ trials with probability at most $1/2d + O(1/d^2)$. $\Box$

Due to Lemma 2.9, we have that the right hand side of (2.7) is
\begin{align*}
\leq \left( 1 - \frac{1}{2d} + O\left( \frac{1}{d^2} \right) \right)^{2d} \left( 1 + O\left( \frac{i}{d} \right) \right) \mathbb{E}\left[ \frac{1}{(1 - \frac{1}{2d} + O\left( \frac{1}{d^2} \right))^B} \right],
\end{align*}
(2.8)

where $\mathbb{E}[z^B] = \sum_{j=0}^{2d-i-1} z^j \binom{2d-i-1}{j} p^j (1-p)^{2d-i-1-j} = (1-p-zp)^{2d-i-1}$.

Hence, (2.8) is
\begin{align*}
\leq e^{-1} \left( 1 + O\left( \frac{i}{d} \right) \right) \left( 1 + O\left( \frac{i}{d} \right) \right) \left( 1 - \frac{1}{2d} + O\left( \frac{1}{d^2} \right) \right)^{2d-i-1} \\
= e^{-1} \left( 1 + O\left( \frac{i}{d} \right) \right)^2 e^{-1} \left( 1 + O\left( \frac{i}{d} \right) \right).
\end{align*}
(2.9)

Lemma 2.10  $q_d(i) \leq O\left( \frac{1}{d^2} \right) + e^{-1} \frac{1}{i!} \left( 1 + O\left( \frac{i}{d} \right) \right)$.

Proof  Due to Lemma 2.7, (2.6) and (2.9), we have
\begin{align*}
q_d(i) &\leq O\left( \frac{1}{d} \right) + \binom{2d}{i} \left( \frac{1}{2d} + O\left( \frac{1}{d^2} \right) \right)^i e^{-1} \left( 1 + O\left( \frac{i}{d} \right) \right) \\
&= O\left( \frac{1}{d} \right) + e^{-1} \frac{2d(2d-1)\ldots(2d-i+1)}{i!} \left( \frac{1}{2d} \right)^i \left( 1 + O\left( \frac{i}{d} \right) \right)^i \\
&\leq O\left( \frac{1}{d} \right) + e^{-1} \frac{1}{i!} \left( 1 + O\left( \frac{i}{d} \right) \right).
\end{align*}

Lemma 2.11  For $k = 1, \ldots, 3$ and distinct $w_1, \ldots, w_k \sim o$, we have
\begin{align*}
\mathbb{P}[\pi(w_i) \text{ passes through } o \text{ for } i = 1, \ldots, k] = \left( \frac{1}{2d} \right)^k + O(d^{-k-1}).
\end{align*}

This lemma can be proved using ideas used to prove Lemma 2.7.

2.4 Proof of the Asymptotic Formula

Proof of Theorem 1.1  We first prove part (i). By Wilson’s algorithm,
\begin{align*}
p_d(i) = \sum_{j=0}^{i} \frac{q_d(j)}{2d - j}.
\end{align*}
Due to Lemmas 2.4 and 2.10, we have

\[ p_d(i) \geq \sum_{j=0}^{i} e^{-1} \frac{1}{j!} \left( 1 + O \left( \frac{i^2}{d^2} \right) \right) = \sum_{j=0}^{i} e^{-1} \frac{1}{j!} + \sum_{j=0}^{i} \frac{1}{j!} O \left( \frac{i^2}{d^2} \right), \]  

(2.10)

and

\[ p_d(i) \leq \sum_{j=0}^{i} \frac{O(\frac{i}{d})}{2d-j} + \sum_{j=0}^{i} \frac{O \left( \frac{j^2}{d} \right)}{2d-j} \]  

(2.11)

Here, using that \( 0 \leq j \leq d^{1/2} \), we have

\[ \sum_{j=0}^{i} \frac{\frac{1}{j!} O \left( \frac{j^2}{d} \right)}{2d-j} \leq \frac{1}{2d-d^{1/2}} O \left( \frac{1}{d} \right) \sum_{j=0}^{i} \frac{j^2}{j!} = O \left( \frac{1}{d^2} \right). \]

Similarly,

\[ \sum_{j=0}^{i} \frac{O \left( \frac{1}{d} \right) + \frac{1}{j!} O \left( \frac{j^2}{d} \right)}{2d-j} \leq \sum_{j=0}^{i} O(d^{-2}) + \sum_{j=0}^{i} \frac{j}{j!} O(d^{-2}) = O(i/d^2). \]

Putting these error bounds together with (2.10) and (2.11), we prove statement (i) of the theorem.

Let us now use that

\[ \frac{1}{2d} e^{-1} \sum_{j=0}^{i} \frac{1}{j!} \leq \sum_{j=0}^{i} e^{-1} \frac{1}{j!} \leq \frac{1}{2d-i} e^{-1} \sum_{j=0}^{i} \frac{1}{j!}. \]

When \( i \leq d^{1/2} \), and \( i, d \to \infty \), we have \( \frac{1}{2d-i} \sim \frac{1}{2d} \) and \( \sum_{j=0}^{i} \frac{1}{j!} \to e \). Hence,

\[ \sum_{j=0}^{i} e^{-1} \frac{1}{j!} \sim \frac{1}{2d}, \quad \text{as } i, d \to \infty. \]

We are left to prove statement (ii). The uniform distribution for \( d^{1/2} \leq i \leq 2d - 1 \) can be obtained from the monotonicity:

\[ p_d(d^{1/2}) \leq p_d(i) \leq p_d(2d-1), \quad d^{1/2} \leq i \leq 2d - 1, \]
if we show that \( p_d(2d - 1) = p_d(d^{1/2}) + O(d^{-3/2}) \).

We write

\[
p_d(2d - 1) = \sum_{j=0}^{2d-1} \frac{q_d(j)}{2d-j} = p_d(d^{1/2}) + \sum_{j=d^{1/2}}^{2d-1} \frac{q_d(j)}{2d-j} \leq p_d(d^{1/2}) + \sum_{j=d^{1/2}}^{2d-1} q_d(j).
\]

Introducing the random variable

\[ X := \left| \{ w \sim o : o \in \pi(w) \} \right|, \]

the last expression in (2.12) equals

\[
p_d(d^{1/2}) + P[X \geq d^{1/2}] \leq p_d(d^{1/2}) + P[X^3 \geq d^{3/2}] \leq p_d(d^{1/2}) + \frac{E[X^3]}{d^{3/2}}.
\]

Therefore, it remains to show that \( E[X^3] = O(1) \). This follows from Lemma 2.11, by summing over \( w_1, \ldots, w_3 \) (not necessarily distinct). The cases \( k = 1, 2 \) of the lemma are used to sum the contributions where one or more of the \( w_i \)'s coincide. \( \Box \)

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. Aldous, D.J.: The random walk construction of uniform spanning trees and uniform labelled trees. SIAM J. Discrete Math. 3(4), 450–465 (1990)
2. Athreya, S.R., Járai, A.A.: Infinite volume limit for the stationary distribution of abelian sandpile models. Commun. Math. Phys. 249(1), 197–213 (2004)
3. Bak, P., Tang, C., Wiesenfeld, K.: Self-organized criticality: an explanation of the 1/f noise. Phys. Rev. Lett. 59(4), 381–384 (1987)
4. Benjamin, I., Lyons, R., Peres, Y., Schramm, O.: Uniform spanning forests. Ann. Probab. 29(1), 1–65 (2001)
5. Dhar, D.: Self-organized critical state of sandpile automaton models. Phys. Rev. Lett. 64(14), 1613–1616 (1990)
6. Dhar, D.: Theoretical studies of self-organized criticality. Physica A 369(1), 29–70 (2006)
7. Dhar, D., Majumdar, S.N.: Abelian sandpile model on the Bethe lattice. J. Phys. A: Math. Gen. 23(19), 4333–4350 (1990)
8. Gamlin, S.L., Járai, A.A.: Anchored burning bijections on finite and infinite graphs. Electron. J. Probab. 19(117), 1–23 (2014)
9. Holroyd, A.E., Levine, L., Mészáros, K., Peres, Y., Propp, J., Wilson, D.B.: Chip-firing and rotor-routing on directed graphs. In: Sidoravicius, V., Vares, M.E. (eds.) In and Out of Equilibrium 2, pp. 331–364. Birkhäuser, Basel (2008)
10. Járai, A.A.: Sandpile models. Probab. Surv. 15, 243–306 (2018)
11. Járai, A.A., Sun, M.: Toppling and height probabilities in sandpiles. J. Stat. Mech. Theory Exp. 2019(11), 113204 (2019). https://doi.org/10.1088/1742-5468/ab2cc6
12. Járai, A.A., Werning, N.: Minimal configurations and sandpile measures. J. Theor. Probab. 27(1), 153–167 (2014)
13. Jeng, M., Piroux, G., Ruelle, P.: Height variables in the abelian sandpile model: scaling fields and correlations. J. Stat. Mech. Theory Exp. 2006(10), P10015 (2006)
14. Kenyon, R.W., Wilson, D.B.: Spanning trees of graphs on surfaces and the intensity of loop-erased random walk on planar graphs. J. Am. Math. Soc. 28(4), 985–1030 (2015)
15. Levine, L., Peres, Y.: Laplacian growth, sandpiles, and scaling limits. Bull. Am. Math. Soc. 54(3), 355–382 (2017)
16. Lyons, R., Peres, Y.: Probability on Trees and Networks. Cambridge University Press, Cambridge (2016)
17. Madras, N., Slade, G.: The Self-Avoiding Walk. Springer, New York (2013)
18. Majumdar, S.N., Dhar, D.: Height correlations in the abelian sandpile model. J. Phys. A: Math. Gen. 24(7), L357–L362 (1991)
19. Majumdar, S.N., Dhar, D.: Equivalence between the abelian sandpile model and the $q \to 0$ limit of the potts model. Physica A 185(1), 129–145 (1992)
20. Papoyan, V.V., Shecherbakov, R.R.: Distribution of heights in the abelian sandpile model on the Husimi lattice. Fractals 4(1), 105–110 (1996)
21. Poghosyan, V.S., Priezzhev, V.B., Ruelle, P.: Return probability for the loop-erased random walk and mean height in the abelian sandpile model: a proof. J. Stat. Mech. Theory Exp. 2011(10), P10004 (2011)
22. Priezzhev, V.B.: Structure of two-dimensional sandpile. I. Height probabilities. J. Stat. Phys. 74(5), 955–979 (1994)
23. Redig, F.: Mathematical aspects of the abelian sandpile model. In: Bovier, A. (ed.) Mathematical Statistical Physics: Lecture Notes of the Les Houches Summer School 2005, pp. 657–729. Elsevier, Amsterdam (2006)
24. Wilson, D.B.: Generating random spanning trees more quickly than the cover time. In: Proceedings of the Twenty-Eighth Annual ACM Symposium on Theory of Computing, STOC ’96. ACM, New York, pp. 296–303 (1996)

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.