Electron. Commun. Probab. **21** (2016), no. 57, 1–12.
DOI: 10.1214/16-ECP4477
ISSN: 1083-589X

ELECTRONIC COMMUNICATIONS in PROBABILITY

Sandpiles and unicycles on random planar maps

Xin Sun* David B. Wilson†

Abstract

We consider the abelian sandpile model and the uniform spanning unicycle on random planar maps. We show that the sandpile density converges to 5/2 as the maps get large. For the spanning unicycle, we show that the length and area of the cycle converges to the exit location and exit time of a simple random walk in the first quadrant. The calculations use the "hamburger-cheeseburger" construction of Fortuin–Kasteleyn random cluster configurations on random planar maps.

Keywords: hamburger-cheeseburger bijection; random planar map; abelian sandpile model; cycle-rooted spanning tree.
AMS MSC 2010: 82B20; 60C05; 05C05.
Submitted to ECP on August 12, 2015, final version accepted on July 25, 2016.

1 Introduction

Random planar maps together with discrete statistical mechanics models (e.g., spanning tree, Ising model, FK model) on them is an active research area (see e.g., [1, 25, 10, 2, 35, 14, 7, 9, 15, 4, 6, 16, 17]). In the annealed distribution of a discrete model on a random planar map of a given class, the joint distribution of the pair (M,Σ), where M is the planar map and Σ is a configuration of the discrete model on M, is just proportional to the weight of Σ. Equivalently, the random map in the class is sampled according to the *partition function* of the discrete model (i.e., the weighted sum of all the configurations on the given map), and then a configuration is sampled according to the weighting rule of the discrete model on the map.

We consider two discrete models on random planar maps: the uniform recurrent sandpile and the uniform spanning unicycle (also known as a cycle-rooted spanning tree). For the sandpile model we show that the sandpile density converges to $5/2$ and concentrates around this value. For the unicycle model, we compute the weak limit of the joint distribution of the length of the cycle and the area inside the cycle. The sandpile model calculations depend on the results for the spanning unicycle.

1.1 Planar maps, spanning trees, and unicycles

Planar graphs are graphs that can be embedded into the sphere. A *planar map* is a connected planar graph, with multiple edges and self loops allowed, embedded on the sphere, considered up to isotopic deformation of the edges, i.e., a planar map contains only information about the combinatorial structure of the embedding. A *rooted planar*

*Massachusetts Institute of Technology. E-mail: xinsun89@math.mit.edu
†Microsoft Research. E-mail: David.Wilson@microsoft.com

map (M, e) is a planar map M with a distinguished directed edge e. We let $\mathcal{M}_n$ denote the set of rooted planar maps with n edges.

Given a finite connected graph $\mathcal{G} = (V, E)$, a *spanning forest* of $\mathcal{G}$ is a subgraph whose vertex set is V and which contains no cycles. A spanning forest that contains k connected components is called a *k-component spanning forest*. A 1-component spanning forest is called a *spanning tree*. A *k-excess subgraph* of $\mathcal{G}$ is the union of a spanning tree and k extra edges in E. A 1-excess subgraph is called a *spanning unicycle*. The planar dual of a k-excess subgraph is a $(k+1)$-component spanning forest. Let $\mathcal{T}(\mathcal{G})$ denote

$$\mathcal{T}(\mathcal{G}) := \left\{\text{spanning trees of } \mathcal{G}\right\}.$$

and $\mathcal{U}_k(\mathcal{G})$ denote

$$\mathcal{U}_k(\mathcal{G}) := \left\{k\text{-excess subgraphs of } \mathcal{G}\right\}.$$

Given a rooted planar map (M, e), since e is a directed edge, there is a unique face f to the right of e, which we call the *outer face*. The outer face allows us to distinguish between the two sides of a cycle: the *outside* is the side containing the outer face, and the *inside* is the side which does not. The *length* of a cycle is the number of edges on the cycle. We define the *area* of a cycle to be twice the number of edges inside the cycle plus its length. Planar maps are in natural bijective correspondence with quadrangulations, and this combinatorial definition of area corresponds to assigning each quadrangle area 2.

1.2 The abelian sandpile model

The abelian sandpile model is a model for self-organized criticality [5] which is defined as follows. (See also [18] for further background.) Suppose $\mathcal{G} = (V, E)$ is a finite connected undirected graph with loops and multiple edges allowed. Let $c(v, w)$ be the number of edges between vertices v and w, where self-loops count twice. For $v \in V$, the degree of v is denoted by $\deg(v) = \sum_{w \in V} c(v, w)$. A sandpile configuration on the graph $\mathcal{G}$, with respect to a distinguished vertex s called the sink, assigns a non-negative integer number of grains of sand to each vertex other than the sink s. If a vertex $v \neq s$ has more sand than its degree, then v is *unstable*, and may *topple*, sending one grain of sand to each neighbor. The sink s never topples. Since every vertex is connected to the sink, we may repeatedly topple unstable vertices until every vertex is stable. The resulting sandpile is called the stabilization of the original sandpile, and is independent of the order in which vertices are toppled (which is the abelian property).

Some sandpile configurations are *recurrent*, meaning that from any sandpile configuration, it is possible to add some amount of sand to the vertices and stabilize to obtain the given configuration. These sandpile configurations are the recurrent states of the Markov chain which at each step adds a grain of sand to a random vertex and then stabilizes the configuration. The stationary distribution of this Markov chain is the uniform distribution on recurrent sandpile configurations.

We let $\mathcal{R}(\mathcal{G}, s)$ denote the set of the recurrent sandpile configurations on a graph $\mathcal{G}$ with sink s. Majumdar and Dhar gave a bijection between $\mathcal{R}(\mathcal{G}, s)$ and $\mathcal{T}(\mathcal{G})$ [27]. In particular, $|\mathcal{R}(\mathcal{G}, s)| = |\mathcal{T}(\mathcal{G})|$.

Given a recurrent sandpile configuration $\sigma \in \mathcal{R}(\mathcal{G}, s)$, where $\sigma(v)$ for $v \neq s$ is the number of grains at v, it is convenient to define $\sigma(s) = \deg(s)$. Then the total amount of sand (i.e., including the sand at the sink) is

$$|\sigma| = \sum_{v \in \mathcal{G}} \sigma(v).$$

With this convention, the distribution of $|\sigma|$ for a random recurrent sandpile $\sigma \in \mathcal{R}(\mathcal{G}, s)$ does not depend on the choice of the sink s [28]. This distribution can be understood in terms of the k-excess subgraphs of $\mathcal{G}$ [28], which we will explain in Section 2.

We define the *sandpile edge density* for $\mathcal{G}$ to be

$$\rho_e(\mathcal{G}) = \frac{1}{|E| \cdot |R(\mathcal{G}, s)|} \sum_{\sigma \in \mathcal{R}(\mathcal{G},s)} |\sigma| \tag{1.1}$$

and the *sandpile vertex density* for $\mathcal{G}$ to be

$$\rho_v(\mathcal{G}) = \frac{1}{|V| \cdot |R(\mathcal{G}, s)|} \sum_{\sigma \in \mathcal{R}(\mathcal{G},s)} |\sigma| \,. \tag{1.2}$$

These sandpile densities $\rho_e(\mathcal{G})$ and $\rho_v(\mathcal{G})$ are independent of the choice of the sink.

1.3 Sandpiles and k-excess subgraphs on random planar maps

We let $\mathbb{P}_n^k$ denote the annealed distribution on k-excess subgraphs of a rooted planar map containing n edges, i.e., a random planar map (M, e) with a k-excess subgraph U_k on it are chosen with probability

$$\mathbb{P}_n^k[M, e, U_k] = \frac{1}{\sum\limits_{(M',e') \in \mathcal{M}_n} |\mathcal{U}_k(M')|} \,. \tag{1.3}$$

We let $\mathbb{E}_n^k$ and Var_n^k denote the expectation and variance with respect to $\mathbb{P}_n^k$.

For the directed edge e, we let $\underline{e}$ denote its source vertex, and $\overline{e}$ denote its destination vertex. Because of the bijection between recurrent sandpile configurations $\mathcal{R}(M, \underline{e})$ and spanning trees $\mathcal{T}(M) = \mathcal{U}_0(M)$, we can interpret P_n^0 as being the annealed distribution of uniform recurrent sandpiles on $\mathcal{M}_n$ with sink at the source of the root edge.

Theorem 1.1. *For the uniform recurrent sandpile on a random planar map with n edges, the sandpile density satisfies*

$$\lim_{n\to\infty} \mathbb{E}_n^0[\rho_v] = 2 \lim_{n\to\infty} \mathbb{E}_n^0[\rho_e] = \frac{5}{2}, \qquad \lim_{n\to\infty} \mathrm{Var}_n^0[\rho_v] = \lim_{n\to\infty} \mathrm{Var}_n^0[\rho_e] = 0. \tag{1.4}$$

This sandpile density computation can be compared to the uniform recurrent sandpile density on $\mathbb{Z}^2$. In 1994 Grassberger conjectured that the (per vertex) sandpile density on $\mathbb{Z}^2$ is $17/8$, based on the numerical integration of singular 4-dimensional integral expressions given by Priezzhev [33] for the sandpile height distribution at a vertex. These integral expressions were greatly simplified by Jeng, Piroux, and Ruelle [20], who verified by numerical integration that the sandpile density for $\mathbb{Z}^2$ is $17/8 \pm 10^{-12}$. An alternative formulation of the sandpile density, relating it to spanning unicyclic graphs and loop-erased random walk, was given by Poghosyan and Priezzhev [31] and Levine and Peres [26], which enabled its rigorous exact evaluation in [32, 23]. The density of $17/8$ corresponds to an edge density of $1 + \frac{1}{16}$, versus $1 + \frac{1}{4}$ for the sandpile on a random planar map. More recently, a simpler proof of the sandpile density on $\mathbb{Z}^2$ was given by Kassel and Wilson [22], who then computed the sandpile density for numerous other lattices as well.

Mullin considered the closely related problem of how many edges of a spanning-tree-decorated planar map with n edges are internally active. Mullin gave an exact formula involving a double summation for the average number of internally active edges, showed that the average is between $n/8$ and $n/2$, and showed that asymptotically the average is $(c + o_n(1))n$ for some constant c [30]. Our results on the sandpile edge density are equivalent to showing $c = \frac{1}{4}$ (this is the $\frac{1}{4}$ of $\rho_e = 1 + \frac{1}{4}$), and that for a random tree-decorated planar map the number of internally active edges is concentrated about its expected value.

Our results for unicycles, and more generally k-excess subgraphs, can be expressed in terms of simple random walk in the first quadrant of $\mathbb{Z}^2$. Suppose (X_t, Y_t) is a simple random walk on $\mathbb{Z}^2$ started from $(1,1)$. Denote

$$t_{\text{quad}} = \inf\{t : X_t Y_t = 0\}, \qquad s_{\text{quad}} = X_{t_{\text{quad}}} + Y_{t_{\text{quad}}}. \tag{1.5}$$

In words, t_{quad} is the exit time from the first quadrant, and s_{quad} is the distance from the origin when the walk exits the quadrant. Random walk in the first quadrant has been studied quite extensively (see e.g., [12, 34, 8, 19]); for an introductory account see [24, Section 8.1.3]. Here we record that

$$\mathbb{P}[t_{\text{quad}} > j] \sim \frac{4}{\pi j} \qquad \text{and} \qquad \mathbb{P}[s_{\text{quad}} > \ell] \sim \frac{4}{\pi \ell^2} \tag{1.6}$$

asymptotically as $j \to \infty$ and $\ell \to \infty$ (see [13, Sec. 3] and [8, Example 3, Sec. 1.3]).

Theorem 1.2. *Consider the uniform k-excess subgraph of a random planar map under the distribution of $\mathbb{P}_n^k$. As $n \to \infty$ with k fixed, with probability $1 - o_n(1)$ the k-excess subgraph has k disjoint unnested loops, unnested in the sense that no loop is separated from the root edge by another loop. Let $L_1, A_1, \ldots, L_k, A_k$ denote the length and area of the k loops. Then the joint distribution of $\{(L_i, A_i)\}_{1 \le i \le k}$ converges weakly to k i.i.d. samples of the random variables $(s_{\text{quad}}, t_{\text{quad}})$ described above.*

The area distribution can be compared to the area of the cycle of a uniform spanning unicycle in $\mathbb{Z}^2$. The moments of the area were computed for $n \times n$ boxes in $\mathbb{Z}^2$ [21], and these moments suggest that $\mathbb{P}[A > j] \asymp 1/j$.

To prove these theorems, we start in Section 2 by explaining how the k^{th} moments of the amount of sand in a uniform recurrent sandpile are related to k-excess subgraphs. Then in Section 3 we review the "hamburger-cheeseburger" bijection, which constructs rooted planar maps with n edges together with a Fortuin–Kasteleyn configuration. (These FK configurations are more general than k-excess subgraphs.) This bijection is due to Mullin [29] in the special case where the FK model is a spanning tree, and to Bernardi and Sheffield [3, 35] in general. We use the formulation in [35] since it is more convenient for our purposes. In Section 4, we do some asymptotic analysis of the hamburger-cheeseburger bijection to prove Theorem 1.2. We conclude the proof of Theorem 1.1 in Section 5.

Acknowledgments: We thank Adrien Kassel for helpful comments. This work was begun while the first author was an intern in the theory group at Microsoft Research Redmond, and was completed at the Isaac Newton Institute for Mathematical Sciences in Cambridge.

2 Sandpiles, spanning trees, and the Tutte polynomial

One approach to computing the sandpile density is via the Tutte polynomial. The Tutte polynomial of an undirected graph $\mathcal{G} = (V, E)$ is a polynomial in two variables defined by

$$T_{\mathcal{G}}(x, y) = \sum_{E' \subset E} (x-1)^{\kappa(E') - \kappa(E)} (y-1)^{\kappa(E') + |E'| - |V|} \tag{2.1}$$

where $\kappa(E')$ is the number of connected components of the spanning subgraph of $\mathcal{G}$ with edge set E', i.e., the subgraph (V, E') of $\mathcal{G}$.

The Tutte polynomial of $\mathcal{G}$ has another formula:

$$T_{\mathcal{G}}(x,y) = \sum_{\text{spanning trees } t \text{ of } \mathcal{G}} x^{\#\text{ externally active edges of } t} \times y^{\#\text{ internally active edges of } t}. \tag{2.2}$$

We do not use this formula or the notion of "active edge" except to compare our results to those of Mullin [30].

For the abelian sandpile model on a finite connected graph $\mathcal{G} = (V,E)$ with sink s, Biggs defined the *level* of a sandpile configuration to be

$$\operatorname{level}(\sigma) = |\sigma| - |E|$$

and showed that $0 \le \operatorname{level}(\sigma) \le |E| - |V| + 1$ and that these bounds are tight. Consequently, the sandpile edge density satisfies

$$1 \le \rho_e \le 2\,.$$

Biggs conjectured and Merino proved [28] that the generating function of recurrent sandpiles by level is the Tutte polynomial evaluated at $(1,y)$:

$$\sum_{\sigma \in \mathcal{R}(\mathcal{G},s)} y^{\operatorname{level}(\sigma)} = T_{\mathcal{G}}(1,y)\,. \tag{2.3}$$

Notice that the generating function is independent of the choice of the sink s.

Since we are assuming that $\mathcal{G}$ is connected, from (2.1) we see

$$T_{\mathcal{G}}(1,y) = \sum_{\substack{E' \subset E \\ (V,E') \text{ connected}}} (y-1)^{1+|E'|-|V|} = \sum_{\ell \ge 0} |\mathcal{U}_\ell(\mathcal{G})| \times (y-1)^\ell\,.$$

We equate this to (2.3), write $y^{\operatorname{level}(\sigma)} = [(y-1)+1]^{\operatorname{level}(\sigma)}$, and extract the coefficient of $(y-1)^\ell$ to obtain

$$\sum_{\text{sandpiles } \sigma \text{ of } \mathcal{G}} \binom{\operatorname{level}(\sigma)}{\ell} = |\mathcal{U}_\ell(\mathcal{G})|\,, \tag{2.4}$$

so for a random recurrent sandpile $\sigma \in \mathcal{R}(\mathcal{G},s)$,

$$\mathbb{E}^{\mathcal{G}}\left[\binom{\operatorname{level}(\sigma)}{\ell}\right] = \frac{|\mathcal{U}_\ell(\mathcal{G})|}{|\mathcal{T}(\mathcal{G})|}\,. \tag{2.5}$$

When the graph $\mathcal{G}$ is itself random, in particular a random planar map on n edges, we have

$$\mathbb{E}_n^0\left[\binom{\operatorname{level}(\sigma)}{\ell}\right] = \mathbb{E}_n^0\left[\frac{|\mathcal{U}_\ell(\mathcal{M}_n)|}{|\mathcal{T}(\mathcal{M}_n)|}\right]\,. \tag{2.6}$$

This equation relates the binomial moments of the sandpile level to quantities that can be evaluated by the hamburger-cheeseburger bijection which we will describe in Section 3.

For any random variable Z, the moments $\mathbb{E}[Z^\ell]$ of Z can be expressed as linear combinations of the binomial moments $\mathbb{E}[\binom{Z}{\ell}]$, and the cumulants $\mathbb{E}[(Z-\mathbb{E}[Z])^\ell]$ can be expressed as linear combinations of products of the form $\mathbb{E}[Z^m]\mathbb{E}[Z]^{\ell-m}$. The variance in particular is $\operatorname{Var}[Z] = 2\,\mathbb{E}[\binom{Z}{2}] + \mathbb{E}[\binom{Z}{1}] - \mathbb{E}[\binom{Z}{1}]^2$. For a random recurrent sandpile σ on a random map,

$$\mathbb{E}_n^0[\operatorname{level}(\sigma(M_n))] = \mathbb{E}_n^0\left[\frac{|\mathcal{U}_1(M_n)|}{|\mathcal{T}(M_n)|}\right] \tag{2.7}$$

$$\operatorname{Var}_n^0[\operatorname{level}(\sigma(M_n))] = 2\,\mathbb{E}_n^0\left[\frac{|\mathcal{U}_2(M_n)|}{|\mathcal{T}(M_n)|}\right] + \mathbb{E}_n^0\left[\frac{|\mathcal{U}_1(M_n)|}{|\mathcal{T}(M_n)|}\right] - \left(\mathbb{E}_n^0\left[\frac{|\mathcal{U}_1(M_n)|}{|\mathcal{T}(M_n)|}\right]\right)^2\,. \tag{2.8}$$

Next we evaluate the terms on the right-hand side using the hamburger-cheesburger bijection.

3 The hamburger-cheeseburger bijection

Sheffield [35, Section 4.1] constructed a bijection called the *hamburger-cheeseburger bijection* between "perfect words" over a five-letter alphabet $\{$ h, $\boxed{\mathsf{H}}$, c, $\boxed{\mathsf{C}}$, $\boxed{\mathsf{F}}$ $\}$ and FK configurations on rooted planar maps. The letters in this alphabet can be interpreted as events at a burger restaurant. h and c indicate that a new hamburger or cheeseburger is produced. New burgers are placed on top of a burger stack, which is initially empty. $\boxed{\mathsf{H}}$ and $\boxed{\mathsf{C}}$ indicate that a customer has ordered a hamburger or cheeseburger, in which case the topmost burger of the appropriate type is removed from the stack and given to the customer. $\boxed{\mathsf{F}}$ indicates a fresh order, where the customer orders whichever burger is on top of the stack, regardless of type. A *perfect word* is a sequence of these letters for which every burger order is fulfilled by a burger already on the stack, and for which the burger stack ends empty. A perfect word of order n is one which contains n burgers and n orders, i.e., has length $2n$.

We recall that when h, $\boxed{\mathsf{H}}$, c, $\boxed{\mathsf{C}}$ are identified with the vectors $(1,0)$, $(-1,0)$, $(0,1)$, $(0,-1)$ respectively, every word W consisting of these four symbols, when read from left to right, becomes a lattice walk on $\mathbb{Z}^2$ started at $(0,0)$. The two coordinates of the walk are called the "net hamburger count" and the "net cheeseburger count". A word W is perfect if and only if both coordinates of its lattice walk remain nonnegative and the walk ends at $(0,0)$.

The hamburger-cheeseburger bijection maps a perfect word of order n to a rooted planar map (M_n, e) with n edges together with a subset E' of the edges in the map. (The edge subset E' is the FK configuration on the map.) In addition to Sheffield's description of the bijection [35, Section 4.1], a nice exposition is given by Chen [6]. We note here a few basic properties of the bijection:

1. The number of edges in E' is the number of $\boxed{\mathsf{H}}$'s plus the number of $\boxed{\mathsf{F}}$'s matching c's.
2. The number of connected components in the subgraph spanned by E' is 1 plus the number of $\boxed{\mathsf{F}}$'s matching h's.
3. Let $(E')^*$ denote the dual edges of $E \setminus E'$ on the dual map M_n^* of M_n. The number of connected components in the dual subgraph spanned by $(E')^*$ is 1 plus the number of $\boxed{\mathsf{F}}$'s matching c's.
4. Suppose that a $\boxed{\mathsf{F}}$ matches a c in a perfect word, and that in between there are ℓ $\boxed{\mathsf{H}}$'s which are fulfilled by h's before the fresh c, and $2m$ other letters (which may include other $\boxed{\mathsf{F}}$'s). Then the $\boxed{\mathsf{F}}$ corresponds the "last edge" of a loop of E' which has length $\ell+1$, the portion of the map M_n inside the loop has area $(1+\ell+2m)$, and the portion of M_n and E' inside of the loop are determined by the subword between the $\boxed{\mathsf{F}}$ and its matching c. (If there are other $\boxed{\mathsf{F}}$'s in this subword, then these correspond to loops that either share edges with or are surrounded by the loop in question.)

Proposition 3.1. *Let*

$$\Theta_n^k = \left\{\textit{perfect words of order } n \textit{ with exactly } k \ \boxed{\mathsf{F}}\textit{'s, which are all fulfilled by } \text{c}\textit{'s}\right\}. \tag{3.1}$$

Under the hamburger-cheeseburger bijection, elements in Θ_n^k correspond to triples (M_n, e, U_k), where $(M_n, e) \in \mathcal{M}_n$ and U_k is a k-excess connected subgraph of M_n. Furthermore,

$$\mathbb{E}_n^0 \left[\frac{|\mathcal{U}_k(M_n)|}{|\mathcal{T}(M_n)|} \right] = \frac{|\Theta_n^k|}{|\Theta_n^0|} \tag{3.2}$$

Proof. Suppose E' has 1 component and its dual $(E')^*$ has $k+1$ components. Then $(E')^*$ contains no cycles, so it is a $(k+1)$-component spanning forest, and consequently E' is a k-excess subgraph of M_n.

Recall that under the distribution $\mathbb{P}_n^0$, each rooted map (M_n, e) occurs with probability $|\mathcal{T}(M_n)|/\sum_{(M_n',e')\in\mathcal{M}_n}|\mathcal{T}(M_n')| = |\mathcal{T}(M_n)|/|\Theta_n^0|$. In the expectation in (3.2), the $|\mathcal{T}(M_n)|$ terms cancel, giving the right-hand side of (3.2). □

4 Asymptotic enumeration of perfect words

In this section we compute the asymptotic number of perfect words in $|\Theta_n^k|$, which together with equations (2.7), (2.8), and (3.2) gives the sandpile edge density. In the course of characterizing Θ_n^k, we also characterize the cycles in k-excess graphs.

Theorem 4.1. *For any fixed nonnegative integer k,*

$$\lim_{n\to\infty}\frac{|\Theta_n^k|}{n^k\,|\Theta_n^0|} = \frac{1}{k!\,4^k}\,. \tag{4.1}$$

To prove Theorem 4.1, we study the canonical injection from Θ_n^k to $\Theta_n^0 \times \binom{[2n]}{k}$: Given a perfect word $W \in \Theta_n^k$, let $\{i_1,\cdots,i_k\} \in \binom{[2n]}{k}$ be the index set of the k $\boxed{\mathsf{F}}$'s in W. We replace each of the k $\boxed{\mathsf{F}}$'s in W with $\boxed{\mathsf{C}}$'s to obtain a perfect word $W' \in \Theta_n^0$. This map $W \mapsto (W', \{i_1,\cdots,i_k\})$ is invertible, so it is an injection. A pair $(W,\{i_1,..i_k\}) \in \Theta_n^0\times\binom{[2n]}{k}$ is in the image of this injection precisely when $W_{i_1},\ldots,W_{i_k}$ are all $\boxed{\mathsf{C}}$'s, and just prior to each of these orders, the top burger in the stack is a ⓒ.

Theorem 4.1 can be interpreted as a statement about the probability that a random element of $\Theta_n^0 \times \binom{[2n]}{k}$ is in the image of the injection. For a random word $W \in \Theta_n^0$ and a random position i, $\mathbb{P}\big[W_i = \boxed{\mathsf{C}}\big] = \frac{1}{4}$, and at a random time, provided the stack is nonempty, the top burger on the stack is a ⓒ with probability $\frac{1}{2}$. For large n, it is plausible that these events at the same random time are approximately uncorrelated. As long as k is not too large ($k \ll \sqrt{n}$), we expect random distinct positions $i_1,\ldots,i_k$ to be far apart, and that consequently these events at the times $i_1,\ldots,i_k$ are nearly independent. Provided that this intution is correct, then $|\Theta_n^k| \approx |\Theta_n^0|\binom{2n}{k}/8^k$, which when $k \ll \sqrt{n}$ would give the theorem. In this section, we justify a more precise version of this intuition to prove the theorem.

To make a more precise statement of this approximate independence, we consider subwords of $W \in \Theta_n^0$. Let $W[a,b]$ denote the subword from positions a through b inclusive. Let w_j be the subword

$$w_j := W[\max(i_j - s + 1, 1), i_j] \tag{4.2}$$

for $j = 1,\ldots,k$, i.e., w_j is the subword of length s which ends at position i_j (unless the position is too close to the front, in which case the length will be less than s). Since the perfect word W corresponds to random walks in the quadrant that start and end at the origin, we expect the subwords $w_1,\ldots,w_k$ to be close in distribution to i.i.d. uniformly random words of length s, so long as both k and s are small enough for the subwords to be disjoint and not to contain enough letters to detect that W is not quite an unbiased random walk.

Theorem 4.2. *Assume $k^3 s^2 \ll n$, and consider the collection of subwords $(w_1,\ldots,w_k)$ defined in (4.2) from a random perfect word in Θ_n^0 and independent random indices $i_1 < \cdots < i_k$ in $\binom{[2n]}{k}$. The subwords are nonoverlapping with probability $1-o(1)$ and have total variation distance $o(1)$ from a list of k i.i.d. uniformly random words of length s which are independent of the indices. (The $o(1)$ terms go to zero as $k^3s^2/n \to 0$.)*

Proof. Let $i_1 < \cdots < i_k$ be a random k-tuple uniformly drawn from $\binom{[2n]}{k}$, and let $\hat{w}_1,\ldots,\hat{w}_k$ be k i.i.d. uniformly random words in $\{ⓗ, ⓒ, \boxed{\mathsf{H}}, \boxed{\mathsf{C}}\}^s$, which are also independent of the $i_1, \dots, i_k$. Our goal is to sample a uniformly random perfect word $W \in \Theta_n^0$ so that the subwords defined by (4.2) using the indices $i_1, \dots, i_k$ coincide with $\hat{w}_1, \dots, \hat{w}_k$. Our strategy is to first sample an independent uniformly random perfect word $X \in \Theta_n^0$, and then modify it, using $i_1, \dots, i_k$ and $\hat{w}_1, \dots, \hat{w}_k$ and some auxillary randomness, to obtain W. This modification procedure defines a Markov chain (which we will run for one step), and we restrict ourselves to modification procedures for which the Markov chain satisfies detailed balance conditional on $i_1, \dots, i_k$, i.e., $\mathbb{P}[X \to W \,|\, i_1, \dots, i_k] = \mathbb{P}[W \to X \,|\, i_1, \dots, i_k]$, so as to ensure that W is uniformly random, and in fact uniformly random even conditional on $i_1, \dots, i_k$. After verifying detailed balance, we argue that with high probability W has the desired subwords: $w_1 = \hat{w}_1, \dots, w_k = \hat{w}_k$.

We use the following modification rule: If $i_1 < s$ or $i_{j+1} - i_j < s$ for some j, then no change is made. Otherwise, all the subwords have length s and are disjoint. We then start by overwriting the relevant positions in X with the values given by $\hat{w}_1, \dots, \hat{w}_k$ to obtain a new word Y. The walk in $\mathbb{Z}^2$ defined by Y is unlikely to return to its start, so some more changes are required to rebalance it.

To describe this rebalancing of the walk, we follow Sheffield [35] in using a pair of coordinates that are rotated $45°$ from the edges of $\mathbb{Z}^2$. In these coordinates, the letters correspond to the following steps:

$$\text{ⓗ} = (+1,+1) \qquad \text{ⓒ} = (+1,-1) \qquad \boxed{\mathsf{C}} = (-1,+1) \qquad \boxed{\mathsf{H}} = (-1,-1)\,.$$

The walk associated with a word starts at $(0,0)$ and is obtained from these steps by reading the word from left to right. The first coordinate is called the "net burger count", and the second coordinate is called the "discrepancy" between hamburgers and cheeseburgers.

We let u and v respectively denote the net burger count and discrepancy of the word Y. If $u > 0$, for instance, then focusing on the first coordinate, we change $u/2$ of the $+1$'s to -1's, while ignoring the second coordinate. The second coordinate can then be rebalanced ignoring the first coordinate. When doing this rebalancing, we only change letters whose position is in the range from $\frac{1}{2}n$ to $\frac{3}{2}n$, and which do not lie within the subwords. For reasons that will become apparent, out of all possible such ways to rebalance the walk Y, we pick one uniformly at random, and let Z denote the resulting walk. (If there are no such ways to rebalance Y, we let $W = Z = X$.)

We will argue later that Z is likely to remain within the quadrant, but if not, then we let $W = X$. Otherwise, Z is a perfect word, and we would like to take $W = Z$, but to ensure detailed balance conditional on $i_1, \dots, i_k$, we consider Z to be a proposal, and use the Metropolis rule to reject this proposal with some probability and instead take $W = X$. The probabilities that X proposes Z and that Z proposes X (conditional on $i_1, \dots, i_k$) are almost the same, but there is a small difference arising from the (likely) possibility that there are different numbers of ways to do the rebalancing in the two cases. If there are r_1 ways to do the rebalancing when going from X to Z, and r_2 ways when going from Z to X, then the probability of accepting the proposed move is $\min(1, r_1/r_2)$.

We have specified the process which produces the uniformly random perfect word W, and at this point we argue that with high probability all the steps succeed so that in the end $w_1 = \hat{w}_1, \dots, w_k = \hat{w}_k$.

Since $k^2 s \ll n$, with probability $1 - o(1)$ the subwords are all disjoint and have length s.

After the positions are overwritten to obtain Y, certainly $|u|, |v| \le 2ks$. Since $ks \ll n$, the number m of letters with position between $\frac{1}{2}n$ and $\frac{3}{2}n$ outside the subwords is certainly $m = (1 + o(1))n$. Let $m_{+,*}$ and $m_{-,*}$ denote the number of these letters that are $+1$ or -1 in the first coordinate respectively, and similarly let $m_{*,+}$ and $m_{*,-}$ denote the number of these letters that are $+1$ or -1 in the second coordinate. Standard large deviation estimates together with the cycle-lemma construction of random Catalan paths [11] imply that

$$\frac{1}{2}m - \alpha\sqrt{m} \le m_{+,*}, m_{-,*}, m_{*,+}, m_{*,-} \le \frac{1}{2}m + \alpha\sqrt{m}$$

with probability tending to 1 as $\alpha \to \infty$. Assuming this event occurs, and $|u|/2, |v|/2 < \frac{1}{2}m - \alpha\sqrt{m}$, then there is a way to rebalance the walk. Indeed, the number of ways to rebalance the walk is

$$r_1 = \binom{m_{\mathrm{sign}(u),*}}{|u|/2}\binom{m_{*,\mathrm{sign}(v)}}{|v|/2}.$$

Provided that the rebalanced walk Z remains in the quadrant, the ratio r_1/r_2 is given by

$$\frac{r_1}{r_2} = \frac{\binom{m_{\mathrm{sign}(u),*}}{|u|/2}}{\binom{m_{\mathrm{sign}(-u),*}+|u|/2}{|u|/2}}\frac{\binom{m_{*,\mathrm{sign}(v)}}{|v|/2}}{\binom{m_{*,\mathrm{sign}(-v)}+|v|/2}{|v|/2}}.$$

Since $|u|, |v| \le 2ks \ll \sqrt{m}$, this ratio r_1/r_2 tends to 1. Thus the proposed move would be almost always accepted.

We are left to argue that the rebalanced walk Z almost always remains in the quadrant. Suppose that between times εn and $(2-\varepsilon)n$ the walk X remains at distance at least h from the boundary of the quadrant. If $\varepsilon \ll 1/k$ and $s \le \varepsilon n$, then with probability $1 - o(1)$ all of the subword regions are contained within the interval from εn to $(2-\varepsilon)n$. There are at most $3ks$ letters that get changed. If $h \ge 6ks$, then we would be guaranteed that the modified walk Z would remain in the quadrant.

For the initial perfect word X, if we eliminate the c and $\boxed{\text{C}}$ letters, it will be a random Catalan path of a random length 2ℓ which is concentrated around $n \pm O(\sqrt{n})$. Consider a uniformly random Catalan path of length 2ℓ. As $\ell \to \infty$, near its endpoints it behaves as a Bessel(3) process. Since a Bessel(3) process always remains positive, between positions $\varepsilon\ell$ and $(2-\varepsilon)\ell$ the Catalan path is likely to be at height at least $\delta\sqrt{\varepsilon\ell}$, with probability tending to 1 as $\delta \to 0$. Setting $\varepsilon = o(1/k)$ and $\delta = o(1)$, we find that X remains distance at least $6ks$ from the boundary of the quadrant with high probability provided $ks \ll \sqrt{n/k}$, i.e., $k^3s^2 \ll n$. □

Using this approximate i.i.d. property of the perfect words in Θ_n^0, we can characterize Θ_n^k to prove both Theorem 4.1 and Theorem 1.2.

Proof of Theorem 4.1 and Theorem 1.2. So long as $k^3s^2 \ll n$, Theorem 4.2 gives a coupling of the subwords $w_1, \ldots, w_k$ of $W \in \Theta_n^0$ at random locations $i_1 < \cdots < i_k \in \binom{[2n]}{k}$ with i.i.d. uniformly random words $\hat{w}_1, \ldots, \hat{w}_k$ of length s, so that w.h.p. the subwords equal the random words. The probability that the last letter of $\hat{w}_j$ is $\boxed{\text{C}}$ is $\frac{1}{4}$. Let $\hat{w}_j'$ be the length $s-1$ prefix of $\hat{w}_j$, i.e., excluding just its last letter. If any suffix of $\hat{w}_j'$ contains more c's than $\boxed{\text{C}}$'s or more h's than $\boxed{\text{H}}$'s, then $\hat{w}_j'$ contains a burger that is unconsumed. The probability of this event is the probability that a random walk of length $s-1$ in $\mathbb{Z}^2$ started from $(0,0)$ remains in the first quadrant, which is $1 - O(1/s)$ as $s \to \infty$. Conditional on $\hat{w}_j'$ containing an unconsumed burger, by symmetry the topmost such burger left on the stack is a c with probability $\frac{1}{2}$. Assuming the coupling between W and the i.i.d. words $\hat{w}_1, \ldots, \hat{w}_k$ holds, with probability $1 - O(k/s)$ the words $\hat{w}_1, \ldots, \hat{w}_k$ determine whether or not $(W, (i_1, \ldots, i_k))$ is in the image of the injection, and conditional on that, the answer is yes with probability 8^{-k}. For fixed k, we may choose s so that $k/s \ll 1$ and $k^3s^2 \ll n$, in which case $(W, (i_1, \ldots, i_k))$ is in the image of the injection with probability $8^{-k}(1 + o(1))$, which proves Theorem 4.1.

Assuming again that $k/s \ll 1$, $k^3 s^2 \ll n$, and that k is fixed, the above coupling gives a characterization for a random perfect word W^k in Θ_n^k. Suppose the fresh orders occur at positions $i_1 < \cdots < i_k$, and $j_1, \ldots, j_k$ are the positions of the corresponding fresh cheeseburgers. Then with high probability these indices alternate, $j_1 < i_1 < j_2 < i_2 < \cdots < j_k < i_k$, so that the loops corresponding to the k $\boxed{\mathsf{F}}$'s given by the hamburger-cheeseburger bijection (recall Item 4 in Section 3) are disjoint and unnested, and the subwords $W^k[j_1{+}1, i_1{-}1], \ldots, W^k[j_k{+}1, i_k{-}1]$ are within $o(1)$ variation distance of k i.i.d. simple random walks of the following type: when reading backwards, the net hamburger and cheeseburger counts (i.e., $\#\text{ⓗ} - \#\boxed{\mathsf{H}}$ and $\#\text{ⓒ} - \#\boxed{\mathsf{C}}$) are always nonpositive, and the net cheeseburger count ends at 0. These walks are of course equivalent to walks in the quadrant, and from the discussion in Section 3, the length of the walk corresponds to the area of the cycle, and the hamburger deficit corresponds to the length of the cycle. This proves Theorem 1.2. □

5 Sandpile density

The sandpile (edge) density is a straightfoward consequence of the above lemmas:

Proof of Theorem 1.1 (edge density). Combining equations (2.7), (2.8), (3.2), and (4.1) gives $\mathbb{E}_n^0[\mathrm{level}(\sigma(M_n))] = (1 + o(1))n/4$ and $\mathrm{Var}_n^0[\mathrm{level}(\sigma(M_n))] = o(n^2)$. The amount of sand is n more than the level, which gives the sandpile density with respect to the number of edges. □

For the random planar map with n edges, the expected number of vertices is $n/2 + 1$. To obtain the sandpile density with respect to the number of vertices, we need to know that the number of vertices is sharply concentrated about its expected value.

Lemma 5.1. *For any fixed $k > 0$, for a random planar map M_n drawn from $(\mathcal{M}_n, \mathbb{P}_n^0)$,*

$$\lim_{n\to\infty} \mathbb{E}_n^0 \left[\left(\frac{|E(M_n)|}{2|V(M_n)|} \right)^k \right] = 1.$$

Proof. Let J denote the number of ⓗ's contained in a uniformly random perfect word in Θ_n^0. With Cat_j denoting the jth Catalan number, we have

$$\begin{aligned} |\Theta_n^0| \times \mathbb{P}[J = j] &= \binom{2n}{2j} \mathrm{Cat}_j \mathrm{Cat}_{n-j} = \frac{(2n)!}{j!\,(j+1)!\,(n-j)!\,(n-j+1)!} \\ &= \binom{n+1}{j}\binom{n+1}{j+1} \frac{(2n)!}{(n+1)!(n+1)!}\,. \end{aligned}$$

As is well known, the binomial coefficients are sharply concentrated, with tails that are at least as small as Gaussian tails. In particular, there are positive constants C and c for which

$$\mathbb{P}\left[\left| J - \tfrac{n+1}{2} \right| > t \right] \le C\, e^{-c\,t^2/n}\,.$$

Recalling the properties of the hamburger-cheeseburger bijection from Section 3, the number of edges in E' is $|J|$, and since E' forms a spanning tree, the map M_n has $|V| = |J| + 1$ vertices, while of course the number of edges is $|E| = n$. Regardless of how atypical $|J|$ may be, we have $1 \le |V| \le n + 1$, so $n/(n+1) \le |E|/|V| \le n$. With, say $t = n^{2/3}$, we obtain

$$\begin{aligned} \mathbb{E}\left[(|E|/|V|)^k \right] &= \mathbb{E}[(|E|/|V|)^k \mathbf{1}_{||V|-n/2|>t}] + \mathbb{E}[(|E|/|V|)^k \mathbf{1}_{||V|-n/2|\le t}] \\ &= n^k e^{-\Theta(n^{1/3})} + 2^k(1 + O(kn^{-1/3})). \end{aligned}$$ □

Proof of Theorem 1.1 (vertex density). It suffices to prove that $\lim_{n\to\infty} \mathbb{E}_n^0[(\rho_v - 5/2)^2] = 0$.

$$\rho_v - \frac{5}{2} = 2\rho_e \left(\frac{|E|}{2|V|} - 1 \right) + 2 \left(\rho_e - \frac{5}{4} \right)$$
$$\left(\rho_v - \frac{5}{2} \right)^2 = 4\rho_e^2 \left(\frac{|E|}{2|V|} - 1 \right)^2 + 8\rho_e \left(\rho_e - \frac{5}{4} \right) \left(\frac{|E|}{2|V|} - 1 \right) + 4 \left(\rho_e - \frac{5}{4} \right)^2$$

Recall that $1 \le \rho_e \le 2$. By Lemma 5.1, $\lim_{n\to\infty} \mathbb{E}_n^0[(|E|/|V| - 2)^k] = 0$ for any fixed integer $k > 0$, so the first two terms above converge to 0 in expectation. From the edge density part of Theorem 1.1, we see that the last term converges to 0 as well. □

References

[1] O. Angel. Growth and percolation on the uniform infinite planar triangulation. *Geom. Funct. Anal.*, 13(5):935–974, 2003. MR-2024412

[2] I. Benjamini. *Coarse geometry and randomness*. Lecture notes in mathematics #2100. Springer, 2013. École d'Été de Probabilités de Saint-Flour XLI – 2011. MR-3156647

[3] O. Bernardi. Tutte polynomial, subgraphs, orientations and sandpile model: new connections via embeddings. *Electron. J. Combin.*, 15(1):Research Paper 109, 53, 2008. MR-2438581

[4] N. Berestycki, B. Laslier, and G. Ray. Critical exponents on Fortuin–Kasteleyn weighted planar maps. 2015. arXiv:1502.00450.

[5] P. Bak, C. Tang, and K. Wiesenfeld. Self-organized criticality. *Phys. Rev. A (3)*, 38(1):364–374, 1988. MR-0949160

[6] L. Chen. Basic properties of the infinite critical-FK random map. 2015. arXiv:1502.01013.

[7] N. Curien and I. Kortchemski. Percolation on random triangulations and stable looptrees. *Probab. Theory Relat. Fields*, 163(1–2):303–337, 2015. arXiv:1307.6818. MR-3405619

[8] D. Denisov and V. Wachtel. Random walks in cones. *Ann. Probab.*, 43(3):992–1044, 2015. arXiv:1110.1254. MR-3342657

[9] B. Duplantier, J. Miller, and S. Sheffield. Liouville quantum gravity as a mating of trees. 2014. arXiv:1409.7055.

[10] B. Duplantier and S. Sheffield. Liouville quantum gravity and KPZ. *Invent. Math.*, 185(2):333–393, 2011. MR-2819163

[11] A. Dvoretzky and T. Motzkin. A problem of arrangements. *Duke Math. J.*, 14:305–313, 1947. MR-0021531

[12] G. Fayolle, R. Iasnogorodski, and V. Malyshev. *Random walks in the quarter-plane*. Applications of Mathematics #40. Springer-Verlag, 1999. MR-1691900 MR-1691900

[13] G. Fayolle and K. Raschel. Some exact asymptotics in the counting of walks in the quarter plane. Discrete Math. Theor. Comput. Sci. Proc., pages 109–124. 2012. arXiv:1201.4152. MR-2957325

[14] A. Guionnet, V. F. R. Jones, D. Shlyakhtenko, and P. Zinn-Justin. Loop models, random matrices and planar algebras. *Comm. Math. Phys.*, 316(1):45–97, 2012. MR-2989453

[15] E. Gwynne, C. Mao, and X. Sun. Scaling limits for the critical Fortuin–Kasteleyn model on a random planar map I: cone times. 2015. arXiv:1502.00546.

[16] E. Gwynne and X. Sun. Scaling limits for the critical Fortuin–Kasteleyn model on a random planar map II: local estimates and empty reduced word exponent. 2015. arXiv:1505.03375.

[17] E. Gwynne and X. Sun. Scaling limits for the critical Fortuin–Kasteleyn model on a random planar map III: finite volume case. 2015. arXiv:1510.06346.

[18] A. E. Holroyd, L. Levine, K. Mészáros, Y. Peres, J. Propp, and D. B. Wilson. Chip-firing and rotor-routing on directed graphs. In *In and out of equilibrium 2*, Progress in Probability #60, pages 331–364. Birkhäuser, 2008. arXiv:0801.3306. MR-2477390

[19] I. Ignatiouk-Robert and C. Loree. Martin boundary of a killed random walk on a quadrant. *Ann. Probab.*, 38(3):1106–1142, 2010. MR-2674995

[20] M. Jeng, G. Piroux, and P. Ruelle. Height variables in the abelian sandpile model: scaling fields and correlations. *J. Stat. Mech.*, P10015, 2006. arXiv:cond-mat/0609284.

[21] A. Kassel, R. Kenyon, and W. Wu. Random two-component spanning forests. *Ann. Inst. Henri Poincaré Probab. Stat.*, 51(4):1457–1464, 2015. arXiv:1203.4858. MR-3414453

[22] A. Kassel and D. B. Wilson. The looping rate and sandpile density of planar graphs. *Amer. Math. Monthly*, 123(1):19–39, 2016. arXiv:1402.4169. MR-3453533

[23] R. W. Kenyon and D. B. Wilson. Spanning trees of graphs on surfaces and the intensity of loop-erased random walk on planar graphs. *J. Amer. Math. Soc.*, 28(4):985–1030, 2015. arXiv:1107.3377. MR-3369907

[24] G. F. Lawler and V. Limic. *Random walk: a modern introduction*. Cambridge Studies in Advanced Mathematics #123. Cambridge Univ. Press, 2010. MR-2677157

[25] J.-F. Le Gall and G. Miermont. Scaling limits of random trees and planar maps. In *Probability and statistical physics in two and more dimensions*, Clay Math. Proc. #15, pages 155–211. Amer. Math. Soc., 2012. MR-3025391

[26] L. Levine and Y. Peres. The looping constant of $\mathbb{Z}^d$. *Random Structures Algorithms*, 45(1):1–13, 2014. arXiv:1106.2226. MR-3231081

[27] S. N. Majumdar and D. Dhar. Height correlations in the abelian sandpile model. *J. Phys. A: Math. Gen.*, 24(7):L357, 1991.

[28] C. Merino López. Chip firing and the Tutte polynomial. *Ann. Comb.*, 1(3):253–259, 1997. MR-1630779 MR-1630779

[29] R. C. Mullin. On the enumeration of tree-rooted maps. *Canad. J. Math.*, 19:174–183, 1967. MR-0205882

[30] R. C. Mullin. On the average activity of a spanning tree of a rooted map. *J. Combinatorial Theory*, 3:103–121, 1967. MR-0211911

[31] V. S. Poghosyan and V. B. Priezzhev. The problem of predecessors on spanning trees. *Acta Polytechnica*, 51(1):59–62, 2011. arXiv:1010.5415.

[32] V. S. Poghosyan, V. B. Priezzhev, and P. Ruelle. Return probability for the loop-erased random walk and mean height in sandpile: a proof. *J. Stat. Mech. Theory Exp.*, P10004, 2011. arXiv:1106.5453.

[33] V. B. Priezzhev. Exact height probabilities in the abelian sandpile model. *Phys. Scr.*, T49:663–669, 1993.

[34] K. Raschel. Counting walks in a quadrant: a unified approach via boundary value problems. *J. Eur. Math. Soc. (JEMS)*, 14(3):749–777, 2012. arXiv:1003.1362. MR-2911883

[35] S. Sheffield. Quantum gravity and inventory accumulation. *Ann. Probab.*, to appear. arXiv:1108.2241.