Inventory accumulation with k products

UROP+ Final Paper, Summer 2014

Cheng Mao, Xin Sun and Tianyou Zhou

Student: Tianyou Zhou
Mentor: Cheng Mao
Project suggested by Scott Sheffield

September 1, 2014

Abstract

There is an extensive literature on random planar maps and their connections to continuous objects in dimension two. However, few results have been formulated or proved in higher dimensions. Recently, an unconventional approach to random planar maps was introduced. It studied inventory accumulation at a last-in-first-out retailer with two products, and presented a bijection between inventory accumulation trajectories and instances of critical Fortuin-Kasteleyn random planar maps.

In this paper, we generalize this inventory accumulation model to k products and prove that the corresponding random walks scale to Brownian motions with appropriate covariance matrices. We observe that a phase transition occurs at a certain critical value. Moreover, we believe that this model leads to a reasonable object in higher dimensions via bijections similar to the two-dimensional case, so our work can be viewed as a small step towards higher-dimensional correspondences of random planar maps.

1 Introduction

Planar maps are connected planar graphs embedded into the two-dimensional sphere defined up to homeomorphisms of the sphere. We allow self loops and multiple edges. Planar maps have been studied extensively in combinatorics (see the seminal work by Tutte [12]) and theoretical physics (for example [7] and [3]). One may restrict attention to planar triangulations (or quadrangulations), namely, one only allows faces in planar maps to have three edges (or four edges), counting multiplicity. In the probabilistic setting, one may choose planar maps *randomly* in a suitable class. For example, one can consider a *random planar map* uniformly distributed over planar triangulations with n faces. See Le Gall and Miermont's work [9] for more about random planar maps and their scaling limits.

In a recent work, Sheffield proposed a new approach to study random planar maps [11]. In this approach, random planar maps are coded by two random walks with certain correlation. More precisely, Sheffield first studied inventory accumulation at a last-in-first-out retailer with two products, called *hamburgers* and *cheeseburgers*. In his model, production of a hamburger, production of a cheeseburger, consumption of a hamburger, consumption of a cheeseburger and consumption of the freshest burger happen with respective probabilities $\frac{1}{4}$, $\frac{1}{4}$, $\frac{1-p}{4}$, $\frac{1-p}{4}$ and $\frac{p}{2}$ at each time point. The freshest burger means the most recently produced burger regardless of type. Then it was proved that the evolution of the two burger inventories scales to a two-dimensional Brownian motion with covariance depending on p. A phase transition happens at $p = 1/2$. In particular, when $p \geq 1/2$, the burger inventory remains balanced as the time goes to infinity, that is, the discrepancy between the two burgers remains small.

This result has its own interest, but more importantly, Sheffield constructed a bijection between the burger inventories (which are two random walks) and instances of the so-called *critical Fortuin-Kasteleyn random planar maps* [6]. Therefore, one can generate a random planar map from two random walks, and study its properties by studying the random walks, that is, the inventory trajectories in the burger model.

One advantage of this approach is that the model can be generalized to higher dimensions naturally, which is the goal of our work. In this paper, we will study inventory accumulation with k products and prove that the corresponding random walks scale to Brownian motions with certain covariance matrices. A phase transition occurs at a $p = 1-1/k$ which generalizes the two-dimensional result.

Furthermore, using bijections similar to the two-dimensional one, we expect to generate a higher-dimensional random object by inventory trajectories (i.e. three or more random walks) and study its properties in subsequent works. So far, few research has been devoted to the analogous theory of random planar maps in higher dimensions, partly due to the difficulty of enumeration and lack of bijective representations. See [1] for an interesting higher-dimensional result among the few. We hope our work can be a first step towards the study of certain higher-dimensional correspondences of random planar maps.

In Section 2, we will describe the model in detail and state the main scaling theorem. The subsequent sections are devoted to the technicalities. Most proofs in this paper are adaptations from the two-product case [11]. We will emphasize the ones that involve new ideas.

Acknowledgement

This project is made possible through the UROP+ program. We thank Pavel Etingof and Slava Gerovitch for directing the program. We thank Scott Sheffield for helpful conversations and suggestions.

2 Model setup and the main theorem

We consider a last-in-first-out retailer with k products, to which we refer as *burger 1*, ..., *burger k*. Following the construction in Section 2 of [11], we define an alphabet of symbols

$$\Theta = \{\textcircled{1}, \textcircled{2}, \dots, \textcircled{k}, \boxed{1}, \boxed{2}, \dots, \boxed{k}, \boxed{F}\}$$

which represent the k types of burgers, the corresponding k types of orders and the "flexible" order which always consumes the most recently produced burger in the remaining burger stack.

A word in the alphabet Θ is a concatenation of symbols in Θ that represents a series of events happened at the retailer. For example, if $W = \textcircled{2}\textcircled{3}\boxed{3}\textcircled{1}\boxed{2}\boxed{F}$, then the word W represents the series of events: a burger 2 is produced, a burger 3 is produced, a burger 3 is ordered, a burger 1 is produced, a burger 2 is ordered and the freshest burger is ordered, which is burger 1 in this case.

To describe the evolution of burger inventory mathematically, we consider the collection $\mathcal{G}$ of (reduced) words in the alphabet Θ modulo the following relations

$$\textcircled{i}\boxed{i} = \textcircled{i}\boxed{F} = \varnothing \quad \text{and} \quad \textcircled{i}\boxed{j} = \boxed{j}\textcircled{i} \tag{2.1}$$

where $1 \leq i, j \leq k$ and $i \neq j$. Intuitively, the first relation means that an order $\boxed{i}$ or $\boxed{F}$ consumes a preceding burger $\textcircled{i}$, and the second means that we move an order one position to the left if it does not consume the immediately preceding burger. For example,

$$W = \textcircled{2}\textcircled{3}\boxed{3}\textcircled{1}\boxed{2}\boxed{F} = \textcircled{2}\textcircled{1}\boxed{2}\boxed{F} = \textcircled{2}\boxed{2}\textcircled{1}\boxed{F} = \varnothing,$$

which is reasonable since every burger is consumed. By the same argument as in the proof of Proposition 2.1 in [11], we see that $\mathcal{G}$ is a semigroup with $\varnothing$ as the identity and concatenation as the binary operation.

Let $X(n)$ be i.i.d. random variables indexed by $\mathbb{Z}$ (i.e. time), each of which takes its value in Θ with respective probabilities

$$\{\frac{1}{2k}, \frac{1}{2k}, \dots, \frac{1}{2k}, \frac{1-p}{2k}, \frac{1-p}{2k}, \dots, \frac{1-p}{2k}, \frac{p}{2}\}.$$

Let μ denote the corresponding probability measure on the space Ω of maps from $\mathbb{Z}$ to Θ. In this paper, we follow the convention that probabilities and expectations are with respect to μ unless otherwise mentioned. For $m \leq n$, we write

$$X(m,n) := \overline{X(m)X(m+1)\cdots X(n)}$$

where $\overline{\cdot}$ means that a word is reduced modulo the relations (2.1). Then $X(m,n)$ describes the events that happen between time m and time n at the retailer.

If a burger is added at time m and consumed at time n, we define $\phi(m) = n$ and $\phi(n) = m$. Proposition 2.2 in [11] remains valid in this k-burger setting:

Proposition 2.1. *It is μ-almost surely that for every $m \in \mathbb{Z}$, $\phi(m)$ is finite. In other words, ϕ is an involution on $\mathbb{Z}$.*

Since a slight modification of the original proof will work here, we only describe the ideas. Let E_i be the event that every burger of type i is ultimately consumed. It can be shown that the union of E_i's has probability one, and since E_i's are translation-invariant, the zero-one law implies that each of them occurs almost surely. A similar argument works for orders, so each $X(m)$ has a correspondence, which is the statement of Proposition 2.1.

Hence we may define

$$Y(n) := \begin{cases} X(n) & X(n) \neq \boxed{\text{F}}, \\ \boxed{\text{i}} & X(n) = \boxed{\text{F}}, X(\phi(n)) = \textcircled{\text{i}}. \end{cases}$$

Proposition 2.1 also enables us to define the *half-infinite burger stack* $X(-\infty,n)$. Namely, $X(-\infty,n)$ is defined to be the sequence of $X(m)$ where $m \leq n$ and $\phi(m) > n$. It contains no orders since each order consumes an earlier burger. It is infinite because otherwise the length of $X(-\infty,n)$ is a simple random walk in n and will be zero at some time, but an order added at that time will consume no burgers.

For a word W in the alphabet Θ, we define $\mathcal{C}^i(W)$ to be the net burger *count* of type i, i.e., the number of $\textcircled{\text{i}}$ minus the number of $\boxed{\text{i}}$. Also, we define $\mathcal{C}(W)$ to be the total burger count, i.e.,

$$\mathcal{C}(W) := \sum_{i=1}^{k} \mathcal{C}^i(W).$$

If W has no $\boxed{\text{F}}$, then for $1 \leq i \neq j \leq k$, we define $\mathcal{D}^{ij}(W)$ to be the net *discrepancy* of burger i over burger j, i.e.,

$$\mathcal{D}^{ij}(W) := \mathcal{C}^i(W) - \mathcal{C}^j(W).$$

Definition 2.2. *Given the infinite $X(n)$ sequence, let $\mathcal{C}^i_n$ be the integer-valued process defined by $\mathcal{C}^i_0 = 0$ and $\mathcal{C}^i_n - \mathcal{C}^i_{n-1} = \mathcal{C}^i(Y(n))$ for all n. Let $\mathcal{C}_n := \sum_{i=1}^k \mathcal{C}^i_n$ and $\mathcal{D}^{ij}_n := \mathcal{C}^i_n - \mathcal{C}^j_n$.*

For any integer n, we define two vector-valued processes A_n and $\widetilde{A}_n$ by

$$A_n := (\mathcal{D}^{12}_n, \mathcal{D}^{23}_n, \ldots, \mathcal{D}^{k-1,k}_n, \mathcal{C}_n) \quad \text{and} \quad \widetilde{A}_n := (\mathcal{C}^1_n, \mathcal{C}^2_n, \ldots, \mathcal{C}^k_n).$$

We extend these definitions to real numbers by piecewise linear interpolation so that $t \mapsto A_t$ and $t \mapsto \widetilde{A}_t$ are infinite continuous paths.

When $n > 0$, we have $\mathcal{C}_n^i = \mathcal{C}^i(Y(1,n))$; when $n < 0$, we have $\mathcal{C}_n^i = \mathcal{C}^i(Y(n+1,0))$; similarly for $\mathcal{C}_n$ and $\mathcal{D}_n^{ij}$. As shorthand notations, we write

$$\mathcal{C}^i(m) = \mathcal{C}^i(Y(m)) \quad \text{and} \quad \mathcal{C}^i(m,n) = \mathcal{C}^i(Y(m,n))$$

for $m \le n$, and we let $\mathcal{C}(m), \mathcal{D}^{ij}(m), \mathcal{C}(m,n)$ and $\mathcal{D}^{ij}(m,n)$ be defined similarly.

Note that the two processes A_n and $\widetilde{A}_n$ actually code the same information about the evolution of the sequence $Y(n)$. More precisely, if we view A_n and $\widetilde{A}_n$ as column vectors, then it follows from Definition 2.2 that $A_n = M\widetilde{A}_n$ where M is an invertible matrix defined by

$$M := \begin{bmatrix} 1 & -1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & -1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & -1 & 0 \\ 0 & \cdots & 0 & 0 & 1 & -1 \\ 1 & 1 & 1 & \cdots & 1 & 1 \end{bmatrix}.$$

It is more natural to describe the evolution of $Y(1,n)$ by $\widetilde{A}_n$, since $\widetilde{A}_n$ is just a random walk on $\mathbb{Z}^k$ (or $\mathbb{R}^k$ if extended linearly) where the ith axis corresponds to the burger count of type i. However, A_n gives one more interesting perspective to view the stack $Y(1,n)$. Consider the line $\mathscr{L}$ through $(0,\dots,0)$ and $(1,\dots,1)$ in $\mathbb{R}^k$. Since $\mathcal{C}_n$ is a simple random walk along $\mathscr{L}$ and is independent of the other $k-1$ coordinates of A_n, we may view A_n as an addition of a one-dimensional simple random walk and an independent walk on the perpendicular $(k-1)$-dimensional hyperplane.

With the linear relation established between A_n and $\widetilde{A}_n$, we are ready to state two equivalent versions of the main scaling limit theorem.

Theorem 2.3. *As $\varepsilon \to 0$, the random variables $\varepsilon A_{t/\varepsilon^2}$ converge in law (with respect to the L^∞ metric on compact intervals) to*

$$(\mathbf{B}^1_{\alpha t}, B^2_t),$$

where $\mathbf{B}^1_t = (W^1_t, \dots, W^{k-1}_t)$ is a $(k-1)$-dimensional Brownian motion with covariance

$$\operatorname{Cov}(W^i_t, W^j_t) = \begin{cases} t & i = j, \\ -\frac{t}{2} & |i-j| = 1, \\ 0 & \textit{otherwise}, \end{cases}$$

B^2_t is a standard one-dimensional Brownian motion independent of $\mathbf{B}^1_t$ and $\alpha := \max\{\frac{2}{k} - \frac{2p}{k-1}, 0\}$.

Theorem 2.4. *As $\varepsilon \to 0$, the random variables $\varepsilon\widetilde{A}_{t/\varepsilon^2}$ converge in law (with respect to the L^∞ metric on compact intervals) to a k-dimensional Brownian motion*

$$\widetilde{\mathbf{B}}_t = (V^1_t, \dots, V^k_t)$$

with covariance

$$\mathrm{Cov}(V_t^i, V_t^j) = \begin{cases} (\frac{1}{k^2} - \frac{\alpha}{2k} + \frac{\alpha}{2})t & i = j, \\ (\frac{1}{k^2} - \frac{\alpha}{2k})t & i < j, \end{cases}$$

where $\alpha := \max\{\frac{2}{k} - \frac{2p}{k-1}, 0\}$.

It can be verified that $(\mathbf{B}^1_{\alpha t}, B^2_t) = M\widetilde{\mathbf{B}}_t$ in distribution, so it is not hard to see that the two theorems are indeed equivalent.

Theorem 2.3 is a direct generalization of Theorem 2.5 in [11]. We will focus on proving this version in later sections. We noted that $\mathcal{C}_n$ is a simple random walk independent of $\mathcal{D}_n^{ij}$, so it scales to B_t^2 which is independent of $\mathbb{B}_t^1$ as in the theorem. Moreover, the value of α suggests that a phase transition happens at $p = 1 - \frac{1}{k}$, which will be further explained in the next section.

To see that the limit in Theorem 2.4 is reasonable, we consider the special case $p = 0$, i.e., there are no "flexible" orders. In this case, $\widetilde{A}_n$ is a simple random walk on $\mathbb{Z}^k$, so we expect the limit to be a k-dimensional Brownian motion. Indeed, if $p = 0$, then $\alpha = 2/k$ and

$$\mathrm{Cov}(V_t^i, V_t^j) = \begin{cases} \frac{1}{k} & i = j, \\ 0 & i < j. \end{cases}$$

3 Computation of the covariance matrix and the critical value

In this section, we calculate the covariance matrix $[\mathrm{Cov}(\mathcal{D}_n^{i,i+1}, \mathcal{D}_n^{j,j+1})]_{ij}$. It determines the value of α, the critical value of p at the phase transition and the covariance matrix of the limiting Brownian motion as in Theorem 2.3.

3.1 First calculations

Following the argument in Section 3.1 of [11], we let J be the smallest positive integer for which $X(-J, -1)$ has at least one burger. We use $|W|$ to denote the length of a reduced word W and let $\chi = \chi(p) = \mathbb{E}[|X(-J, -1)|]$.

The orders in $X(-J, -1)$ are of types different from the one burger in $X(-J, -1)$. In particular, we have that

$$|\mathcal{D}^{ij}(-J, -1)| \leq |X(-J, -1)| = -\mathcal{C}(-J, -1) + 2. \tag{3.1}$$

Since $\mathcal{C}(-n, -1)$ is a martingale in n, the optional stopping theorem applied to the stopping time $J \wedge n$ implies that

$$0 = \mathbb{E}[\mathcal{C}(-1, -1)] = \mathbb{E}[\mathcal{C}(-J, -1)\mathbb{1}_{J \leq n}] + \mathbb{E}[\mathcal{C}(-n, -1)\mathbb{1}_{J > n}]. \tag{3.2}$$

For $J > n$, $\mathcal{C}(-n,-1) \le 0$, so $\mathbb{E}[\mathcal{C}(-J,-1)\mathbb{1}_{J\le n}] \ge 0$. Letting $n \to \infty$, we see that $\mathbb{E}[\mathcal{C}(-J,-1)] \ge 0$. On the other hand, $\mathbb{E}[\mathcal{C}(-J,-1)] \le 1$, so by (3.1),

$$\chi = \mathbb{E}[|X(-J,-1)|] \in [1,2]. \tag{3.3}$$

Notice that $\chi = 2$ if and only if $\mathbb{E}[\mathcal{C}(-J,-1)] = 0$. Therefore, as $n \to \infty$ in (3.2), we deduce that

$$\chi = 2 \text{ if and only if } \lim_{n\to\infty} \mathbb{E}[\mathcal{C}(-n,-1)\mathbb{1}_{J>n}] = 0. \tag{3.4}$$

By (3.1), (3.3) and symmetry, $\mathbb{E}[\mathcal{D}^{ij}(-J,-1)]$ exists and equals zero. Moreover, since $|\mathcal{D}^{ij}(-n,-1)| \le -\mathcal{C}(-n,-1)$ for $n < J$, by (3.4),

$$\chi = 2 \text{ implies that } \lim_{n\to\infty} \mathbb{E}[|\mathcal{D}^{ij}(-n,-1)|\mathbb{1}_{J>n}] = 0. \tag{3.5}$$

It turns out that there is a dichotomy between $\chi = 2$ and $1 \le \chi < 2$, which corresponds exactly to the phase transition at $p = 1 - 1/k$. In this section, we focus on the case $\chi = 2$ and show that $p \le 1 - 1/k$. We leave the case $1 \le \chi < 2$ to the following sections.

3.2 Computation of $\mathbb{E}[\mathcal{D}^{ij}(0)\mathcal{D}^{lm}(-J,-1)]$

In preparation for computing $\mathrm{Cov}(\mathcal{D}^{ij}_n, \mathcal{D}^{lm}_n) = \mathbb{E}[\mathcal{D}^{ij}_n\mathcal{D}^{lm}_n]$ for any $i \ne j$ and $l \ne m$, we first calculate $\mathbb{E}[\mathcal{D}^{ij}(0)\mathcal{D}^{lm}(-J,-1)]$.

If i, j, l and m are distinct, then $\mathcal{D}^{ij}(0)$ is independent of $\mathcal{D}^{lm}(-J,-1)$, so by symmetry

$$\mathbb{E}[\mathcal{D}^{ij}(0)\mathcal{D}^{lm}(-J,-1)] = 0. \tag{3.6}$$

Next, we evaluate $\mathbb{E}[\mathcal{D}^{ij}(0)\mathcal{D}^{ij}(-J,-1)]$ for $i \ne j$. On the event $X(0) \ne \boxed{\mathrm{F}}$, $\mathcal{D}^{ij}(0)$ is determined by $X(0)$ independently of $\mathcal{D}^{ij}(-J,-1)$, so $\mathbb{E}[\mathcal{D}^{ij}(0)\mathcal{D}^{ij}(-J,-1)] = 0$ by symmetry.

On the event $X(0) = \boxed{\mathrm{F}}$, we have $\phi(0) = -J$. Suppose $Y(0) = \boxed{\mathrm{i}}$. Then for any $j \ne i$, $\mathcal{D}^{ij}(0) = -1$, and for any other j, l, $\mathcal{D}^{jl}(0) = 0$. Because $X(-J,-1)$ contains a burger i and (possibly) orders of types other than i, it follows that

$$\begin{aligned}|X(-J,-1)| + k - 2 &= \sum_{j\ne i} \mathcal{D}^{ij}(-J,-1)\\ &= -\sum_{j\ne i} \mathcal{D}^{ij}(0)\mathcal{D}^{ij}(-J,-1) = -\frac{1}{2}\sum_{j\ne l} \mathcal{D}^{jl}(0)\mathcal{D}^{jl}(-J,-1).\end{aligned}$$

Taking the expectation of the above equation which does not depend on i, we see that conditioned on $X(0) = \boxed{\mathrm{F}}$,

$$\chi + k - 2 = -\frac{1}{2}\sum_{j\ne l} \mathbb{E}[\mathcal{D}^{jl}(0)\mathcal{D}^{jl}(-J,-1)] = -\frac{k(k-1)}{2}\mathbb{E}[\mathcal{D}^{jl}(0)\mathcal{D}^{jl}(-J,-1)] \tag{3.7}$$

by symmetry, where $j \ne l$ are arbitrary. Together with the case $X(0) \ne \boxed{\mathrm{F}}$, (3.7) implies that for any $i \ne j$,

$$\mathbb{E}[\mathcal{D}^{ij}(0)\mathcal{D}^{ij}(-J,-1)] = -\frac{p(\chi + k - 2)}{k(k-1)}, \tag{3.8}$$

since $X(0) = \boxed{\mathrm{F}}$ with probability $p/2$.

It remains to compute $\mathbb{E}[\mathcal{D}^{ij}(0)\mathcal{D}^{il}(-J,-1)]$ for distinct i, j and l. On the event $X(0) \neq \boxed{\mathrm{F}}$, because of the independence of $\mathcal{D}^{ij}(0)$ and $\mathcal{D}^{il}(-J,-1)$, we have $\mathbb{E}[\mathcal{D}^{ij}(0)\mathcal{D}^{il}(-J,-1)] = 0$ as before. On the event $X(0) = \boxed{\mathrm{F}}$ and $Y(0) \neq \boxed{\mathrm{i}}$ or $\boxed{\mathrm{j}}$, we have $\mathcal{D}^{ij}(0) = 0$, so $\mathbb{E}[\mathcal{D}^{ij}(0)\mathcal{D}^{il}(-J,-1)] = 0$. On the event $X(0) = \boxed{\mathrm{F}}$ and $Y(0) = \boxed{\mathrm{j}}$, we have $\mathcal{D}^{ij}(0) = 1$, so $\mathbb{E}[\mathcal{D}^{ij}(0)\mathcal{D}^{il}(-J,-1)] = 0$. Finally, on the event $X(0) = \boxed{\mathrm{F}}$ and $Y(0) = \boxed{\mathrm{i}}$, we observe that $\mathcal{D}^{ij}(0) = \mathcal{D}^{il}(0) = -1$, so $\mathbb{E}[\mathcal{D}^{ij}(0)\mathcal{D}^{il}(-J,-1)] = \mathbb{E}[\mathcal{D}^{il}(0)\mathcal{D}^{il}(-J,-1)]$. Summarizing the cases above, we obtain that

$$\mathbb{E}[\mathcal{D}^{ij}(0)\mathcal{D}^{il}(-J,-1)] = \mathbb{E}[\mathcal{D}^{il}(0)\mathcal{D}^{il}(-J,-1)\mathbb{1}_{X(0)=\boxed{\mathrm{F}},Y(0)=\boxed{\mathrm{i}}}]. \tag{3.9}$$

Since $\mathcal{D}^{il}(0)\mathcal{D}^{il}(-J,-1) = \mathcal{D}^{li}(0)\mathcal{D}^{li}(-J,-1)$ and $\mathcal{D}^{il}(0) = 0$ if $Y(0) \neq \boxed{\mathrm{i}}$ or $\boxed{\mathrm{l}}$,

$$\begin{aligned}\mathbb{E}[\mathcal{D}^{il}(0)\mathcal{D}^{il}(-J,-1)\mathbb{1}_{X(0)=\boxed{\mathrm{F}},Y(0)=\boxed{\mathrm{i}}}] &= \mathbb{E}[\mathcal{D}^{il}(0)\mathcal{D}^{il}(-J,-1)\mathbb{1}_{X(0)=\boxed{\mathrm{F}},Y(0)=\boxed{\mathrm{l}}}] \\ &= \frac{1}{2}\mathbb{E}[\mathcal{D}^{il}(0)\mathcal{D}^{il}(-J,-1)\mathbb{1}_{X(0)=\boxed{\mathrm{F}}}] \\ &= \frac{1}{2}\mathbb{E}[\mathcal{D}^{il}(0)\mathcal{D}^{il}(-J,-1)].\end{aligned}$$

Together with (3.9) and (3.8), this implies that

$$\mathbb{E}[\mathcal{D}^{ij}(0)\mathcal{D}^{il}(-J,-1)] = \frac{1}{2}\mathbb{E}[\mathcal{D}^{ij}(0)\mathcal{D}^{ij}(-J,-1)] = -\frac{p(\chi+k-2)}{2k(k-1)}. \tag{3.10}$$

3.3 The covariance matrix and the phase transition

By the same argument as in Section 3.1 of [11], we have

$$\mathbb{E}[\mathcal{D}^{ij}(0)\mathcal{D}^{lm}(-n,-1)] = \mathbb{E}[\mathcal{D}^{ij}(0)\mathcal{D}^{lm}(-J,-1)\mathbb{1}_{J\le n}] + \mathbb{E}[\mathcal{D}^{ij}(0)\mathcal{D}^{lm}(-n,-1)\mathbb{1}_{J>n}] \tag{3.11}$$

where the rightmost term tends to zero as $n \to \infty$ if $\chi = 2$ because of (3.5). Therefore, summarizing (3.6), (3.8) and (3.10), we see that for $i \neq j$, $l \neq m$,

$$\chi = 2 \text{ implies } \lim_{n\to\infty} \mathbb{E}[\mathcal{D}^{ij}(0)\mathcal{D}^{lm}(-n,-1)] = \begin{cases} -\frac{p}{k-1} & i = l, j = m, \\ -\frac{p}{2(k-1)} & i = l, j \neq m, \\ 0 & i, j, l, m \text{ distinct.}\end{cases} \tag{3.12}$$

Moreover, $\mathcal{D}^{ij}(0)^2 = 1$ if $Y(0)$ is of type i or j, and $\mathcal{D}^{ij}(0)\mathcal{D}^{il}(0) = 1$ if $Y(0)$ is of type i, so

$$\mathbb{E}[\mathcal{D}^{ij}(0)\mathcal{D}^{lm}(0)] = \begin{cases} \frac{2}{k} & i = l, j = m, \\ \frac{1}{k} & i = l, j \neq m, \\ 0 & i, j, l, m \text{ distinct.}\end{cases} \tag{3.13}$$

Now we evaluate $\mathrm{Cov}(\mathcal{D}^{ij}_n, \mathcal{D}^{lm}_n) = \mathbb{E}[\mathcal{D}^{ij}_n\mathcal{D}^{lm}_n]$. Using

$$\mathcal{D}^{ij}_r\mathcal{D}^{lm}_r = \mathcal{D}^{ij}(r)\mathcal{D}^{lm}(r) + \mathcal{D}^{ij}(r)\mathcal{D}^{lm}_{r-1} + \mathcal{D}^{ij}_{r-1}\mathcal{D}^{lm}(r) + \mathcal{D}^{ij}_{r-1}\mathcal{D}^{lm}_{r-1}$$

recursively for $2 \le r \le n$ and applying the translation invariance of the law of Y_m, we deduce that when $\chi = 2$,

$$\begin{aligned}
&\mathrm{Cov}(\mathcal{D}_n^{ij}, \mathcal{D}_n^{lm}) \\
=&\sum_{r=1}^{n} \mathbb{E}[\mathcal{D}^{ij}(r)\mathcal{D}^{lm}(r)] + \sum_{r=2}^{n} \mathbb{E}[\mathcal{D}^{ij}(r)\mathcal{D}_{r-1}^{lm} + \mathcal{D}_{r-1}^{ij}\mathcal{D}^{lm}(r)] \\
=&n\mathbb{E}[\mathcal{D}^{ij}(0)\mathcal{D}^{lm}(0)] + \sum_{r=2}^{n} \big(\mathbb{E}[\mathcal{D}^{ij}(0)\mathcal{D}^{lm}(1-r,-1)] + \mathbb{E}[\mathcal{D}^{lm}(0)\mathcal{D}^{ij}(1-r,-1)]\big) \\
=&\begin{cases} \frac{2n}{k} - \frac{2np}{k-1} + o(n) & i = l, j = m, \\ \frac{n}{k} - \frac{np}{k-1} + o(n) & i = l, j \neq m, \\ o(n) & i, j, l, m \text{ distinct}, \end{cases}
\end{aligned} \tag{3.14}$$

where the last equation follows from (3.12) and (3.13).

For $i = l$ and $j = m$, the variance is nonnegative, so

$$\chi = 2 \text{ implies } p \le 1 - \frac{1}{k}. \tag{3.15}$$

We remark that (3.14) and (3.15) suggest that the phase transition happens at the critical value $p = 1 - \frac{1}{k}$. Let $\alpha = \max\{\frac{2}{k} - \frac{2p}{k-1}, 0\}$. When $\chi = 2$ and $p \le 1 - \frac{1}{k}$, it follows immediately from (3.14) that

$$\mathrm{Cov}(\mathcal{D}_n^{i,i+1}, \mathcal{D}_n^{j,j+1}) = \begin{cases} \alpha n + o(n) & i = j, \\ -\frac{\alpha n}{2} + o(n) & |i - j| = 1, \\ o(n) & \text{otherwise}. \end{cases} \tag{3.16}$$

This explains why the limiting Brownian motion should have the covariance matrix as in Theorem 2.3. In the following sections, we will take care of the case $\chi < 2$ and prove that the convergence indeed happens.

4 Excursion words revisited

This section generalizes the discussion of excursion words in Section 3.3 of [11] to the k-burger case. First, we quote two results from Section 3.2 of [11] directly.

Lemma 4.1. *Let $Z_1, Z_2, Z_3, \ldots$ be i.i.d. random variables on some measure space and ψ a measurable function on that space such that $\mathbb{E}[\psi(Z_1)] < \infty$. Let T be stopping time of the process $Z_1, Z_2, \ldots$ and $\mathbb{E}[T] < \infty$. Then $\mathbb{E}[\sum_{j=1}^{T} \psi(Z_j)] < \infty$.*

Lemma 4.2. *Let $Z_1, Z_2, \ldots$ be i.i.d. random variables on some measure space and let $\mathcal{Z}_n$ be a non-negative integer-valued process adapted to the filtration of the Z_n (i.e., each $\mathcal{Z}_n$ is a function of $Z_1, Z_2, \ldots, Z_n$) that has the following properties:*

1. **Bounded initial expectation:** $\mathbb{E}[\mathcal{Z}_1] < \infty$.

2. **Positive chance to hit zero when close to zero:** *For each $k > 0$ there exists a positive chance p_k such that conditioned on any choice of $Z_1, Z_2, \dots, Z_n$ for which $\mathcal{Z}_n = k$, the conditional probability that $\mathcal{Z}_{n+1} = 0$ is at least p_k.*

3. **Uniformly negative drift when far from zero:** *There exist positive constants C and c such that if we condition on any choice of $Z_1, Z_2, \dots, Z_n$ for which $\mathcal{Z}_n \geq C$, the conditional expectation of $\mathcal{Z}_{n+1} - \mathcal{Z}_n$ is less than $-c$.*

4. **Bounded expectation when near zero:** *There further exists a constant b such that if we condition on any choice of $Z_1, Z_2, \dots, Z_n$ for which $\mathcal{Z}_n < C$, then the conditional expectation of $\mathcal{Z}_{n+1}$ is less than b.*

Then $\mathbb{E}[\min\{n : \mathcal{Z}_n = 0\}] < \infty$.

Recall some definitions in Section 3.3 of [11]:

Definition 4.3. *We define E to be the excursion word $X(1, K)$ where K is the smallest integer such that $\mathcal{C}_{K+1} < 0$.*

If i is positive, let V_i be the symbol corresponding to the ith record minimum of $\mathcal{C}_n$, counting forward from zero. If i is negative, let V_i be the $-i$th record minimum of $\mathcal{C}_n$, counting backward from zero. Denote by E_i the reduced word in between V_{i-1} and V_i (or in between 0 and V_i if $i = 1$). Note that $E = E_1$.

We still have the assertions that E almost surely contains no $\boxed{\text{F}}$ symbols and there are always as many burgers as orders in the word E. Also, the E_i's and E are i.i.d. excursion words. The following analogy to Lemma 3.5 in [11] remains true:

Lemma 4.4. *If p is such that $\chi < 2$, then the expected word length $\mathbb{E}[|E|]$ is finite, and hence the expected number of symbols in E of each type in $\{$①$, \dots,$ ⓚ$, \boxed{1}, \dots, \boxed{\text{k}}\}$ is $\mathbb{E}[|E|]/(2k)$.*

Since E is balanced between burgers and orders, the second statement follows from the first immediately by symmetry. For the first statement, it suffices to prove that the expected number of burgers in E_{-1} is finite, since E and E_{-1} have the same distribution. The original proof still works, so we omit it.

For a variant of Lemma 3.6 in [11], we consider the following sequences:

1. m**th empty order stack**: O_m is the mth smallest value of $j \geq 0$ with the property that $X(-j, 0)$ has an empty order stack.

2. m**th empty burger stack**: B_m is the mth smallest value of $j \geq 1$ with the property that $X(1, j)$ has an empty burger stack.

3. m**th left record minimum**: $L_m = L_m^0$ is the smallest value of $j \geq 0$ such that $\mathcal{C}(-j, 0) = m$. Thus, $X(-L_m, 0) = \overline{V_{-m}E_{-m} \dots V_{-1}E_{-1}}$.

4. m**th right record minimum**: $R_m = R_m^0$ is the smallest value $j \geq 1$ such that $\mathcal{C}(1, j) = -m$. Thus, $X(1, R_m) = \overline{E_1V_1 \dots E_mV_m}$.

5. **mth left minimum with no orders of type** $1,2,\dots,i$: for $1 \le i \le k$, L^i_m is the mth smallest value of $j \ge 0$ with the property that $j = L_{m'}$ for some m' and $X(-j,0)$ has no orders of type $1,2,\dots,i$.

6. **mth right minimum with no burgers of type** $1,2,\dots,i$: for $1 \le i \le k$, R^i_m is the mth smallest value of $j \ge 1$ with the property that $j = R_{m'}$ for some m' and $X(1,j)$ has no burgers of type $1,2,\dots,i$.

We observe that all these record sequences have the property that the word between two consecutive records are i.i.d.. Moreover, for $1 \le i \le k$, each L^i_m is equal to $L^{i-1}_{m'}$ for some m' by definition. Thus we can write each $X(-L^i_m, -L^i_{m-1}-1)$ as a product of consecutive words of the form $X(-L^{i-1}_{m'}, -L^{i-1}_{m'-1}-1)$. We have the following lemma:

Lemma 4.5. *The following are equivalent:*

1. $\mathbb{E}[|E|] < \infty$.
2. $\mathbb{E}[|X(-L^i_1,0)|] < \infty$ *where* $0 \le i \le k$.
3. $\mathbb{E}[|X(-O_1,0)|] < \infty$.
4. $\mathbb{E}[|X(1,R^i_1)|] < \infty$ *where* $0 \le i \le k$.
5. $\mathbb{E}[|X(1,B_1)|] < \infty$.

Proof. **1 implies 2:** Note that for $i=0$, $L^0_1 = L_1$ and $X(-L^0_1,0) = \overline{V_{-1}E_{-1}}$. Since E_{-1} and E have the same law, 2 follows immediate from 1 when $i=0$. To prove 2 for $1 \le i \le k$, we use induction.

Assume 2 holds for $i-1$. Let $H(m)$ be the number of orders of type i in $X(-L^{i-1}_m,0)$. If we can apply Lemma 4.2 with $Z_m = X(-L^{i-1}_m, -L^{i-1}_{m-1}-1)$ and $\mathcal{Z}_m = H(m)$, then $\mathbb{E}[\min\{m : H(m)=0\}] < \infty$. That means the expected number of $X(-L^{i-1}_m, -L^{i-1}_{m-1}-1)$ concatenated to produce $X(-L^i_1,0)$ is finite. Since $X(-L^{i-1}_m, -L^{i-1}_{m-1}-1)$ are identically distributed as $X(-L^{i-1}_1,0)$ which has finite expected length by inductive hypothesis, Lemma 4.1 implies that $X(-L^i_1,0)$ also has finite expected length.

Therefore it remains to check the four assumptions of Lemma 4.2. For any $m > 1$,

$$H(m) = \max\{H(m-1) - h_m, 0\} + o_m,$$

where h_m is the number of burger i in $X(-L^{i-1}_m, -L^{i-1}_{m-1}-1)$ and o_m is the number of order i in it. The expected number of burger i equals the expected number of order i in E_{-m} by Lemma 4.4, while the expected number of burger i in V_{-m} is $1/k$, which has no orders. Hence $\mathbb{E}[h_m] \ge \mathbb{E}[o_m] + 1/k$ since $X(-L^{i-1}_m, -L^{i-1}_{m-1}-1)$ is a concatenation of at least one $\overline{V_{-m'}E_{-m'}}$. Then following the same argument as in the proof of Lemma 3.6 of [11], we can verify the negative drift assumption. The other three assumptions follow easily from the construction of the sequence and the inductive hypothesis.

2 implies 3: By definition, $X(-O_1, 0)$ corresponds to the first time that the stack contains only burgers, while $X(-L_1^k, 0)$ corresponds to the first time that the stack contains only burgers and increases in length, it follows easily that $|X(-O_1, 0)| \le |X(-L_1^k, 0)|$, so the expectation is finite.

3 implies 1: The number of burgers in $X(-O_1, 0)$ is at least the number of burgers in E_{-1}, which accounts for half of its length, so $\mathbb{E}[|E_{-1}|] < \infty$. Thus the same holds for E.

The equivalence of 1, 4 and 5 are proved similarly. $\square$

Resembling Lemma 3.7 in [11], we have:

Lemma 4.6. *If $\mathbb{E}[|E|] < \infty$, then as $n \to \infty$ the fraction of* ⓘ *symbols among the rightmost n elements of $X(-\infty, 0)$ tends to $1/k$ almost surely for any i. Also, as $n \to \infty$ the fraction of* $\boxed{\text{i}}$ or $\boxed{\text{F}}$ *symbols among the leftmost n elements of $X(1, \infty)$ tends to some positive constant almost surely.*

On the other hand if $\mathbb{E}[|E|] = \infty$, then as $n \to \infty$ the fraction of $\boxed{\text{F}}$ *symbols among the leftmost n elements of $X(1, \infty)$ tends to zero almost surely.*

Proof. The proof is almost the same as in [11]. If $\mathbb{E}[|E|] < \infty$, then by Lemma 4.5, the words $X(-O_m, -O_{m-1}-1)$ are i.i.d. with finite expectation, so $X(-\infty, 0)$ is an i.i.d. concatenation of words $X(-O_m, -O_{m-1} - 1)$. The law of large numbers implies that the number of each type of burgers in $X(-O_m, 0)$ is given by $Cm + o(m)$ almost surely for some constant C. By symmetry, these constants are all equal to $\mathbb{E}[|X(-O_1, 0)|]/k$. The first statement then follows, and the second is proved analogously.

For the last statement, we note that $X(1, \infty)$ is an i.i.d. concatenation of burger-free words $X(B_{m-1}+1, B_m)$, and an $\boxed{\text{F}}$ symbol can be added only when the burger stack is empty. Hence the number of $\boxed{\text{F}}$ symbols in $X(1, B_m)$ grows like a constant times m. If $\mathbb{E}[|E|] = \infty$, Lemma 4.5 implies that $\mathbb{E}[|X(1, B_1)|] = \infty$. Thus the number of orders in $X(1, B_m)$ grows faster than any constant multiple of m almost surely, so the fraction of $\boxed{\text{F}}$ symbols tends to zero almost surely. $\square$

5 Bounded increments and large deviation estimates

We fix a semi-infinite stack $S_0 = X(-\infty, 0)$ and let $X(1), X(2), \ldots$ be chosen according to μ. Lemma 3.10 in [11] still holds in this case:

Lemma 5.1. *For $N > 0$, $\mathbb{E}[\mathcal{D}_N^{ij} | X(l) : 1 \le l \le n]$ and $\mathbb{E}[\mathcal{D}_N^{ij} | X(l) : 1 \le l \le n, \mathcal{C}_l : l \le N]$ are both martingales in n with increments of magnitude at most two.*

To prove the lemma, [11] discussed monotonicity of stacks in the two-dimensional case, which does not have a correspondence in higher dimensions. Instead, Scott Sheffield suggested us to use the notion of *neighboring stacks* here.

Definition 5.2. *Two semi-infinite stacks S_0 and S_1 are called neighbors if S_1 can be achieved from S_0 by removing an arbitrary burger from S_0, or vice versa.*

For example, $S_0 = \cdots ②①①③②②③$ and $S_1 = \cdots ②①①②②③$ are neighbors, because one can get from S_0 to S_1 by removing the fourth burger from the right.

Lemma 5.3. *If S_0 and S_1 are neighbors, then for any word W, $\overline{S_0W}$ and $\overline{S_1W}$ are still neighbors.*

Proof. Assume that we get S_1 from S_0 by deleting a ⓙ. By induction, we may also assume that W contains a single element.

If W is a burger, then for $\sigma = 1, 2$, $\overline{S_\sigma W}$ is achieved by adding W onto S_σ. If $W = \boxed{\text{F}}$, then $\overline{S_\sigma W}$ is achieved by deleting the rightmost burger from S_σ. If $W = \boxed{\text{i}}$, then $\overline{S_\sigma W}$ is achieved by deleting the rightmost ⓘ from S_σ. Hence in these three cases, it is easily seen that the resulting two stacks are still neighbors.

If $W = \boxed{\text{j}}$ and there is a ⓙ in S_0 to the right of the ⓙ which we deleted to get S_1, then $\overline{S_\sigma W}$ is achieved by deleting the rightmost ⓙ from S_σ. Hence the resulting two stacks are neighbors. Otherwise, the ⓙ deleted to get S_1 is the rightmost ⓙ in S_0, so $\overline{S_0W} = S_1$. Hence $\overline{S_0W}$ and $\overline{S_1W}$ are neighbors. □

Proof of Lemma 5.1. Since the two conditional expectations are clearly martingales in n, we only need to prove that the increments are bounded. To this end, it suffices to show that changing $X(l)$ for a single $1 \leq l \leq N$ only changes $\mathcal{D}_N^{ij}$ by at most two.

Suppose that $X(l)$ is changed to $X(l)'$. Here we make the convention that a product of words is always reduced. It is easy to see that $X(-\infty, l)$ and $X(-\infty, l-1)X(l)'$ have a common neighbor $X(-\infty, l-1)$. Lemma 5.3 then implies that $X(-\infty, N)$ and $X(-\infty, l-1)X(l)'X(l+1, N)$ have a common neighbor $X(-\infty, l-1)X(l+1, N)$. Since the ij-discrepancy differs by at most one between neighbors, we see that $\mathcal{D}_N^{ij}$ changes by at most two if we change a single $X(l)$. □

The following large deviation estimates are modifications of Lemma 3.12 and 3.13 in [11].

Lemma 5.4. *Fix any $p \in [0, 1]$ and a semi-infinite stack $S_0 = X(-\infty, 0)$. There exist positive constants C_1 and C_2 such that for any choice of S_0, $a > 0$, $n > 1$ and any i, j,*

$$\mathbb{P}(\max_{1 \leq l \leq n} |\mathcal{C}_l| > a\sqrt{n}) \leq C_1 e^{-C_2 a} \quad \textit{and} \quad \mathbb{P}(\max_{1 \leq l \leq n} |\mathcal{D}_l^{ij}| > a\sqrt{n}) \leq C_1 e^{-C_2 a}.$$

The original proof carries over almost verbatim. The idea is to use Lemma 5.1 to give bounded increments, and then apply a pre-established large deviation estimate of martingales with bounded increments.

We remark that it is an important technique to estimate the deviation of martingales with bounded jumps. See [4] for more interesting results.

Lemma 5.5. *Fix any $p \in [0, 1]$. There exist positive constants C_1 and C_2 such that for any $a > 0$ and $n > 1$,*

$$\mathbb{P}(|X(1, n)| > a\sqrt{n}) \leq C_1 e^{-C_2 a}.$$

Proof. Let the semi-infinite stack S_0 be rotating among ①, ..., ⓚ. Suppose that $\mathcal{C}_l$ and all $\mathcal{D}_l^{ij}$ fluctuate by at most $a\sqrt{n}/(4k-1)$ for $1 \le l \le n$.

Claim that no burger in S_0 expect the rightmost $a\sqrt{n}(2k-1)/(4k-1)$ burgers will be consumed in the first n steps. Assume the opposite. If the first such burger is consumed at step l and is an ⓜ, then at this moment all burgers to the right are of types different from ⓜ. Since $\mathcal{C}_l \ge -a\sqrt{n}/(4k-1)$, there are at least $a\sqrt{n}(2k-2)/(4k-1)$ burgers above the ⓜ. Among them there are at least $2a\sqrt{n}/(4k-1)$ burgers of some type $m' \ne m$. Hence $|\mathcal{D}_l^{mm'}| > a\sqrt{n}/(4k-1)$, which is a contradiction.

It follows from the claim that there are at most $a\sqrt{n}(2k-1)/(4k-1)$ orders in $X(1,n)$. Since $\mathcal{C}_l$ fluctuates by at most $a\sqrt{n}/(4k-1)$, there are at most $2ka\sqrt{n}/(4k-1)$ burgers in $X(1,n)$. Therefore, $|X(1,n)| \le a\sqrt{n}$.

Thus, to have $|X(1,n)| > a\sqrt{n}$, $\mathcal{C}_l$ or at least one $\mathcal{D}_l^{ij}$ must fluctuate by more than $a\sqrt{n}/(4k-1)$. An application of Lemma 5.4 then completes the proof. □

6 The case $\chi < 2$

In this section, we will resolve the remaining case from Section 3, i.e., the case $\chi < 2$. We will use the results from Section 4 and 5 to prove that when $\chi < 2$, the scaling limit of A_n on a compact interval has the law of a one-dimensional Brownian motion. This means that the total burger count $\mathcal{C}_n$ dominates. As we remarked after the statement of Theorem 2.3, $\mathcal{C}_n$ is a simple random walk and thus scales to a Brownian motion, so it suffices to show that $\mathcal{D}_n^{ij}$ scales to 0 in law on compact intervals.

In addition to the statement above, we will show that $\chi < 2$ implies that $p > 1 - 1/k$. Together with (3.15), this gives the dichotomy mentioned in Section 3.1, namely,

$$\chi < 2 \iff p > 1 - 1/k \quad \text{and} \quad \chi = 2 \iff p \le 1 - 1/k. \tag{6.1}$$

Thus this section proves Theorem 2.3 in the case $p > 1 - 1/k$. We divide the proof into three lemmas.

Lemma 6.1. *If $\mathbb{E}[|E|] < \infty$ (which holds when $\chi < 2$), then* $\mathrm{Var}[\mathcal{D}_n^{ij}] = o(n)$ *for all pairs (i,j).*

Proof. First, we prove that the random variables $n^{-1/2}\mathcal{D}_n^{ij}$ converge to 0 in probability. To do this, we consider the following events:

1. $|X(1,n)| < a\sqrt{n}$.
2. The top $2ka\sqrt{n}$ burgers in stack $X(-\infty,0)$ are well balanced among all burger types with error $\varepsilon\sqrt{n}$, i.e., the number of burgers of any type is between $(2a-\varepsilon)\sqrt{n}$ and $(2a+\varepsilon)\sqrt{n}$.
3. The top b burgers in the stack $X(-\infty,n)$ are well balanced among all burger types with error $\varepsilon\sqrt{n}$ for all $b > (2k-1)a\sqrt{n}$.

We assert that if all three events happen, then $|n^{-1/2}\mathcal{D}_n^{ij}| < 4\varepsilon$. First, 1 and 2 together imply that all the orders in $X(1,n)$ are fulfilled by the top $2ka\sqrt{n}$ burgers in $X(-\infty,0)$, so the burgers below height $-2ka\sqrt{n}$ in $X(-\infty,0)$ are not affected by $X(1,n)$. Hence the stacks $X(-\infty,0)$ and $X(-\infty,n)$ are identical below height $-2ka\sqrt{n}$. On the other hand, $|X(1,n)| < a\sqrt{n}$ implies that $|\mathcal{C}_n| < a\sqrt{n}$, so the number of burgers in $X(-\infty,n)$ above height $-2ka\sqrt{n}$ is at least $(2k-1)a\sqrt{n}$. By 2 and 3, the discrepancies between two burger types above height $-2ka\sqrt{n}$ are less than $2\varepsilon\sqrt{n}$ for both stacks, so $|\mathcal{D}_n^{ij}|$ is at most $4\varepsilon\sqrt{n}$, as desired.

Next, we observe that all three events happen with high probability if we choose a and n properly. For fixed $\varepsilon > 0$, we first choose a large enough so that 1 happens with high probability using Lemma 5.4. Then by Lemma 4.6, we choose n large enough so that 2 and 3 happen with high probability.

Thus we conclude that $\lim_{n\to\infty} \mathbb{P}[|n^{-1/2}\mathcal{D}_n^{ij}| > \varepsilon] = 0$ for all $\varepsilon > 0$, i.e., $n^{-1/2}\mathcal{D}_n^{ij}$ converge to 0 in probability.

It remains to check that $\mathrm{Var}[n^{-1/2}\mathcal{D}_n^{ij}] = \mathbb{E}[n^{-1}(\mathcal{D}_n^{ij})^2]$ tends to 0 as $n \to \infty$. This follows from the fact that $n^{-1}(\mathcal{D}_n^{ij})^2$ tends to 0 in probability together with the uniform bounds on the tails given by Lemma 5.4. □

The following two lemmas are proved in exactly the same way as in [11], so we omit the proofs.

Lemma 6.2. *If* $\mathrm{Var}[\mathcal{D}_n^{ij}] = o(n)$*, then* $n^{-1/2}\max\{|\mathcal{D}_l^{ij}| : 1 \le l \le nt\}$ *converges to zero in probability as* $n \to \infty$ *for any fixed* $t > 0$.

The trick of the proof is to first divide the time interval into small subintervals, then observe the convergence at the end points, and finally use approximation to complete the proof. Note that by Lemma 6.2, we immediately obtain that A_n converges in law to a one-dimensional Brownian motion on compact intervals.

Lemma 6.3. *If* $\chi < 2$ *and* $\mathrm{Var}[\mathcal{D}_n^{ij}] = o(n)$*, then*

$$\lim_{n\to\infty} \mathbb{E}[|\mathcal{D}^{ij}(-n,-1)|\mathbb{1}_{J>n}] = 0.$$

Interested readers may refer to the proof of the original lemma which involves introducing new measures via Radon-Nikodym derivatives and recentering the sequence. The original proof also uses the fact that one-dimensional random walk conditioned to stay positive scales to a three-dimensional Bessel process, which is explained by [10].

Letting $n \to \infty$ in (3.11) and using Lemma 6.3 and (3.8), we deduce that

$$\lim_{n\to\infty} \mathbb{E}[\mathcal{D}^{ij}(0)\mathcal{D}^{lm}(-n,-1)] = \mathbb{E}[\mathcal{D}^{ij}(0)\mathcal{D}^{lm}(-J,-1)] = -\frac{p(\chi+k-2)}{k(k-1)}.$$

Following the same computation as in (3.14), we obtain that

$$\mathrm{Var}(\mathcal{D}_n^{ij}) = \frac{2n}{k} - \frac{2np(\chi+k-2)}{k(k-1)} + o(n).$$

By Lemma 6.1, we must have $\frac{2n}{k} = \frac{2np(\chi+k-2)}{k(k-1)}$, i.e., $p = \frac{k-1}{\chi+k-2}$. Hence $\chi < 2$ implies that $p > 1 - 1/k$, which gives us the promised dichotomy (6.1).

7 The case $\chi = 2$

It finally remains to prove the main theorem in the case $\chi = 2$. First, if $p = 1 - 1/k$, then $\mathrm{Var}[\mathcal{D}_n^{ij}] = o(n)$ by (3.14), so the convergence follows from our argument in Section 6.

Next, we may assume $p < 1 - 1/k$, so that $\mathrm{Var}[\mathcal{D}_n^{ij}] \neq o(n)$. By the contrapositive of Lemma 6.1, we must have $\mathbb{E}[|E|] = \infty$. Then we can apply the second part of Lemma 4.6, which asserts that the number of $\boxed{\text{F}}$ symbols in $X(1, n)$ is small relative to the total number of orders in $X(1, n)$ as n gets large. To be more precise, the number of $\boxed{\text{F}}$ in $X(1, \lfloor tn \rfloor)$ is $o(\sqrt{n})$ with probability tending to one as $n \to \infty$ by Lemma 4.6 and Lemma 5.5. Therefore, for $t_1 + t_2 = t_3$, the laws of $A_{\lfloor t_1 n \rfloor}$ and $A_{\lfloor t_2 n \rfloor}$ add to the law of $A_{\lfloor (t_1+t_2) n \rfloor}$ up to an error of $o(\sqrt{n})$ with high probability.

On the other hand, since the variances of the random variables $n^{-1/2}A_{tn}$ converge to constants as $n \to \infty$ for fixed t, at least subsequentially the random variables $n^{-1/2}A_{tn}$ converge in law to a limit. Moreover, if we choose a finite collection of t values, namely $0 < t_1 < t_2 < \cdots < t_m < \infty$, the joint law of $(n^{-1/2}A_{\lfloor t_1 n \rfloor}, n^{-1/2}A_{\lfloor t_2 n \rfloor}, \ldots, n^{-1/2}A_{\lfloor t_m n \rfloor})$ also converges subsequentially to a limit law.

Now we combine the two observations above. We have that the law of $n^{-1/2}A_{\lfloor tn \rfloor}$ is equal to the law of the sum of l independent copies of $n^{-1/2}A_{\lfloor tn/l \rfloor}$ plus a term which is $o(1)$ with high probability (since we have multiplied by $n^{-1/2}$). Hence, the subsequential weak limit of $n^{-1/2}A_{\lfloor tn \rfloor}$ must equal the sum of l i.i.d. random variables. In particular, since l is arbitrary, the limit law has to be infinitely divisible. Notice that the process $n^{-1/2}A_{\lfloor tn \rfloor}$ is almost surely continuous in t, so we conclude that the subsequential limit discussed above has to be a Gaussian with mean zero. We refer to [8] and [2] for more background on infinitely divisible processes, Lévy processes and Gaussian processes.

The covariance matrix of $n^{-1/2}A_n$ is already given by our calculation in Section 3, and Lemma 5.4 guarantees that $n^{-1/2}A_{\lfloor tn \rfloor}$ are tight, so the subsequential limit has the correct covariance matrix. We conclude that the limit indeed has the Gaussian distribution given in Theorem 2.3. Moreover, our argument implies that any subsequence of $n^{-1/2}A_{tn}$ has a further subsequence converging in law to this Gaussian distribution, so the whole sequence converges to this law. See [5] for more details.

The same is true if we choose a finite collection of t_i's, so the finite-dimensional joint law of $(n^{-1/2}A_{\lfloor t_1 n \rfloor}, n^{-1/2}A_{\lfloor t_2 n \rfloor}, \ldots, n^{-1/2}A_{\lfloor t_m n \rfloor})$ converges to a limit law, which is exactly the law of $(\mathbf{W}_{t_1}, \mathbf{W}_{t_2}, \ldots, \mathbf{W}_{t_m})$, where $\mathbf{W}_t$ is the k-dimensional Brownian motion $(\mathbf{B}^1_{\alpha t}, B^2_t)$ described in Theorem 2.3.

The transition from a discrete collection of t_i's to a compact interval follows in the same way as in the proof of Lemma 6.2. As the maximum gap between t_i's gets smaller, the probability that (the norm of) the fluctuation in some interval $[t_i, t_{i+1}]$ exceeds ε tends to zero as $n \to \infty$ for both $n^{-1/2}A_{\lfloor tn \rfloor}$ and $\mathbf{W}_t$ where $t \in [0, t_m]$. Hence the two processes are uniformly close on the interval $[0, t_m]$ with probability tending to one as $n \to \infty$. Therefore, Theorem 2.3 is fully proved.

References

[1] Itai Benjamini and Nicolas Curien. On limits of graphs sphere packed in euclidean space and applications. *European Journal of Combinatorics*, 32(7):975–984, 2011.

[2] Jean Bertoin. *Lévy processes*, volume 121. Cambridge university press, 1998.

[3] Edouard Brézin, Claude Itzykson, Giorgio Parisi, and Jean-Bernard Zuber. Planar diagrams. *Communications in Mathematical Physics*, 59(1):35–51, 1978.

[4] Amir Dembo. Moderate deviations for martingales with bounded jumps. *Elect. Comm. in Probab*, 1:11–17, 1996.

[5] Rick Durrett. *Probability: theory and examples.* Cambridge university press, 2010.

[6] Cornelius Marius Fortuin and Piet W Kasteleyn. On the random-cluster model: I. introduction and relation to other models. *Physica*, 57(4):536–564, 1972.

[7] Gerard Hooft. A planar diagram theory for strong interactions. *Nuclear Physics B*, 72(3):461–473, 1974.

[8] Ioannis Karatzas and Steven E Shreve. *Brownian motion and stochastic calculus*, volume 113. springer, 1991.

[9] Jean-François Le Gall, Grégory Miermont, et al. Scaling limits of random trees and planar maps. *Probability and statistical physics in two and more dimensions*, 15:155–211, 2012.

[10] James W Pitman. One-dimensional brownian motion and the three-dimensional bessel process. *Advances in Applied Probability*, pages 511–526, 1975.

[11] Scott Sheffield. Quantum gravity and inventory accumulation. *arXiv preprint math.PR/1108.2241v1*, 2011.

[12] WT Tutte. A census of planar maps. *Canad. J. Math*, 15(2):249–271, 1963.