Lecture notes on Spatial Random Permutations

Ron Peled*

June 7, 2015

Abstract

These notes accompany a sequence of lectures given at the Warsaw Probability Summer School on Spatial Random Permutations. Topics include random uniform and Ewens permutations, the interchange model and its analysis on the complete graph and on trees, a continuum model for spatial random permutations in $\mathbb{R}^d$ related to the Feynman-Kac representation of the ideal Bose gas and a discussion of models of spatial random permtuations in one dimension which are collectively referred to as band permutations. The notes are not in final state and any comments or corrections are welcome.

1 Introduction

A *spatial random permutation* is a random permutation which is biased towards the identity in some underlying geometry. For instance, given a finite graph G, can sample a random permutation π of the vertices of G with probability proportional to $\exp(-d(\pi, \mathrm{Id}))$, where $d(\pi, \mathrm{Id})$ is a measure of distance between π and the identity permutation Id which respects the geometry of G, such as $\sum_v d_G(\pi(v), v)$ or $\sum_v d_G(\pi(v), v)^2$. Later we will see other examples, paying special attention to the so-called *interchange process*. The study of such random permutations stems from physics, where they are related to the phenomenon of Bose-Einstein condensation and to properties of quantum models such as the quantum Heisenberg ferromagnet. We will not discuss the physical theory here and present only some aspects of the mathematical study (but see Daniel Ueltschi's talk for a discussion of the physical connections). Our main focus will be on the cycle structure of spatial random permutations and specifically on the question of whether macroscopic cycles appear.

2 Uniform and Ewens permutations

We start by discussing the cycle structure of non-spatial random permutations.

Notation. We denote by S_n the *permutation group* on n elements. For a permutation $\pi \in S_n$ and $1 \le i \le n$ we write $\ell_i(\pi)$ for the *length of the cycle* which contains i in π. We

*School of Mathematical Sciences, Tel Aviv University, Tel Aviv, Israel. E-mail: peledron@post.tau.ac.il.

write $r_j(\pi)$ for the number of cycles of π whose length is exactly j, so that $\sum_{j=1}^n jr_j(\pi) = n$ for all $\pi \in S_n$, and set $\mathbf{r}(\pi) := (r_1(\pi), r_2(\pi), \ldots, r_n(\pi))$. We write $C(\pi)$ for the number of cycles of π.

A *uniform permutation* π in S_n is a random permutation having equal chance to be any of the $n!$ permutations of S_n. When π is uniform, is it more likely for the point 1 to be a fixed point of π or to be a member of a giant cycle spanning all the n elements? It may be surprising at first to learn that the cycle lengths of uniform permutations are themselves uniform, i.e., that

$$\mathbb{P}(\ell_i(\pi) = t) = \frac{1}{n}, \quad 1 \le i, t \le n. \tag{1}$$

This fact is most easily proved by a direct calculation. It implies that a uniform permutation has cycles with length of order n. The main question we shall pursue in this course is whether such macroscopic cycles appear also for other distributions on permutations.

It will sometimes be useful for us to focus only on the lengths of cycles in a permutation, forgetting the precise numbers lying in each cycle. To this end, the following combinatorial exercise is useful.

Exercise 2.1. *Let π be a uniform permutation on n elements. For every $\mathbf{r} = (r_1, r_2, \ldots, r_n)$ such that $r_j \ge 0$, $\sum_{j=1}^n jr_j = n$ we have*

$$\mathbb{P}(\mathbf{r}(\pi) = \mathbf{r}) = \frac{1}{\prod_{j=1}^n j^{r_j} \cdot r_j!}.$$

It will be of interest to embed the uniform permutation as one instance of a wider model. The *Ewens distribution* on permutations with parameter $\theta > 0$ is the probability measure on S_n in which

$$\mathbb{P}(\pi) = \frac{\theta^{C(\pi)}}{Z_{n,\theta}}, \tag{2}$$

where $Z_{n,\theta}$ is an appropriate normalization constant (given explicitly by (3) below). The case $\theta = 1$ corresponds to the uniform distribution. The Ewens distribution was introduced by Ewens in 1972 [17] (see also [21]) in the study of a mathematical biology model for the genealogy of a population undergoing mutation. In this context, one is interested solely in the induced distribution on $\mathbf{r}(\pi) = (r_1(\pi), r_2(\pi), \ldots, r_n(\pi))$ The definition (2) of the Ewens distribution and Exercise 2.1 imply that

$$\mathbb{P}(\mathbf{r}(\pi) = \mathbf{r}) = \frac{n!}{Z_{n,\theta}} \frac{\theta^{C(\pi)}}{\prod_{j=1}^n j^{r_j} \cdot r_j!}$$

when π is sampled from the Ewens distribution, a result known as the *Ewens sampling formula* (see also (3) below).

The Ewens distribution is amenable to analysis via the following algorithm for generating a sample from the distribution, which is termed the *Chinese restaurant process*. Consider a restaurant with circular tables to which n customers enter. The first customer sits at one of the tables. Then, inductively, the k'th customer decides either to sit immediately to the right of one of the previous $k-1$ customers or to sit alone at a new table, with the probability to sit to the right of each customer being $\frac{1}{\theta+k-1}$ and the probability to open a new table being

$\frac{\theta}{\theta+k-1}$. After all n customers sit, their positions determine a permutation π whose cycles are exactly the tables. As an illustration, the probability that this process will generate the permutation 4371265, whose cycle structure is (14)(2375)(6), is

$$\mathbb{P}(\pi = 4371265) = \frac{\theta}{\theta+1} \cdot \frac{1}{\theta+2} \cdot \frac{1}{\theta+3} \cdot \frac{1}{\theta+4} \cdot \frac{\theta}{\theta+5} \cdot \frac{1}{\theta+6} = \frac{\theta^3}{\prod_{k=1}^{n}(\theta+k-1)}.$$

Similarly, one sees that the probability of obtaining any permutation σ under this process is

$$\mathbb{P}(\pi = \sigma) = \frac{\theta^{C(\pi)}}{\prod_{k=1}^{n}(\theta+k-1)}.$$

We conclude that the Chinese restaurant process indeed generates samples from the Ewens distribution and, in addition, that the normalizing constant $Z_{n,\theta}$ of the Ewens distribution satisfies

$$Z_{n,\theta} = \prod_{k=1}^{n}(\theta+k-1). \tag{3}$$

It follows also that the number of cycles $C(\pi)$ in an Ewens random permutation satisfies

$$C(\pi) = X_1 + X_2 + \cdots + X_n$$

with the (X_k) independent Bernoulli random variables such that $\mathbb{P}(X_k = 1) = \frac{\theta}{\theta+k-1}$. Thus, in particular,

$$\begin{aligned}
\mathbb{E}(C(\pi)) &= \theta\left(\frac{1}{\theta} + \cdots + \frac{1}{\theta+n-1}\right) \sim \theta\log(n), \quad n\to\infty,\\
\mathrm{Var}(C(\pi)) &= \theta\left(\frac{1}{\theta}\left(1-\frac{1}{\theta}\right) + \cdots + \frac{1}{\theta+n-1}\left(1-\frac{1}{\theta+n-1}\right)\right) \sim \theta\log(n), \quad n\to\infty,
\end{aligned}$$

where we write $a_n \sim b_n$ as $n\to\infty$ to denote that $\lim_{n\to\infty}\frac{a_n}{b_n} = 1$. It follows that $C(\pi)$ tends to infinity as n grows in the sense that $\mathbb{P}(C(\pi)\le k)\to 0$ for each fixed k. Moreover, $C(\pi)$ satisfies a central limit theorem in that $\frac{C(\pi)-E(C(\pi))}{\sqrt{\mathrm{Var}(C\pi)}}$ converges to the standard Gaussian distribution. Another conclusion that we may immediately draw from the process is that

$$\mathbb{P}(\text{points } i \text{ and } j \text{ are in the same cycle of } \pi) = \frac{1}{\theta+1}.$$

This follows as, by symmetry, we may take $i=1, j=2$, in which case the Chinese restaurant process immediately tells us that the chance that customer 2 sits at the table of customer 1 is $\frac{1}{\theta+1}$.

To obtain information on the $\ell_i(\pi)$, the length of the cycle containing i, a different algorithm for sampling from the Ewens distribution is handy. Recall that we may write π in cycle notation with numbers and parentheses, e.g., the permutation 4371265 is written as (14)(2375)(6). In the algorithm we sample this cycle notation. As the first step we write (1 to denote the beginning of the cycle notation. After k steps we have already put the first k numbers in the notation, e.g., when $k=5$ for the example permutation then we have written (14)(237. Then we append the next number to the current cycle, giving probability $\frac{1}{\theta+n-k}$ for it to be any of the remaining $n-k$ numbers, or, with probability $\frac{\theta}{\theta+n-k}$ we close the current cycle and open a new one by appending)(j to the cycle notation, with j being the first number which has not already been put down. After the n'th step we finish the notation by appending). It is simple to check that the probability of a permutation π under this algorithm is proportional to $\theta^{C(\pi)}$ and hence this algorithm also generates samples from the Ewens distribution. The algorithm implies that

$$\mathbb{P}(\ell_i(\pi) = t) = \frac{n-1}{\theta+n-1} \cdot \frac{n-2}{\theta+n-2} \cdot \cdots \cdot \frac{n-(t-1)}{\theta+n-(t-1)} \cdot \frac{\theta}{\theta+n-t}, \quad 1 \le i, t \le n. \quad (4)$$

This is straightforward from the algorithm for $i = 1$ and follows for other i by symmetry. In particular, taking $\theta = 1$ we recover the uniform distribution (1) for cycle lengths of a uniform permutation. In addition, if we sort the cycles of π in terms of their minimal element and let $(\ell'_1, \ell'_2, \cdots, \ell'_{C(\pi)})$ be the lengths of cycles in this sorted order then we may conclude from the algorithm that, for any $k \ge 1$, the distribution of ℓ'_k conditioned on $C(\pi) \ge k$ and $\ell'_1, \ldots, \ell'_{k-1}$ is given by the distribution (4) with n replaced by $n - \sum_{i=1}^{k-1} \ell'_i$. With a bit of effort we also obtain the following limit theorem.

Exercise 2.2. *Prove that when π is sampled from the Ewens distribution then $\frac{\ell_1(\pi)}{n}$ converges in distribution to the* Beta$(1,\theta)$ *distribution (for $\theta = 1$ this is the uniform distribution on* $[0, 1]$*). Precisely,*

$$\mathbb{P}(\ell_1(\pi) \le xn) \to \theta \int_0^x (1-x)^{\theta-1} dx, \quad \text{as } n \to \infty \text{ with } 0 \le x \le 1 \text{ and } \theta \text{ fixed.}$$

The preceding remarks now imply that the normalized lengths of the cycles of π sorted by their minimal element converge in distribution to the GEM(θ) *distribution* named after Griffiths, Engen and McCloskey. Precisely, that if $X_1, X_2, \ldots$ is a sequence of independent Beta$(1,\theta)$ random variables then for any fixed $k \ge 1$,

$$\left(\frac{\ell'_1(\pi)}{n}, \frac{\ell'_2(\pi)}{n}, \ldots, \frac{\ell'_k(\pi)}{n}\right) \xrightarrow{d} (X_1, (1-X_1)X_2, \ldots, (1-X_1)(1-X_2)\cdots(1-X_{k-1})X_k).$$

The GEM(θ) distribution is sometimes called the *stick breaking construction* as we may think of generating a sample from it by starting with a stick of length 1, breaking it in two at a location which is Beta$(1,\theta)$ distributed, then continuing to break the second piece of the stick at a location which is Beta$(1,\theta)$ distributed relative to the length of that stick and continuing in this manner, each time breaking the last piece of the stick at a location which is Beta$(1,\theta)$ distributed relative to its length.

Lastly, we mention that the limiting distribution for the normalized cycle lengths in *sorted order* is also known explicitly and called the *Poisson-Dirichlet distribution* with parameter θ, or PD(θ) for short. This is the distribution on infinite sequences $(a_1, a_2, \ldots)$ with $a_1 \ge a_2 \ge \cdots$, $\sum a_i = 1$ which is obtained by sorting the infinite GEM(θ) sequence

$$(X_1, (1-X_1)X_2, (1-X_1)(1-X_2)X_3, \ldots), \quad (X_i) \text{ IID Beta}(1,\theta).$$

2.1 The random transposition shuffle

Consider a deck of n cards which is initially sorted. In order to shuffle the deck, one samples uniformly at random two (possibly equal) positions $1 \le i, j \le n$ and transposes (exchanges) the cards at positions i and j in the deck (if $i = j$ then nothing is done). How many such transposition steps are required for the state of the deck to be approximately uniform? To study this question rigorously we introduce the following terms. A state of the deck is a permutation $\pi \in S_n$. We write π_t, $t \ge 0$, for the state of the deck after exactly t transposition steps, with π_0 being the initial state of the deck which we take to be the identity permutation. We measure the distance of the distribution of π_t from uniform with the *total variation distance*, given by

$$d_t := \frac{1}{2}\sum_{\sigma\in S_n}\left|\mathbb{P}(\pi_t = \sigma) - \frac{1}{n!}\right| = \sup_{A\subseteq S_n}\left(\mathbb{P}(\pi_t \in A) - \frac{|A|}{n!}\right). \tag{5}$$

Thus d_t measures the maximum discrepancy in the probability of some event under the distribution of π_t and under the uniform distribution. In a beautiful paper, Diaconis and Shashahani [14] proved that d_t undergoes an abrupt transition from being close to 1 when $t \le \frac{1}{2}n\log n - Cn$ to being close to 0 when $t \ge \frac{1}{2}n\log n + Cn$. This was the first example of the so-called *cutoff phenomenon*, later observed in many Markov chains, in which the distance to stationarity remains roughly at 1 for a long time and then drops to near 0 in a much shorter time scale.

Theorem 2.3 (Diaconis-Shashahani [14]). *1. For each $\varepsilon > 0$ there exists $C(\varepsilon) > 0$ such that if $t \le \frac{1}{2}n\log n - C(\varepsilon)n$ and $n \ge C(\varepsilon)$ then*

$$d_t \ge 1 - \varepsilon. \tag{6}$$

2. There exists a constant $b > 0$ such that for all $C > 0$ and $n \ge 10$, if $t \ge \frac{1}{2}n\log n + Cn$ then

$$d_t \le b\exp(-2C). \tag{7}$$

The lower bound (6) is the easy part of the theorem and will be shown below. It was proved in [14], using a similar approach, in a more precise form when t is close to $\frac{1}{2}n\log n$. The upper bound (7) is the heart of the theorem and was proved in [14] using the representation theory of the symmetric group, interpreting the random transposition shuffle as a random walk on S_n.

Proof of part (1) *of Theorem 2.3.* Fix $\varepsilon > 0$. In order to lower bound d_t we construct a suitable event A to use in (5). The event A is the event that the permutation has many fixed points,

$$A := \{\sigma \in S_n \colon \text{the number of fixed points of } \sigma \text{ is at least } m\},$$

where

$$m := \left\lceil\frac{2}{\varepsilon}\right\rceil$$

and we assume throughout that $n > m$. The expected number of fixed points in a uniform permutation is 1, as follows from (1). Thus, by Markov's inequality,

$$\frac{|A|}{n!} \le \frac{1}{m} \le \frac{\varepsilon}{2}. \tag{8}$$

It remains to lower bound $\mathbb{P}(\pi_t \in A)$. Let N_t be the number of cards which were not selected either as i or j in the first t transpositions of the random transposition shuffle. We clearly have that the number of fixed points in π_t is at least N_t, so that

$$\mathbb{P}(\pi_t \in A) \ge \mathbb{P}(N_t \ge m). \tag{9}$$

Analyzing N_t leads to the *coupon collector problem.* Think about each step in the random transposition shuffle as consisting of 2 half steps, one for choosing the first card i and the other for choosing the second card j (after which the transposition takes place). Let T_k be the number of such half steps until exactly k cards have been chosen as either i or j in some half step. Then

$$\mathbb{P}(N_t \ge m) = \mathbb{P}(T_{n-m+1} > 2t). \tag{10}$$

In addition, $T_1 := 1$ and $T_k - T_{k-1}$, $k \ge 2$, are independent, with $T_k - T_{k-1} \sim \text{Geom}\left(\frac{n-k+1}{n}\right)$. In particular,

$$\mathbb{E}\left(T_{n-m+1}\right) = \sum_{k=1}^{n-m+1} \frac{n}{n-k+1} = n\sum_{k=m}^{n} \frac{1}{k} \ge n\log n - n\log m,$$
$$\text{Var}\left(T_{n-m+1}\right) = n\sum_{k=1}^{n-m+1} \frac{k-1}{(n-k+1)^2} = n\sum_{k=m}^{n} \frac{n-k}{k^2} \le n^2\sum_{k=1}^{\infty} \frac{1}{k^2} \le 2n^2.$$

Lastly, Chebyshev's inequality implies that

$$\mathbb{P}\left(T_{n-m+1} > n\log n - n\left(\log m + \frac{2}{\sqrt{\varepsilon}}\right)\right) \ge 1 - \mathbb{P}\left(|T_{n-m+1} - \mathbb{E}(T_{n-m+1})| \ge \frac{2n}{\sqrt{\varepsilon}}\right) \ge 1 - \frac{\varepsilon}{2}.$$

Putting the last inequality together with (8), (9) and (10) finishes the proof. □

The last proof shows that a main obstacle to the fast mixing of the random transposition shuffle is the number of fixed points, i.e., cycles of length 1, in the permutation. A main goal of the research that we shall discuss is to show that, in a sense, the structure of the long cycles of the permutation mixes much faster, for this shuffle and in many other situations.

3 Spatial random permutations and the interchange process

As explained in the introduction, a spatial random permutation is a random permutation which is biased towards the identity in some underlying geometry. One example, which will be a prime example in our course, is the interchange process which we proceed to define.

3.1 The interchange process

Given a (finite or countably infinite) graph $G = (V, E)$, we denote by S_V the set of permutations on the vertices of G, where we mean that an element $\pi \in S_V$ is a one-to-one and onto function $\pi : V \to V$. The *interchange process* (also called the *stirring process* in some of the literature) gives a dynamics on permutations in S_V which is associated to the structure of the graph. Precisely, the interchange process is a permutation-valued stochastic process (π_t), $t \geq 0$, with each $\pi_t \in S_V$. The initial state π_0 is the identity permutation. Each edge of the graph is endowed with an independent Poisson process of rate 1. We say that an edge $e \in E$ *rings* at time t if an event occurs at time t for the Poisson process associated with e. When an edge $e = \{u, v\}$ rings, the current permutation is updated by multiplying it on the left by the transposition (u, v). Graphically, the dynamics may be visualized by starting with particles at each vertex of V, with the particle at v being labeled by v, and when an edge $e = \{u, v\}$ rings swapping the particles at u and v. This process is well defined when G is a finite graph but may be ill defined for an infinite graph. We do not discuss here conditions under which the process is well defined but mention that it suffices that the graph has bounded degree (this may be deduced from knowing that $p_c(G) > 0$, see the discussion around (11)), which will be the case in all our examples. One consequence of the definition which is worth noting already is that $(\pi_t(v))$, $t \geq 0$, is a simple random walk on G in continuous time for each fixed $v \in V$. These random walks are, however, generally dependent for different v's.

Our main object of study for the interchange process will be the cycle structure of π_t at a given time t and specifically whether *macroscopic cycles* arise. Here, a macroscopic cycle means one of two things: When G is finite it means a cycle whose length is a fixed proportion of all vertices in the graph (this notion will be used for a sequence of finite graphs G_n having $|V(G_n)| \to \infty$) and when G is infinite it means an infinite orbit, i.e., an infinite sequence $(v_i) \subseteq V$ of *distinct* vertices with $\pi_t(v_i) = v_{i+1}$.

Exercise 3.1. *Let G be an infinite bounded-degree graph. Use Kolmogorov's zero-one law to prove that for each fixed t,*

$$\mathbb{P}(\pi_t \textit{ has an infinite orbit}) \in \{0, 1\}.$$

In studying the interchange process it is useful to introduce an associated percolation process. For a given time t, declare an edge e to be open if e rings at least once by time t. Otherwise declare that e is closed. Thus, each edge is open independently with probability $p = p(t) = 1 - \exp(-t)$. We focus on the connected components of open edges in this percolation process. A moment's thought reveals that the cycle in π_t containing a given vertex v must be contained in the connected component of v in the percolation. This fact is quite useful in showing that no macroscopic cycles occur for small time t. For instance, if G is infinite with bounded degree we may define

$$\begin{aligned} p_c(G) := \sup\{t\colon & \text{ all connected components in an edge} \\ & \text{percolation on } G \text{ with parameter } p \text{ are finite almost surely}\} \end{aligned} \tag{11}$$

and obtain that

$$\mathbb{P}(\pi_t \text{ has an infinite orbit}) = 0, \quad \text{for all } t < \log\left(\frac{1}{1 - p_c(G)}\right). \tag{12}$$

We note also that in such case $p_c(G) \geq \frac{1}{\Delta-1}$ where Δ is the maximal degree in G (this is a standard exercise which follows by estimating the expected number of open simple paths of a given length which emanate from a vertex). This implies, for instance, that there are never any infinite cycles when $G = \mathbb{Z}$, the one-dimensional lattice graph. While this approach is generally useful for small t, understanding the cycle structure of π_t for large t is a major challenge. The following conjecture of Bálint Tóth (see [30] for context) is especially tantalizing.

Conjecture 3.2 (Tóth's conjecture)**.** *Let G be the hyper-cubic lattice $\mathbb{Z}^d$.*

1. *If $d = 2$ then*
$$\mathbb{P}(\pi_t \text{ has an infinite orbit}) = 0, \quad \text{for all } t \geq 0.$$
2. *If $d \geq 3$ then there exists a critical time t_c such that*
$$\mathbb{P}(\pi_t \text{ has an infinite orbit}) = \begin{cases} 0, & t < t_c \\ 1, & t > t_c \end{cases}.$$

This conjecture is at present wide open and even the existence of a single pair of d and t for which the interchange process on $\mathbb{Z}^d$ has a macroscopic cycle at time t is unknown. Tóth made this conjecture in the context of studying quantum statistical mechanical models [30]. He discovered, in particular, that the question of existence of macroscopic cycles for a variant of the interchange process (involving an Ewens-type bias factor $2^{C(\pi)}$) is equivalent to the existence of spontaneous magnetization for the quantum Heisenberg ferromagnet.

Rigorous mathematical results for the interchange process are so far limited mostly to the cases that G is the complete graph, a regular infinite tree, or the one-dimensional lattice $\mathbb{Z}$ (with recent progress on the hypercube graph by Kotecký, Miłoś and Ueltschi). In the next two sections we describe some of the known results for the complete graph and for regular trees.

3.1.1 The interchange process on the complete graph

We start by discussing percolation on the complete graph. The standard Erdős-Rényi random graph model is an edge percolation with parameter p on the complete graph with n vertices. It is usually denoted by $G(n,p)$. Define a function $\theta : (1,\infty) \to (0,1)$, the survival probability of a Galton-Watson tree with Poisson(c) offspring distribution (see also Section 3.1.2 below), implicitly as the unique positive solution of

$$1 - \theta(c) = \exp(-c\theta(c)), \quad c > 1.$$

The following facts are well known and we refer the reader to Alon and Spencer [2] for a proof (see also Krivelevich and Sudakov [22] for a recent short proof of the existence of the phase transition, a fact which already suffices to obtain a version of Theorem 3.5 below).

Theorem 3.3. *Let $p = \frac{c}{n}$ for a fixed $c > 0$. Then with probability tending to 1 as n tends to infinity, the Erdős-Rényi model $G(n,p)$ satisfies that*

1. *If $c < 1$ then all connected components have size at most $O(\log n)$.*

2. *If $c = 1$ then the largest connected component has size of order $n^{2/3}$ (with a non-trivial limiting distribution for the normalized size).*

3. *If $c > 1$ then there exists a connected component (called the 'giant component') of size $(\theta(c) + o(1))n$ while all other connected components have size at most $O(\log n)$.*

(the functions implicit in the O and o notation are deterministic but depend on c).

Let us now consider the interchange process (π_t), $t \geq 0$, on the complete graph with n vertices. Note that π_t is naturally coupled with the random transposition shuffle which we previously discussed, in which the number of transpositions is a Poisson random variable with mean $\binom{n}{2}t$. It follows that π_t becomes close to uniform in total variation distance when $t > \frac{\log n}{n} + \Omega(n)$. As we have seen, one parameter slowing down the convergence is the number of fixed points in π_t. In contrast, we now present the result of Schramm [26], proving a conjecture of Aldous stated in [6], which shows that the structure of the long cycles in π_t converges much faster, already for t of order $\frac{1}{n}$.

Recall the associated percolation process with parameter $p = 1 - \exp(-t)$ of all edges which ring at least once by time t. Let $x_1(t)$ denote the (random) size of the largest connected component in this percolation process. As the largest cycle of the interchange process has length at most $x_1(t)$, the previous theorem implies that when $t = \frac{c}{n}$ for a fixed $c < 1$ then all cycles are of length $O(\log n)$ and when $t = \frac{1}{n}$ then all cycles have length of order at most $n^{2/3}$. In both cases the cycles are not macroscopic (i.e., not constituting a fraction of all vertices). Now denote by $\ell^{(1)}(t) \geq \ell^{(2)}(t) \geq \cdots \geq \ell^{(C(\pi_t))}(t)$ the sorted lengths of cycles in π_t.

Theorem 3.4 (Schramm [26])**.** *Let $t = \frac{c}{n}$ for a fixed $c > 1$. Then $\left(\frac{\ell^{(1)}(t)}{x_1(t)}, \frac{\ell^{(2)}(t)}{x_1(t)}, \ldots\right)$ converges to the Poisson-Dirichlet distribution with parameter 1 as n tends to infinity.*

The convergence in the theorem means that for any fixed k, the first k cycle lengths, in sorted order and divided by $x_1(t)$, converge in distribution to the first k coordinates in the PD(1) distribution (in particular, $\mathbb{P}(C(\pi_t) \geq k) \to 1$ as $n \to \infty$). As $x_1(t) \sim \theta(c)n$ by Theorem 3.3 we see that macroscopic cycles exist and their structure is as in a uniform permutation. The theorem also implies that there are only $x_1(t)$ vertices in macroscopic cycles. The remaining $n - x_1(t)$ vertices must belong to cycles whose lengths are at most $O(\log n)$ by Theorem 3.3. The proof of Schramm for this theorem proceeds in two steps: First it is proved that macroscopic cycles exist. Then it is proved by a coupling argument that their structure quickly becomes close to that of the PD(1) distribution. We will not present the full proof and content ourselves with presenting a weaker statement, showing only the emergence of a macroscopic cycle, with a clever short argument of Berestycki found after the work of Schramm.

Theorem 3.5. *(Berestycki [5]) Let $t = \frac{c}{n}$ for a fixed $c > 1$. Then with probability tending to one as n tends to infinity, there exists some $0 \leq s \leq t$ such that $\ell^{(1)}(s) \geq \frac{\theta(c)^2}{8}n$.*

Proof. Fix a $c > 1$ and let $t := \frac{c}{n}$. For each $s \geq 0$ let G_s be the associated percolation process at time s, i.e., the subgraph of edges which rang at least once by time s. Recall that $C(\pi_s)$ denotes the number of cycles in π_s and let $C(G_s)$ denote the number of connected components in G_s. Observe that both $C(\pi_s)$ and $C(G_s)$ may only change at the times that edges ring, with $C(G_s)$ either remaining the same or decreasing by 1 and $C(\pi_s)$ either increasing or decreasing by 1 at each such ring. To prove the theorem we will establish (formal versions of) the following statements: (a) $C(G_s)$ does not decrease too rapidly for s near t, (b) $C(\pi_s)$ is rather close to $C(G_s)$ for all s, (c) In order for $C(\pi_s)$ not to decrease too rapidly there must exist a macroscopic cycle.

As the first step we prove that

$$\begin{aligned}&\text{if } t - s = n^{-\alpha} \text{ for some } 1 < \alpha < 2 \text{ then for any fixed } \varepsilon > 0,\\ &\mathbb{P}\left(C(G_s) - C(G_t) \leq \binom{n}{2}(t-s)(1-\theta(c)^2+\varepsilon)\right) \to 1 \text{ as } n \to \infty.\end{aligned} \tag{13}$$

Denote by $x_1(u) \geq x_2(u) \geq \cdots$ the sorted sizes of the connected components in G_u. Let

$$q(u) := 1 - \frac{\sum_i \binom{x_i(u)}{2}}{\binom{n}{2}}$$

and observe that $q(u)$ is a (random) non-increasing function of u. When an edge is added to G_u, it decreases the number of connected components if and only if it connects two distinct components. As the first edge to ring after time u is uniformly chosen among the $\binom{n}{2}$ edges, it follows that the conditional probability given G_u that it will decrease the number of connected components equals $q(u)$. As $q(u)$ is non-increasing,

$$\begin{aligned}&\text{on the event } \{q(s) \leq p\},\\ &C(G_s) - C(G_t) \text{ is stochastically dominated by a Poisson } \left(\binom{n}{2}(t-s)p\right) \text{ variable.}\end{aligned} \tag{14}$$

Fix $\varepsilon > 0$. We note that

$$\mathbb{P}(q(s) \leq 1 - \theta(c)^2 + \varepsilon) \to 1, \quad \text{as } n \to \infty, \tag{15}$$

by Theorem 3.3, using the fact that $t - s = o(\frac{1}{n})$. In addition, for any fixed $0 < p < 1$, if N has the Poisson$(\binom{n}{2}(t-s)p)$ distribution then

$$\mathbb{E}(N) = \mathrm{Var}(N) = \binom{n}{2}(t-s)p$$

which implies, by Chebyshev's inequality, that $\mathbb{P}(N > \binom{n}{2}(t-s)(p+\varepsilon)) \to 0$ as $n \to \infty$, using the fact that $t - s = \omega(\frac{1}{n^2})$. As ε is arbitrary, this fact together with (14) and (15) implies (13).

As a second step, we prove that

$$\text{for any } s \geq 0,\ \mathbb{E}|C(\pi_s) - C(G_s)| \leq 2n^{3/2}s + \sqrt{n}. \tag{16}$$

Fix an $s \geq 0$. Recall again that the cycle in π_s of a vertex v is contained in the connected component of v in G_s. Thus we trivially have $C(\pi_s) \geq C(G_s)$ and need only prove that $C(\pi_s) - C(G_s)$ is not too large with high probability. Call a connected component of G_s *cyclic* if all its vertices belong to a single cycle in π_s and otherwise call it *fragmented*. Call a cycle in π_s *short* if its length is at most $\sqrt{n}$ and if it is contained in a fragmented connected component of G_s. Let A_s denote the number of short cycles in π_s. Observe that

$$C(\pi_s) - C(G_s) \leq A_s + \sqrt{n} \tag{17}$$

as there are at most $\sqrt{n}$ cycles which are longer than $\sqrt{n}$ in π_s. Observe next that each edge $e = (u, v)$ which rings causes a coagulation or fragmentation event in the current permutation. Precisely, if σ is the current permutation and u and v are in different cycle of σ, then the ring causes these two cycles to merge. If u and v are in the same cycle of σ, whose length is ℓ, and $v = \sigma^j u$, then the ring causes this cycle to fragment into the two cycles $(u, \sigma u, \ldots, \sigma^{j-1}u)$ and $(v = \sigma^j u, \sigma^{j+1}u, \ldots, \sigma^{\ell-1}u)$. Now suppose that $\mathcal{C}$ is a short cycle in the fragmented connected component V of G_s. Let E be the set of edges with an endpoint in $\mathcal{C}$ which had a fragmentation event by time s. As V is fragmented, it follows that E is non-empty. As $\mathcal{C}$ is short, we conclude that the last fragmentation event of an edge in E by time s must have created a cycle of length at most $\sqrt{n}$ (and possibly two such cycles). We have proven that

$$A_s \leq 2B_s \tag{18}$$

where B_s is the number of fragmentation events which occurred by time s in which one or both of the resulting cycles had length at most $\sqrt{n}$. Lastly, observe that for any permutation σ, there are at most $2n^{3/2}$ edges whose ring would cause such a fragmentation event (at most n choices for the first endpoint of the edge and at most $2\sqrt{n}$ choices for the second endpoint). As there are on average $\binom{n}{2}s$ rings by time s and as, given their number, the ringing edges are chosen uniformly among all edges, we conclude that

$$\mathbb{E}B_s \leq \frac{2n^{3/2}}{\binom{n}{2}} \cdot \binom{n}{2}s = 2n^{3/2}s.$$

We thus obtain (16) by combining the fact that $C(\pi_s) \geq C(G_s)$, with (17), (18) and the last inequality.

We proceed to use (13) and (16) to finish the proof of the theorem. Let $s := t - n^{-5/4}$. We have

$$C(\pi_s) - C(\pi_t) \leq C(G_s) - C(G_t) + |C(G_s) - C(\pi_s)| + |C(G_t) - C(\pi_t)|.$$

As

$$\mathbb{P}\left(C(G_s) - C(G_t) \leq \binom{n}{2}n^{-5/4}\Big(1 - \frac{1}{2}\theta(c)^2\Big)\right) \to 1 \text{ as } n \to \infty$$

by (13) and

$$\mathbb{E}|C(G_u) - C(\pi_u)| \leq (2c+1)\sqrt{n}, \quad u \in \{s, t\}$$

by (16), we conclude using Markov's inequality that

$$\mathbb{P}\left(C(\pi_s) - C(\pi_t) \leq \binom{n}{2}n^{-5/4}\Big(1 - \frac{1}{3}\theta(c)^2\Big)\right) \to 1 \text{ as } n \to \infty.$$

However, letting $N_{s,t}$ be the number of rings between times s and t and letting $F_{s,t}$ be the number of these which cause cycles to fragment we have

$$C(\pi_s) - C(\pi_t) = N_{s,t} - 2F_{s,t}.$$

As $N_{s,t}$ has the Poisson($\binom{n}{2}n^{-5/4}$) distribution we conclude (with a similar application of Chebyshev's inequality as before) that

$$\mathbb{P}\left(F_{s,t} \geq \frac{1}{7}\binom{n}{2}n^{-5/4}\theta(c)^2\right) \to 1 \text{ as } n \to \infty. \tag{19}$$

Lastly, let $k := \frac{1}{7}\binom{n}{2}n^{-5/4}\theta(c)^2$ and $\delta := \frac{\theta(c)^2}{8}$. If the longest cycle of a permutation is shorter than δn then the probability that applying a uniformly chosen transposition fragments a cycle is at most $\frac{n}{\delta n}\binom{\delta n}{2}/\binom{n}{2} \leq \delta$. Letting $F'_{s,t}$ denote a Poisson($\binom{n}{2}n^{-5/4}\delta$) random variable, we conclude that

$$\mathbb{P}(F_{s,t} \geq k) \leq \mathbb{P}(F'_{s,t} \geq k) + \mathbb{P}(\exists\, s \leq u \leq t, \ell^{(1)}(u) \geq \delta n).$$

Comparing the last expression with (19) yields that

$$\mathbb{P}(F'_{s,t} \geq k) + \mathbb{P}(\exists\, s \leq u \leq t, \ell^{(1)}(u) \geq \delta n) \to 1 \text{ as } n \to \infty \tag{20}$$

and the theorem follows since, again using Chebyshev's inequality, $\mathbb{P}(F'_{s,t} \geq k) \to 0$ for our choice of k and δ. $\square$

3.1.2 The interchange process on trees

In this section we describe results of Angel and Hammond [3, 19, 20] on the interchange process on tree graphs. Again, we start by familiarizing ourselves with the associated percolation process. Our trees will be rooted and we denote their root by o. The *parent* of a vertex $v \neq$ o is the unique vertex w adjacent to v and closer to o than v (in the graph metric). The root vertex has no parent. The children of a vertex v are all vertices adjacent to v except the parent of v. The (infinite) *d-ary tree* $\mathcal{T}_d$, $d \geq 2$, is the tree satisfying that each vertex has exactly d children, i.e., the tree having degree $d+1$ at all vertices except the root, which has degree d. As mentioned above, on any graph G of maximal degree Δ one has $p_c(G) \geq \frac{1}{\Delta-1}$. For the d-ary tree this turns out to be sharp, that is $p_c(\mathcal{T}_d) = \frac{1}{d}$. This is a special case (with binomial offspring distribution) of the standard theorem for Galton-Watson trees which we now discuss. We call μ an *offspring distribution* if μ is a probability measure on the non-negative integers. A Galton-Watson tree with offspring distribution μ is a random tree formed by the following process: The root has a random number of children distributed μ. Then each child of the root has, independently, a random number of children distributed μ, and so on and so forth, with each new child having, independently, a random number of children distributed μ. The basic question regarding Galton-Watson trees is whether they are infinite with positive probability. The following classical theorem gives the answer.

Theorem 3.6. *Let $\mathcal{T}$ be a Galton-Watson tree with offspring distribution μ. Denote by m the expectation of μ and let p be the probability that $\mathcal{T}$ is infinite.*

1. *If $m \le 1$ then $p = 0$, unless $m = 1$ and $\mu(1) = 1$ (in which case p is clearly 1).*

2. *If $m > 1$ then $p > 0$.*

One standard approach to proving this theorem proceeds via analysis of generating functions. The next exercise, which also expands on further aspects of Galton-Watson trees, uses an approach based on martingale theory.

Exercise 3.7. *Let $\mathcal{T}$ be a Galton-Watson tree with offspring distribution μ. Let X be a random variable distributed as μ and set $m := \mathbb{E}(X)$. We assume $m < \infty$ (it is not difficult to reduce the general case to this one) and also that $\mathbb{P}(X = 1) < 1$. For $n \ge 0$, let Z_n be the number of vertices in $\mathcal{T}$ at distance exactly n from the root. Let E be the event that $\mathcal{T}$ is finite (the event of extinction), that is,*

$$E := \{\text{there exists some } n \ge 1 \text{ for which } Z_n = 0\}.$$

1. *Prove that the process (M_n), $n \ge 0$, defined by $M_n := Z_n/m^n$ is a martingale. Deduce that $\mathbb{P}(E) = 1$ if $m \le 1$.*

2. *Define $f(s) := \mathbb{E}(s^X)$ for $0 \le s \le 1$ (where we use the convention that $f(0) = \mathbb{P}(X = 0)$ so that f is real analytic on $[0, 1]$). Suppose there exists some $0 \le \rho < 1$ satisfying $f(\rho) = \rho$. Prove that the process (G_n), $n \ge 0$, defined by $G_n := \rho^{Z_n}$ is a martingale. Deduce that $\mathbb{P}(E) = \rho$ and $\mathbb{P}(Z_n \to \infty) = 1 - \rho$. Infer also that the equation $f(\rho) = \rho$ has at most one solution in $[0, 1)$.*

3. *Observe that $f'(s) = \mathbb{E}(Xs^X)$ and in particular $f'(1) = m$. Deduce that $P(E) < 1$ if $m > 1$.*

4. *Suppose that $m > 1$. Write $M_\infty = \lim_{n\to\infty} M_n$ (why does it exist?). Since (M_n) is a martingale one may speculate that Z_n grows as m^n on the event of non-extinction, i.e., that $M_\infty > 0$ on E^c. The Kesten-Stigum theorem shows that the sharp condition for this to occur is $\mathbb{E}X \log(X + 1) < \infty$. We will instead prove it here under the stronger condition that*
$$\mathbb{E}(X^2) < \infty. \tag{21}$$
Prove that $\mathbb{E}M_\infty = 1$ under the assumption (21).

   Hint: Bound $\mathbb{E}(M_n^2)$.

5. *Still under the assumptions $m > 1$ and (21), observe that $\theta := \mathbb{P}(M_\infty = 0)$ satisfies $f(\theta) = \theta$ and deduce that $\mathbb{P}(M_\infty = 0) = \mathbb{P}(E)$.*

   Hint: Condition on Z_1.

Returning to the interchange process (π_t), $t \ge 0$, on a tree $\mathcal{T}$ we may apply the relation (12) with percolation to deduce the following. If each vertex in $\mathcal{T}$ has at least d children then

$$\mathbb{P}(\pi_t \text{ has an infinite orbit}) = 0, \quad \text{for all } t < \log\left(\frac{1}{1 - \frac{1}{d}}\right) = \frac{1}{d} + \frac{1}{2d^2} + O(d^{-3}) \text{ as } d \to \infty.$$

Angel was the first to prove that the interchange process on a d-ary tree has infinite cycles for certain t, establishing in particular that the above bound is rather tight for large d.

Theorem 3.8 (Angel [3]). *Let (π_t) be the interchange process on the d-ary tree for $d \geq 4$. Then there exists a non-empty interval of times $I_d \subset [0, \infty)$ such that*

$$\mathbb{P}(\pi_t \text{ has an infinite orbit}) = 1, \quad \text{for all } t \in I_d.$$

Moreover, for any $\varepsilon > 0$ there exists $d_0(\varepsilon)$ such that if $d \geq d_0(\varepsilon)$ then we may take $I_d = \left[\frac{1}{d} + (\frac{7}{6} + \varepsilon)\frac{1}{d^2}, \log(3) - \varepsilon\right]$.

We note that Angel's theorem leaves open the possibility that occurrence of infinite orbits is non-monotone in t. That is, that for some $s > t$, π_t has infinite cycles with probability one while π_s does not have them with probability one. Indeed, monotonicity, while natural, is not known in general for the interchange process. Angel's theorem was expanded upon by Hammond who established, among other results, the monotonicity for trees of sufficiently high degree.

Theorem 3.9 (Hammond [19, 20]). *Let (π_t) be the interchange process on a tree $\mathcal{T}$.*

1. *If each vertex of $\mathcal{T}$ has at least two children then there exists a t_0 such that for each $t \geq t_0$, π_t has infinite orbits almost surely. Moreover, for each $d \geq 55$, if each vertex of $\mathcal{T}$ has at least d children then we may take $t_0 = \frac{101}{d}$.*

2. *Suppose $\mathcal{T} = \mathcal{T}_d$ for some $d \geq 764$. Then there exists a $t_c = t_c(d)$ such that*

$$\mathbb{P}(\pi_t \text{ has an infinite orbit}) = \begin{cases} 0 & t < t_c \\ 1 & t > t_c \end{cases}.$$

   In addition, $t_c \in \left[\frac{1}{d} + \frac{1}{2d^2}, \frac{1}{d} + \frac{2}{d^2}\right]$.

We content ourselves with explaining the main ideas in the proof of Angel's theorem, Theorem 3.8, and do not enter into the more involved details of Hammond's results.

A useful tool in thinking of the interchange process is given by *cyclic-time random walk* (CTRW). Recall that the interchange process permutation π_T is constructed via independent Poisson processes of rate 1 associated to each edge of the graph which are run up to time T. Consider extending these processes to run for all positive time in a cyclic manner, by repeating the events at $[0, T)$ to $[T, 2T)$, $[2T, 3T)$, etc.. Now, given these extended processes, define the CTRW (X_t) as a walk on the graph in which X_0 is a given vertex v of the graph and which is defined for all positive time via the rule that if an edge incident to the current position of the walker rings, then the walker switches position to the other endpoint of that edge. Because the processes on each edge have been extended in a periodic way, it is simple to see that if $X_{kT} = v$ for some integer $k > 0$, then the walk will repeat itself periodically for all $t > kT$. In fact, it is a simple exercise to verify that the range of the walk (that is, the set of vertices visited by the walk) exactly equals the orbit of v in the interchange process permutation π_T. In particular, the walk is *transient* in the sense that it visits infinitely many vertices if and only if the orbit of v in π_T is infinite.

The argument of Angel is based on finding local conditions on the ringing times at each vertex such that the set of vertices which satisfy these conditions is a Galton-Watson sub-tree of $\mathbb{T}_d$ and if this tree is infinite then the interchange process has infinite orbits. Fix $T > 0$.

We consider the Poisson processes on the edges of the tree as being extended periodically from $[0,T)$ to $[0,\infty)$ as described above (and still refer to events of the extended processes as rings). We denote the associated CTRW started from o by (X_t), $t\ge 0$. We proceed to discuss the local conditions in Angel's argument. Say that a vertex v in $\mathbb{T}_d$, other than the root o, is *good* if the edge connecting v with its parent vertex rings exactly once in $[0,T)$. Write t_v for this unique ringing time. The root o is always said to be good and we set $t_{\mathrm{o}} := 0$. Suppose that $v\neq \mathrm{o}$ is a good vertex with a good parent u. We say that a sibling v' of v (that is, a child of u other than v) *covers* v if the ringing times of the edge (u,v') separate t_u and t_v cyclically modulo T. Precisely, if $t_u<t_v$ this means that there are ringing times both in (t_u,t_v) and in (t_v,t_u+T) and if $t_u>t_v$ this means that there are ringing times both in (t_v,t_u) and in (t_u,t_v+T). We say that v is *uncovered* if it is not covered by any of its siblings.

Now suppose that $v\neq \mathrm{o}$ is a good vertex with a good parent u and that v is uncovered. We leave it as an exercise to check that if $X_t=u$ for some time t then necessarily either X_t is transient or $X_s=v$ at some later time $s>t$ (or both). We explain briefly the rational behind this claim. The first time t that $X_t=u$ must satisfy $t\equiv t_u$ modulo T since u is good. Following this time and before going to v, the CTRW may proceed to some sibling v' of v. If it does so and ever returns, then the time s of its return must satisfy that it lies after t_u and before t_v when thinking of the times modulo T, by the definition of v being uncovered. From this one may deduce that if the CTRW returns from all its visits to siblings of v then it must arrive at v.

It follows from the previous exercise that if there is an infinite simple path of good and uncovered vertices in $\mathbb{T}_d$ starting from some child of the root then the CTRW (X_t) is transient (which, as explained above, is equivalent to having an infinite orbit in π_T). We now proceed to show that the probability of this event is positive. Let N_u be the number of good and uncovered children of a good vertex u. Observe that the distribution of N_u does not depend on t_u and hence is the same for all good vertices u. Moreover, one sees simply that N_u is independent of $(N_v)_v$ where v goes over all other good vertices. It follows that the connected component of the root of good and uncovered vertices, together with the root itself, forms a Galton-Watson tree. Thus we need only show that $\mathbb{E}N_u>1$ for a good vertex u. Note that $\mathbb{E}N_u$ equals d times the probability that a specific child v of u is good and uncovered. The probability that v is good (that is, that the edge (u,v) rings exactly once in $[0,T)$) equals $T\exp(-T)$. Conditioned on this, assuming without loss of generality that $t_u=0$ and writing $t_v=a$, the probability that v is covered by a sibling v' of it equals $(1-\exp(-a))(1-\exp(-(T-a)))$ and the covering events are independent between the different siblings. Thus we conclude that

$$\begin{aligned}\mathbb{E}N_u &= d\,Te^{-T}\int_0^T \left(1-\left(1-e^{-a}\right)\left(1-e^{-(T-a)}\right)\right)^{d-1}\frac{da}{T} =\\ &= d\,e^{-T}\int_0^T \left(e^{-a}+e^{-(T-a)}-e^{-T}\right)^{d-1}da =\\ &= 2d\,e^{-T}\int_0^{T/2} \left(e^{-a}+e^{-(T-a)}-e^{-T}\right)^{d-1}da.\end{aligned}$$

Angel's theorem follows from a careful analysis of this integral. Here we present a less precise result, noting only that

$$\mathbb{E}N_u \geq 2d\, e^{-T} \int_0^{T/2} e^{-(d-1)a} da = \frac{2d}{d-1}(e^{-T} - e^{-(d+1)T/2}) \geq 2(e^{-T} - e^{-(d+1)T/2})$$

which (as one may verify) is greater than 1 if d is sufficiently large and $T \in \left[\frac{2}{d+1}, \frac{1}{2}\right]$. Thus we have proven that for d sufficiently large and T in this range, the orbit of o in π_T is infinite with positive probability (which means that π_T has infinite orbits almost surely by Exercise 3.1).

3.2 Spatial random permutations in the continuum

In this section we consider a different model of spatial random permutations, in which the permutation is of a finite set in $\mathbb{R}^d$ which is itself random. This model, which may seem less natural at first sight, is well-motivated from physics as it relates to the phenomenon of Bose-Einstein condensation. As it turns out, the model is more amenable to analysis than the models on graphs considered so far and allows a precise determination of the structure of macroscopic cycles. We follow the work of Betz and Ueltschi [7, 8, 9] who were themselves continuing works of Sütő [28, 29] and Buffet and Pulé [13].

Let $\Lambda \subseteq \mathbb{R}^d$ be a cubic box of side length L. The space of configurations for the model is $\Omega_{\Lambda,N} := \Lambda^N \times S_N$, i.e., a choice of N points in Λ and a permutation on them. Given a *potential* function $\xi : \mathbb{R}^d \to \mathbb{R}$, $\xi(x) = \xi(-x)$, we introduce a *Hamiltonian* function on configurations by

$$H(x, \pi) := \sum_{i=1}^{N} \xi(x_i - x_{\pi_i}), \quad x = (x_1, x_2, \ldots, x_N) \in \Lambda^N,\ \pi \in S_N.$$

We focus on the example that $\xi(x) = \|x\|_2^2$ (corresponding to the Feynman-Kac representation of the ideal Bose gas) but other examples are possible as explained in [9] (see also the mathscinet reference of the paper for additional restrictions). Without loss of generality we assume that

$$\int_{\mathbb{R}^d} e^{-\xi(x)} dx = 1 \tag{22}$$

(by adding a constant to ξ if necessary). The main assumption (but not the only one) on ξ is that $e^{-\xi}$ has a positive Fourier transform. This allows to introduce a function $\varepsilon : \mathbb{R}^d \to [0, \infty)$ by

$$e^{-\varepsilon(k)} := \int_{\mathbb{R}^d} e^{-2\pi i k\cdot x} e^{-\xi(x)} dx.$$

When $\xi(x) = \|x\|_2^2 + c$ we have that $\varepsilon(k) = a\|k\|_2^2 + b$ for certain constants a, b. The measure on configurations should be thought of as having a density proportional to $\exp(-H(x, \pi))$ with respect to Lebesgue measure dx on x and the counting measure on S_N. However, for technical reasons it is useful to introduce 'periodized' versions of the potential and Hamiltonian, defined by

$$e^{-\xi_\Lambda(x)} := \sum_{z\in\mathbb{Z}^d} e^{-\xi(x-Lz)}, \tag{23}$$

$$H_\Lambda(x,\pi) := \sum_{i=1}^{N} \xi_\Lambda(x_i - x_{\pi_i}).$$

We finally define the probability measure on configurations as the measure with density

$$\frac{1}{N!Y} e^{-H_\Lambda(x,\pi)} \tag{24}$$

where Y is the normalization constant,

$$Y := \frac{1}{N!}\sum_{\pi\in S_n}\int_{\Lambda^N} e^{-H_\Lambda(x,\pi)}dx.$$

The parameter here which is the analog of time in the interchange model is the particle density ρ defined by

$$\rho := \frac{N}{|\Lambda|}$$

where $|\Lambda| := L^d$ is the volume of Λ. The idea is that the potential suppresses particles from jumping far (that is, $|x_i - x_{\pi_i}|$ is typically small for most i), so that long cycles are only possible at high particle density. In the sequel, whenever we take the limit $\Lambda, N \to \infty$ we mean that N and the side length L of Λ both tend to infinity in such a way that the particle density ρ remains fixed. We introduce the *critical density* by

$$\rho_c := \int_{\mathbb{R}^d} \frac{1}{e^{\varepsilon(k)}-1}dk.$$

The critical density may be finite or infinite according to the choice of potential and the dimension d. In the case that $\xi(x) = \|x\|_2^2 + c$, the critical density is finite for $d \geq 3$. Although we focus only on this case, we mention that there exist potentials for which the critical density is finite in lower dimensions, e.g., $e^{-\xi(x)} = c(|x|+1)^{-\gamma}$, $1 < \gamma < 2$, in dimension $d = 1$. Recall that for a permutation π, $\ell^{(1)}(\pi) \geq \ell^{(2)}(\pi) \geq \cdots \geq \ell^{(C(\pi))}(\pi)$ stand for the sorted list of cycle lengths in π. One observable for which ρ_c is the critical density is the fraction of points in infinite cycles, given by

$$\nu := \lim_{K\to\infty}\liminf_{\Lambda,N\to\infty}\mathbb{E}\left(\frac{1}{N}\sum_{i:\,\ell^{(i)}(\pi)>K}\ell^{(i)}(\pi)\right).$$

We are now ready to describe a special case of the theorem of Betz and Ueltschi.

Theorem 3.10 (Betz and Ueltschi [9])**.** *Suppose $d \geq 3$ and $\xi(x) = \|x\|_2^2 + c$ with c chosen to satisfy* (22)*.*

1. *The fraction of points in infinite cycles satisfies $\nu = \max\left(0, 1-\frac{\rho_c}{\rho}\right)$.*

2. *If $\rho > \rho_c$, so that $\nu > 0$, the cycle structure converges in distribution to the Poisson-Dirichlet distribution, i.e.,*

$$\left(\frac{\ell^{(1)}(\pi)}{\nu N}, \frac{\ell^{(2)}(\pi)}{\nu N}, \dots\right) \xrightarrow{d} \mathrm{PD}(1), \quad \textit{as } \Lambda, N \to \infty.$$

The theorem of Ueltschi and Betz holds for more general potentials satisfying the assumptions described above, along with several additional assumptions. Moreover, more general Hamiltonians are allowed in [9] in which cycle weights are introduced. Precisely, (periodized versions of) Hamiltonians of the type

$$H(x, \pi) = \sum_{i=1}^{N} \xi(x_i - x_{\pi(i)}) + \sum_{\ell \geq 1} \alpha_\ell r_l(\pi),$$

where (α_ℓ) are given parameters and, as before, $r_\ell(\pi)$ denotes the number of cycles of length ℓ in π. The analysis of such Hamiltonians relies on earlier work of Betz, Ueltschi and Velenik [10]. The presence of the cycle weights alters the form of the critical density. In addition, when the α_ℓ converge to a constant α sufficiently fast, the cycle structure converges to that of an Ewens permutation, i.e., to the $\mathrm{PD}(e^{-\alpha})$ distribution. If instead the α_ℓ grow logarithmically with ℓ then the Poisson-Dirichlet distribution is replaced with a single giant cycle. That is, all points belonging to macroscopic cycles belong to a single cycle. We shall not deal with these extensions here but mention that they are analyzed via similar techniques to those that we discuss.

3.3 Ideas of proof of Theorem 3.10

From now on we suppose $d \geq 3$ and $\xi(x) = \|x\|_2^2 + c$ with c chosen to satisfy (22). We introduce the notation

$$\lambda^* := \frac{1}{L}\mathbb{Z}^d$$

for the dual space to Λ with respect to Fourier transform. Recall that for $\pi \in S_N$, $r_j(\pi)$ stands for the number of cycles of π whose length is exactly j, so that $\sum_{j=1}^{N} j r_j(\pi) = N$, and $\mathbf{r}(\pi) = (r_1(\pi), r_2(\pi), \dots, r_N(\pi))$. We start by calculating the distribution of $\mathbf{r}(\pi)$ for the above distribution of random permutations.

Lemma 3.11. *If (x, π) is sampled according to the density* (24) *then*

$$\mathbb{P}(\mathbf{r}(\pi) = (r_1, r_2, \dots, r_N)) = \frac{1}{Y}\prod_{j=1}^{N}\left[\frac{1}{r_j!}\left(\frac{1}{j}\sum_{k \in \Lambda^*} e^{-j\varepsilon(k)}\right)^{r_j}\right]. \tag{25}$$

Proof. We first calculate the marginal probability of (x, π) on π. For each $\sigma \in S_N$ we have

$$\mathbb{P}(\pi = \sigma) = \frac{1}{N!Y}\int_{\Lambda^N} e^{-H_\Lambda(x,\sigma)} dx = \frac{1}{N!Y}\int_{\Lambda^N} e^{-\sum_{i=1}^{N}\xi_\Lambda(x_i - x_{\sigma_i})} dx_1 \cdots dx_N.$$

The last integral factorizes as a product of integrals according to the cycles of σ, with a cycle of length j contributing the factor, with the notation $y_{j+1} := y_1$,

$$
\begin{aligned}
&\int_{\Lambda^j} e^{-\sum_{i=1}^j \xi_\Lambda(y_i - y_{i+1})} dy_1 \cdots dy_j \overset{\text{by (23)}}{=} \\
&= \int_{\Lambda^j} \sum_{z_1,\dots,z_j \in \mathbb{Z}^d} e^{-\sum_{i=1}^j \xi(y_i - y_{i+1} - Lz_i))} dy_1 \cdots dy_j = \\
&= \int_{\Lambda} dy_1 \sum_{w \in \mathbb{Z}^d} \int_{\mathbb{R}^{d(j-1)}} e^{-\xi(y_1 - y_2 + Lw) - \sum_{i=2}^j \xi(y_i - y_{i+1})} dy_2 \cdots dy_j = \\
&= |\Lambda| \sum_{w \in \mathbb{Z}^d} \int_{\mathbb{R}^{d(j-1)}} e^{-\xi(Lw - y_2) - \sum_{i=2}^{j-1} \xi(y_i - y_{i+1}) - \xi(y_j)} dy_2 \cdots dy_j = |\Lambda| \sum_{w \in \mathbb{Z}^d} \left(e^{-\xi}\right)^{*j}(Lw),
\end{aligned}
$$

where we write f^{*j} to denote the convolution of f with itself j times. Thus,

$$
\mathbb{P}(\pi = \sigma) = \frac{1}{N!Y} \prod_{j=1}^N \left(|\Lambda| \sum_{w \in \mathbb{Z}^d} \left(e^{-\xi}\right)^{*j}(Lw) \right)^{r_j(\sigma)}.
$$

By the Poisson summation formula, for any sufficiently good function f,

$$
\sum_{w \in \mathbb{Z}^d} f(Lw) = \frac{1}{L^d} \sum_{k \in \Lambda^*} \hat{f}(k).
$$

Putting the above calculations together and recalling that $|\Lambda| = L^d$, that $e^{-\varepsilon}$ is the Fourier transform of $e^{-\xi}$ and that Fourier transform turns convolution into multiplication, we see that

$$
\mathbb{P}(\pi = \sigma) = \frac{1}{N!Y} \prod_{j=1}^N \left(\sum_{k \in \Lambda^*} e^{-j\varepsilon(k)} \right)^{r_j(\sigma)}.
$$

As the probability of a permutation depends only on its cycle structure, it remains only to plug in the result of Exercise 2.1 which counts the number of permutations with a given cycle structure. □

The usefulness of the previous result stems from the fact that it allows us to introduce a new model on permutations in which the induced distribution on the cycle structure is the same as that of our model. This allows us to analyze the new model and transfer the results to the original one. The new model is in fact a model on a sequence of permutations in a way that we now define. Denote by $\mathbf{n} = (n_k)$, $k \in \Lambda^*$, a sequence of non-negative numbers, called *occupation numbers*, and let $\mathcal{N}_N$ be the set of such sequences which sum to N. A sequence $\boldsymbol{\pi} = (\pi_k)$, $k \in \Lambda^*$, is compatible with $\mathbf{n} \in \mathcal{N}_N$ if π_k is a permutation in S_{n_k} for each $k \in \Lambda^*$. Let $\mathcal{M}_N$ be the set of all pairs $(\mathbf{n}, \boldsymbol{\pi})$, with $\mathbf{n} \in \mathcal{N}_N$ and $\boldsymbol{\pi}$ compatible with $\mathbf{n}$. We introduce a probability measure on $\mathcal{M}_N$ by

$$
\mathbb{P}((\mathbf{n}, \boldsymbol{\pi})) = \frac{1}{Y} \prod_{k \in \Lambda^*} \frac{1}{n_k!} e^{-n_k \varepsilon(k)}. \tag{26}
$$

Here Y is the same constant as in (24) and a consequence of Lemma 3.12 below is that this is indeed a probability distribution. Given an element $(\mathbf{n}, \boldsymbol{\pi}) \in \mathcal{M}_N$ we may think of constructing a permutation by concatenating all the permutations in $\boldsymbol{\pi}$. Such a concatenation would give rise to a cycle structure, which we denote by $\mathbf{r}(\boldsymbol{\pi}) = (r_1(\boldsymbol{\pi}), r_2(\boldsymbol{\pi}), \ldots, r_N(\boldsymbol{\pi}))$, which is given by $r_j(\boldsymbol{\pi}) = \sum_{k \in \Lambda^*} r_j(\pi_k)$.

Lemma 3.12. *If $(\mathbf{n}, \boldsymbol{\pi}) \in \mathcal{M}_N$ is sampled from the distribution (26) then the distribution of $\mathbf{r}(\boldsymbol{\pi})$ is identical to the distribution given in (25).*

The proof of the lemma is a calculation, which is not overly difficult, and which we leave as an exercise to the reader.

We continue by studying the marginal on $\mathbf{n}$ in the distribution on $(\mathbf{n}, \boldsymbol{\pi}) \in \mathcal{M}_N$ given by (26). It is straightforward that the marginal distribution is given by

$$\mathbb{P}(\mathbf{n}) = \frac{1}{Y} \prod_{k \in \Lambda^*} e^{-\varepsilon(k) n_k}, \quad \mathbf{n} \in \mathbb{N}_N.$$

Ueltschi and Betz now proceed to prove the following three properties:

- $\frac{n_0}{N}$ converges in probability to $\max\left(0, 1 - \frac{\rho_c}{\rho}\right)$.
- $\frac{1}{N} \sum_{0 < \|k\| < \delta} n_k$ is small when δ is small with high probability.
- For all $\delta > 0$, $\frac{1}{N} \sum_{\|k\| \geq \delta} n_k 1_{n_k > M}$ is small when M is large with high probability.

These properties imply that in a typical $\mathbf{n}$ there are either no coordinates of order N, when $\rho \leq \rho_c$, or the only coordinate of order N is n_0, when $\rho > \rho_c$, and its value is approximately $\left(1 - \frac{\rho_c}{\rho}\right) N$. Theorem 3.10 follows from these facts, due to the observation that if $(\mathbf{n}, \boldsymbol{\pi})$ are distributed according to (26) then given $\mathbf{n}$, the elements of $\boldsymbol{\pi}$ are uniform permutations. Thus, given an $\mathbf{n}$ which satisfies the above three properties, there are no macroscopic cycles in the cycle structure $\mathbf{r}(\boldsymbol{\pi})$ when $\rho \leq \rho_c$, and the only macroscopic cycles when $\rho > \rho_c$ are obtained from the cycle structure $\mathbf{r}(\pi_0)$. As π_0 is a uniform permutation of size approximately $\left(1 - \frac{\rho_c}{\rho}\right) N$, we get the convergence to the Poisson-Dirichlet distribution which is stated in the theorem.

4 Band permutations and longest increasing subsequences

In this section we discuss various models of spatial random permutations in one dimension.

4.1 Cycle structure

Let us consider the interchange process (π_t), $t \geq 0$, on the integer lattice $G = \mathbb{Z}$. It is clear that π_t has only finite cycles for all t, almost surely, as at any time there will be infinitely many edges which have not rang even once. Still, one may seek to quantify this fact, asking for instance for the expected size of the cycle containing 1 in π_t. Results of this kind have been obtained by Kozma and Sidoravicius who prove, in a work in preparation, that this expected length is of order $\min(t, n)$.

An important feature of the one-dimensional interchange process is its 'band structure'. As for each i, $(\pi_t(i))$ performs a simple random walk on $\mathbb{Z}$, we see that

$$\mathbb{E}|\pi_t(i) - i| \sim C\sqrt{t}, \quad \text{as } t \to \infty, \tag{27}$$

for some $C > 0$. Thus 'most' particles will have a relatively small displacement. This results in the fact that the permutation matrix (i.e., the graph $(i, \pi_t(i))$, $i \in \mathbb{Z}$) is close to being a band matrix with only about order $\sqrt{t}$ diagonals. There are several other models with this property, which we collectively refer to as *band permutations*, and which may well have many properties in common with each other. We focus next on one such model, the *Mallows model*.

Given a parameter $0 < q \leq 1$ and integer $n \geq 1$, the Mallows distribution on S_n with parameter q is the probability measure defined by

$$\mathbb{P}(\pi) = \frac{q^{\mathrm{Inv}(\pi)}}{Z_{n,q}}, \tag{28}$$

where $\mathrm{Inv}(\pi)$ measures the number of inversions of π, that is,

$$\mathrm{Inv}(\pi) = \{(i, j) \colon i < j, \pi(i) > \pi(j)\}.$$

One may also define the Mallows distribution according to (28) with $q > 1$. However, this does not lead to an essentially new distribution as one may easily check that if π is sampled from the distribution (28) with a $q = r$ then the permutation σ defined by $\sigma(i) := n+1-\pi(i)$ has the distribution (28) with $q = 1/r$.

It is well known that $\mathrm{Inv}(\pi)$ also equals the distance of π from the identity in adjacent transposition. I.e., the minimal number of transpositions of the form $(i, i+1)$ which one needs to multiply π by in order to reach the identity. Thus, the Mallows distribution describes a random spatial permutation for which the probability of a permutation π is proportional to $\exp(-\beta d(\pi, \mathrm{Id}))$, where Id stands for the identity permutation, $\exp(-\beta) = q$ and d is the adjacent transposition distance. The following lemma shows that Mallows permutations satisfy the above 'band property'.

Theorem 4.1. *There exists a $c > 0$ such that for all $0 < q \leq 1$ and integer $n \geq 1$, if π is sampled from the distribution* (28) *then*

$$c \min\left(\frac{q}{1-q}, n-1\right) \leq \mathbb{E}|\pi(i) - i| \leq \min\left(\frac{2q}{1-q}, n\right).$$

Results of this kind have been obtained by several authors and one may consult, for instance, [11][Theorem 1.1] for a proof. Thus, recalling that displacements of the interchange process satisfy (27) it makes sense to compare the Mallows distribution with parameter q, for q close to 1, with the interchange process with parameter

$$t = \frac{1}{(1-q)^2}. \tag{29}$$

In work in preparation with Alexey Gladkich, we obtain the following result on the cycle structure of a Mallows permutation. The result is analogous to the result of Kozma and Sidoravicius for the interchange model under the identification (29) and thus lend further support to the idea that band permutations share universal properties.

Theorem 4.2. *There exist $C, c > 0$ such that for all $0 < q \leq 1$ and integer $n \geq 1$, if π is sampled from the distribution* (28) *then for all $1 \leq i \leq n$,*

$$c \min\left(\frac{1}{(1-q)^2}, n\right) \leq \mathbb{E}\ell_i(\pi) \leq C \min\left(\frac{1}{(1-q)^2}, n\right),$$

where $\ell_i(\pi)$ denotes the length of the cycle containing $\imath$ in π.

The proof of this result, as well as that of Theorem 4.1, rely on the fact that the Mallows model is *integrable* in a certain sense. A sample from the Mallows distribution may be formed as a simple function of independent random variables. These facts will be elaborated upon in the talk of Gladkich.

4.2 Longest increasing subsequence

Another observable which has been studied intensively for random permutations with a one-dimensional structure is the *longest increasing subsequence*. For a permutation $\pi \in S_n$, the longest increasing subsequence of π, denoted $\mathrm{LIS}(\pi)$, is defined as

$$\mathrm{LIS}(\pi) := \max\{k \colon \exists 1 \leq i_1 < i_2 < \cdots < i_k,\ \pi(i_1) < \pi(i_2) < \cdots < \pi(i_k)\}.$$

Stanisław Ulam asked in 1961 to determine the asymptotic behavior of $\mathbb{E}(\mathrm{LIS}(\pi))$ when π is a uniform permutation. It is a nice exercise to find explicit constants $C, c > 0$ so that

$$c \leq \frac{\mathbb{E}(\mathrm{LIS}(\pi))}{\sqrt{n}} \leq C \quad \text{for all } n$$

and show additionally, using Fekete's subadditivity lemma, that there exists a constant $c_0 > 0$ so that

$$\frac{\mathbb{E}(\mathrm{LIS}(\pi))}{\sqrt{n}} \to c_0 \quad \text{as } n \to \infty.$$

However, determining the precise value of c_0 is a surprisingly difficult question and it was not until 1977 that Vershik-Kerov [31] and independently Logan-Shepp [23] managed to show that in fact $c_0 = 2$. Both proofs relied on the *RSK algorithm* (a different proof was given later by Aldous and Diaconis [1]). This algorithm, discovered by Robinson and Schensted and later extended by Knuth, is a remarkable mapping connecting two rather different objects. It is a bijection from the permutation group S_n onto pairs of standard Young tableaux of the same shape. The algorithm has the property that the length of the first row of the common Young diagram in the image of a permutation π exactly equals $\mathrm{LIS}(\pi)$. When the RSK algorithm is applied to a uniform distribution on permutations the induced distribution on the common Young diagram in the image is called the *Plancherel measure* and has many connections with the representation theory of the symmetric group. The analysis of $\mathbb{E}(\mathrm{LIS}(\pi))$ proceeds by describing the asymptotic behavior of the Plancherel measure. This and much more on the problem can be found in the book of Dan Romik titled 'The Surprising Mathematics of Longest Increasing Subsequences' [25].

Following the solution of Ulam's problem, a natural question was to estimate also the variance of $\mathrm{LIS}(\pi)$, for a uniform permutation π, and find its limiting distribution. This turned out to be considerably more challenging and was solved only in 1999 by Baik, Deift and Johansson [4].

Theorem 4.3 (Baik-Deift-Johansson)**.** *Let π be a uniform permutation. Then*

$$\frac{\mathrm{LIS}(\pi) - 2\sqrt{n}}{n^{1/6}} \xrightarrow{d} TW,$$

where TW stands for the Tracy-Widom *distribution.*

This theorem was a breakthrough result, proving for the first time that the variance of $\mathrm{LIS}(\pi)$ had the unusual scaling $n^{1/3}$ and linking the topic with the Tracy-Widom distribution which was found earlier in the study of the largest eigenvalue of random matrices. The analysis again uses the RSK algorithm and proceeds with an asymptotic analysis which links the problem with random matrix theory.

We mention that the problem of understanding $\mathrm{LIS}(\pi)$ for a uniform permutation π may be seen as a limiting case of the study of *last passage percolation*. In this problem, one assigns IID positive random variables (X_{ij}) to the vertices of an $n \times n$ grid. To a simple path in this grid one then assigns a weight which is the sum of all the X_{ij} along the path. The problem is then to determine the maximal weight of a path which starts at $(1, 1)$ and ends at (n, n) and takes only right and up steps. One expects this last-passage weight to have a similar behavior as that discovered in the Baik-Deift-Johansson theorem. This, however, has only been proved in very special cases including the cases that the (X_{ij}) have an exponential and geometric distributions. In all other cases our understanding is rather lacking and, in particular, obtaining good bounds on the variance of the last-passage weight is a well-known open problem.

The study of the longest increasing subsequence for random band permutations is relatively new. It was asked in a paper of Borodin, Diaconis and Fulman [12] "Picking a permutation randomly from $P_\theta(\cdot)$, what is the distribution of the cycle structure, longest increasing subsequence, ...?", where $P_\theta(\cdot)$ refers to the Mallows distribution and more general distributions with a similar structure. Starr [27] considered the limiting empirical measure for a random Mallows permutation. Specifically, in the regime

$$q = 1 - \frac{\beta_n}{n}, \quad \beta_n \to \beta \quad \text{as } n \to \infty, \text{ for some } 0 \le \beta < \infty, \tag{30}$$

he found that the empirical measure $\sum_{i=1}^n \delta_{(i/n, \pi(i)/n)}$ tends weakly to a limiting measure on $[0, 1]^2$ which is absolutely continuous and found an explicit expression for its density as a function of β. Mueller and Starr [24] were the first to consider the longest increasing subsequence of a random Mallows permutation. Following the work [27] of Starr and results of Deuschel and Zeitouni [15, 16], they proved that when q is in the regime (30) we have

$$\frac{\mathrm{LIS}(\pi)}{\sqrt{n}} \to f(\beta), \quad \text{in probability,}$$

where f is an explicitly described function. This work was complemented by the work of Bhatnagar and the author [11], who considered the Mallows measure in the regime,

$$n(1-q) \to \infty \quad \text{and} \quad q \to 1,$$

and proved that

$$\frac{\mathrm{LIS}(\pi)}{n\sqrt{1-q}} \to 1, \quad \text{in probability and in } L^p,\ 0 < p < \infty.$$

The work [11] provides additional information on the Mallows distribution, including large deviation results for the length of the longest increasing subsequence and the identification of *five* different regimes in terms of n and q for the length of the longest *decreasing* subsequence.

It seems rather challenging to bring our understanding of the length of the longest increasing subsequence of a Mallows permutation to the level of the Baik-Deift-Johansson theorem, Theorem 4.3, as the RSK algorithm, an important tool in the analysis of [4], does not seem as well suited to the study of the Mallows measure. The work [11] provides a simple bound for the variance of the longest increasing subsequence, proving that when π has the Mallows distribution then

$$\mathrm{Var}(\mathrm{LIS}(\pi)) \le n - 1, \quad \text{for all } n \ge 1 \text{ and } 0 < q \le 1$$

with an accompanying concentration inequality. This bound, however, is not expected to be of the correct order of magnitude unless q is constant as n tends to infinity.

References

[1] D. Aldous and P. Diaconis, Hammersley's interacting particle process and longest increasing subsequences, Probab. Theory Related Fields **103** (1995), no. 2, 199–213.

[2] N. Alon and J. H. Spencer, *The probabilistic method*, third edition, Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley, Hoboken, NJ, 2008.

[3] O. Angel, Random infinite permutations and the cyclic time random walk, in *Discrete random walks (Paris, 2003)*, 9–16 (electronic), Discrete Math. Theor. Comput. Sci. Proc., AC, Assoc. Discrete Math. Theor. Comput. Sci., Nancy.

[4] J. Baik, P. Deift and K. Johansson, On the distribution of the length of the longest increasing subsequence of random permutations, J. Amer. Math. Soc. **12** (1999), no. 4, 1119–1178.

[5] N. Berestycki, Emergence of giant cycles and slowdown transition in random transpositions and k-cycles, Electron. J. Probab. **16** (2011), no. 5, 152–173.

[6] N. Berestycki and R. Durrett, A phase transition in the random transposition random walk, Probab. Theory Related Fields **136** (2006), no. 2, 203–233.

[7] V. Betz and D. Ueltschi, Spatial random permutations and infinite cycles, Comm. Math. Phys. **285** (2009), no. 2, 469–501.

[8] V. Betz and D. Ueltschi, Spatial random permutations with small cycle weights, Probab. Theory Related Fields **149** (2011), no. 1-2, 191–222.

[9] V. Betz and D. Ueltschi, Spatial random permutations and Poisson-Dirichlet law of cycle lengths, Electron. J. Probab. **16** (2011), no. 41, 1173–1192.

[10] V. Betz, D. Ueltschi and Y. Velenik, Random permutations with cycle weights, Ann. Appl. Probab. **21** (2011), no. 1, 312–331.

[11] N. Bhatnagar and R. Peled, Lengths of monotone subsequences in a Mallows permutation, Probab. Theory Related Fields **161** (2015), no. 3-4, 719–780.

[12] A. Borodin, P. Diaconis and J. Fulman, On adding a list of numbers (and other one-dependent determinantal processes), Bull. Amer. Math. Soc. (N.S.) **47** (2010), no. 4, 639–670.

[13] E. Buffet and J. V. Pulé, Fluctuation properties of the imperfect Bose gas, J. Math. Phys. **24** (1983), no. 6, 1608–1616.

[14] P. Diaconis and M. Shahshahani, Generating a random permutation with random transpositions, Z. Wahrsch. Verw. Gebiete **57** (1981), no. 2, 159–179.

[15] J.-D. Deuschel and O. Zeitouni, Limiting curves for i.i.d. records, Ann. Probab. **23** (1995), no. 2, 852–878.

[16] J.-D. Deuschel and O. Zeitouni, On increasing subsequences of I.I.D. samples, Combin. Probab. Comput. **8** (1999), no. 3, 247–263.

[17] W. J. Ewens, The sampling theory of selectively neutral alleles, Theoret. Population Biology **3** (1972), 87–112; erratum, ibid. **3** (1972), 240; erratum, ibid. **3** (1972), 376.

[18] R. P. Feynman, Atomic theory of the λ transition in Helium, Phys. Rev. 91, 1291-1301 (1953)

[19] A. Hammond, Infinite cycles in the random stirring model on trees, Bull. Inst. Math. Acad. Sin. (N.S.) **8** (2013), no. 1, 85–104.

[20] A. Hammond, Sharp phase transition in the random stirring model on trees, Probab. Theory Related Fields **161** (2015), no. 3-4, 429–448.

[21] S. Karlin and J. McGregor, Addendum to a paper of W. Ewens, Theoret. Population Biology **3** (1972), 113–116.

[22] M. Krivelevich and B. Sudakov, The phase transition in random graphs: a simple proof, Random Structures Algorithms **43** (2013), no. 2, 131–138.

[23] B. F. Logan and L. A. Shepp, A variational problem for random Young tableaux, Advances in Math. **26** (1977), no. 2, 206–222.

[24] C. Mueller and S. Starr, The length of the longest increasing subsequence of a random Mallows permutation, J. Theoret. Probab. **26** (2013), no. 2, 514–540.

[25] D. Romik, The surprising mathematics of longest increasing subsequences, Cambridge University Press, 2015. See also `www.math.ucdavis.edu/~romik/book/`.

[26] O. Schramm, Compositions of random transpositions, Israel J. Math. **147** (2005), 221–243.

[27] S. Starr, Thermodynamic limit for the Mallows model on S_n, J. Math. Phys. **50** (2009), no. 9, 095208, 15 pp.

[28] A. Sütő, Percolation transition in the Bose gas, J. Phys. A **26** (1993), no. 18, 4689–4710.

[29] A. Sütő, Percolation transition in the Bose gas. II, J. Phys. A **35** (2002), no. 33, 6995–7002.

[30] B. Tóth, Improved lower bound on the thermodynamic pressure of the spin 1/2 Heisenberg ferromagnet, Lett. Math. Phys. **28** (1993), no. 1, 75–84.

[31] A. M. Veršik and S. V. Kerov, Asymptotic behavior of the Plancherel measure of the symmetric group and the limit form of Young tableaux, Dokl. Akad. Nauk SSSR **233** (1977), no. 6, 1024–1027.