Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2017, Vol. 53, No. 4, 1934–1951
DOI: 10.1214/16-AIHP777

ANNALES DE L'INSTITUT HENRI POINCARÉ
PROBABILITÉS ET STATISTIQUES
www.imstat.org/aihp

Limit theorems for longest monotone subsequences in random Mallows permutations

Riddhipratim Basu[a] and Nayantara Bhatnagar[b]

[a]*Department of Mathematics, Stanford University, Stanford, CA 94305, USA. E-mail:* rbasu@stanford.edu
[b]*Department of Mathematical Sciences, University of Delaware, Newark, DE 19716, USA. E-mail:* nayantara.bhatnagar@gmail.com

Received 23 January 2016; revised 29 June 2016; accepted 30 June 2016

Abstract. We study the lengths of monotone subsequences for permutations drawn from the Mallows measure. The Mallows measure was introduced by Mallows in connection with ranking problems in statistics. Under this measure, the probability of a permutation π is proportional to $q^{\mathrm{inv}(\pi)}$ where q is a positive parameter and $\mathrm{inv}(\pi)$ is the number of inversions in π.

In our main result we show that when $0 < q < 1$, then the limiting distribution of the longest increasing subsequence (LIS) is Gaussian, answering an open question in (*Probab. Theory Related Fields* **161** (2015) 719–780). This is in contrast to the case when $q = 1$ where the limiting distribution of the LIS when scaled appropriately is the GUE Tracy–Widom distribution. We also obtain a law of large numbers for the length of the longest decreasing subsequence (LDS) and identify the limiting constant, answering a further open question in (*Probab. Theory Related Fields* **161** (2015) 719–780).

Résumé. Nous étudions les longueurs des sous-suites monotones de permutations aléatoires tirées sous la mesure de Mallows. La mesure de Mallows a été introduite par Mallows dans le contexte des problèmes de classement en statistique. Sous cette mesure la probabilité d'une permutation π est proportionnelle à $q^{\mathrm{inv}(\pi)}$ où q est un paramètre positif et $\mathrm{inv}(\pi)$ est le nombre d'inversions de π.

Notre résultat principal montre que lorsque $0 < q < 1$, la loi de la plus longue sous-suite croissante est Gaussienne, répondant ainsi à une question posée dans (*Probab. Theory Related Fields* **161** (2015) 719–780). Notons le contraste avec le cas $q = 1$, où la loi limite de la plus longue sous-suite croissante proprement normalisée est la distribution du GUE Tracy–Widom. Nous obtenons aussi une loi des grands nombres pour la longueur de la plus longue sous-suite décroissante et identifions la limite, répondant ainsi à une autre question posée dans (*Probab. Theory Related Fields* **161** (2015) 719–780).

MSC: 60C05; 60F05

Keywords: Mallows permutations; Longest increasing subsequence; Central limit theorem

1. Introduction

Random permutations are well-studied objects in combinatorics and probability. Whereas different statistics of a uniform random permutation have been extensively studied, some non-uniform measures on permutations have, in recent years, generated attention as well. Among the non-uniform models of permutations, the following exponential family of distributions on $\mathcal{S}_n$, the set of permutations on $[n] := \{1, 2, \dots, n\}$, introduced by Mallows [29], has been one of the popular choices. Let q be a positive parameter. A random permutation $\Pi = \Pi_{n,q}$ on $\mathcal{S}_n$ is said to be drawn from the Mallows(q) measure if

$$\mathbb{P}(\Pi = \pi) = \frac{q^{\mathrm{inv}(\pi)}}{Z_{n,q}}$$

for every permutation π of $[n]$ where

$$\mathrm{inv}(\pi) := \#\{(i,j) : 1 \le i < j \le n, \pi(i) > \pi(j)\}$$

is the number of inversions in π and $Z_{n,q}$ is a normalizing constant. Observe that for $q = 1$, this reduces to the uniform measure on permutations whereas for $q > 1$ and $q < 1$, permutations with more and less inversions are favored respectively.

In this paper, we study the lengths of longest monotone subsequence of a Mallows(q) permutation. This is a classically studied question for uniform permutations and the study for general Mallows measures was initiated following a question raised by Borodin et al. in [10] and has received significant attention of late [9,31]; see Section 1.1 for more details. For a permutation π in $\mathcal{S}_n$, we say that $1 \le i_1 < i_2 < \cdots < i_k$ is an *increasing subsequence* of π with *length* k if $\pi(i_1) < \pi(i_2) < \cdots < \pi(i_k)$. An increasing subsequence of maximum length is called a *longest increasing subsequence* (LIS). Define the length of any such subsequence as the *length of a longest increasing subsequence* of π. Analogously, we say that $1 \le i_1 < i_2 < \cdots < i_k$ is a *decreasing subsequence* of π with *length* k if $\pi(i_1) > \pi(i_2) > \cdots > \pi(i_k)$. A decreasing subsequence of maximum length is called a *longest decreasing subsequence* (LDS) and the common length of such subsequences is defined to be the *length of a longest decreasing subsequence* of π. For a random Mallows(q) permutation Π as defined above, let $L_n = L_n(q)$ and $L_n^{\downarrow} = L_n^{\downarrow}(q)$ denote the length of a longest increasing subsequence and the length of a longest decreasing subsequence of Π respectively.

Weak laws of large numbers have been established for $L_n(q)$ for different ranges of $q = q(n)$. For $q = 1$, it is a classical result that L_n scales as $2\sqrt{n}$. It is also not hard to see via a subadditive argument (see e.g. [9]) that L_n scales linearly with n when $q < 1$. The growth rate of L_n as well the limiting constant has been identified in intermediate regimes in [9,31]. However, so far, scaling limits for L_n have not been established in any case except when $q = 1$. In this paper we consider $q \in (0,1)$ fixed. Our main result is a central limit theorem for $L_n(q)$, which is the first result identifying the limiting distribution for the Mallows model in a case when $q \ne 1$. We prove that L_n is asymptotically Gaussian with linear variance, confirming a conjecture in [9]. This provides an instance of a phase transition in the scaling limit of the length of LIS between Gaussian and GUE Tracy–Widom distribution as q varies; see Section 1.1 for more details.

Theorem 1. *Fix $0 < q < 1$. Then exist constants $\sigma = \sigma(q) > 0$ and $a = a(q) > 0$ such that for $L_n = L_n(q)$ defined as above, we have*

$$\frac{L_n - an}{\sigma\sqrt{n}} \Rightarrow \mathcal{N}(0,1)$$

as $n \to \infty$ where $\Rightarrow$ denotes convergence in distribution and $\mathcal{N}(0,1)$ denotes the standard Normal distribution with mean 0 and variance 1.

We do not have explicit formulae for $a(q)$ and $\sigma(q)$ in the above theorem. However, in Section 5, we derive representations of a and σ as certain statistics of a measure on the space of permutations with variable length. Observe that $a = \lim_{n\to\infty} \frac{\mathbb{E}L_n}{n}$, where the existence of the limit is guaranteed by a subadditive argument. Bounds on a were derived in [9], but evaluating this constant was left as one of the open problems there.

Our second main result is a law of large numbers for $L_n^{\downarrow}(q)$ for $0 < q < 1$. The order of the rate of growth of $L_n^{\downarrow}(q)$ was found in [9] and the identification of the limiting constant was left as an open question, which we answer in the following theorem.

Theorem 2. *Fix $0 < q < 1$. For $L_n^{\downarrow} = L_n^{\downarrow}(q)$ defined as above, we have*

$$\frac{L_n^{\downarrow}\sqrt{\log q^{-1}}}{\sqrt{2\log n}} \to 1$$

in probability as $n \to \infty$.

The reversal π^R of a permutation π is given by $\pi^R(i) := n + 1 - \pi(i)$ for $1 \le i \le n$. It is easy to check (see e.g. [9]) that the reversal of a Mallows(q) distributed permutation is a Mallows($1/q$) permutation. Since the length of an LIS of a permutation is equal to the length of an LDS of its reversal, Theorem 2 can also be interpreted as a law of large numbers of $L_n(q)$ when $q > 1$.

1.1. *Background and related works*

The Mallows model was originally introduced motivated by ranking problems in statistics. The Mallows distribution can be defined more generally with respect to a reference permutation π_0 through the Kendall–Tau distance

$$d(\pi, \pi_0) := \sum_{1 \le i < j \le n} \mathbf{1}_{\{\pi_0(i) < \pi_0(j)\}} \mathbf{1}_{\{\pi(i) > \pi(j)\}}$$

which reduces to $\mathrm{inv}(\pi)$ when π_0 is the identity permutation. Mallows was interested in the problem of determining an unknown "true" ordering or "reference permutation" on the elements π_0 given samples of orderings penalized according to the number of pairs out of order compared to the reference permutation. Generalized Mallows models using a variety of metrics on permutations are used to model ranked and partially ranked data in machine learning and social choice theory [13,17–19,27,30].

The Mallows(q) measure is related to representations of the Iwahori–Hecke algebra [18] and to a natural q-deformation of exchangeability studied by Gnedin and Olshanski [21,22]. It also arises in connection with the stationary measures of the biased adjacent transpositions shuffle on $\mathcal{S}_n$ and of the nearest-neighbor asymmetric exclusion process on an interval where particles jump to the left with probability $1 - p$ and to the right with probability p and $q = p/(1 - p)$ [8]. Further, by the correspondence between the ASEP and the XXZ quantum spin system [11,36], the ground state of the XXZ model is a projection of the Mallows model on permutations.

The normalizing constant $Z_n(q)$ in the Mallows distribution has a closed form formula which Diaconis and Ram [18] observed to be the Poincaré polynomial

$$Z_{n,q} := \sum_{\pi \in \mathcal{S}_n} q^{\mathrm{inv}(\pi)} = \prod_{i=1}^{n} \frac{q^i - 1}{q - 1} = [n]_q! = [n]_q \cdots [1]_q, \quad \text{where } [i]_q = \frac{q^i - 1}{q - 1}. \tag{1}$$

The formula (1) for $Z_{n,q}$ implies a straightforward method for generating a random Mallows distributed permutation.

As mentioned before, the question of determining the length of the longest increasing subsequence of permutations drawn from a Mallows model for general q was raised in [10]. When $q = 1$, i.e. in the case of uniform random permutations, the asymptotics of L_n (known as Ulam's problem) have been extensively studied. Vershik and Kerov [38] and Logan and Shepp [28] showed that $\mathbb{E}L_n/\sqrt{n} \to 2$ (see also [1] for a proof using Hammersley's interacting particle system). Mueller and Starr [31] first studied L_n under the Mallows measure for $q \neq 1$. In the regime that $n(1 - q)$ tends to a constant β, they established a weak law of large numbers showing that $L_n/\sqrt{n}$ converges in probability to a constant $C(\beta)$ which they determined explicitly. Their arguments rely on a Boltzmann–Gibbs formulation of a continuous version of the Mallows measure and the probabilistic approach of Deuschel and Zeitouni for analyzing the longest increasing subsequence of i.i.d. random points in the plane [14,15].

Subsequently, in [9], Bhatnagar and Peled established the leading order behavior of L_n in the regime that $n(1 - q) \to \infty$ and $q \to 1$. They analyzed an insertion process which they called the Mallows process for randomly generating Mallows distributed permutations and showed that $L_n/n\sqrt{1 - q} \to 1$ in L_p for $0 < p < \infty$ as $n \to \infty$. They established the order of $L_n^{\downarrow}$ and showed that it grows at different rates for different regimes of q as a function of n (in particular they showed $\mathbb{E}L_n^{\downarrow}(q) = \Theta(\sqrt{\log n/\log q^{-1}})$ when $q \in (0, 1)$ is fixed) and proved large deviation bounds for L_n and $L_n^{\downarrow}$. They also established a linear upper bound on the variance of L_n and left open the questions of determining the precise variance and the distribution of L_n for all regimes of n and q.

Recently, further progress been made in the analysis of the empirical measure of points corresponding to the Boltzmann–Gibbs measure of Mueller and Starr. Mukherjee [32] determined the large deviation rate function of the empirical measure of points and Starr and Walters recently showed that the large deviation principle has a unique optimizer [37].

The case $q = 1$ is special because it is one of the exactly solvable models that belong to the so-called KPZ universality class. In this case, the longest increasing subsequence problem can also be represented as a directed last passage percolation problem in a Poissonian environment. The results of [28,38] follow from an asymptotic analysis of exact formulae that can be obtained for $\mathbb{E}L_n$ through a combinatorial bijection between $\mathcal{S}_n$ and Young tableaux known as the Robinson–Schensted–Knuth (RSK) bijection [26,33,34]. The RSK bijection can be further used to obtain the order of fluctuations and scaling limit of L_n in this case. In their breakthrough work, Baik, Deift and Johansson showed that for uniformly random permutations L_n has fluctuations of the order of $n^{1/6}$ and the limiting distribution of $n^{-1/6}(L_n - 2\sqrt{n})$ is the GUE Tracy–Widom distribution from random matrix theory [5].

When $q \neq 1$, the integrable structure is lost, and the powerful combinatorial, algebraic and analytic tools used in [5] are no longer available for finding the limiting distribution of L_n. Indeed, as Theorem 1 shows, for q bounded away from 1, we get a different scaling limit, namely Gaussian with diffusive scaling, in contrast with the Tracy–Widom limit with subdiffusive scaling one gets for $q = 1$.

Theorem 1 indicates that as $q \to 1$, there is a phase transition in the limiting distribution of L_n between Gaussian and Tracy–Widom. There are other models where such transitions have been shown to occur. The so-called BBP transition [4] for the spiked complex Wishart model, a model for non-null covariance matrices [25], is a prime example. Connections with random matrices have been exploited to show similar transitions in exactly solvable models of last passage percolation with external sources [6,7]. Much less is understood in the absence of the exactly solvable machinery. Chatterjee and Dey [12] show that the first passage time across thin cylinders obeys a central limit theorem. Dey, Joseph and Peled [16] consider the similar problem of directed last passage percolation in a Poissonian environment restricted to band of width n^γ around the diagonal. Relying on finer estimates available for exactly solvable models they establish a sharp transition in γ for the limiting law of the length of a maximal path, showing that when $\gamma < 2/3$, the limiting distribution is Gaussian. It is known that if $\gamma > 2/3$ the limit is again Tracy–Widom. Another recent example of a Gaussian scaling limit in a last passage percolation model is obtained by Houdré and Işlak [24] for the length of the longest common subsequence of random words.

The Mallows model is not known to be exactly solvable and unlike many of the exactly solvable models described above, the location of the transition (or indeed, if there is only one transition or multiple ones) is currently not known. It can be shown (see Remark 2 in Section 7) that if $q \to 1$ sufficiently fast as $n \to \infty$ that L_n, properly centered and scaled, indeed has a Tracy–Widom distribution. To understand this transition(s) is a fascinating problem. By making the calculations in our proof of Theorem 1 quantitative, it may be possible to show that a central limit theorem also holds for q tending to 1 sufficiently slowly as n goes to infinity. However, in order to avoid technical complications, we do not pursue this direction in this paper.

Our analysis hinges on identifying a regenerative process associated with an infinite Mallows permutation. In [21], Gnedin and Olshanski study a notion of q-exchangeability in infinite Mallows permutations on $\mathbb{Z}$. As in the finite case, there is a natural insertion process for generating infinite Mallows permutations. One key observation in our work is that we can define a certain $\mathbb{N}$-valued Markov chain associated with this process. The times when the chain reaches 0 form a set of regeneration times which enables us to view L_n as a sum of longest increasing subsequences restricted to the permutation defined by the interval between the renewal times; see Section 3 for details. This view allows us to apply results from renewal theory directly to establish Theorem 1 and makes the analysis rather clean. A more refined analysis of the same Markov chain allows us to establish Theorem 2.

One could hope that this or similar constructions might be useful in studying other properties of the Mallows measure. Indeed, shortly before completing this work, we learnt that Gladkich and Peled use a construction of Mallows permutations and an associated Markov chain similar to what we define to analyze the cycle structure of random Mallows distributed permutations [20]. For the chain in [20], Gladkich and Peled also analyze the excursions away from zero, and indeed, the two chains share the same return times to zero and they obtain estimates similar to ours.

Organization of the paper

The rest of this paper is organized as follows. In Section 2 we recall the construction of a Mallows permutation via Mallows' process. In Section 3 we obtain a regenerative process representation of a Mallows process. In Section 4 we obtain estimates on the renewal time of the regenerative process by relating it to a return time of a certain Markov chain. Using these estimates we complete proofs of Theorem 1 and Theorem 2 in Section 5 and Section 6 respectively. We conclude with some open questions in Section 7.

2. Constructing Mallows permutations

Fix $0 < q < 1$. Gnedin and Olshanski [21,22] constructed an *infinite* Mallows(q) *permutation* on $\mathbb{N}$ via an insertion process, which we will also refer to as Mallows(q) process, often dropping q from the argument. The process is as follows. Start with an i.i.d. sequence $\{Z_i\}_{i\geq 1}$ of Geom$(1-q)$ variables. Construct a permutation $\widetilde{\Pi}$ of the natural numbers inductively as follows: Set $\widetilde{\Pi}(1) = Z_1$. For $i > 1$, set $\widetilde{\Pi}(i) = k$ where k is the Z_ith number in the increasing order from the set $\mathbb{N} \setminus \{\widetilde{\Pi}(j) : 1 \leq j < i\}$. For example, suppose that the realizations of the first five independent geometrics are $z_1 = 4, z_2 = 1, z_3 = 6, z_4 = 2$, and $z_5 = 3$. Then we have $\widetilde{\Pi}(1) = 4, \widetilde{\Pi}(2) = 1, \widetilde{\Pi}(3) = 8, \widetilde{\Pi}(4) = 3$ and $\widetilde{\Pi}(5) = 6$. We represent the process step-by-step below. Note that in each step, the new element i is placed in the z_ith unassigned position among the currently unassigned positions.

$$
\begin{array}{cccccccccc}
\_ & \_ & \_ & 1 & \_ & \_ & \_ & \_ & \_ & \cdots \\
2 & \_ & \_ & 1 & \_ & \_ & \_ & \_ & \_ & \cdots \\
2 & \_ & \_ & 1 & \_ & \_ & \_ & 3 & \_ & \cdots \\
2 & \_ & 4 & 1 & \_ & \_ & \_ & 3 & \_ & \cdots \\
2 & \_ & 4 & 1 & \_ & 5 & \_ & 3 & \_ & \cdots
\end{array}
$$

Let Π_n be the permutation on $[n]$ induced by $\widetilde{\Pi}$, i.e., $\Pi_n(i) = j$ if $\widetilde{\Pi}(i)$ has rank j when the set $\{\widetilde{\Pi}(k) : k \in [n]\}$ is arranged in an increasing order. Consider in the above example $n = 4$. Then we have $\Pi_4(1) = 3$, $\Pi_4(2) = 1$, $\widetilde{\Pi}(3) = 4$ and $\Pi_4(4) = 2$. In short, we shall write this permutation as $\Pi_4 = 2413$, with the interpretation that Π_4 takes the element 1 to the position 3 and so on. In general, with this representation we can read off Π_n from the above array by restricting to the elements in $[n]$ in the representation above. The following lemma is essentially contained in [21] although it is not explicitly spelt out there. We provide a proof for completeness.

Lemma 2.1. *Let* $\widetilde{\Pi}$ *be an infinite* Mallows(q) *permutation and let* Π_n *be the induced permutation on* $[n]$, *as defined above. Then* Π_n *is a* Mallows(q) *permutation on* $[n]$.

Proof. Fix the values of $\widetilde{\Pi}(i)$ for $i \in [n]$ and let the set of these values be $A = A_n = \{a_i : i \in [n]\}$. Suppose that $\pi \in \mathcal{S}_n$. Given A and π, there is a uniquely determined set of values $\{z_i\}_{i=1}^n$ for the first n geometric random variables so that $\Pi_n = \pi$. Moreover,

$$\mathbb{P}(\Pi_n = \pi, A) = (1-q)^n \prod_{i=1}^n q^{z_i - 1}.$$

For any $1 \leq i \leq n$, let $\pi' = (i, i+1) \circ \pi$, that is, π' is the permutation obtained from π by exchanging the positions of the elements i and $i+1$. Letting $\{z_i'\}_{i=1}^n$ denote the corresponding geometrics, it is simple to verify that, depending on whether $\pi(i) < \pi(i+1)$ or not, either $z_i' = z_{i+1} + 1$ and $z_{i+1}' = z_i$ or $z_i' = z_{i+1}$ and $z_{i+1}' = z_i - 1$ while $z_j = z_j'$ for all $j \notin \{i, i+1\}$. Thus

$$\frac{\mathbb{P}(\Pi_n = \pi', A)}{\mathbb{P}(\Pi_n = \pi, A)} = \begin{cases} q & \text{if } \pi(i) < \pi(i+1), \\ \frac{1}{q} & \text{if } \pi(i) > \pi(i+1). \end{cases}$$

Thus the distribution of Π_n conditioned on A is Mallows(q) since the transpositions generate the group $\mathcal{S}_n$. Summing over all possibilities for A completes the proof of the claim. $\square$

Notice that the construction of $\widetilde{\Pi}$ and Π_n as described in Section 2 generates a family of Mallows(q) permutations on $[n]$ for $n \in \mathbb{N}$ on the same probability space. Henceforth when we talk about a Mallows(q) permutation Π_n on $[n]$, it will be assumed that Π_n is constructed from an infinite Mallows(q) permutation as described above.

3. The regenerative process representation

A stochastic process $\mathbf{X} = \{X(t) : t \geq 0\}$ is said to be a *regenerative process* if there exist *regeneration times* $0 \leq T_0 < T_1 < T_2 < \cdots$ such that for each $k \geq 1$, the process $\{X(T_k + t) : t \geq 0\}$ has the same distribution as $\{X(T_0 + t) : t \geq 0\}$

and is independent of $\{X(t) : 0 \le t < T_k\}$. Below we define a regenerative process associated with the Mallows(q) process described above.

Recall the sequential construction of $\widetilde{\Pi}$ and the induced permutation Π_n. Suppose $m \in \mathbb{N}$ is such that we have $\widetilde{\Pi}([m]) = [m]$, i.e. the permutation $\widetilde{\Pi}$ restricted to $[m]$ defines a bijection from $[m]$ to $[m]$. Define the permutation $\widetilde{\Pi}^* : \mathbb{N} \to \mathbb{N}$ by $\widetilde{\Pi}^*(i) = \widetilde{\Pi}(i+m) - m$. It is clear from the construction of $\widetilde{\Pi}$ that $\widetilde{\Pi}^*$ and $\widetilde{\Pi}$ have the same law. Together with the independence of the geometric variables $\{Z_i\}$ this implies that $\{\widetilde{\Pi}(i) - i\}_{i \in \mathbb{N}}$ is a regenerative process with regeneration times $0 = T_0 < T_1 < T_2 < \cdots$ where for $i > 1$ we have,

$$T_i = \min\bigl\{j > T_{i-1} : \bigl\{\widetilde{\Pi}(k) : k \in [j] \setminus [T_{i-1}]\bigr\} = [j] \setminus [T_{i-1}]\bigr\}.$$

We illustrate by an example. Suppose that in the sequential construction for $\widetilde{\Pi}$, the values of the first 8 geometrics are $z_1 = 1, z_2 = 2, z_3 = 3, z_4 = 1, z_5 = 2, z_6 = 3, z_7 = 1$ and $z_8 = 1$. This corresponds to the permutation

$$1\ 4\ 2\ 7\ 3\ 5\ 8\ 6\ 9\ 10\ \_\ \_\ \cdots$$

and we see that $T_1 = 1, T_2 = 8, T_3 = 9$ and $T_4 = 10$ are the regeneration times. We may also view the graphical representation of the permutation by plotting the points $(i, \widetilde{\Pi}(i))$ in $\mathbb{R}^2$ for each $i \in \mathbb{N}$. This is illustrated in Figure 1 for the first 50 elements of $\widetilde{\Pi}$ which is a random Mallows(0.55) permutation. The regeneration times are marked by the corners of the squares which lie on the diagonal $y = x$. The figure illustrates that the points can be partitioned into such squares, which are minimal in the sense that no smaller square with its corners on the diagonal can contain a strict subset of the points in a box.

Set $X_i = T_i - T_{i-1}$ for $i \ge 1$. Clearly, X_i are independent and identically distributed. Let $\Sigma_j(i) := \widetilde{\Pi}(i + T_{j-1}) - T_{j-1}$ for $i \in \{T_{j-1}+1, T_{j-1}+2, \dots, T_j\}$. Then Σ_j is a permutation of $[X_j]$ and furthermore the $\{\Sigma_j\}_{j\ge1}$ are i.i.d. Let $S_n := \min\{j : T_j \ge n\}$.

Recall that L_n (resp. $L_n^{\downarrow}$) is the length of the longest increasing (resp. decreasing) subsequence in Π_n. The following two lemmas connect L_n and $L_n^{\downarrow}$ with the corresponding statistics defined in the permutations $\{\Sigma_j\}_{j\ge1}$.

Lemma 3.1. *For $j \ge 1$, let Y_j denote length of a longest increasing subsequence of Σ_j. Then we have,*

$$\sum_{j=1}^{S_n-1} Y_j < L_n \le \sum_{j=1}^{S_n} Y_j.$$

Proof. By Lemma 2.1, the LIS of Π_n is distributed as L_n. Observe that any subsequence in Σ_j corresponds, in an obvious way, to a subsequence in $\widetilde{\Pi}$ and conversely, any subsequence of $\widetilde{\Pi}$ contained in $[T_j] \setminus [T_{j-1}]$ corresponds

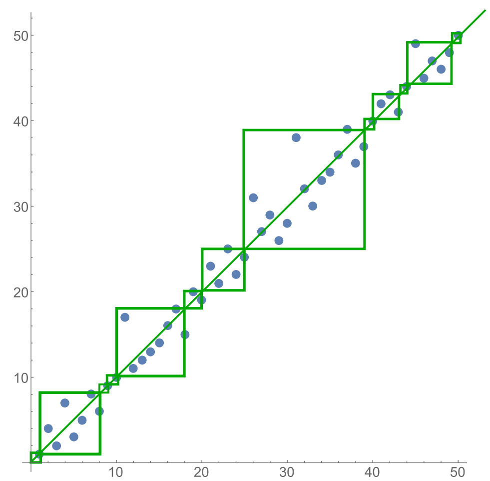

Fig. 1. The regeneration times $T_0 < T_1 < \cdots$ are marked by the corners of the squares.

uniquely to a subsequence in Σ_j. An increasing subsequence of Π_n when restricted to $[T_j] \setminus [T_{j-1}]$ for $1 \le j \le S_n$ corresponds to an increasing subsequence of Σ_j, which implies the upper bound. On the other hand, any union of increasing subsequences in the Σ_j for $1 \le j \le S_n - 1$ corresponds to an increasing subsequence in Π_n, which implies the lower bound. □

Lemma 3.2. *Let $Y_i^{\downarrow}$ denote the length of the LDS in Σ_i. We have*

$$\max_{i \le S_n - 1} Y_i^{\downarrow} \le L_n^{\downarrow} \le \max_{i \le S_n} Y_i^{\downarrow}. \tag{2}$$

Proof. The lemma follows by observing that any decreasing subsequence of Π must be contained in $[T_i] \setminus [T_{i-1}]$ for some i. We omit the details. □

4. Renewal time estimates via a Markov chain

Our objective in this section is to prove that the inter-renewal times X_i as defined in the previous section has finite first and second moments. These are the conditions we require to apply results from the theory of regenerative processes to show the central limit theorem for L_n. We define a Markov chain such that the X_i's can be represented as the excursion lengths of this Markov chain. Kac's formula for the moments of return times from the theory of recurrent Markov chains then implies that the moments of X_i are finite.

For convenience, let X denote a random variable with the distribution same as the common one of X_i's. First we show that X has the same law as the return time of a certain Markov chain which we define below.

Let $\{M_n\}_{n \ge 0}$ denote a Markov chain with the state space $\Omega = \mathbb{N} \cup \{0\}$ and the one step transition defined as follows: set $M_n = \max\{M_{n-1}, Z_n\} - 1$ where $\{Z_i\}$ is a sequence of i.i.d. Geom$(1-q)$ variables. Let R_0^+ denote the first return time to 0 of this chain, i.e.

$$R_0^+ = \min\{k > 0 : M_k = 0\}.$$

Lemma 4.1. *For the Markov chain M_n started at $M_0 = 0$, the return time $R_0^+ \stackrel{d}{=} T_1$. In particular X has the same law as R_0^+.*

Proof. Couple the Markov chain M_n with $M_0 = 0$ with the Mallows' process by using the same sequence $\{Z_i\}$ of random variables to run both processes. Under this coupling, it is easy to verify that for each n, by definition

$$M_n = \max_{1 \le j \le n} \{\widetilde{\Pi}(j)\} - n.$$

The claim thus follows immediately from the definitions of R_0^+ and T_1. □

We analyze the Markov chain M_n and the return time R_0^+ in the next few lemmas.

Lemma 4.2. *The Markov chain M_n is a positive recurrent Markov chain whose unique stationary distribution $\mu = (\mu_j)_{j \ge 0}$ is given by*

$$\mu_j = \left(1 + \sum_{j=1}^{\infty} \frac{q^j}{\prod_{k=1}^{j}(1-q^k)}\right)^{-1} \frac{q^j}{\prod_{k=1}^{j}(1-q^k)}; \quad j \ge 0.$$

Proof. Let $\boldsymbol{P} = \{P_{i,j}\}_{i,j \ge 0}$ denote the transition matrix of the chain and let Z denote a Geom$(1-q)$ random variable. It is clear from the definition of the chain that for $i \ge 0$ and $j \ge i$ we have $P_{i,j} = \mathbb{P}(Z = j+1) = q^j(1-q)$; for $i \ge 1$ we have $P_{i,i-1} = \mathbb{P}(Z \le i) = 1 - q^i$ and for all other pairs (i, j) we have $P_{i,j} = 0$. Clearly the chain is irreducible. It is known from elementary Markov chain theory (see e.g. [2]) that a stationary distribution exists and is unique if and only if there exists a unique probability vector (i.e., vector with non-negative entries whose co-ordinates sum up to 1) μ solving the set of linear equations $\mu \boldsymbol{P} = \mu$. The equation corresponding to the jth column of the matrix $\boldsymbol{P}$ is given by

$$\mu_0(1-q) + \mu_1(1-q) = \mu_0 \tag{3}$$

for $j = 0$ and

$$\sum_{k=0}^{j} \mu_k q^j (1-q) + \mu_{j+1}\big(1 - q^{j+1}\big) = \mu_j \tag{4}$$

for $j > 0$. It is easy to check that any solution of this set of equations must satisfy $\mu_{j+1} = \frac{q}{1-q^{j+1}} \mu_j$, and hence we must have

$$\mu_j = \frac{q^j}{\prod_{k=1}^{j}(1-q^k)} \mu_0.$$

Since

$$1 + \sum_{j=1}^{\infty} \frac{q^j}{\prod_{k=1}^{j}(1-q^k)} = \mathcal{Z}(q) < \infty$$

a unique probability vector μ satisfying the above conditions does indeed exist and is given by

$$\mu_j = \frac{1}{\mathcal{Z}(q)} \frac{q^j}{\prod_{k=1}^{j}(1-q^k)}.$$

Since the chain is irreducible and has a stationary distribution, it is positive recurrent (see e.g. [2], Theorem 13.4). $\square$

Remark 1. In fact, as an anonymous referee has pointed out, the expression for $\mathcal{Z}(q)$ can be simplified and is given by

$$\mathcal{Z}(q) = \frac{1}{\prod_{k=1}^{\infty}(1-q^k)}.$$

The existence of first and second moments of R_0^+ follows from the above lemma and is proved in Lemma 4.5. We begin with the following preliminary lemmas. Let R_i denote the time for the chain to reach state i. We shall denote by $\mathbb{E}_i$ (resp. $\mathbb{P}_i$) the expectation (resp. the probability measure) with respect to the chain started at the state i and $\mathbb{E}_\mu$ shall denote the expectation with respect to the chain started at stationarity.

Lemma 4.3. *For all $i \geq 1$, $\mathbb{E}_i R_{i-1} \geq \mathbb{E}_{i+1} R_i$.*

Proof. If two copies of the chain are both started at $k \geq i$ and coupled using the same set of geometric variables, so that they are identical, then $R_{i-1} > R_i$ and hence $\mathbb{E}_k R_{i-1} \geq \mathbb{E}_k R_i$. Suppose now that we couple two copies of the chain, one started at i and the other at $i+1$ using the same geometric variables. Let Z be a $\mathrm{Geom}(1-q)$ variable and suppose that we make one move according to Z in both chains. If $Z \leq i$, then the chain started at i reaches $i-1$ and the chain started at $i+1$ reaches i. If $Z \geq i+1$, then both chains go to the state $Z-1$. Thus,

$$\mathbb{E}_i R_{i-1} = \mathbb{P}(Z \leq i) + \sum_{j=i+1}^{\infty} \mathbb{P}(Z = j) \mathbb{E}_{j-1} R_{i-1} \geq \mathbb{P}(Z \leq i) + \sum_{j=i+1}^{\infty} \mathbb{P}(Z = j) \mathbb{E}_{j-1} R_i = \mathbb{E}_{i+1} R_i.$$

$\square$

Lemma 4.4. *For the Markov chain M_n, $\mathbb{E}_\mu R_0 < \infty$.*

Proof.

$$\begin{aligned}\mathbb{E}_\mu R_0 &= \sum_{j=1}^{\infty} \mu_j \mathbb{E}_j R_0 \\ &= \sum_{j=1}^{\infty} \mu_j \sum_{k=1}^{j} \mathbb{E}_k R_{k-1} \\ \text{(by Lemma 4.3)} &\le \sum_{j=1}^{\infty} j \mu_j \mathbb{E}_1 R_0 < \infty.\end{aligned}$$

The last inequality can be justified as follows. The positive recurrence of the chain M_n implies that $E_1 R_0$ is finite. Further, Lemma 4.2 shows that $\sup_j \mu_j / q^j$ is finite (this can be verified by either noting that $\frac{1}{\prod_{k=1}^{\infty}(1-q^k)} > 0$ or by using the formula in Remark 1) and hence $\sum_j j\mu_j$ is finite. □

Lemma 4.5. *Let R_0^+ be as defined in Lemma 4.1. Then we have $\mathbb{E}_0 R_0^+ < \infty$ and $\mathbb{E}_0 (R_0^+)^2 < \infty$.*

Proof. It is a basic fact about Markov chains that (see e.g. [2]) that $\mathbb{E}_0 R_0^+ = \mu_0^{-1}$. From the proof of Lemma 4.2, we have $\mu_0^{-1} = \mathcal{Z}(q) < \infty$. For the second moment we argue as follows. Let R_0 denote the time of the first visit of the Markov chain to the state $\{0\}$. It is a consequence of Kac's formula (Corollary 2.24, [2]) that (see (2.21) in [2])

$$\mathbb{E}_0 \big(R_0^+\big)^2 = \frac{2\mathbb{E}_\mu R_0 + 1}{\mu_0}. \tag{5}$$

The claim follows by Lemma 4.4 since $\mathbb{E}_\mu R_0 < \infty$. □

We shall also need the following tail estimates for the return time to prove Theorem 2.

Proposition 4.6. *Let $0 < q < 1$ and consider the Markov chain M_n as defined above. There exist positive constants $A = A(q)$ and $c = c(q)$ such that for all $t \ge 0$ and $s \ge 0$, we have*

$$\mathbb{P}_t\big[R_0^+ > 10t + s\big] \le Ae^{-cs}. \tag{6}$$

More generally, denoting the first return time to or below $v \ge 0$ by R_v^+ we have

$$\mathbb{P}_{t+v}\big[R_v^+ > 10t + s\big] \le Ae^{-cs}.$$

First we show that it suffices to only prove the first statement in the above proposition. To see this notice the following. If we couple two copies of the chain M_n and M_n' with $M_0 = t$ and $M_0' = t + v$ using the same sequence $\{Z_i\}$ of Geometric variables, then we have that $M_n' - M_n$ does not increase with n and in particular, $\min\{n \ge \ell \ge 1 : M_\ell\} \ge \min\{n \ge \ell \ge 1 : M_\ell' - v\}$. Hence the return time to 0 in M_n is at least as large as the return time to v in M_n' showing that it is sufficient to establish (6). We shall prove (6) using the following estimates.

Lemma 4.7. *Fix $0 < q < 1$. Let $C_1 = C_1(q)$ be sufficiently large such that $\frac{q^{C_1}}{1-q} < \frac{1}{10}$. There exist positive constants $A = A(q) > 1$ and $c = c(q)$ such that for any $t \ge C_1$ and $s \ge 0$*

$$\mathbb{P}_t\big[R_{C_1}^+ > 10t + s/2\big] \le Ae^{-cs}.$$

Proof. Consider the Markov chain M_n with $M_0 = t$. Let $\alpha > 0$ be a constant that we will choose to be sufficiently small. Observe that we have for $qe^\alpha < 1$,

$$\begin{aligned}\mathbb{E}\big(e^{\alpha M_{\ell+1}} \mid M_\ell\big) &= \big(1-q^{M_\ell}\big)e^{\alpha(M_\ell - 1)} + \frac{q^{M_\ell}(1-q)}{1-qe^\alpha}e^{\alpha M_\ell} \\ &= e^{\alpha M_\ell}\left(e^{-\alpha}\big(1-q^{M_\ell}\big) + \frac{q^{M_\ell}(1-q)}{1-qe^\alpha}\right).\end{aligned}$$

Since $M_\ell > C_1$ on the event $\{\ell < R^+_{C_1}\}$ it follows by using this that $\frac{q^{C_1}}{1-q} < \frac{1}{10}$, by choosing α sufficiently small one can make the quantity in the parenthesis above less than $e^{-\alpha/10}$ (for small α, it is asymptotically $e^{-9\alpha/10}$). It follows that

$$\mathbb{E}\big(e^{\alpha M_{10t+s/2}}\mathbf{1}_{\{R^+_{C_1} > 10t+s/2\}} \mid M_0 = t\big) \le e^{\alpha t}e^{-\alpha(10t+s/2)/10}.$$

The lemma now follows. □

Let $\mathcal{L}_x(t)$ denote the time the chain M_n spends at or below x up to time t. We have the following estimate.

Lemma 4.8. *Fix $r \le C_1$ where C_1 is as above. Then there exist constants $C_2 = C_2(q) > 0$ and $c > 0$ such that*

$$\mathbb{P}_r\left[\mathcal{L}_{C_1}(s/2) < \frac{s}{C_2(q)}\right] \le A_1 e^{-cs}.$$

Proof. For two copies of the chain M_n and M'_n started at a and b respectively with $a \le b$, the chains can be coupled so that $M_n \le M'_n$, and hence without loss of generality we may assume $r = C_1$. Now let $\xi_1, \xi_2, \dots$ denote the lengths of a sequence of independent excursions above C_1. Hence it suffices to show that for C_2 sufficiently large

$$\mathbb{P}\left[\sum_{i=1}^{s/C_2}\xi_i > s/2\right] \le Ae^{-cs}.$$

This in turn follows by observing that by Lemma 4.7 we have $\mathbb{E}e^{\alpha\xi_i} < \infty$ for α sufficiently small. □

Now we are ready to prove Proposition 4.6.

Proof of Proposition 4.6. Let C_1, C_2 be as in the above two lemmas. It follows from Lemma 4.7 that it suffices to prove

$$\mathbb{P}_{C_1}\left[R^+_0 > \frac{s}{2}\right] \le Ae^{-cs}$$

for some positive constants A and c for s sufficiently large. For $i = 1, 2, \dots, s/4C_1$, denote the interval $[(2i-2)C_1, (2i-1)C_1)$ (resp. $[(2i-1)C_1, 2iC_1)$) by J_i (resp. J^*_i). For each i, let $Z^{(i)}_j Z^{(i,*)}_j$, $j = 1, 2, \dots, C_1$ denote independent sequences of i.i.d. Geom$(1-q)$ variables. For the chain M_n, define

$$\tau_i = \min\{n \in J_i : M_n \le C_1\}; \qquad \tau^*_i = \min\big\{n \in J^*_i : M_n \le C_1\big\}.$$

Let $\mathcal{A}_i$ denote the event

$$\mathcal{A}_i = \left\{\{i : \tau_i < \infty\} < \frac{s}{4C_1C_2}\right\}$$

and define the event $\mathcal{A}_i^*$ similarly by replacing τ_i by τ_i^*. Observe that from Lemma 4.8 it follows that

$$\mathbb{P}_{C_1}\big[\mathcal{A}_i \cap \mathcal{A}_i^*\big] \le Ae^{-cs}.$$

Let B (resp. B^*) denote the event that $R_0^+ > \frac{s}{2}$ together with the complement of $\mathcal{A}_i$ (resp. the complement of $\mathcal{A}_i^*$). Clearly using the above display it suffices to show

$$\mathbb{P}[B] + \mathbb{P}\big[B^*\big] \le Ae^{-cs}$$

for some positive constants A and c for s sufficiently large. Run the chain M_n as follows. Let

$$\tau_i = \min\{n \in J_i : M_n \le C_1\}.$$

If $\tau_i < \infty$, then use the variables $Z_j^{(i)}$ to run the chain for the next C_1 steps, and use independent external randomness to run the chain for other steps. Call i *good* if $Z_j^{(i)} = 1$ for all $j = 1, 2, \dots, C_1$. Clearly if for some i, $\tau_i < \infty$ and i is good then $R_0^+ \le s/2$. Now observe that $\mathbb{P}[i \text{ is good}] = (1-q)^{C_1} = d(q) > 0$ and also observe that on the complement of $\mathcal{A}_i$ we have

$$\#\{i : \tau_i < \infty\} > \frac{s}{4C_1C_2}.$$

It follows that

$$\mathbb{P}_{C_1}[B] \le \big(1 - d(q)\big)^{s/4C_1C_2}.$$

Arguing similarly with replacing J_i, τ_i and $Z_j^{(i)}$ by J_i^*, τ_i^* and $Z_j^{(i,*)}$ respectively gives us the same upper bound for $\mathbb{P}_{C_1}[B^*]$. The proposition now follows by noting

$$\mathbb{P}_{C_1}\left[R_0^+ > \frac{s}{2}\right] \le \mathbb{P}_{C_1}[B] + P_{C_1}\big[B^*\big] + \mathbb{P}_{C_1}\big[\mathcal{A}_i \cap \mathcal{A}_i^*\big]. \qquad \square$$

5. Anscombe's Theorem and a CLT for the length of the LIS

In this section we complete the proof of Theorem 1 by invoking a central limit theorem for a random sum due to Anscombe. Let $X_1, X_2 \cdots$ be i.i.d. random variables with finite mean and variance $\sigma^2 > 0$ and let $N(t)$ be an integer-valued process defined on the same probability space as the X_i. Anscombe's Theorem [3] says that if the partial sums Q_n for the $\{X_i\}$ obey a central limit theorem and do not fluctuate too much, then the random sum $Q_{N(t)}$ also obeys the central limit theorem.

Theorem 5.1 (Anscombe's Theorem, e.g. [23]). *Let $X, X_1, X_2, \dots$ be independent, identically distributed random variables with mean 0 and positive, finite variance σ^2. For $n \ge 1$, let $Q_n = \sum_{i=1}^n X_i$. Suppose $\{N(t), t \ge 0\}$ is a family of positive, integer values random variables such that for some $0 < c < \infty$,*

$$\frac{N(t)}{t} \xrightarrow{p} c \quad \text{as } t \to \infty.$$

Then,

$$\frac{Q_{N(t)}}{\sqrt{t}} \xrightarrow{d} \mathcal{N}\big(0, c\sigma^2\big) \quad \text{as } t \to \infty.$$

To apply Anscombe's Theorem in our context, we make use of the following concentration result. Recall the regenerative process from Section 3 with inter-renewal times X_i. Recall $S_n = \min\{j : \sum_{i=1}^j X_i \ge n\}$.

Lemma 5.2. *For μ_0 as defined in the previous section,*

$$\frac{S_n}{n} \overset{a.s.}{\to} \mu_0.$$

Proof. Observe that

$$\frac{\sum_{j=1}^{S_n-1} X_j}{S_n} \le \frac{n}{S_n} \le \frac{\sum_{j=1}^{S_n} X_j}{S_n}.$$

As $n \to \infty$, by strong law, both the left and right hand sides of the above inequality converges to μ_0^{-1}, hence the lemma. □

Using Theorem 5.1 and Lemma 5.2. we can show the following regenerative version of the Central Limit Theorem (see e.g. [35], Chapter 2, Theorem 65), we omit the proof.

Theorem 5.3 (Regenerative CLT). *Let $(X_i, Y_i)_{i\ge 1}$ and S_n be as defined in Section 3. Define $a := \mu_0 \mathbb{E} Y_1 < \infty$. Suppose further that $\eta^2 := \mathrm{Var}(Y_1 - aX_1)$ is positive and finite. Set $Q_n = \sum_{j=1}^{S_n} Y_j$. Then we have*

$$\frac{Q_n - an}{\sqrt{n}} \Rightarrow \mathcal{N}\big(0, \mu_0 \eta^2\big).$$

We need the following to complete the proof of Theorem 1.

Lemma 5.4. *In the context of Theorem 5.3*, $0 < \eta^2 < \infty$.

Proof. Observe that, since $1 \le Y_1 \le X_1$, we have $|Y_1 - aX_1| \le (1+a)X_1$ and hence $\eta^2 < \infty$ using Lemmas 4.1 and 4.5. To see $\eta^2 > 0$, note that

$$\mathrm{Var}(Y_1 - aX_1) = \mathbb{E}\big((Y_1 - aX_1)^2\big) \ge \mathbb{E}\big(\mathrm{Var}(Y_1 \mid X_1)\big)$$

and hence it suffices to prove that for some $j \in \mathbb{N}$ with $\mathbb{P}(X_1 = j) > 0$, we have $\mathrm{Var}(Y_1 \mid X_1 = j) > 0$. To see this consider $j = 3$; it is straightforward to see that $\mathbb{P}(X_1 = 3) > 0$, and conditioned on $\{X_1 = 3\}$, notice that Σ_1 can be both the permutations (3 2 1) and (3 1 2) with positive probability. It then follows that $\mathrm{Var}(Y_1 \mid X_1 = 3) > 0$ and the proof is complete. □

Now we are in a position to complete the proof of Theorem 1. We make use of the following basic result.

Lemma 5.5. *Let $W_1, W_2, \dots,$ be an i.i.d. sequence of non-negative random variables with $\mathbb{E}W_i^2 < \infty$. Then we have for all constants $C > 0$*

$$\frac{\max_{1\le i\le Cn} W_i}{\sqrt{n}} \to 0$$

in probability.

Proof. Fix $C > 0$. For every $\epsilon > 0$ we have

$$\begin{aligned}\mathbb{P}\Big(\max_{1\le i\le Cn} W_i \ge \epsilon\sqrt{n}\Big) &= 1 - \left(1 - \mathbb{P}\left(\frac{W_1^2}{\epsilon^2} \ge n\right)\right)^{Cn} \\ &\to 0\end{aligned}$$

as $n \to \infty$. This follows from the fact that $n\mathbb{P}(W_1^2/\epsilon^2 \ge n) \to 0$ as $n \to \infty$ since $\mathbb{E}[W_1^2/\epsilon^2] < \infty$. This completes the proof. □

Proof of Theorem 1. It follows from Lemma 3.1 that

$$\frac{Q_n - an}{\sqrt{n}} - \frac{\max_{i \le S_n} Y_i}{\sqrt{n}} \le \frac{L_n - an}{\sqrt{n}} \le \frac{Q_n - an}{\sqrt{n}}.$$

Note that $\mathbb{E}Y_i^2 < \infty$ since $Y_i \le X_i$ and $\mathbb{E}(X_i^2) < \infty$ by Lemma 4.5. Using this and Lemma 5.5 it follows that

$$\frac{\max_{i \le 2\mu_0 n} Y_i}{\sqrt{n}} \xrightarrow{p} 0.$$

Using Lemma 5.2 now gives

$$\frac{\max_{i \le S_n} Y_i}{\sqrt{n}} \xrightarrow{p} 0.$$

Hence setting $\sigma = \mu_0^{1/2}\eta$ and using Theorem 5.3 we have

$$\frac{L_n - an}{\sigma\sqrt{n}} \Rightarrow \mathcal{N}(0,1).$$

This completes the proof. □

6. Law of large numbers for the length of the LDS

In this section we establish Theorem 2, a weak law for the length of the longest decreasing subsequence $L_n^{\downarrow}$ of a Mallows(q) permutation, or equivalently, L_n for a Mallows($1/q$) permutation for $0 < q < 1$. Our proof makes use of the Markov chain defined in Section 3. Along the way, we show large deviations estimates for the longest decreasing subsequence that improve upon some of the results in [9], Theorem 1.7, and simplify the proofs.

Recall the regenerative process from Section 3. Let Σ denote a random permutation having the same distribution as Σ_i. Let $Y^{\downarrow}$ denote the length of LDS of Σ. Theorem 2 follows from the following proposition and Lemma 3.2.

Proposition 6.1. $\mathbb{P}(Y^{\downarrow} \ge k) = q^{k^2/2(1+o(1))}$ *as* $k \to \infty$.

We postpone the proof of Proposition 6.1 and assuming it, prove the theorem. We use Lemma 5.2, the fact that S_n is concentrated.

Proof of Theorem 2. Fix $\varepsilon > 0$. Since the $Y_i^{\downarrow}$ are independent and identically distributed, using Proposition 6.1 it can be verified that as $n \to \infty$,

$$\mathbb{P}\left(\max_{i \le (1-\varepsilon)\mu_0 n} Y_i^{\downarrow} < (1-2\varepsilon)\sqrt{\frac{2\log n}{\log q^{-1}}}\right) \le \left(1 - \frac{1}{n^{(1+o(1))(1-2\varepsilon)^2}}\right)^{(1-\varepsilon)\mu_0 n} \to 0 \tag{7}$$

and

$$\mathbb{P}\left(\max_{i \le (1+\varepsilon)\mu_0 n} Y_i^{\downarrow} > (1+2\varepsilon)\sqrt{\frac{2\log n}{\log q^{-1}}}\right) \le (1+\varepsilon)\mu_0 n \frac{1}{n^{(1+o(1))(1+2\varepsilon)^2}} \to 0. \tag{8}$$

By Lemma 5.2, with probability going to 1 as $n \to \infty$, for every $\varepsilon > 0$, $(1-\varepsilon)\mu_0 n \le S_n \le (1+\varepsilon)\mu_0 n$. The result thus follows from equations (2), (7) and (8). □

We break the proof of Proposition 6.1 into two parts, proved in the following lemmas.

Lemma 6.2. *We have* $\mathbb{P}(Y^{\downarrow} = k) \ge q^{k^2/2(1+o(1))}$ *as* $k \to \infty$.

Proof. The probability $\mathbb{P}(Y^{\downarrow} = k)$ can be lower bounded by the probability that Σ is the permutation $(k, k-1, \dots, 2, 1)$. Further,

$$\mathbb{P}\big(\Sigma \text{ is the permutation } (k, k-1, \dots, 2, 1)\big) = (1-q)^k q^{\sum_{i=1}^{k-1} i} = (1-q)^k q^{\frac{k(k-1)}{2}} \geq q^{k^2/2(1+o(1))}. \qquad \square$$

Lemma 6.3. *We have* $\mathbb{P}(Y^{\downarrow} \geq k) \leq q^{k^2/2(1+o(1))}$ *as* $k \to \infty$.

To prove Lemma 6.3 we need the following lemmas. For the rest of this section, we shall consider the coupling between the Markov chain M_n and the Mallows' process used in construction of Σ as described in Lemma 4.1.

Lemma 6.4. *Suppose* $\ell_1 < \ell_2 < \cdots < \ell_k$ *are such that* $(\ell_k, \ell_{k-1}, \dots, \ell_1)$ *is a decreasing subsequence in* Σ. *Then* $Z_{\ell_1} > Z_{\ell_2} > \cdots > Z_{\ell_k}$. *Further, for* $i \geq 2$, $M_{\ell_i} = M_{\ell_i - 1} - 1$, *and finally*, $\min_{\ell_1 \leq t < \ell_k} M_t > Z_{\ell_k} - 1$.

Proof. For $(\ell_k, \ell_{k-1}, \dots, \ell_1)$ to be a decreasing subsequence we must have that ℓ_i is placed to the left of ℓ_{i-1} for all $1 < i \leq k$. By construction $M_{\ell_{i-1}} \leq Z_{\ell_{i-1}} - 1$ for all i. So at step ℓ_{i-1} there are at most $Z_{\ell_{i-1}} - 1$ many empty spots to the left of the spot where ℓ_{i-1} is placed. So for ℓ_i to be placed in one of these spots we must have $Z_{\ell_i} < Z_{\ell_{i-1}}$. This proves the first assertion of the lemma.

For the second assertion suppose that for some $i \geq 2$, we have $M_{\ell_i} \geq M_{\ell_i - 1}$. Then one must have $Z_{\ell_i} > M_{\ell_i - 1}$. This implies that ℓ_i is placed to the right of all elements placed so far, in particular to the right of ℓ_{i-1}, which contradicts the assumption that $(\ell_k, \ell_{k-1}, \dots, \ell_1)$ is a decreasing subsequence.

For the last assertion, observe that when any $\ell_1 \leq t < \ell_k$ is assigned to its position, there must be at least Z_{ℓ_k} empty positions to the left of ℓ_1. This is because it must be the case that ℓ_k is assigned to the left of ℓ_1 since $(\ell_k, \dots, \ell_1)$ is a decreasing subsequence. Thus, it cannot be the case that $M_t \leq Z_{\ell_k} - 1$ since M_t counts the total number of unassigned positions to the left of the rightmost assigned position. $\square$

Lemma 6.5. *Let* $\{Z_i\}_{i \geq 1}$ *be a sequence of i.i.d.* $\mathrm{Geom}(1-q)$ *random variables. Consider the Markov chain* $\{M_t\}_{t \geq 0}$ *defined by* $M_{t+1} = \max\{M_t, Z_{t+1}\} - 1$ *started from* $M_0 = m$. *Fix* $0 < \ell_1 < \ell_2 < \cdots < \ell_k$ *such that* $M_{\ell_i} = M_{\ell_i - 1} - 1$ *for all* i. *Let* $S = \{\ell_1, \ell_2, \dots, \ell_k\}$. *Consider the chain* M' *started from* m *which is run using the same sequence of geometric random variables* $\{Z_i\}$ *except that the* ℓ_i*th steps are censored for each* $1 \leq i \leq k$, *i.e.* $M'_{t+1} = \max\{M'_t, Z'_{t+1}\} - 1$ *where* $Z'_i = Z_{f(i)}$ *where* $f(i)$ *is the* i*th number when* $\mathbb{N} \setminus S$ *is arranged in increasing order. Then*

$$\min_{t \in [\ell_k - k]} M'_t \geq \min_{t \in [\ell_k]} M_t.$$

Lemma 6.5 is an immediate consequence of the following lemma using induction on k.

Lemma 6.6. *In the set-up of Lemma 6.5, suppose* $Z_1 \leq m_1 \leq m_2$. *Consider running two copies of the chain* M *and* M' *with* $M_0 = m_1$ *and* $M'_0 = m_2$. *Let* M *evolve using the sequence* $\{Z_i\}_{i \geq 1}$ *and* M' *evolve using the sequence* $\{Z_i\}_{i \geq 2}$. *Then* $M'_t \geq M_{t+1}$ *for all* t.

Proof. Since $Z_1 \leq m_1$, it follows that $M_1 = m_1 - 1 \leq m_2$. The result now follows by induction and the definitions of the chains. By definition, $M'_t = \max\{M'_{t-1}, Z_{t+1}\} - 1$ and $M_{t+1} = \max\{M_t, Z_{t+1}\} - 1$. Thus if $M'_{t-1} \geq M_t$, then $M'_t \geq M_{t+1}$. $\square$

We are now ready to prove Lemma 6.3.

Proof of Lemma 6.3. Observe that if $Y^{\downarrow} \geq k$, there must exist a sequence $\ell_1 < \ell_2 < \cdots < \ell_k$ such that $(\ell_k, \ell_{k-1}, \dots, \ell_1)$ is a decreasing subsequence in Σ and there does not exist $\ell_0 < \ell_1$ such that ℓ_0 can be added to the sequence to make a longer decreasing subsequence. Let $\mathbf{l} = \{\ell_1 < \ell_2 < \cdots < \ell_k\}$ and $\mathbf{h} = \{h_1 > h_2 > \cdots > h_k\}$. Let $\mathcal{A}_{\mathbf{l},\mathbf{h}}$ denote the event that $(\ell_k, \ell_{k-1}, \dots, \ell_1)$ is a decreasing subsequence of Σ satisfying the above property that

ℓ_1 is as small as possible, and $Z_{\ell_i} = h_i$ for all i. Clearly

$$\mathbb{P}\big[Y^{\downarrow} \geq k\big] = \sum_{\mathbf{l},\mathbf{h}} \mathbb{P}[\mathcal{A}_{\mathbf{l},\mathbf{h}}]. \tag{9}$$

Now, it is easy to observe using Lemma 6.4 that

$$\mathcal{A}_{\mathbf{l},\mathbf{h}} \subseteq \mathcal{E}_{\ell_1} \cap \mathcal{F}_{\mathbf{h}} \cap \mathcal{D}_{h_1,h_k,\mathbf{l}} \tag{10}$$

where

$$\mathcal{E}_{\ell_1} = \Big\{\min_{t\in[\ell_1-1]} M_t > 0\Big\};$$

$$\mathcal{F}_{\mathbf{h}} = \big\{\forall i \in [k], Z_{\ell_i} = h_i\big\};$$

$$\mathcal{D}_{h_1,h_k,\mathbf{l}} = \Big\{M_{\ell_1} = h_1 - 1; \forall i \geq 2, M_{\ell_i} = M_{\ell_i - 1} - 1; \min_{\ell_1 \leq t < \ell_k} M_t > h_k - 1\Big\}.$$

Let M'_t be the chain started at $h_1 - 1$ so that $M'_0 = h_1 - 1$, $M'_t = \max\{Z'_t, M'_{t-1}\} - 1$ and $\{Z'_i\}$ is the sequence of geometrics restricted to $\{Z_i\}_{i=\ell_1+1}^{\infty}$ omitting the sequence $\{Z_{\ell_i}\}_{i=2}^{k}$. Let $\mathcal{G}_{h_1,h_k,\mathbf{l}}$ denote the event that

$$\min_{t\in[\ell_k-\ell_1-k+1]} M'_t > h_k - 1.$$

By Lemma 6.5,

$$\mathcal{D}_{h_1,h_k,\mathbf{l}} \subseteq \mathcal{G}_{h_1,h_k,\mathbf{l}}. \tag{11}$$

Now observe that $\mathcal{E}_{\ell_1}$, $\mathcal{F}_{\mathbf{h}}$ and $\mathcal{G}_{h_1,h_k,\mathbf{l}}$ are independent. Combining equations (10) and (11), we have that

$$\mathbb{P}[\mathcal{A}_{\mathbf{l},\mathbf{h}}] \leq \mathbb{P}_0\big[R_0^+ > \ell_1 - 1\big]\mathbb{P}_{h_1-1}\big[R_{h_k-1}^+ > \ell_k - \ell_1 - k + 1\big]\big(q^{-1}(1-q)\big)^k q^{\sum_i h_i}.$$

Using Proposition 4.6 we have

$$\mathbb{P}_0\big[R_0^+ > \ell_1 - 1\big] \leq Ae^{-c(\ell_1-1)}$$

and

$$\mathbb{P}_{h_1-1}\big[R_{h_k-1}^+ > \ell_k - \ell_1 - k + 1\big] \leq Ae^{-c(\max\{\ell_k-\ell_1-k-10(h_1-h_k),0\})}.$$

Observe that $\sum_{i=2}^k h_i \geq k(k-1)/2 = k^2/2(1+o(1))$. Now we split the sum over $\mathbf{l}$ and $\mathbf{h}$ in the right hand side of (9) into a few cases. Let $\mathcal{C}_1$ denotes the set of all $\mathbf{l}$ such that $\ell_k \leq k^{3/2}$. Then we have

$$\begin{aligned}\sum_{\mathbf{l}\in\mathcal{C}_1,\mathbf{h}} \mathbb{P}[A_{\mathbf{l},\mathbf{h}}] &\leq q^{k^2/2(1+o(1))} \sum_{h_1} \binom{k^{3/2}}{k}\binom{h_1}{k} q^{h_1} \\ &\leq q^{k^2/2(1+o(1))} \sum_{h\geq k} h^k q^h = q^{k^2/2(1+o(1))}.\end{aligned}$$

Let $\mathcal{C}_2$ denote the set of all $\mathbf{l}$ such that $\ell_k > k^{3/2}$ and $\ell_1 > \ell_k/2$. Then we have

$$\begin{aligned}\sum_{\mathbf{l}\in\mathcal{C}_2,\mathbf{h}} \mathbb{P}[A_{\mathbf{l},\mathbf{h}}] &\leq q^{k^2/2(1+o(1))} \sum_{\ell_k \geq k^{3/2}} \sum_{h_1} \binom{\ell_k}{k}\binom{h_1}{k} e^{-c\ell_k/3} q^{h_1} \\ &\leq q^{k^2/2(1+o(1))} \sum_{\ell\geq k^{3/2}} \sum_{h\geq k} \ell^k e^{-c\ell/3} h^k q^h = q^{k^2/2(1+o(1))}.\end{aligned}$$

To aid the reader attempting to verify the calculations, we note that above, as well as in the following estimate, we have not attempted to optimize the constant in the exponent of the bound.

Let $\mathcal{C}_3$ denote all the pairs $(\mathbf{l},\mathbf{h})$ such that $\mathbf{l}\notin\mathcal{C}_1\cup\mathcal{C}_2$ and $h_1<\ell_k/200$. Then we have

$$\begin{aligned}\sum_{(\mathbf{l},\mathbf{h})\in\mathcal{C}_3}\mathbb{P}[A_{\mathbf{l},\mathbf{h}}] &\le q^{k^2/2(1+o(1))}\sum_{\ell_k\ge k^{3/2}}\sum_{h_1}\binom{\ell_k}{k}\binom{h_1}{k}q^{h_1}e^{-c\ell_k/10}\\ &\le q^{k^2/2(1+o(1))}\sum_{\ell\ge k^{3/2}}\sum_{h\ge k}\ell^k e^{-c\ell/10}h^k q^h\\ &= q^{k^2/2(1+o(1))}.\end{aligned}$$

Finally let $\mathcal{C}_4$ denote all the pairs $(\mathbf{l},\mathbf{h})$ such that $\mathbf{l}\notin\mathcal{C}_1\cup\mathcal{C}_2$ and $h_1\ge\ell_k/200$. In this case we have

$$\begin{aligned}\sum_{(\mathbf{l},\mathbf{h})\in\mathcal{C}_4}\mathbb{P}[A_{\mathbf{l},\mathbf{h}}] &\le q^{k^2/2(1+o(1))}\sum_{h_1}\sum_{\ell_k\le 200h_1}\binom{\ell_k}{k}\binom{h_1}{k}q^{h_1}\\ &\le q^{k^2/2(1+o(1))}\sum_{h\ge 200k^{3/2}}(200h)^{k+1}h^k q^h\\ &= q^{k^2/2(1+o(1))}.\end{aligned}$$

Combining these four cases we complete the proof of the lemma. □

7. Concluding remarks and open questions

In this paper, based on a regenerative representation of the Mallows process and analysis of an associated Markov chain, we established some limit theorems for the lengths of longest increasing and decreasing subsequences of a Mallows(q) permutation for a fixed $q\in(0,1)$. Many interesting open questions remain. We conclude with a discussion of a few of them.

1. For which regime of q is the limiting distribution of L_n Tracy–Widom? If $q\to 1$ sufficiently fast as $n\to\infty$ the limiting distribution is Tracy–Widom, but how fast does $(1-q)$ need to decay for this conclusion to hold? Does there exist a range of q where the limiting distribution is neither Gaussian nor Tracy–Widom?

   Remark 2. Let us parameterize $q=1-\delta$. For $\delta=o(n^{-2})$ it is straightforward to couple a Mallows(1) permutation and a Mallows$(1-\delta)$ permutation to agree with high probability so that the total variation distance goes to 0 as $n\to\infty$. Clearly, in this case L_n has the Tracy–Widom distribution when scaled appropriately. Our observation is that it is possible to improve the bound for the regime with Tracy–Widom limit to $\delta=o(n^{-4/3})$. For $q=1-o(n^{-4/3})$, using Lemma 4.2 of [31], it is possible to stochastically sandwich $L_n(q)$, between the length of LIS of two uniform random permutations of sizes $N_1(n)$ and $N_2(n)$, where N_1 and N_2 are such that both these when centered by $2\sqrt{n}$ and scaled by $n^{1/6}$ converges weakly to Tracy–Widom distribution.

2. How does the variance of L_n grow for different rates of $q\to 1$? There is a general linear upper bound on variance available from [9]. We expect the variance to go from linear in n to the order of $n^{1/3}$ as $q\to 1$, but it would be interesting to understand the dependence on q.
3. Can one prove a law of large numbers for $L_n^{\downarrow}$ for some range of q going to one? It is shown in [9] that $\mathbb{E}L_n^{\downarrow}(q)=\Theta(\sqrt{\log n/\log q^{-1}})$ for $q\to 1$ sufficiently slowly, but showing the existence and identification of a limiting constant remains open.

Acknowledgements

The authors are grateful to Ron Peled for many helpful conversations and letting us know of the results in [20] and [16]. We thank Allan Sly for useful discussions and Shannon Starr for explaining some of the details of the results in [31]. We also thank two anonymous referees for an extremely careful reading of the manuscript and many useful comments and suggestions that helped improve the paper. We also thank one of the anonymous referees for pointing out how the expression for the stationary distribution of the Markov chain we used can be simplified (see Remark 1).

References

[1] D. J. Aldous and P. Diaconis. Hammersley's interacting particle process and longest increasing subsequences. *Probab. Theory Related Fields* **103** (1995) 199–213. MR1355056
[2] D. J. Aldous and J. Fill. *Reversible Markov Chains and Random Walks on Graphs*. Available at https://www.stat.berkeley.edu/~aldous/RWG/book.pdf.
[3] F. J. Anscombe. Large-sample theory of sequential estimation. *Math. Proc. Cambridge Philos. Soc.* **48** (1952) 600–607. MR0051486
[4] J. Baik, G. Ben Arous and S. Péché. Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices. *Ann. Probab.* **33** (2005) 1643–1697. MR2165575
[5] 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) 1119–1178. MR1682248
[6] J. Baik, E. M. Rains. The asymptotics of monotone subsequences of involutions. *Duke Math. J.* **109** (2) (2001) 205–281. MR1845180
[7] G. Ben Arous, I. Corwin. Current fluctuations for TASEP: A proof of the Prähofer–Spohn conjecture. *Ann. Probab.* **39** (1) (2011) 104–138. MR2778798
[8] I. Benjamini, N. Berger, C. Hoffman and E. Mossel. Mixing times of the biased card shuffling and the asymmetric exclusion process. *Trans. Amer. Math. Soc.* **357** (8) (2005) 3013–3029. MR2135733
[9] N. Bhatnagar and R. Peled. Lengths of monotone subsequences in a Mallows permutation. *Probab. Theory Related Fields* **161** (3–4) (2015) 719–780. MR3334280
[10] A. Borodin, P. Diaconis and J. Fulman. On adding a list of numbers (and other one-dependent determinantal processes). *Bull. Amer. Math. Soc.* **47** (4) (2010) 639–670. MR2721041
[11] P. Caputo. Energy gap estimates in XXZ ferromagnets and stochastic particle systems. *Markov Process. Related Fields* **11** (2005) 189–210. MR2150138
[12] S. Chaterjee and P. S. Dey. Central limit theorem for first-passage percolation time across thin cylinders. *Probab. Theory Related Fields* **156** (2013) 613–663. MR3078282
[13] D. Critchlow. *Metric Methods for Analyzing Partially Ranked Data*. *Lecture Notes in Statistics* **34**. Springer, Berlin, 1985. MR0818986
[14] J.-D. Deuschel and O. Zeitouni. Limiting curves for i.i.d. records. *Ann. Probab.* **23** (1995) 852–878. MR1334175
[15] J.-D. Deuschel and O. Zeitouni. On increasing subsequences of i.i.d. samples. *Combin. Probab. Comput.* **8** (3) (1999) 247–263. MR1702546
[16] P. Dey, M. Joseph and R. Peled. Longest increasing path within the critical strip. Preprint, 2016.
[17] P. Diaconis. *Group Representations in Probability and Statistics*. *Lecture Notes – Monograph Series* **11**. Institute of Mathematical Statistics, Hayward, CA, 1988. MR0964069
[18] P. Diaconis and A. Ram. Analysis of systematic scan Metropolis algorithms using Iwahori–Hecke algebra techniques. *Michigan Math. J.* **48** (1) (2000) 157–190. MR1786485
[19] M. Figner and J. Verducci. *Probability Models and Statistical Analyses for Ranking Data*. *Lecture Notes in Statistics* **80**. Springer, New York, 1993. MR1237197
[20] A. Gladkich and R. Peled. On the cycle structure of random Mallows permutation. Preprint, 2016. Available at https://arxiv.org/abs/1601.06991.
[21] A. Gnedin and G. Olshanski. q-Exchangeability via quasi-invariance. *Ann. Probab.* **38** (2010) 2103–2135. MR2683626
[22] A. Gnedin and G. Olshanski. The two-sided infinite extension of the Mallows model for random permutations. *Adv. in Appl. Math.* **48** (5) (2012) 615–639. MR2920835
[23] A. Gut. *Probability: A Graduate Course*. Springer, Berlin, 2006. MR2977961
[24] C. Houdré and Ü. Işlak. A central limit theorem for the length of the longest common subsequence in random words. Preprint, 2014. Available at arXiv:1408.1559.
[25] I. M. Johnstone. On the distribution of the largest eigenvalue in principal components analysis. *Ann. Statist.* **29** (2001) 295–327. MR1863961
[26] D. E. Knuth. Permutations, matrices and generalized Young tableaux. *Pacific J. Math.* **34** (3) (1970) 709–727. MR0272654
[27] G. Lebanon and J. Lafferty. Cranking: Combining rankings using conditional probability models on permutations. In *Proceedings of the Nineteenth International Conference on Machine Learning (ICML'02)* 363–370. Morgan Kaufmann Publishers Inc., San Francisco, CA, 2002.
[28] B. F. Logan and L. A. Shepp. A variational problem for random Young tableaux. *Adv. Math.* **26** (1977) 206–222. MR1417317
[29] C. L. Mallows. Non-null ranking models I. *Biometrika* **44** (1–2) (1957) 114–130. MR0087267
[30] J. Marden. *Analyzing and Modeling Ranking Data*. *Monographs on Statistics and Applied Probability* **64**. Chapman & Hall, London, 1993. MR1346107
[31] C. Mueller and S. Starr. The length of the longest increasing subsequence of a random Mallows permutation. *J. Theoret. Probab.* **26** (2013) 514–540. MR3055815

[32] S. Mukherjee. Estimation of parameters in non-uniform models on permutations. Preprint, 2013. Available at arXiv:1307.0978.
[33] G. de B. Robinson. On representations of the symmetric group. *Amer. J. Math.* **60** (1938) 745–760. MR1507943
[34] C. Schensted. Longest increasing and decreasing subsequences. *Canad. J. Math.* **13** (1961) 179–191. MR0121305
[35] R. Serfozo. *Basics of Applied Stochastic Processes*. Springer, Berlin, 2009. MR2484222
[36] S. Starr. Thermodynamic limit for the mallows model on S_n. *J. Math. Phys.* **50** (9) (2009) 095208. MR2566888
[37] S. Starr and M. Walters. Phase uniqueness for the Mallows measure on permutations. Preprint, 2015. Available at http://arxiv.org/abs/1502.03727v3.
[38] A. M. Vershik and S. V. Kerov. Asymptotics of the Plancherel measure of the symmetric group and the limiting form of Young tableaux. *Dokl. Akad. Nauk SSSR (Sov. Math. Dokl.)* **233** (6) (1977) 1024–1027. MR0480398