Cycles in Mallows random permutations

He, Jimmy; Müller, Tobias; Verstraaten, Teun W.

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2023

[Link to publication in University of Groningen/UMCG research database](#)

WILEY

Cycles in Mallows random permutations

Jimmy He[1] | Tobias Müller[2] | Teun W. Verstraaten[2]

[1]Department of Mathematics, MIT, Cambridge, Massachusetts, USA
[2]Bernoulli Institute, Groningen University, Groningen, The Netherlands

Correspondence
Tobias Müller, Bernoulli Institute, Groningen University, Groningen, The Netherlands.
Email: tobias.muller@rug.nl

Abstract

We study cycle counts in permutations of $1, \ldots, n$ drawn at random according to the Mallows distribution. Under this distribution, each permutation $\pi \in S_n$ is selected with probability proportional to $q^{\text{inv}(\pi)}$, where $q > 0$ is a parameter and $\text{inv}(\pi)$ denotes the number of inversions of π. For ℓ fixed, we study the vector $(C_1(\Pi_n), \ldots, C_\ell(\Pi_n))$ where $C_i(\pi)$ denotes the number of cycles of length i in π and Π_n is sampled according to the Mallows distribution. When $q = 1$ the Mallows distribution simply samples a permutation of $1, \ldots, n$ uniformly at random. A classical result going back to Kolchin and Goncharoff states that in this case, the vector of cycle counts tends in distribution to a vector of independent Poisson random variables, with means $1, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{\ell}$. Here we show that if $0 < q < 1$ is fixed and $n \to \infty$ then there are positive constants m_i such that each $C_i(\Pi_n)$ has mean $(1 + o(1)) \cdot m_i \cdot n$ and the vector of cycle counts can be suitably rescaled to tend to a joint Gaussian distribution. Our results also show that when $q > 1$ there is a striking difference between the behavior of the even and the odd cycles. The even cycle counts still have linear means, and when properly rescaled tend to a multivariate Gaussian distribution. For the odd cycle counts on the other hand, the limiting behavior depends on the parity of n when $q > 1$. Both $(C_1(\Pi_{2n}), C_3(\Pi_{2n}), \ldots)$ and $(C_1(\Pi_{2n+1}), C_3(\Pi_{2n+1}), \ldots)$ have discrete limiting distributions—they do not need to be renormalized—but the two limiting distributions are distinct for all $q > 1$. We describe these limiting distributions in terms of Gnedin and Olshanski's bi-infinite extension of the Mallows model. We investigate these limiting distributions further, and study

the behavior of the constants involved in the Gaussian limit laws. We for example show that as $q \downarrow 1$ the expected number of 1-cycles tends to $1/2$—which, curiously, differs from the value corresponding to $q = 1$. In addition we exhibit an interesting "oscillating" behavior in the limiting probability measures for $q > 1$ and n odd versus n even.

KEYWORDS

random permutations, mallows distribution, cycle counts

1 | INTRODUCTION AND STATEMENT OF MAIN RESULTS

Let S_n denote the set of permutations of $[n] := \{1, \ldots, n\}$. For a permutation $\pi \in S_n$ the ordered pair $(i,j) \in [n]^2$ is an inversion of π if $i < j$ and $\pi(i) > \pi(j)$. We denote the number of inversions of a permutation π by $\text{inv}(\pi)$. For $n \in \mathbb{N}$ and $q > 0$, the Mallows distribution Mallows(n, q) samples a random element Π_n of S_n in such a way that each $\pi \in S_n$ has probability proportional to $q^{\text{inv}(\pi)}$. That is,

$$\mathbb{P}(\Pi_n = \pi) = \frac{q^{\text{inv}(\pi)}}{\sum_{\sigma \in S_n} q^{\text{inv}(\sigma)}}. \tag{1}$$

This distribution on S_n was introduced in the late fifties by Mallows [27] in the context of statistical ranking models. It has since been studied in connection with a diverse range of topics, including mixing times of Markov chains [4, 11], finitely dependent colorings of the integers [22], stable matchings [2], random binary search trees [1], learning theory [8, 34], q-analogs of exchangeability [16, 17], determinantal point processes [7], statistical physics [32, 33], and genomics [12].

Aspects of the Mallows distribution that have been studied include the longest increasing subsequence [3, 5, 28], longest common subsequences [24], pattern avoidance [9, 10, 29], the number of descents [20], and the cycle structure [15].

In the special case when $q = 1$ the Mallows distribution coincides with the uniform distribution on S_n. A classical result going back to Gontcharoff [18] and Kolchin [25] states that in this case, for every fixed ℓ:

$$(C_1(\Pi_n), \ldots, C_\ell(\Pi_n)) \xrightarrow[n \to \infty]{d} (X_1, \ldots, X_\ell),$$

where $C_i(\pi)$ denotes the number of cycles of length i in the permutation π, and $X_1, \ldots, X_\ell$ are independent and X_i is Poisson distributed with mean $1/i$ for each $i = 1, \ldots, \ell$. In spite of the long history and considerable attention received by the Mallows distribution, until very recently the problem of determining analogues of this result for the Mallows(n, q) distribution with $q \neq 1$ seems to have escaped attention. In a recent paper, Gladkich and Peled [15] studied the cycle structure of the Mallows distribution when $q = q(n)$ depends on n and approaches one as $n \to \infty$. Here we will focus instead on the limiting distribution of the cycle counts when $q \neq 1$ is fixed and n tends to infinity.

Our first result shows that for $0 < q < 1$ each $C_i(\Pi_n)$ has a mean that is linear in n, and that for every fixed ℓ, the vector $(C_1(\Pi_n), \ldots, C_\ell(\Pi_n))$ can be suitably rescaled so that it tends to a jointly normal limiting distribution.

Theorem 1.1. *Fix* $0 < q < 1$ *and let* $\Pi_n \sim \mathrm{Mallows}(n, q)$. *There exist positive constants* $m_1, m_2, \ldots$ *and an infinite matrix* $P \in \mathbb{R}^{\mathbb{N} \times \mathbb{N}}$ *such that for all* $\ell \geq 1$ *we have*

$$\frac{1}{\sqrt{n}} \left(C_1(\Pi_n) - m_1 n, \ldots, C_\ell(\Pi_n) - m_\ell n\right) \xrightarrow[n \to \infty]{d} \mathcal{N}_\ell(\underline{0}, P_\ell),$$

where $\mathcal{N}_\ell(\cdot, \cdot)$ *denotes the* ℓ-*dimensional multivariate normal distribution and* P_ℓ *is the submatrix of* P *on the indices* $[\ell] \times [\ell]$.

As it happens, for $q > 1$, there is a major difference between the behavior of even cycles and odd cycles. For even cycle counts we have a result analogous to the previous theorem.

Theorem 1.2. *Fix* $q > 1$ *and let* $\Pi_n \sim \mathrm{Mallows}(n, q)$. *There exists positive constants* $\mu_2, \mu_4, \ldots$ *and an infinite matrix* $Q \in \mathbb{R}^{\mathbb{N} \times \mathbb{N}}$ *such that for all* $\ell \geq 1$ *we have*

$$\frac{1}{\sqrt{n}}(C_2(\Pi_n) - \mu_2 n, \ldots, C_{2\ell}(\Pi_n) - \mu_{2\ell} n) \xrightarrow[n \to \infty]{d} \mathcal{N}_\ell(\underline{0}, Q_\ell),$$

where $\mathcal{N}_\ell(\cdot, \cdot)$ *denotes the* ℓ-*dimensional multivariate normal distribution and* Q_ℓ *is the submatrix of* Q *on the indices* $[\ell] \times [\ell]$.

We will describe the limiting distributions for odd cycles in the case when $q > 1$ in terms of the bi-infinite analogue of the Mallows distribution that was introduced by Gnedin and Olshanski [17]. This is a random bijection $\Sigma : \mathbb{Z} \to \mathbb{Z}$, whose distribution we'll denote by $\mathrm{Mallows}(\mathbb{Z}, q)$. See Section 2 for more discussion and relevant facts.

Throughout the article r, ρ denote the bijections of $\mathbb{Z}$ defined by $r(i) := -i$ and $\rho(i) := 1 - i$.

Theorem 1.3. *Let* $q > 1$ *and* $\Pi_n \sim \mathrm{Mallows}(n, q)$ *and* $\Sigma \sim \mathrm{Mallows}(\mathbb{Z}, 1/q)$. *We have*

$$(C_1(\Pi_{2n+1}), C_3(\Pi_{2n+1}), \ldots) \xrightarrow[n \to \infty]{d} (C_1(r \circ \Sigma), C_3(r \circ \Sigma), \ldots),$$

and

$$(C_1(\Pi_{2n}), C_3(\Pi_{2n}), \ldots) \xrightarrow[n \to \infty]{d} (C_1(\rho \circ \Sigma), C_3(\rho \circ \Sigma), \ldots).$$

Moreover, the two limiting distributions above are distinct for all $q > 1$. *The permutations* $r \circ \Sigma$ *and* $\rho \circ \Sigma$ *almost surely have only finitely many odd cycles.*

Next, we study the properties of the constants $m_1, m_2, \ldots$ occurring in Theorem 1.1. The first part of the next result gives an interpretation of these constants in terms of the $\mathrm{Mallows}(\mathbb{Z}, q)$ distribution.

Theorem 1.4. *Let* $0 < q < 1$ *and* $\Sigma \sim \mathrm{Mallows}(\mathbb{Z}, q)$, *and let* $m_1, m_2, \ldots$ *be as provided by Theorem 1.1.*

(i) For $i = 1, 2, \ldots$ *we have*

$$m_i = (1/i) \cdot \mathbb{P}(0 \text{ lies in an } i\text{-cycle of } \Sigma).$$

In particular

$$m_1 = \mathbb{P}(\Sigma(0) = 0).$$

(ii) $\sum_{i=1}^{\infty} i \cdot m_i = 1$.

(iii) We have

$$(m_1, m_2, m_3, \dots) \to (1, 0, 0, \dots) \qquad \text{as } q \downarrow 0,$$
$$(m_1, m_2, m_3, \dots) \to (0, 0, 0, \dots) \qquad \text{as } q \uparrow 1,$$

where the convergence can be taken with respect to the L_1-norm. Moreover,

$$m_1 = 1 - 2q + o(q) \qquad \text{as } q \downarrow 0,$$
$$m_1 = \frac{1-q}{4} + o(1-q) \qquad \text{as } q \uparrow 1.$$

Combining Part (i) of the above theorem with Theorem 5.1 of [17], we can write

$$m_1 = {}_0\phi_1(-; q; q, q^3) \cdot (1-q) \cdot \prod_{i=1}^{\infty} (1 - q^i),$$

where ${}_r\phi_s$ denotes the q-hypergeometric function—see [14] for the definition and background. An alternative expression, based on the work of Gladkich and Peled [15], for m_1 is given in Lemma 6.1 below. Figure 1 shows a plot of m_1 versus q together with the results of computer simulations.

The next result provides similar results for the constants appearing in Theorem 1.2.

Theorem 1.5. *Let $q > 1$, let $\mu_2, \mu_4, \dots$ be as provided by Theorem 1.2, and let $\Sigma, \Sigma' \sim \text{Mallows}(\mathbb{Z}, 1/q)$ be independent.*

(i) For $i = 1, 2, \dots$ we have

$$\mu_{2i} = \frac{1}{2i} \cdot \mathbb{P}(0 \text{ is in an } i\text{-cycle of } \Sigma' \circ \Sigma),$$

and in particular

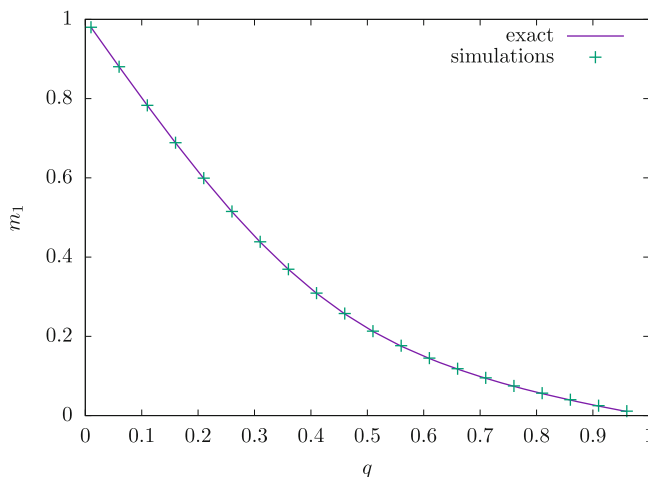$$\mu_2 = \frac{1}{2} \cdot \sum_{i \in \mathbb{Z}} \mathbb{P}[\Sigma(0) = i]^2.$$

FIGURE 1 A plot of m_1 versus q. The crosses correspond to the average number of 1-cycles in 10 000 samples of the Mallows(1000, q) distribution.
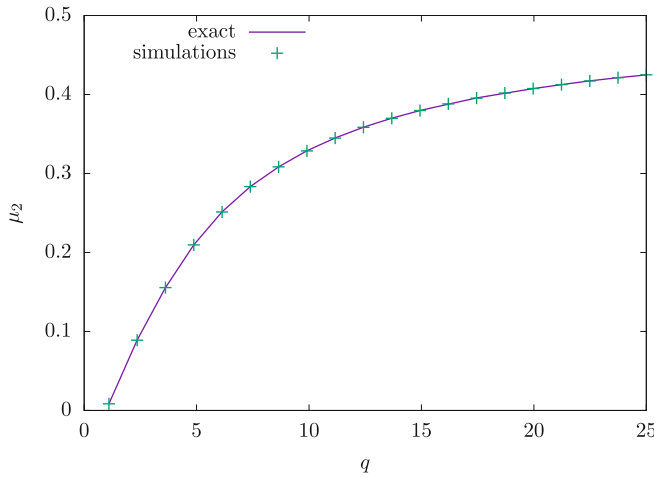
FIGURE 2 A plot of μ_2 versus q. The simulations were done sampling a Mallows$(1000, q)$ distribution 10 000 times, and taking the average number of 2-cycles.

(ii) $\sum_{i=1}^{\infty} 2i \cdot \mu_{2i} = 1$.

(iii) We have

$$(\mu_2, \mu_4, \mu_6, \ldots) \to (0, 0, 0, \ldots) \qquad \text{as } q \downarrow 1,$$
$$(\mu_2, \mu_4, \mu_6, \ldots) \to (1/2, 0, 0, \ldots) \qquad \text{as } q \to \infty,$$

where the convergence can be taken with respect to the L_1-norm.

Again, combining Part (i) with Theorem 5.1 in [17] gives an expression for μ_2 as an explicit function of q. Figure 2 provides a plot of μ_2 versus q together with the results of computer simulations. We mention that Pitman and Tang [30, Proposition 3.3] give a result for so-called regenerative random permutations, that is closely related to parts (i) and (ii) of Theorems 1.4 and 1.5.

Next we provide some results on the asymptotic expected number of 1-cycles when $q > 1$. We remind the reader that a "1-cycle" is the same as a fixed point. For notational convenience let us write

$$c_e := \mathbb{E} C_1(\rho \circ \Sigma), \quad c_o := \mathbb{E} C_1(r \circ \Sigma), \tag{2}$$

where again $\Sigma \sim \text{Mallows}(\mathbb{Z}, 1/q)$ and r, ρ are given by $r(i) = -i, \rho(i) = 1 - i$.

Theorem 1.6. *Let $q > 1$ and c_e, c_o as given by (2) and $\Sigma \sim \text{Mallows}(\mathbb{Z}, 1/q)$.*

(i) We have

$$c_e = \mathbb{P}[\Sigma(0) \text{ odd}] \qquad \text{and} \qquad c_o = \mathbb{P}[\Sigma(0) \text{ even}].$$

(ii) We have

$$\lim_{q \downarrow 1} c_e = \lim_{q \downarrow 1} c_o = \frac{1}{2},$$
$$\lim_{q \to \infty} c_e = 0,$$
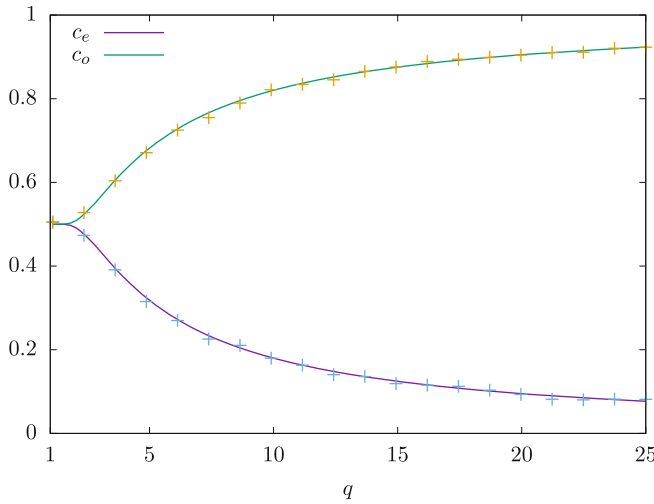$$\lim_{q \to \infty} c_o = 1.$$

FIGURE 3   The graph of c_o and c_e for $q > 1$. Simulations were done for $n = 1000$ and $n = 1001$, each sampled 10 000 times.

Moreover, as $q \rightarrow \infty$ we have

$$c_o = 1 - 2/q + o(1/q) \quad \text{and} \quad c_e = 2/q + o(1/q) \quad \text{as} \quad q \rightarrow \infty.$$

We note that Part (i) of the above theorem gives that in particular

$$c_e + c_o = 1,$$

for all $q > 1$. Again, Theorem 5.1 in [17] allows us to convert the probabilities given in Part (i) of the above theorem into explicit functions of q. Plots of c_e and c_o as a function of q, together with the results of computer simulations are shown in Figure 3.

As mentioned previously, when $q = 1$ we retrieve the uniform distribution on S_n, for which the expected number of 1-cycles equals one. So the fact that the limits for $q \downarrow 1$ of c_e, c_o equal $1/2$ is pretty curious. Of course there is no contradiction, since our results apply to the situation where $q > 1$ is fixed and we send n to infinity. Our results however do suggest something interesting must be going on in the "phase change" when $q = q(n)$ is a function of n that approaches one from above as $n \rightarrow \infty$.

Our final (main) result highlights an interesting "oscillating" behavior in the probability measures corresponding to the limit of $C_1(\Pi_{2n})$, respectively $C_1(\Pi_{2n+1})$, when $q > 1$. The probability that Π_n has at least m one cycles is a lot larger when the parities of m and n agree than when they do not (for m large but fixed and $n \rightarrow \infty$).

Theorem 1.7. *For $0 < q < 1$ and $\Sigma \sim \text{Mallows}(\mathbb{Z}, q)$ we have, as $k \rightarrow \infty$*

$$\mathbb{P}\left(C_1(\rho \circ \Sigma) \geq 2k\right) \ll \mathbb{P}\left(C_1(r \circ \Sigma) \geq 2k\right),$$
$$\mathbb{P}\left(C_1(\rho \circ \Sigma) \geq 2k + 1\right) \gg \mathbb{P}\left(C_1(r \circ \Sigma) \geq 2k + 1\right).$$

(The notation $g(k) \ll f(k)$ means that the ratio $g(k)/f(k)$ tends to zero as $k \rightarrow \infty$.)

Sketches of some ideas used in the proofs

The proofs of Theorems 1.1 and 1.2 are adaptations of a proof technique developed by Basu and Bhatnagar [3] to prove a Gaussian limit law for the length of the longest monotone subsequence of a Mallows permutation, and in fact Theorem 1.1 closely follows the original proof. The intuition behind it is that if $\Pi_n \sim \text{Mallows}(n, q)$ with $0 < q < 1$ then given that $\Pi_n[\{1, \dots, j\}] = \{1, \dots, j\}$ the remainder of the permutation behaves like a Mallows random permutation of length $n - j$. As it turns out, there will typically be linearly many such j.

A very rough sketch of the argument giving Theorem 1.1 is as follows. If $T_1 < \cdots < T_k = n$ are such that $\Pi_n[\{1, \dots, T_i\}] = \{1, \dots, T_i\}$ then each cycle must lie completely in $\{T_{i-1} + 1, \dots, T_i\}$ for some i (setting $T_0 = 0$). This allows us to show that the cycle counts behave approximately like a stopped two-dimensional random walk. This refers to the situation where $(X_1, Y_1), (X_2, Y_2), \dots$ are i.i.d. and we are interested in $\sum_{i=1}^{\tau} Y_i$ where $\tau := \inf\{k : X_1 + \cdots + X_k > n\}$. Here the X_i correspond to $T_i - T_{i-1}$ and Y_i counts the number of cycles contained in the interval $\{T_{i-1} + 1, \dots, T_i\}$. A convenient result of Gut and Janson [19] allows us to derive that the mentioned sum is approximately Gaussian after suitable rescaling. The same argument applies to arbitrary linear combinations of the cycle counts, so that we can employ the Cramer-Wold device to deduce that the—suitably rescaled—vector of cycle counts is multivariate Gaussian.

The proof of Theorem 1.2 goes along the same lines. Now it turns out that there are linearly many $T_1 < \cdots < T_k$ such that $\Pi_n[\{1, \dots, T_i\}] = \{n + 1 - T_i, \dots, n\}$ and $\Pi_n[\{n + 1 - T_i, \dots, n\}] = \{1, \dots, T_i\}$. (Almost) every even cycle must then be contained in some set $\{T_{i-1} + 1, \dots, T_i\} \cup \{n + 1 - T_i, \dots, n - T_{i-1}\}$, and we can adapt the proof strategy that gave Theorem 1.1 to work also here. We mention that Theorem 1.2 can also be proved by starting from the center of the permutation, rather than the sides, see [21].

In the proof of Theorem 1.3, we rely on results of Gnedin and Olshanski [17] that show that "locally," for $0 < q < 1$, the finite Mallows permutation resembles the bi-infinite Mallows permutation Σ defined and analyzed in [17]. An elementary, but crucial, observation is that $\Pi_n \overset{d}{=} \text{Mallows}(n, q)$ if and only if $r_n \circ \Pi_n \overset{d}{=} \text{Mallows}(n, 1/q)$ where $r_n(i) := n + 1 - i$. (See the next section for the explanation.) Note that if n is odd then r_n leaves $(n + 1)/2$ invariant, but when n is even no element of $\{1, \dots, n\}$ is invariant ($n/2$ and $n/2 + 1$ are flipped). For $q > 1$, the relation with the $\text{Mallows}(n, 1/q)$ distribution translates to $\Pi_n \sim \text{Mallows}(n, q)$ being "approximated" by $r \circ \Sigma$ with $r(i) := -i$ when n is odd; and $\rho \circ \Sigma$ with $\rho(i) := 1 - i$ when n is even. In particular we for instance have that the number of 1-cycles (fixed points) of Π_n approximately behaves like the number of $i \in \mathbb{Z}$ such that $\Sigma(i) = -i$ when n is odd, and the number of $i \in \mathbb{Z}$ such that $\Sigma(i) = 1 - i$ when n is even.

For the proof of Theorem 1.4 we again use that the $\text{Mallows}(n, q)$ distribution locally looks like the bi-infinite Mallows model. The main intuition is the elementary observation that the number of i-cycles equals $1/i$ times the number of points that are in i-cycles. That $\sum i m_i = 1$ is then more or less immediate from the observation that, almost surely, all cycles of Σ have finite length. (As can for instance be seen from the regenerative nature of Mallows permutations for fixed q.) The statements about the limits as $q \downarrow 0$ and $q \uparrow 1$ can be derived by using explicit expressions for the expected number of 1-cycles that follow by combining our work with results of Gnedin and Olshanski [17] and Gladkich and Peled [15].

The idea behind the proof of Theorem 1.5 is very similar, but more technical. When $q > 1$ then the $\text{Mallows}(n, q)$ model is approximated well locally by the composition of two independent bi-infinite Mallows models.

The first part of Theorem 1.6 will follow from the aforementioned fact that, when $q > 1$, the expected number of 1-cycles is well-approximated by the number of $i \in \mathbb{Z}$ such that $\Sigma(i) = -i$ when

n is odd, and the number of $i \in \mathbb{Z}$ such that $\Sigma(i) = 1 - i$ when n is even. The limit of $1/2$ for the expected number of 1-cycles when $q \downarrow 1$ will follow from the relatively elementary observation that $1/q < \mathbb{P}(\Sigma(0) = j+1)/\mathbb{P}(\Sigma(0) = j) < q$ for all $j \in \mathbb{Z}$. For the other limits we again analyze the various explicit expressions in q.

The proof of Theorem 1.7 is technically involved, but the intuition behind it is relatively easy to explain. When k is large the "most likely" way in which $r \circ \Sigma$ will have at least $2k + 1$ fixed points is if $\Sigma(-k) = k, \ldots, \Sigma(k) = -k$, or some minor perturbation of this configuration. For $\rho \circ \Sigma$ the "most likely" way to have $2k+1$ fixed points is something like $\Sigma(-k) = 1+k, \Sigma(-k+1) = k, \ldots, \Sigma(k) = 1-k$, or some minor perturbation of that situation. However, as shown by Gnedin and Olshanski [17], Σ is almost surely balanced: the number of $i < 0$ with $\Sigma(i) \geq 0$ is finite and equals the number of $i \geq 0$ with $\Sigma(i) < 0$. This forces the existence of one more $i \in \mathbb{Z}$ with $|\Sigma(i) - i| = \Omega(k)$—which makes the probability exponentially smaller. (We remind the reader that $\Omega(k)$ denotes a quantity that is at least as large as a positive constant times k.) The intuition for $2k$ fixed points is similar.

Remark 1.8. Let us note that Theorems 1.1–1.3 all hold for a more general class of permutation statistics.

Let $w \in S_n$, and suppose that $w = w_1 w_2$ where w_1 sends $[1, i]$ to itself and $[i + 1, n]$ to itself. Say that a function $f : \cup_n S_n \to \mathbf{R}^d$ is *additive* if $f(w) = f(w_1) + f(w_2)$ for all $w = w_1 w_2$ decomposing in this way, where $f(w)$ for w a permutation on an interval $[i, j]$ is defined by shifting down the permutation to $[1, j - i + 1]$. Then Theorem 1.1 holds for any (non-trivial) additive function satisfying $|f(w)| \leq Cn^k$ for $w \in S_n$, where C and k are constants. The proof is exactly the same, with the key being that $f(w)$ decomposes into a sum of independent pieces in the same way as the number of cycles. The other assumptions are to ensure that the moments are finite, and that the variance is non-zero.

Let $w \in S_n$, and suppose that $w = w_1 w_2$, where w_1 sends $[i, n - i]$ to itself and w_2 sends $[1, i]$ to $[n - i + 1, n]$ and vice versa. Say that $f : \cup_n S_n \to \mathbf{R}^d$ is anti-additive if $f(w) = f(w_1) + f(w_2)$ for all $w = w_1 w_2$ decomposing in this way. Then Theorems 1.2 and 1.3 holds for any (non-trivial) anti-additive function satisfying $|f(w)| \leq Cn^k$ for $w \in S_n$, where C and k are constants. Again, the proof is the same, writing $f(w)$ as a sum of independent pieces.

Here, non-trivial means that f, when restricted to permutations of size n for which no non-trivial decomposition of the form $w = w_1 w_2$ exists, is non-constant. This assumption is needed to ensure that the variances β_{ii} and β'_{ii} are non-zero. Almost any reasonable function satisfies this, but note in particular that $f(w) = n$ the size of the permutation does not satisfy this.

Note that if $f(w)$ is either additive or anti-additive, then so is $f(w^{-1})$. Thus, the theorems also apply to joint statistics for a Mallows permutation and its inverse, giving another proof of Theorem 1.2 of [20].

2 | NOTATION AND PRELIMINARIES

Here we collect some notation, definitions, and results from the literature that we will use in our proofs. Throughout the article, we use $[n] := \{1, \ldots, n\}$ to denote the set consisting of the first n natural numbers, and $[a, b] := \{a, \ldots, b\}$ for $a < b$. If $f(n), g(n)$ are two functions depending on the parameter n, we will use $f(n) = o(g(n))$ to denote that $f(n)/g(n) \to 0$, we will use $f(n) = O(g(n))$ to denote that there exists a constant $C > 0$ such that $f(n) \leq C \cdot g(n)$, we will use $f(n) = \Omega(g(n))$ to denote that there exist a constant $c > 0$ such that $f(n) > c \cdot g(n)$, and $f(n) = \Theta(g(n))$ to denote

that $f(n) = O(g(n))$ and $f(n) = \Omega(g(n))$. We will use $\text{Bi}(n, p)$ to denote the binomial distribution with parameters n and p and we use $\text{Geo}(p)$ to the note the geometric distribution with parameter p. So $X \sim \text{Geo}(p)$ means that

$$\mathbb{P}(X = k) = p(1 - p)^{k-1},$$

for all $k \in \mathbb{N}$. We use $\text{TruncGeo}(n, p)$ to denote the *truncated* geometric distribution, truncated at n. That is, if $Y \sim \text{TruncGeo}(p)$ and $X \sim \text{Geo}(p)$ then

$$\mathbb{P}(Y = k) = \mathbb{P}(X = k | X \leq n) = \frac{p(1 - p)^{k-1}}{1 - (1 - p)^n} \quad (k = 1, \dots, n).$$

As is usual in the literature on the Mallows distribution, we denote by

$$Z(n, q) := \sum_{\sigma \in S_n} q^{\text{inv}(\sigma)},$$

the denominator in (1). By a standard result in enumerative combinatorics (see Corollary 1.3.13 in [31]) we have

$$Z(n, q) = \prod_{i=1}^{n} \frac{1 - q^i}{1 - q}.$$

An elementary observation is that the indices $i, j \in [n]$ form an inversion for $\pi \in S_n$ if and only if $\pi(i), \pi(j)$ form an inversion for π^{-1}. In particular

$$\text{inv}(\pi^{-1}) = \text{inv}(\pi).$$

Similarly, letting $r_n \in S_n$ denote the "reversal map" given by $r_n(i) := n + 1 - i$, we have that $i, j \in [n]$ are an inversion in π if and only if they are not an inversion in $r_n \circ \pi$. The same holds true for $\pi \circ r_n$. In other words

$$\text{inv}(r_n \circ \pi) = \text{inv}(\pi \circ r_n) = \binom{n}{2} - \text{inv}(\pi),$$

and hence also

$$\text{inv}(r_n \circ \pi \circ r_n) = \text{inv}(\pi).$$

As a direct consequence of these observations and the definition of the Mallows probability measure, we have:

Corollary 2.1. *Let $q > 1$ and $\Pi_n \sim \text{Mallows}(n, q)$ and let r_n be given by $r_n(i) = n + 1 - i$. The following hold.*

(i) $\Pi_n^{-1} \stackrel{d}{=} \Pi_n$, *and;*
(ii) $r_n \circ \Pi_n \circ r_n \stackrel{d}{=} \Pi_n$, *and;*
(iii) $r_n \circ \Pi_n \stackrel{d}{=} \text{Mallows}(n, 1/q)$, *and;*
(iv) $\Pi_n \circ r_n \stackrel{d}{=} \text{Mallows}(n, 1/q)$.

(To see the third and fourth parts of the lemma, note that $\mathbb{P}(r_n \circ \Pi_n = \pi) = \mathbb{P}(\Pi_n = r_n \circ \pi) = q^{\binom{n}{2} - \text{inv}(\pi)}/Z(n, q)$ is proportional to $(1/q)^{\text{inv}(\pi)}$, and similarly for $\Pi_n \circ r_n$.) The last two parts of the

corollary provide a way to express the Mallows distribution with $q > 1$ in terms of the Mallows distribution with $0 < q < 1$. We will rely on this a lot in our proofs of the results for $q > 1$.

For $0 < q < 1$, there is an iterative procedure for generating $\Pi_n \in \text{Mallows}(n, q)$, going back to the work of Mallows [27]. We let $Z_1, \ldots, Z_n$ be independent with $Z_i \sim \text{TruncGeo}(n + 1 - i, 1 - q)$. We now set

$$\Pi_n(1) = Z_1,$$
$$\Pi_n(i) = \text{the } Z_i\text{-th smallest number in the set } [n] \setminus \{\Pi(1), \ldots, \Pi(i-1)\}, \qquad \text{for } 1 < i \leq n.$$

Put differently, having determined $\Pi_n(1), \ldots, \Pi_n(i-1)$, we determine $\Pi_n(i)$ by writing $[n] \setminus \{\Pi_n(1), \ldots, \Pi_n(s-1)\}$ in increasing order as $\{j_1, j_2, \ldots, j_{n-s+1}\}$, and setting set $\Pi_n(i) := j_{Z_i}$. To see that this procedure indeed generates a random element of S_n chosen according to the Mallows(n, q) distribution, we can argue as follows. We first note that for each $\pi \in S_n$ there is exactly one choice of $(k_1, \ldots, k_n) \in [n] \times [n-1] \times \cdots \times [1]$ such that setting $Z_1 = k_1, \ldots, Z_n = k_n$ results in $\Pi_n = \pi$ (and vice versa each choice of $k_1, \ldots, k_n$ determines a unique element of S_n). In particular

$$\mathbb{P}(\Pi_n = \pi) = \prod_{i=1}^{n} \mathbb{P}(Z_i = k_i) \propto q^{\sum_{i=1}^{n}(k_i - 1)},$$

where the symbol $\propto$ denotes "proportional to," and hides a multiplicative term not depending on $k_1, \ldots, k_n$. We now note that for each $i \in \mathbb{N}$ we must have

$$|\{j : i < j \text{ and } \pi(i) > \pi(j)\}| = k_i - 1.$$

In other words, $\text{inv}(\pi) = k_1 + \cdots + k_n - n$, which shows that we've indeed sampled according to the Mallows(n, q)-distribution.

There is a natural extension of the Mallows model to random functions $\Pi : \mathbb{N} \to \mathbb{N}$, called the *Mallows process* by some authors. Similarly to the iterative procedure for generating $\Pi_n \sim$ Mallows(n, q) described above, we let $Z_1, Z_2, \ldots$ be an infinite sequence of i.i.d. *Geom*$(1 - q)$ random variables and we iteratively construct an infinite sequence Π of natural numbers by setting

$$\Pi(1) = Z_1,$$
$$\Pi(i) = \text{the } Z_i\text{-th smallest number in the set } \mathbb{N} \setminus \{\Pi(1), \ldots, \Pi(i-1)\}, \qquad \text{for } i > 1.$$

We denote the probability distribution of Π generated in this manner by Mallows$(\mathbb{N}, q)$ (see also Figure 4).

For a non-empty, finite $A \subseteq \mathbb{R}$ and a bijection $\pi : A \to A$ we define $\text{inv}(\pi)$ in exactly the same way as for bijections from $[n]$ to itself. The distribution of the random bijection $\Pi_A : A \to A$ satisfying

$$\mathbb{P}(\Pi_A = \pi) = \frac{q^{\text{inv}(\pi)}}{Z(n, q)},$$

will be denoted by Mallows(A, q). For $\sigma : B \to B$ a bijection with $B \subseteq \mathbb{R}$ and $A \subseteq B$ finite (but B may be infinite) we denote by σ_A the bijection we obtain by setting

$$\sigma_A(a) := a_{(i)} \text{ if } \sigma(a) \text{ is the } i\text{th smallest element of } \sigma[A],$$

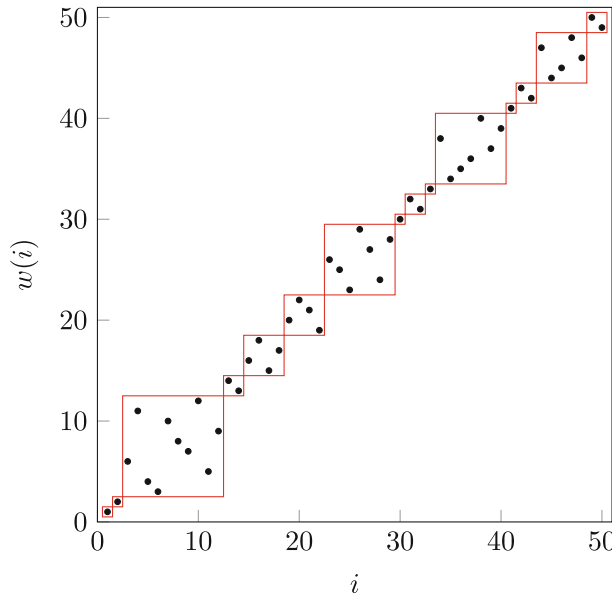where $a_{(i)}$ is the ith smallest element of A.

FIGURE 4 Sample of Mallows($\mathbb{N}, q$). The red squares indicate the first time that an interval is sent to itself, and the lengths and contents of the red squares are independent and identically distributed.

As shown by Basu and Bhatnagar [3, Lemma 2.1] and independently Crane and DeSalvo [9, Lemma 5.2], if $\Pi \sim$ Mallows($\mathbb{N}, q$) with $0 < q < 1$ and $I = \{a, \dots, a + n\} \subseteq \mathbb{N}$ is a finite "interval" of consecutive integers then

$$\Pi_I \overset{d}{=} \text{Mallows}(I, q).$$

Moreover, as is easily seen from the definitions, if $I = \{a + 1, \dots, a + n\} \subseteq \mathbb{Z}$ is such an interval then

$$\Sigma \sim \text{Mallows}(I, q) \quad \text{if and only if} \quad s^{(a)} \circ \Sigma \circ s^{(-a)} \overset{d}{=} \text{Mallows}(n, q).$$

Here and in the rest of the article s denotes the *shift map* given by $i \mapsto i + 1$ and $f^{(n)}$ denotes the n-fold composition of the function f with itself and $f^{(-n)}$ denotes the n-fold composition of the inverse f^{-1} of f with itself. (So in particular $s^{(k)}$ is the map $i \mapsto i + k$ and $s^{(-k)}$ is the map $i \mapsto i - k$.)

We next discuss Gnedin and Olshanski's bi-infinite Mallows model. For every $0 < q < 1$, this is a random bijection of $\mathbb{Z}$, whose distribution we will denote by Mallows($\mathbb{Z}, q$). The work of Gnedin and Olshanski provides several definitions, but all of them are rather involved. So we refer the reader to the original paper [17] for the precise definition and mention only the properties and relevant facts we will be using in what follows.

We will need the following notion of convergence. For a sequence $\sigma_1 : I_1 \to I_1, \sigma_2 : I_2 \to I_2, \dots$ of bijections of subsets of $\mathbb{Z}$ with the property that $\mathbb{Z} = \bigcup_n I_n$, we write $\sigma_n \to \sigma$ if for every $i \in \mathbb{Z}$ there is an $n = n(i)$ such that $n' \geq n$ implies $\sigma_n(i) = \sigma(i)$.

As in [17], we call a permutation π of $\mathbb{Z}$ balanced if

$$|\{(i, \pi(i)) : i < 1 \leq \pi(i)\}| = |\{(i, \pi(i)) : \pi(i) < 1 \leq i\}| < \infty.$$

As noted in [17], we may replace the 1 above by any fixed number in $\mathbb{Z}$ and obtain an equivalent definition of balanced permutations. The random permutation $\Sigma \sim \mathrm{Mallows}(\mathbb{Z}, q)$ is almost surely balanced.

The q-Pochhammer symbols $(a; q)_n$ and $(a; q)_\infty$ are defined as

$$(a; q)_n = \prod_{i=1}^{n}\left(1 - aq^{i-1}\right) \qquad \text{and} \qquad (a; q)_\infty = \prod_{i=1}^{\infty}\left(1 - aq^{i-1}\right),$$

Recall that, throughout the article, we will use s, r, ρ to denote the maps given by

$$s(i) := i + 1, \quad r(i) := -i, \quad \rho(i) := 1 - i.$$

The following lemma lists the facts on the $\mathrm{Mallows}(\mathbb{Z}, q)$ distribution we will be relying on in our proofs below.

Lemma 2.2. *Let $0 < q < 1$ and $\Sigma \sim \mathrm{Mallows}(\mathbb{Z}, q)$. We have*

(i) If $I \subseteq \mathbb{Z}$ is a finite set of consecutive integers then $\Sigma_I \stackrel{d}{=} \mathrm{Mallows}(I, q)$ [17, comments below Theorem 6.1];

(ii) If $I_1 \subseteq I_2 \subseteq \dots$ are finite sets of consecutive integers with $\bigcup_n I_n = \mathbb{Z}$, then $\Sigma_{I_n} \to \Sigma$ almost surely [17, Proposition 7.6];

(iii) $\Sigma^{-1} \stackrel{d}{=} \mathrm{Mallows}(\mathbb{Z}, q)$ [17, Corollary 3.4];

(iv) $s \circ \Sigma \circ s^{-1} \stackrel{d}{=} \mathrm{Mallows}(\mathbb{Z}, q)$ [17, Lemma 4.4];

(v) $\rho \circ \Sigma \circ \rho \stackrel{d}{=} \mathrm{Mallows}(\mathbb{Z}, q)$ [17, Corollary 3.5];

(vi) $\mathbb{P}[|\Sigma(0)| > m] = \Theta(q^m)$ [17, Remark 5.2];

(vii) For each $d \in \mathbb{Z}$ we have

$$\mathbb{P}[\Sigma(0) = d] = (1 - q)(q; q)_\infty \sum_{\substack{r, l \geq 0; \\ r - l = d}} \frac{q^{rl + r + l}}{(q; q)_r (q; q)_l}.$$

[17, Theorem 5.1].

We remark that it follows from Part (iv) that also

$$s^{(k)} \circ \Sigma \circ s^{(-k)} \stackrel{d}{=} \Sigma,$$

for all $k \in \mathbb{Z}$. It now also follows that

$$\Sigma(i) - i \stackrel{d}{=} \Sigma(0),$$

for all $i \in \mathbb{Z}$. Similarly, since $\rho = r \circ s^{-1} = s \circ r$, we can combine (iv) and (v) to derive that

$$r \circ \Sigma \circ r = s^{-1} \circ \rho \circ \Sigma \circ \rho \circ s \stackrel{d}{=} s^{-1} \circ \Sigma \circ s \stackrel{d}{=} \Sigma.$$

Finally, let us also remark that the $\mathrm{Mallows}(\mathbb{Z}, q)$ process can be thought of as a stationary version of the $\mathrm{Mallows}(\mathbb{N}, q)$ process, as described in [30]. In particular, we will use the following result on renewal processes, applied to the Mallows process [13, eq. 5.72].

Proposition 2.3. *Let X_i be a sequence of independent and identically distributed random variables taking values in $\mathbb{N}$, with $\mathbb{E}X_i = \mu > 0$, and let $\tau(n) = \inf\{t : \sum_{i=1}^{t} X_i > n\}$. Then $X_{\tau(n)} \to X^*$ in distributed, where X^* has the size-bias distribution of the X_i, which means $\mathbb{P}(X^* = x) = x\mathbb{P}(X_i = x)/\mu$.*

We will also use some of the tools developed in [15]. In particular we will use the arc chain $\{\kappa_t\}_{t=0}^{n}$ corresponding to $\Pi_n \sim \text{Mallows}(n, q)$ with $0 < q < 1$, defined by

$$\kappa_t := |\{i \in [t] : \Pi_n(i) > t\}|.$$

We speak of the (n, q) arc chain. We have the following.

Lemma 2.4 ([15], Proposition 3.3). *Let $\Pi_n \sim \text{Mallows}(n, q)$ with $0 < q < 1$. The arc chain $(\kappa_t)_{t=0,\ldots,n}$ of Π_n is a time-inhomogeneous Markov chain with transition probabilities*

$$\mathbb{P}(\kappa_{t+1} = j | \kappa_t = k) = \begin{cases} \left(\frac{1-q^k}{1-q^{n-t}}\right)^2 & \text{if } j = k-1, \\ \frac{q^k - q^{n-t}}{1-q^{n-t}} \cdot \frac{2 - q^k - q^{k+1}}{1-q^{n-t}} & \text{if } j = k, \\ \frac{q^k - q^{n-t}}{1-q^{n-t}} \cdot \frac{q^{k+1} - q^{n-t}}{1-q^{n-t}} & \text{if } j = k+1, \\ 0 & \text{otherwise.} \end{cases}$$

Lemma 2.5 ([15, Lemma 3.4]). *Let $\Pi_n \sim \text{Mallows}(n, q)$ and κ be its arc chain. Then*

$$\mathbb{P}[\Pi_n(t+1) = t+1 \mid \kappa_t = k] = \frac{(q^k - q^{k+1})(q^k - q^{n-t})}{(1 - q^{n-t})^2}, \qquad 0 \le t \le n.$$

(We mention that we have slightly adapted the statements from [15] in the above two lemmas.)

Analogously to the arc chain κ_t for $\Pi_n \sim \text{Mallows}(\mathbb{N}, q)$, we can define the arc-chain for $\Pi \sim \text{Mallows}(\mathbb{N}, q)$ with $0 < q < 1$ setting

$$\hat{\kappa}_t := |\{i \in [t] : \Pi(i) > t\}|.$$

We speak of the (∞, q)-arc chain. It is straightforward to verify that $(\hat{\kappa}_t)_{t \ge 0}$ forms a Markov chain with the following transition probabilities.

$$\mathbb{P}(\hat{\kappa}_{t+1} = j | \hat{\kappa}_t = k) = \begin{cases} (1-q^k)^2 & \text{if } j = k-1, \\ 2q^k - q^{2k} - q^{2k+1} & \text{if } j = k, \\ q^{2k+1} & \text{if } j = k+1, \\ 0 & \text{otherwise.} \end{cases}$$

Alternatively, this can be seen by combining Lemma 2.4 with Lemma 3.2 below. As a side remark, we mention that Gladkich and Peled [15] *defined* the (∞, q)-arc chain directly via the transition probabilities given here, without explicitly mentioning the connection to the definition we give here.

As explained in [15, Section 3.2], the Markov chain $\hat{\kappa}_t$ has a unique stationary distribution ν given by

$$\nu_s := \frac{\prod_{i=1}^{s} \frac{q^{2i-1}}{(1-q^i)^2}}{\sum_{t \geq 0} \prod_{i=1}^{t} \frac{q^{2i-1}}{(1-q^i)^2}}. \tag{3}$$

Here, as usual, an empty product equals 1.

The following result provides a useful link between the (n, q) and the (∞, q) arc chains.

Proposition 2.6 ([15], Proposition 3.8). *Set $t = t(n)$. If both $t \to \infty$ and $n - t \to \infty$ then the law of κ_t converges to the stationary distribution ν_s as n tends to infinity with q fixed.*

Given two discrete probability distributions μ_1 and μ_2 on a countable set Ω, their total variation distance is defined as

$$d_{TV}(\mu_1, \mu_2) = \max_{A \subseteq \Omega} |\mu_1(A) - \mu_2(A)|.$$

A useful alternative expression is

$$d_{TV}(\mu_1, \mu_2) = \frac{1}{2} \sum_{x \in \Omega} |\mu_1(x) - \mu_2(x)| = \sum_{x : \mu_1(x) > \mu_2(x)} \mu_1(x) - \mu_2(x).$$

(For a proof, see for instance Proposition 4.2 in [26].) As is common, we will interchangeably use the notation $d_{\text{TV}}(X, Y) := d_{\text{TV}}(\mu, \nu)$ if $X \sim \mu_1$ and $Y \sim \mu_2$.

A *coupling* of two probability measures μ, ν is a joint probability measure for a pair of random variables (X, Y) satisfying $X \overset{d}{=} \mu, Y \overset{d}{=} \nu$. We will also speak of a coupling of X, Y as being a probability space for (X', Y') with $X' \overset{d}{=} X, Y' \overset{d}{=} Y$. Another useful characterization of the total variational distance is as follows.

Lemma 2.7. *Let μ and ν be two probability distributions on the same countable set Ω. Then*

$$d_{TV}(\mu, \nu) = \inf\{\mathbb{P}[X \neq Y] : (X, Y) \text{ is a coupling of } \mu \text{ and } \nu\}.$$

There is a coupling that attains this infimum.

(For a proof, see for instance [26], Proposition 4.7 and Remark 4.8.)

For the proofs of the normal limiting laws in Theorems 1.1 and 1.2 we will make use of a result on stopped two-dimensional random walks by Gut and Janson [19] that seems tailor made for our purposes. Here we consider an i.i.d. sequence $(X_1, Y_1), (X_2, Y_2), \ldots$ and for $t > 0$ we define $\tau(t)$ as the first k such that $X_1 + \cdots + X_k$ exceeds t:

$$\tau(t) = \inf\{k \geq 1 : X_1 + \cdots + X_k > t\}. \tag{4}$$

The result of Gut and Janson we'll use states that:

Theorem 2.8 ([19], Theorem 3). *Let $(X_1, Y_1), (X_2, Y_2), \ldots$ be an i.i.d. sequence and let $\tau(n)$ be as given by (4). Suppose that $\mathbb{E}X_1 > 0$, that $\text{Var}(X_1), \text{Var}(Y_1) < \infty$ and*

$\mathrm{Var}(Y_1 \mathbb{E} X_1 - X_1 \mathbb{E} Y_1) > 0$. *Then*

$$\frac{\sum_{i=1}^{\tau(t)} Y_i - \left(\frac{\mathbb{E} Y_1}{\mathbb{E} X_1}\right) t}{\sqrt{t}} \xrightarrow[t \to \infty]{d} \mathcal{N}\left(0, \frac{\mathrm{Var}(Y_1 \mathbb{E} X_1 - X_1 \mathbb{E} Y_1)}{(\mathbb{E} X_1)^3}\right).$$

For the proofs of the normal limiting laws in Theorems 1.1 and 1.2 it will be convenient to use the Cramer-Wold device. A proof can for instance be found in [6, Theorem 29.4].

Theorem 2.9 (Cramer-Wold device)**.** *For random vectors $X_n = (X_{n,1}, \ldots, X_{n,k})$ and $Y = (Y_1, \ldots, Y_k)$, a necessary and sufficient condition for $X_n \xrightarrow{d} Y$ is that $\sum_{u=1}^{k} t_u X_{n,u} \xrightarrow{d} \sum_{u=1}^{k} t_u Y_u$ for each $(t_1, \ldots, t_k) \in \mathbb{R}^k$.*

Several times, we are going to rely on the following result of Basu and Bhatnagar [3].

Lemma 2.10 ([3, Lemma 5.5])**.** *Let $W_1, W_2, \ldots$ be an i.i.d. sequence of random variables with $\mathbb{E} W_i^2 < \infty$. Then*

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

in probability.

We will also make use of the following fact. Even though it seems pretty standard we have not been able to find a convenient reference. We therefore provide a short proof.

Lemma 2.11. *Suppose that $(W_t)_{t \geq 0}$ is a Markov chain with state space $\{0\} \cup \mathbb{N}$, started in state $W_0 = 0$, and whose transition probabilities satisfy $p_{i,j} = 0$ if and only if $|i - j| \neq 1$ and $\liminf_{i \to \infty} p_{i,i-1} > 1/2$. Let*

$$T := \inf\{t \geq 1 : W_t = 0\}.$$

Then $\mathbb{E} T^k < \infty$ for all $k \in \mathbb{N}$.

Proof. Let T_i denote the number of steps to reach $i - 1$, in the chain starting from $W_0 = i$. Let $i_0 \in \mathbb{N}, p > 1/2$ be such that $p_{i,i-1} > p$ for all $i \geq i_0$. We have

$$\mathbb{P}(T_i > t) \leq \mathbb{P}(\mathrm{Bi}(t, p) \leq t/2) = \exp\left[-\Omega(t)\right],$$

for each $i \geq i_0$, using the Chernoff inequality (see for instance [23], Corollary 2.3). This implies

$$\mathbb{E} T_i^k \leq \sum_t t^k \cdot \mathbb{P}(T_i \geq t) < \infty,$$

for all $i \geq i_0$ and $k \in \mathbb{N}$.

Starting from $W_0 = 0$, we of course move to state 1 with probability one in the first step, giving

$$\mathbb{E} T^k = \mathbb{E}(1 + T_1)^k = \sum_{j=0}^{k} \binom{n}{j} \mathbb{E} T_1^j,$$

for all $k \in \mathbb{N}$. In particular, it suffices to show $\mathbb{E} T_1^k < \infty$ for all $k \in \mathbb{N}$.

Similarly, by considering the first step of the chain we see for each $i \geq 1$ and $k \in \mathbb{N}$:

$$
\begin{aligned}
\mathbb{E} T_i^k &= p_{i,i-1} + p_{i,i-1} \cdot \mathbb{E}\left(1 + T_i' + T_{i+1}'\right)^k \\
&= p_{i,i-1} + p_{i,i+1} \cdot \sum_{\substack{0 \leq k_1, k_2 \leq k, \\ k_1 + k_2 \leq k}} \binom{k}{k_1, k_2, k - (k_1 + k_2)} \mathbb{E} T_i^{k_1} \mathbb{E} T_{i+1}^{k_2},
\end{aligned}
\tag{5}
$$

where we take $T_i' \overset{d}{=} T_{i-1}, T_{i+1}' \overset{d}{=} T_{i+1}$ independent in the first line. (To see the first inequality, note that in the first step we move to $i - 1$ with probability $p_{i,i-1}$. If, on the other hand, we move to $i + 1$ in the first step then we first have to wait until we reach state i again, and then we have to wait until we reach $i - 1$ from i). Rewriting (5), we obtain

$$
\mathbb{E} T_i^k = \frac{1}{1 - p_{i,i-1}} \cdot \left(p_{i,i-1} + p_{i,i+1} \cdot \sum_{\substack{0 \leq k_1, k_2 \leq k, \\ k_1 + k_2 \leq k, \\ k_1 \neq k}} \binom{k}{k_1, k_2, k - (k_1 + k_2)} \mathbb{E} T_i^{k_1} \mathbb{E} T_{i+1}^{k_2} \right).
$$

We can thus apply induction on k to show that $\mathbb{E} T_{i_0-1}^k < \infty$ for all $k \in \mathbb{N}$. Repeating the argument, we also have $\mathbb{E} T_i^k < \infty$ for $i = i_0 - 2, i_0 - 3, \dots$ and so on until $i = 1$. ∎

3 | THE PROOF OF THEOREM 1.1

We define a sequence of *regeneration times* $T_0 < T_1 < T_2 < \cdots$ as follows:

$$
\begin{aligned}
T_0 &:= 0, \\
T_i &:= \inf\{j > T_{i-1} \text{ s.t. } \Pi([j]) = [j]\} \qquad (i = 1, 2, \dots).
\end{aligned}
$$

In Section 4 of [3], Basu and Bhatnagar show that T_1 has finite second moment.

Lemma 3.1. $\mathbb{E} T_1^2 < \infty$.

(We combine Lemmas 4.1 and 4.5 of [3].) We also define the interarrival times

$$
X_i = T_i - T_{i-1},
$$

and

$$
\tau(n) = \inf\{t : T_t > n\}.
$$

Looking at the description of Π in Section 2, it is not difficult to see that conditional on the event $T_1 = t$, the bijection $i \mapsto \Pi(i + t) - t$ is distributed like Π. It follows that the interarrival times $X_1, X_2, \dots$ are i.i.d. Moreover, writing $\mathcal{X}_i := \{T_{i-1} + 1, \dots, T_i\}$ we see that Π maps $\mathcal{X}_i$ bijectively onto $\mathcal{X}_i$, and in fact the permutations $\Sigma_1 : [X_1] \to [X_1], \Sigma_2 : [X_2] \to [X_2], \dots$ given by

$$
\Sigma_i(j) := \Pi(T_{i-1} + j) - T_{i-1} \quad \text{for } j = 1, \dots, X_i,
$$

are i.i.d. as well.

With this regenerative structure, the following lemma follows.

Lemma 3.2. *For $0 < q < 1$ and $n \in \mathbb{N}$, let $\Pi_n \sim \mathrm{Mallows}(n, q)$ and $\Pi \sim \mathrm{Mallows}(\mathbb{N}, q)$. There exists a coupling of Π_n and Π satisfying*

$$\mathbb{P}(\Pi_n(i) = \Pi(i) \ \text{for all} \ 1 \leq i \leq n - \log^2 n) = 1 - o(1),$$

and in fact, $\log^2 n$ *can be replaced with any function going to ∞ with n.*

Proof. Let $\Pi \sim \mathrm{Mallows}(\mathbb{N}, q)$, and recall that $\Pi_{[n]} \sim \mathrm{Mallows}(n, q)$. We claim that this coupling works. Indeed, it's clear that $\Pi(i) = \Pi_{[n]}(i)$ up until $T_{\tau(n)-1}$. But $n - T_{\tau(n)-1} \leq X_{\tau(n)}$, and $X_{\tau(n)}$ converges to a limiting distribution by Proposition 2.3. In particular, the probability that it is larger than $\log^2 n$ (or any function going to ∞) goes to 0. ∎

By this last lemma, with probability $1 - o(1)$, the number of i-cycles in Π_n differs by at most $2\log^2 n$ from the number of i-cycles of Π that are completely contained in $[n]$ (for each $i = 1, \dots, \ell$).

Fix an $\ell \in \mathbb{N}$, and let $a_1, \dots, a_\ell$ be a sequence of real numbers, not all zero. For π a permutation, we define $\varphi(\pi) := \sum_{j=1}^{\ell} a_j C_j(\pi)$ and let $Y_i = \varphi(\Sigma_i)$.

We plan to apply Theorem 2.8 to the i.i.d. sequence $(X_1, Y_1), (X_2, Y_2), \dots$. For this we first need to establish the conditions of that theorem are met.

Lemma 3.3. *We have $\mathbb{E}X_1 > 0$ and $\mathrm{Var}(X_1), \mathrm{Var}(Y_1) < \infty$.*

Proof. Since $X_1 \geq 1$ by definition, we trivially have $\mathbb{E}X_1 > 0$. As $|\varphi(\Sigma_1)| \leq \max_i |a_i| \cdot X_1$, it suffices to show that $\mathrm{Var}(X_1) < \infty$. But this has already been established by Lemma 3.1. ∎

Lemma 3.4. $\mathrm{Var}(Y_1 \mathbb{E}X_1 - X_1 \mathbb{E}Y_1) > 0$.

Proof. We first note that, for each $i \in \mathbb{N}$, there is a positive probability that Σ_1 consists of a single i-cycle. (This happens for instance when $Z_1 = i, Z_2 = 1, \dots Z_i = i - 1$.)

Aiming for a contradiction, assume that $Y_1 \mathbb{E}X_1 - X_1 \mathbb{E}Y_1$ is almost surely constant. Whenever Σ_1 consists of a single $> \ell$ cycle, we have $Y_1 = 0$. In particular $Y_1 \mathbb{E}X_1 - X_1 \mathbb{E}Y_1$ can equal both $(\ell + 1)\mathbb{E}Y_1$ and $(\ell + 2)\mathbb{E}Y_1$ with positive probability. The quantity $Y_1 \mathbb{E}X_1 - X_1 \mathbb{E}Y_1$ being an almost sure constant now implies $\mathbb{E}Y_1 = 0$.

Let $1 \leq i \leq \ell$ be such that $a_i \neq 0$. There is a positive probability that $Y_1 \mathbb{E}X_1 = 0$ and a positive probability that $Y_1 \mathbb{E}X_1 = a_i \mathbb{E}X_1$. But that implies $a_i = 0$, contradicting the choice of i.

It follows that $Y_1 \mathbb{E}X_1 - X_1 \mathbb{E}Y_1$ is not almost surely constant. In other words, $\mathrm{Var}(Y_1 \mathbb{E}X_1 - X_1 \mathbb{E}Y_1) > 0$. ∎

Having established Lemmas 3.3 and 3.4, we can apply Theorem 2.8 to conclude that

$$\frac{\sum_{i=1}^{\tau(n)} Y_i - \left(\frac{\mathbb{E}Y_1}{\mathbb{E}X_1}\right) n}{\sqrt{n}} \xrightarrow[n \to \infty]{d} \mathcal{N}(0, (\mathbb{E}X_1)^{-3} \mathrm{Var}(Y_1 \mathbb{E}X_1 - X_1 \mathbb{E}Y_1)).$$

By Lemma 3.2 and the definition of $\tau(n)$ we have

$$\left| \varphi(\Pi_n) - \sum_{j=1}^{\tau(n)} Y_j \right| \leq \max_{i=1,\dots,\ell} |a_i| \cdot \left(X_{\tau(n)} + \log^2(n) \right), \tag{6}$$

with probability $1-o(1)$, under the coupling provided by Lemma 3.2. Moreover, applying Lemmas 2.10 and 3.1 we have that, with probability $1 - o(1)$, the RHS of (6) is $o(\sqrt{n})$. We can conclude:

$$\frac{\varphi(\Pi_n) - \left(\frac{\mathbb{E}Y_1}{\mathbb{E}X_1}\right)n}{\sqrt{n}} \quad \xrightarrow[n \to \infty]{d} \quad \mathcal{N}\left(0, \frac{\mathrm{Var}(Y_1\mathbb{E}X_1 - X_1\mathbb{E}Y_1)}{(\mathbb{E}X_1)^3}\right).$$

Recalling that $Y_1 = \sum_{i=1}^{\ell} a_i C_i(\Sigma_1)$ and setting $m_i := \frac{\mathbb{E}C_i(\Sigma_1)}{\mathbb{E}X_1}$, we can write

$$\frac{\varphi(\Pi_n) - \left(\frac{\mathbb{E}Y_1}{\mathbb{E}X_1}\right)n}{\sqrt{n}} = a_1 \frac{C_1(\Pi_n) - m_1 n}{\sqrt{n}} + \cdots + a_\ell \frac{C_\ell(\Pi_n) - m_\ell n}{\sqrt{n}}.$$

Setting

$$U_i := \frac{C_i(\Sigma_1)\mathbb{E}X_1 - X_1\mathbb{E}C_i(\Sigma_1)}{(\mathbb{E}X_1)^{3/2}},$$

we see that

$$\frac{Y_1\mathbb{E}X_1 - X_1\mathbb{E}Y_1}{(\mathbb{E}X_1)^{3/2}} = a_1 U_1 + \cdots + a_\ell U_\ell.$$

Therefore, if we set $P_{ij} := \mathrm{Cov}(U_i, U_j)$ then

$$\frac{\mathrm{Var}(Y_1\mathbb{E}X_1 - X_1\mathbb{E}Y_1)}{(\mathbb{E}X_1)^3} = \sum_{i=1}^{\ell}\sum_{j=1}^{\ell} a_i a_j P_{ij}.$$

This shows that if $(N_1, \ldots, N_\ell) \sim \mathcal{N}_\ell(\underline{0}, P_\ell)$ then

$$a_1 N_1 + \cdots + a_\ell N_\ell \stackrel{d}{=} \mathcal{N}\left(0, \frac{\mathrm{Var}(Y_1\mathbb{E}X_1 - X_1\mathbb{E}Y_1)}{(\mathbb{E}X_1)^3}\right).$$

We've thus shown that

$$a_1 \frac{C_1(\Pi_n) - m_1 n}{\sqrt{n}} + \cdots + a_\ell \frac{C_\ell(\Pi_n) - m_\ell n}{\sqrt{n}} \quad \xrightarrow[n \to \infty]{d} \quad a_1 N_1 + \cdots + a_\ell N_\ell,$$

for all $a_1, \ldots, a_\ell$. An application of Theorem 2.9 now allows us to conclude

$$\left(\frac{C_1(\Pi_n) - m_1 n}{\sqrt{n}}, \ldots, \frac{c_\ell(\Pi_n) - m_\ell n}{\sqrt{n}}\right) \quad \xrightarrow[n \to \infty]{d} \quad (N_1, \ldots, N_\ell),$$

completing the proof of Theorem 1.1.

4 | THE PROOF OF THEOREM 1.2

The proof is very similar to the proof of Theorem 1.1. We first introduce a two-sided sampling procedure in the case $0 < q < 1$ for a Mallows(n, q) distributed permutation Π_n taking $\left\lfloor \frac{n}{2} \right\rfloor$ iterations.

During iteration $i \geq 1$ we determine the images of i and $n - i + 1$. Again we take $Z_1, \ldots, Z_n$ independent with $Z_i \sim \text{TruncGeo}(n + 1 - i, q)$. In the first iteration we set

$$\Pi_n(1) = Z_1,$$
$$\Pi_n(n) = \text{the } Z_2\text{-th largest number in the set } [n] \setminus \{\Pi(1)\}.$$

In the i-th iteration we set

$$\Pi_n(i) = \text{the } Z_{2i-1}\text{-th smallest element of } [n] \setminus \{\Pi_n(j) : j < i \text{ or } j > n + 1 - i\},$$
$$\Pi_n(n + 1 - i) = \text{the } Z_{2i}\text{-th largest element of } [n] \setminus \{\Pi_n(j) : j \leq i \text{ or } j > n + 1 - i\}.$$

(If n is odd then after $\left\lfloor \frac{n}{2} \right\rfloor$ iterations, the image of $\lceil \frac{n}{2} \rceil$ has formally speaking not yet been determined, but of course there will be only one possible element of $[n]$ left.)

That this adapted procedure indeed produces a random permutation sampled according to the Mallows(n, q) measure follows analogously the the corresponding argument for the original sampling procedure: For every $\pi \in S_n$ there is a choice of $(k_1, \ldots, k_n) \in [n] \times [n-1] \times \cdots \times [1]$ such that $\{\Pi_n = \pi\} = \{Z_1 = k_1, \ldots, Z_n = k_n\}$. Again $\mathbb{P}(Z_1 = k_1, \ldots, Z_n = k_n) \propto q^{k_1 + \cdots + k_n - n}$. We also again have $\text{inv}(\pi) = k_1 + \cdots + k_n - n$, because when we are determining $\Pi_n(i)$ with $i \leq n/2$ then the number of $i < j < n + 1 - i$ such that i, j form an inversion is precisely $Z_{2i-1} - 1$, and similarly for $\Pi_n(n + 1 - i)$.

Recall that we use r_n to denote the map $i \mapsto n + 1 - i$. Analogously to Lemma 3.2, we have

Lemma 4.1. *Let $0 < q < 1$ and $\Pi_n \sim \text{Mallows}(n, q)$ and let $\Pi, \Pi' \sim \text{Mallows}(\mathbb{N}, q)$ be independent. There is a coupling for Π_n, Π, Π' such that*

$$\mathbb{P}\left(\begin{array}{l} \Pi_n(i) = \Pi(i) \text{ for all } 1 \leq i \leq n/2 - \log^2 n, \text{ and} \\ \Pi_n(i) = (r_n \circ \Pi' \circ r_n)(i) \text{ for all } n/2 + \log^2 n \leq i \leq n \end{array} \right) = 1 - o(1).$$

Moreover, the $\log^2 n$ can be replaced with any function going to ∞ with n.

Proof. The proof is similar to the proof of Lemma 3.2. By Lemma 3.2 applied to Π and Π', we can couple Π, Π' with independent Mallows$(n/2, q)$ permutations, $\Pi_{n/2}, \Pi'_{n/2}$ (for simplicity, we write $n/2$ even if n is odd, where it should be rounded either up or down as needed), such that $\mathbb{P}(\Pi_{n/2}(i) = \Pi(i), \Pi'_{n/2}(i) = \Pi'(i) \text{ for all } 1 \leq i \leq n/2 - \log^2 n) = 1 - o(1)$. We now claim that we can couple Π_n with $\Pi_{n/2}, \Pi'_{n/2}$ such that

$$\mathbb{P}\left(\begin{array}{l} \Pi_n(i) = \Pi_{n/2}(i) \text{ for all } 1 \leq i \leq n/2 - \log^2 n, \text{ and} \\ \Pi_n(i) = (r_n \circ \Pi'_{n/2} \circ r_n)(i) \text{ for all } n/2 + \log^2 n \leq i \leq n \end{array} \right) = 1 - o(1),$$

and this would immediately finish the proof, since any coupling with these bivariate marginals would satisfy the lemma.

To see the claim, note that $(\Pi_n)_{[n/2]}$ and $(\Pi_n)_{[n/2+1,n]}$ are independent Mallows$(n/2, q)$ (see e.g., Lemma 2.3 of [20]), and so can be coupled to perfectly agree with $\Pi_{n/2}$ and $r_n \circ \Pi'_{n/2} \circ r_n$ on the intervals $[n/2]$ and $[n/2, n]$. Now if $\Pi_n([k]) = [k]$ for $k \leq n/2$, then Π_n and $\Pi_{n/2}$ agree on that interval, since then $\Pi_n(i) = (\Pi_n)_{[1,n/2]}(i) = \Pi_{n/2}(i)$ for $i \in [k]$, and similarly for $r_n \circ \Pi'_{n/2} \circ r_n$ and intervals of the form $[k, n]$. Thus, the number of i with disagreements is bounded by the length of the smallest interval $[a, b]$ with $a \leq n/2$,

$b \geq n/2$, and $\Pi_n([a]) = [a]$ and $\Pi_n([b, n]) = [b, n]$. However, this random variable is stochastically dominated by $X_{\tau(n/2)}$, since if we couple Π_n with a Mallows$(\mathbb{N}, q)$ process, then the length of the interval is $b - a = X_{\tau(n/2)}$ unless $X_{\tau(n/2)} > n$, in which case $X_{\tau(n/2)}$ is strictly larger. But now we are done, since $X_{\tau(n/2)}$ converges to a limiting distribution by Proposition 2.3, and so the probability that Π_n disagrees with either $\Pi_{n/2}$ or $r_n \circ \Pi'_{n/2} \circ r_n$ for a growing number of locations goes to 0. ∎

Recall that $r_n \circ \Pi_n \overset{d}{=}$ Mallows$(n, 1/q)$ if $\Pi_n \sim$ Mallows(n, q). We have the following corollary.

Corollary 4.2. *Let* $q > 1$, $\Pi_n \sim$ *Mallows*(n, q) *and* $\Pi, \Pi' \sim$ *Mallows*$(\mathbb{N}, 1/q)$ *be independent. Let* $m = \left\lfloor \frac{n}{2} - (\log n)^2 \right\rfloor$. *There exists a coupling between* Π_n *and* Π, Π' *such that that*

$$\mathbb{P}\left[(r_n \circ \Pi_n, \Pi_n \circ r_n) \text{ and } (\Pi, \Pi') \text{ agree on } 1, \dots, m\right] = 1 - o(1).$$

Proof. We let $\Pi_n^* := r_n \circ \Pi_n$. Then $\Pi_n^* \overset{d}{=}$ Mallows$(n, 1/q)$ and we can couple Π_n^* with $\Pi, \Pi' \sim$ Mallows$(\mathbb{N}, 1/q)$ independent as in the previous lemma. With probability $1 - o(1)$, we have

$$(r_n \circ \Pi_n)(i) = \Pi_n^*(i) = \Pi(i),$$

and

$$(r_n \circ \Pi_n \circ r_n)(i) = \Pi_n^*(r_n(i)) = (r_n \circ \Pi' \circ r_n)(r_n(i)) = (r_n \circ \Pi')(i),$$

for all $i \leq m$. Now notice that $(r_n \circ \Pi_n \circ r_n)(i) = (r_n \circ \Pi')(i)$ if and only if $(\Pi_n \circ r_n)(i) = \Pi'(i)$. ∎

For $q > 1$, we let $\Pi_n \sim$ Mallows(n, q) and $\Pi, \Pi' \sim$ Mallows$(\mathbb{N}, 1/q)$ be coupled as in Corollary 4.2. We define

$$\begin{aligned} T_0 &= 0, \\ T_i &= \inf\{j > T_{i-1} : \Pi([j]) = \Pi'([j]) = [j]\} \qquad (i = 1, 2, \dots), \end{aligned}$$

We next show the analogue of Lemma 3.1.

Lemma 4.3. $\mathbb{E}T_1^2 < \infty$.

Proof. Let $\hat{\kappa}, \hat{\kappa}'$ denote the arc-chains for Π, respectively Π', as defined in Section 2. Note that $\Pi([j]) = [j]$ if and only if $\hat{\kappa}_j = 0$ and $\Pi'([j]) = [j]$ if and only if $\hat{\kappa}'_j = 0$. Thus T_1 is the first return to the origin of the two-dimensional random walk $(\hat{\kappa}_t, \hat{\kappa}'_t)$, started at $(\hat{\kappa}_0, \hat{\kappa}'_0) = (0, 0)$. For convenience, let us write $\lambda_t := \max(\hat{\kappa}_t, \hat{\kappa}'_t)$. We observe that

$$\mathbb{P}(\lambda_{t+1} = i - 1 | \lambda_t = i, \lambda_{t-1} = i_{t-1}, \dots, \lambda_1 = i_1) \geq \mathbb{P}(\hat{\kappa}_{t+1} = i - 1 | \hat{\kappa}_t = i)^2 = \left(1 - (1/q)^i\right)^4,$$

for all $i \geq 1$ and $i_1, \dots, i_{t-1} \in \{0\} \cup \mathbb{N}$ and all t. So, while $(\lambda_t)_{t \geq 0}$ is itself not necessarily a Markov chain, it is stochastically dominated by the chain $(W_t)_{t \geq 0}$ on $\{0\} \cup \mathbb{N}$ with starting

state $W_0 = 0$, and transition probabilities

$$\mathbb{P}(W_{t+1} = j | W_t = i) = \begin{cases} \left(1 - (1/q)^i\right)^4 & \text{if } j = 1 - 1 \text{ and } i \geq 1, \\ 1 - \left(1 - (1/q)^i\right)^4 & \text{if } j = 1 + 1 \text{ and } i \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

The result now immediately follows from Lemma 2.11. ∎

We also define $X_i := T_i - T_{i-1}$ and $\tau(t) = \inf\{j : T_j > t\}$ for all $t > 0$. Again it can be easily seen from the iterative procedure generating Π and Π' than $X_1, X_2, \ldots$ are i.i.d. Moreover, if we define the maps $\Sigma_1, \Sigma_1' : [X_1] \to [X_1], \Sigma_2, \Sigma_2' : [X_2] \to [X_2], \ldots$ by setting

$$\Sigma_i(j) := \Pi(T_{i-1} + j) - T_{i-1}, \quad \Sigma_i'(j) := \Pi'(T_{i-1} + j) - T_{i-1},$$

then $(\Sigma_1, \Sigma_1'), (\Sigma_2, \Sigma_2'), \ldots$ are i.i.d. as well.

We write $\mathcal{X}_i := \{T_{i-1} + 1, \ldots, T_i\}$. Observe that, with probability $1 - o(1)$, for each i such that $T_i < n/2 - \log^2 n$ we have

$$\Pi_n [\mathcal{X}_i] = r_n [\Pi [\mathcal{X}_i]] = r_n [\mathcal{X}_i],$$

$$\Pi_n [r_n [\mathcal{X}_i]] = \Pi' [\mathcal{X}_i] = \mathcal{X}_i,$$

by Corollary 4.2. In other words, writing $\mathcal{Y}_i := \mathcal{X}_i \cup r_n [\mathcal{X}_i]$, we have

$$\Pi_n [\mathcal{Y}_i] = \mathcal{Y}_i,$$

for each i such that $T_i < n/2 - \log^2 n$. In particular, every cycle of Π_n is either completely contained in one of $\mathcal{Y}_1, \ldots, \mathcal{Y}_{\tau(n/2)-1}$ or it contains some number between $\min(n/2 - \log^2 n, T_{\tau(n/2)-1})$ and $\max(n/2 + \log^2 n, n + 1 - T_{\tau(n/2)})$. We observe that the number of cycles of Π_n length $2i$ contained in $\mathcal{Y}_i$ equals the number of cycles of $\Pi_n \circ \Pi_n$ of length i contained in $\mathcal{X}_i$. Now note that on $\mathcal{X}_i$ we have $\Pi_n \circ \Pi_n = (\Pi_n \circ r_n) \circ (r_n \circ \Pi_n) = \Pi' \circ \Pi$. In particular, the number cycles of $\Pi_n \circ \Pi_n$ of length i contained in $\mathcal{X}_i$ equals the number of cycles of $\Sigma_i' \circ \Sigma_i$ of length i.

We fix $a_1, \ldots, a_\ell \in \mathbb{R}$, not all zero, and set

$$\varphi(\pi) := \sum_{i=1}^{\ell} a_i C_{2i}(\pi), \quad \psi(\pi) := \sum_{i=1}^{\ell} a_i C_i(\pi),$$

$$Y_i := \psi(\Sigma_i' \circ \Sigma_i).$$

By the observations in the previous paragraph

$$\left| \varphi(\Pi_n) - \sum_{i=1}^{\tau(n/2)} Y_i \right| \leq 2\max_i |a_i| \cdot \left(X_{\tau(n/2)} + \log^2 n \right) = o\left(\sqrt{n}\right), \tag{7}$$

with probability $1 - o(1)$ (where the last equality holds because $X_{\tau(n/2)} = o(\sqrt{n})$ with probability $1 - o(1)$ by Lemmas 2.10 and 4.3). Completely analogously to Lemmas 3.3 and 3.4 we have

Lemma 4.4. $\mathbb{E}X_1 > 0$ *and* $\mathrm{Var}X_1, \mathrm{Var}Y_1 < \infty$.

Lemma 4.5. $\mathrm{Var}(Y_1\mathbb{E}X_1 - X_1\mathbb{E}Y_1) > 0$.

We can thus conclude from Theorem 2.8 that

$$\frac{\sum_{i=1}^{\tau(n/2)}Y_i - \left(\frac{\mathbb{E}Y_1}{\mathbb{E}X_1}\right)\cdot(n/2)}{\sqrt{n/2}} \xrightarrow[n\to\infty]{d} \mathcal{N}\left(0, \frac{\mathrm{Var}(Y_1\mathbb{E}X_1 - X_1\mathbb{E}Y_1)}{(\mathbb{E}X_1)^3}\right).$$

Setting $\mu_{2i} := \frac{\mathbb{E}C_i(\Sigma_1' \circ \Sigma_1)}{2\mathbb{E}X_1}$ and using (7) this gives

$$a_1 \cdot \frac{C_2(\Pi_n) - \mu_2 n}{\sqrt{n}} + \cdots + a_\ell \cdot \frac{C_{2\ell}(\Pi_n) - \mu_{2\ell} n}{\sqrt{n}} \xrightarrow[n\to\infty]{d} a_1 N_1 + \cdots + a_\ell N_\ell,$$

where $(N_1, \dots, N_\ell) \stackrel{d}{=} \mathcal{N}(\underline{0}, P)$ with

$$P_{ij} := \mathrm{Cov}\left(\frac{C_i(\Sigma_1' \circ \Sigma_1)\mathbb{E}X_1 - X_1\mathbb{E}C_i(\Sigma_1' \circ \Sigma_1)}{\sqrt{2}\cdot(\mathbb{E}X_1)^{3/2}}, \frac{C_j(\Sigma_1' \circ \Sigma_1)\mathbb{E}X_1 - X_1\mathbb{E}C_j(\Sigma_1' \circ \Sigma_1)}{\sqrt{2}\cdot(\mathbb{E}X_1)^{3/2}}\right).$$

Again the result follows by an application of the Cramer-Wold device.

5 | THE PROOF OF THEOREM 1.3

Part (ii) of Lemma 2.2 says that, almost surely, $\Sigma_{I_n} \to \Sigma$ for $I_n := \{-n, \dots, n\}$. It however leaves open how fast the convergence is. The following lemma shows that in fact, with high probability, for the vast majority of elements of I_n, the values of Σ and Σ_{I_n} agree. This will be very helpful for us.

Lemma 5.1. *Let* $0 < q < 1$ *and* $\Sigma \sim \mathrm{Mallows}(\mathbb{Z}, q)$ *and* $I_n := \{-n, \dots, n\}, J_n := \{-n+1, \dots, n\}$. *Almost surely there exists a (random)* $N \in \mathbb{N}$ *such that* $\Sigma_{I_n}(i) = \Sigma_{J_n}(i) = \Sigma(i)$ *for all* $n \geq N$ *and* i *with* $|i| \leq n - \log^2 n$.

Proof. We define the event B_n by

$$B_n := \left\{\begin{array}{ll} |\Sigma(i) - i| < 5\log_{1/q} n & \text{for all } i \in I_n, \text{ and} \\ |\Sigma(j) - j| \leq 5\log_{1/q} n + |j| - n & \text{for all } j \notin I_n \end{array}\right\}.$$

By Lemma 2.2, Part (vi) we have

$$\mathbb{P}\left[B_n^c\right] = O\left((2n+1)q^{5\log_{1/q} n} + 2\sum_{j>n}q^{5\log_{1/q} n+j-n}\right) = O(n^{-4}).$$

We will show that B_n implies the conclusion of the lemma. Let $a \geq 0$ and consider the intervals I_{n+a}. Let i be such that $|i| < n - 10\log_{1/q} n$. If B_n holds, then for all $j < -n$ we

have $\Sigma(j) < -n + 5\log_{1/q} n \leq \Sigma(i)$. Similarly $j > n$ implies $\Sigma(j) > n - 5\log_{1/q} n \geq \Sigma(i)$. Thus B_n implies

$$|\{j \in I_{n+a} : \Sigma(j) < \Sigma(i)\}| = a + |\{j \in I_n : \Sigma(j) < \Sigma(i)\}|.$$

As

$$\Sigma_{I_{n+a}}(i) = -n - a + |\{j \in I_{n+a} : \Sigma(j) < \Sigma(i)\}|,$$

the sequence $\Sigma_{I_{n+a}}(i)$ is constant for all $a \geq 0$. By Lemma 2.2, Part (ii), with probability one there is some n' such that for all $n'' \geq n'$ we have $\Sigma_{I_{n''}}(i) = \Sigma(i)$. There is some $a \geq 0$ such that $n' \leq n + a$, so that in particular we must have $\Sigma_{I_n}(i) = \Sigma(i)$. Similarly, $\Sigma_{J_n}(i) = \Sigma(i)$. By the Borel–Cantelli lemma B_n holds for all but finitely many n. ∎

We also require the following Markov chain representation for the times T_i. Let $\Pi \sim$ Mallows($\mathbb{N}, q$). Consider the process

$$M_n = \max_{1 \leq i \leq n} \Pi(i) - n$$

on $\mathbb{N}$. This is a positive recurrent Markov process—see [3]. The Markov process can be described in terms of the geometric random variables defining the Mallows process. Specifically, the walk can be described as moving from M_n to $M_{n+1} = \max(M_n, Z_n) - 1$ where the Z_n are independent geometric random variables. Let R_i denote the hitting time of i and let R_i^+ denote the return time at i. Then if the chain is started from 0, R_0^+ is distributed as the size of an excursion in the Mallows process. This Markov chain was introduced in [3] to study the moments of the T_i. We are now ready for the proof of Theorem 1.3.

We start by considering Π_{2n+1}. We let $\Sigma \sim$ Mallows($\mathbb{Z}, 1/q$), and let S_0 be the smallest integer (we will show that this exists) such that $\Sigma([-S_0, S_0]) = [-S_0, S_0]$, and then let S_i be defined inductively as the smallest number larger than S_{i-1} such that Σ preserves $[-S_i, S_i]$. To see that all these values are finite almost surely, we note that the times T_i for $i \in \mathbb{Z}$ for which $\Sigma([T_i+1, T_{i+1}]) = [T_i+1, T_i]$, with the convention that T_0 contains 0, forms a stationary renewal process, with the T_i for $i \neq 0$ the same as for the Mallows($\mathbb{N}, 1/q$) process, and T_0 having its size-bias distribution (see Theorem 3.2 of [30]). Then T_0 is finite almost surely, and given T_0, the two sides are independent and behave like Mallows($\mathbb{N}, 1/q$) processes. Then the S_i correspond to a simultaneous renewal on both sides of this process, with S_0 being the first time this occurs, which are the return times in a product of two independent copies of the positive recurrent Markov chain M_i defined above, which is thus also positive recurrent. Thus, all S_i are finite almost surely. By definition, $\Sigma \circ r$ preserves $[-S_0, S_0]$ and exchanges $[S_{i-1} + 1, S_i]$ and $[-S_i, -S_{i-1} - 1]$.

We then immediately see that there are no infinite cycles in $\Sigma \circ r$, and the odd cycles must be contained in the interval $[-S_0, S_0]$. Thus, the $C_{2i+1}(\Pi_{[-n,n]}) \to C_{2i+1}(\Pi_{[-S_0, S_0]})$ almost surely. The result for Π_{2n+1} follows by noting that $C_{2i+1}(\Pi_{2n+1}) \overset{d}{=} C_{2i+1}(\Pi_{[-n,n]})$ by Lemma 2.2 and the observations that precede it.

The proof of the result for Π_{2n} follows in exactly the same manner, except centered at $\frac{1}{2}$ rather than 0, and using ρ instead of r.

Technically speaking, the proof of Theorem 1.3 is not yet complete, as we have not yet shown the random vectors $(C_1(r \circ \Sigma), C_3(r \circ \Sigma), \ldots)$ and $(C_1(\rho \circ \Sigma), C_3(\rho \circ \Sigma), \ldots)$ have distinct probability distributions. This will however follow immediately from Theorem 1.7, which we will prove in a little while.

We also note the following consequence which will be useful later.

Lemma 5.2. *Let* $\Sigma \sim \text{Mallows}(\mathbb{Z}, q)$ *for* $0 < q < 1$. *Almost surely*, Σ *has no cycles of infinite length.*

Proof. Any cycle must be contained in an interval $[-S_0, S_0]$, $[S_{i-1} + 1, S_i]$, or $[-S_i, -S_{i-1} - 1]$, all of which are finite almost surely. ∎

6 | THE PROOF OF THEOREM 1.4

6.1 | The proof of Part (i) of Theorem 1.4

Let $0 < q < 1$, let $\Pi_n \sim \text{Mallows}(n, q)$ and $\Sigma \sim \text{Mallows}(\mathbb{Z}, q)$ and let $i \in \mathbb{N}$ be fixed. Using that $0 \leq C_i(\Pi_n)/n \leq 1$, it follows from Theorem 1.1 that

$$\lim_{n \to \infty} \frac{\mathbb{E}C_i(\Pi_n)}{n} = m_i. \tag{8}$$

Setting $I_n := \{-n, \ldots, n\}$, by Part (i) of Lemma 2.2, we have $C_i(\Pi_{2n+1}) \stackrel{d}{=} C_i(\Sigma_{I_n})$. We define the events

$$E := \{\Sigma_{I_n}(j) = \Sigma(j) \text{ for all } -n + \log^2 n \leq j \leq n - \log^2 n\},$$
$$F := \{|\Sigma(j) - j| \leq \log^2 n \text{ for all } j \in I_n\}.$$

Then $\mathbb{P}(E) = 1 - o(1)$ by Lemma 5.1. By Part (vii) of Lemma 2.2 and the remarks that follow that lemma $\mathbb{P}(F^c) = O\left(n \cdot q^{\log^2 n}\right) = o(1)$.

It follows that for all j with $|j| < n - i \cdot \log^2 n$ we have

$$\left| \mathbb{P}(j \text{ in an } i\text{-cycle of } \Sigma_{I_n}) - \mathbb{P}(j \text{ in an } i\text{-cycle of } \Sigma) \right| \leq \mathbb{P}(E^c) + \mathbb{P}(F^c) = o(1).$$

The number of elements of I_n in an i-cycle of Σ_{I_n} equals $i \cdot C_i(\Sigma_{I_n})$. We see that

$$i \cdot \mathbb{E}C_i(\Pi_{2n+1}) = \sum_{-n \leq j \leq n} \mathbb{P}(j \text{ in an } i\text{-cycle of } \Sigma_{I_n})$$
$$= \sum_{-n \leq j \leq n} \mathbb{P}(j \text{ in an } i\text{-cycle of } \Sigma) + o(n)$$
$$= (2n + 1) \cdot \mathbb{P}(0 \text{ in an } i\text{-cycle of } \Sigma) + o(n), \tag{9}$$

where we use Part (iv) of Lemma 2.2 (together with the remarks following the lemma) for the last line. Dividing the LHS and RHS of (9) by $2n + 1$ and sending $n \to \infty$, and recalling (8), proves the result.

6.2 | The proof of Part (ii) of Theorem 1.4

By Lemma 5.2, $\Sigma \sim \text{Mallows}(\mathbb{Z}, q)$ for $0 < q < 1$ almost surely has no infinite cycles. By Part (i) of Theorem 1.4 we have

$$\sum_{i=1}^{\infty} im_i = \sum_{i=1}^{\infty} \mathbb{P}\left[0 \text{ in an } i\text{-cycle of } \Sigma\right] = \mathbb{P}\left[0 \text{ lies in a finite cycle of } \Sigma\right] = 1,$$

the last equality due to Lemma 5.2.

6.3 | The proof of Part (iii) of Theorem 1.4

We start by giving an alternative expression for m_1 by employing the tools developed by Gladkich and Peled [15].

Lemma 6.1.

$$m_1 = \sum_{s \geq 0} \nu_s q^{2s} (1 - q), \tag{10}$$

with ν as given in (3).

Proof. Let $\Pi_{2n+1} \sim \text{Mallows}(2n + 1, q)$. By Lemma 5.1 we have that

$$\mathbb{P}\left[\Pi_{2n+1}(n + 1) = n + 1\right] \to \mathbb{P}\left[\Sigma(0) = 0\right], \qquad \text{as } n \to \infty,$$

where $\Sigma \sim \text{Mallows}(\mathbb{Z}, q)$. Now,

$$\lim_{n \to \infty} \mathbb{P}\left[\Pi_{2n+1}(n + 1) = n + 1\right]$$
$$=$$
$$\lim_{n \to \infty} \sum_{s \geq 0} \mathbb{P}\left[\kappa_{n+1} = s\right] \cdot \mathbb{P}\left[\Pi_{2n+1}(n + 1) = n + 1 \mid \kappa_{n+1} = s\right]. \tag{11}$$

By Lemma 2.5 and Proposition 2.6 we have for all $s \geq 0$ that

$$\lim_{n \to \infty} \mathbb{P}\left[\Sigma_{2n+1}(n + 1) = n + 1 \mid \kappa_{n+1} = s\right] = q^{2s}(1 - q),$$
$$\lim_{n \to \infty} \mathbb{P}\left[\kappa_{n+1} = s\right] = \nu_s.$$

So the summands on the right hand side of (11) converge pointwise to $\nu_s q^{2s}(1 - q)$ as $n \to \infty$. Moreover, by Lemma 2.5, we have the bound

$$\mathbb{P}\left[\kappa_{n+1} = s\right] \cdot \mathbb{P}\left[\Pi_{2n+1}(n + 1) = n + 1 \mid \kappa_{n+1} = s\right] \leq \frac{(q^s - q^{s+1})(q^s - q^{n+1})}{(1 - q^{n+1})^2},$$
$$\leq q^{2s}(1 - q)\frac{1}{(1 - q)^2}.$$

Let $g(s) = q^{2s}/(1-q)$, we observe that $\sum_{s \geq 0} g(s) < \infty$. We conclude that the summands in (11) are bounded by the integrable function $g(s)$, so by the dominated convergence theorem we have

$$m_1 = \mathbb{P}\left[\Sigma(0) = 0\right] = \sum_{s \geq 0} \nu_s q^{2s}(1 - q).$$

∎

Next, we will show that $m_1 = 1 - 2q + O(q^2)$ as $q \downarrow 0$ be analyzing (10). Let K_q denote the denominator in the expression for ν given in (3). Define $t_s = \prod_{i=1}^{s} q^{2i-1}/(1 - q^i)^2$. The t_s satisfy for $s \geq 1$ the recursion relation

$$t_{s+1} = t_s \cdot \frac{q^2}{(1 - q^{s+1})^2} \leq \frac{q^2 t_s}{(1 - q)^2} \leq t_1\left(\frac{q^2}{(1 - q)^2}\right)^{s-1} = \frac{q}{(1 - q)^2}\left(\frac{q^2}{(1 - q)^2}\right)^{s-1}.$$

Thus

$$K_q = 1 + \sum_{s \geq 1} t_s \leq 1 + \frac{q}{(1-q)^2} \sum_{s \geq 1} \left(\frac{q^2}{(1-q)^2} \right)^{s-1}.$$

If $q < 1/2$ then $q^2/(1-q)^2 < 1$, in which case the above equals

$$1 + \frac{q}{(1-q)^2} \cdot \frac{1}{1 - \frac{q^2}{(1-q)^2}} = 1 + \frac{q}{(1-q)^2 - q^2} = \frac{1-q}{1-2q}.$$

Then

$$v_0 = \frac{1}{K_q} \geq \frac{1-2q}{1-q},$$

so that by $m_1 \geq v_0(1-q)$ we obtain

$$m_1 \geq v_0(1-q) \geq 1 - 2q. \tag{12}$$

We also have the simple bound

$$K_q \geq 1 + t_1 = 1 + \frac{q}{(1-q)^2},$$

by which

$$m_1 \leq \frac{1-q}{1 + \frac{q}{(1-q)^2}} + \sum_{s \geq 1} q^{2s}(1-q) = \frac{(1-q)^3}{1-q+q^2} + O(q^2). \tag{13}$$

The function $(1-q)^3/(1-q+q^2)$ is infinitely differentiable at $q = 0$, where its first derivative equals -2. Thus

$$\frac{(1-q)^3}{1-q+q^2} = 1 - 2q + O(q^2).$$

Together with (12) and (13) this completes the the proof that $m_1 = 1 - 2q + O(q^2)$.

We now proceed to show that $m_1 = (1-q)/4 + o(1-q)$ as $q \uparrow 1$. Fix some $0 < \varepsilon < \frac{1}{2}$, let $0 < q < 1$ be so close to 1 that $\frac{1}{q} \leq (1+\varepsilon)$ and let $S = \{s \in \mathbb{N}_{\geq 0} : |q^s - \frac{1}{2}| < \varepsilon\}$. If $q^s < \frac{1}{2} - \varepsilon$, then by (3) we have the recursion

$$v_{s+1} = v_s \frac{1}{q} \left(\frac{q^s}{1-q^s} \right)^2 \leq v_s \delta \qquad \text{where} \qquad \delta \leq (1+\varepsilon) \left(\frac{\frac{1}{2} - \varepsilon}{\frac{1}{2} + \varepsilon} \right)^2 < 1.$$

Then

$$\sum_{s > \max(S)} v_s \leq v_{\max(S)} \sum_{s > \max(S)} \delta^{s - \max(S)} = v_{\max(S)} \frac{\delta}{1-\delta}.$$

We now claim that $v_{\max(S)} = o(1)$ as $q \uparrow 1$. Let $s^* = \left\lceil \log_q(\frac{1}{2}(1-\varepsilon)) \right\rceil$, then for $k \geq 0$ we have

$$v_{s^*+k+1} \leq v_{s^*} \frac{1}{q} \left(\frac{q^{s^*+k+1}}{1-q^{s^*+k+1}} \right) \leq v_{s^*}(1+\varepsilon) \left(\frac{1-\varepsilon}{1+\varepsilon} \right)^2 < v_{s^*}.$$

So

$$1 \geq \sum_{s=s^*}^{\max(S)} v_s \geq (\max(S) - s^*) v_{\max(S)}.$$

By

$$\max(S) - s^* = \left\lfloor \log_q \left(\frac{1}{2} + \varepsilon \right) \right\rfloor - \left\lceil \log_q \left(\frac{1}{2} + \frac{\varepsilon}{2} \right) \right\rceil \to \infty \qquad \text{as } q \uparrow 1,$$

we indeed have $v_{\max(S)} = o(1)$ as $q \uparrow 1$.

Similarly, if $q^s > \frac{1}{2} + \varepsilon$, then $v_{s+1} \geq v_s \tilde{\delta}$, where $\tilde{\delta} \geq \frac{\frac{1}{2} + \varepsilon}{\frac{1}{2} - \varepsilon} > 1$. Then

$$\sum_{0 \leq s < \min(S)} v_s \leq v_{\min(S)} \sum_{0 \leq s < \min(S)} \tilde{\delta}^{\min(S)-s} \leq v_{\min(S)} \frac{\tilde{\delta}}{1 - \tilde{\delta}}.$$

By the same reasoning as before, $v_{\min(S)}$ is also $o(1)$ as $q \uparrow 1$. We conclude by Lemma 6.1 that

$$m_1 = \sum_{s \in S} (1-q)q^{2s} v_s + \sum_{s \notin S} v_s q^{2s}(1-q) = \sum_{s \in S} (1-q)q^{2s} v_s + o(1-q), \quad \text{as } q \uparrow 1.$$

By the definition of S we also have

$$(1-\varepsilon)(1-q)\frac{1}{4}\sum_{s \in S} v_s \leq \sum_{s \in S}(1-q)q^{2s} v_s \leq (1+\varepsilon)(1-q)\frac{1}{4}\sum_{s \in S} v_s.$$

Now $\sum_{s \in S} v_s = 1 - o(1)$ as $q \uparrow 1$, the result follows.

We now show the limits

$$(m_1, m_2, \dots) \to (0, 0, \dots) \qquad \text{as } q \uparrow 1,$$
$$(m_1, m_2, \dots) \to (1, 0, \dots) \qquad \text{as } q \downarrow 0.$$

By Part (ii) of Theorem 1.4 and the previous we have for $q \downarrow 0$ the expansion

$$\sum_{i=2}^{\infty} m_i \leq \sum_{i=2}^{\infty} i m_i = 1 - m_1 = 2q + O(q^2).$$

So necessarily all the m_i converge to 0 for $i \geq 2$ as $q \downarrow 0$.

Let $\Pi_n \sim \text{Mallows}(n, q)$. For the limit as $q \uparrow 1$ we count the number of times during the sampling algorithm that i is the maximum element of an r-cycle for $r \geq 1$. So suppose that during iteration i of the algorithm there is an element $i' \leq i$ such that having $\Pi_n(i) = i'$ would create an r-cycle. Then there is some $k \geq 0$ such that the probability that $\Pi_n(i) = i'$ is

$$\frac{(1-q)q^{k-1}}{1 - q^{n-s+1}} \leq \frac{1-q}{1 - q^{n-s+1}}.$$

For $i \leq n - \sqrt{n}$ the above is not more than $2(1-q)$ for n large enough. So the expected number of elements in Π_n that are the maximum element of an r-cycle is not more than $O(\sqrt{n}) + 2(1-q)n$. Thus $m_r = \lim_{n \to \infty} \mathbb{E}[C_r(\Pi_n)/n] \leq 2(1-q)$ so that $\lim_{q \uparrow 1} m_r = 0$.

This shows the pointwise convergence of the two sequences in Part (iii) of Theorem 1.4. That the pointwise convergence implies convergence wrt. the L_1 norm follows from the fact that $\sum_{j \geq i} m_j \leq \frac{1}{i}$ for all i by Part (ii) of Theorem 1.4. So, for every $\varepsilon > 0$ we can take $i > 2/\varepsilon$ and find a $\delta > 0$ so that $m_1, \ldots, m_i < \varepsilon/(2i)$ whenever $1 - \delta < q < 1$. If then follows $|(m_1, m_2, \ldots) - (0, 0, \ldots)|_1 \leq \varepsilon$ for all $0 < q < \delta$. The case when $q \downarrow 0$ can be dealt with analogously.

7 | THE PROOF OF THEOREM 1.5

7.1 | The proof of Part (i) of Theorem 1.5

We start by proving the existence of a coupling in the same spirit as the couplings used in previous proofs.

Lemma 7.1. *Let $0 < q < 1$ and $\Pi \sim \text{Mallows}(\mathbb{N}, q)$ and $\Sigma \sim \text{Mallows}(\mathbb{Z}, q)$. There exists a coupling between Π and Σ such that*

$$\mathbb{P}\left[\Pi(i) = \Sigma(i) \text{ for all } i \geq \log^2 n\right] = 1 - o(1).$$

Moreover, the $\log^2 n$ can be replaced with any function going to ∞ with n.

Proof. We can run both Π and Σ until both processes regenerate simultaneously, in the sense that $\Pi(i) \geq k$ for all $i \geq k$ and $\Sigma(i) \geq k$ for all $i \geq k$. After this time, both processes will have the same distribution, and so can be coupled to be equal. The first time that both processes regenerate is equal in distribution to the hitting time of $(0, 0)$ for the Markov chain defined by taking two independent copies of M_i, where we start by first running both chains until Σ regenerates (so the Markov chain starts from a random state $(0, X)$, where X is the state of the second copy of the Markov chain stopped at the first time that Σ regenerates). Since the product chain is still positive recurrent, this hitting time is finite almost surely, and the result follows. ∎

Let $q > 1$ and $\Pi_n \sim \text{Mallows}(n, q)$. Since $0 \leq C_{2i}(\Pi_n)/n \leq 1$, Theorem 1.2 implies

$$\mu_{2i} = \lim_{n \to \infty} \frac{\mathbb{E}C_{2i}(\Pi_n)}{n}.$$

Let $\Pi, \Pi' \sim \text{Mallows}(\mathbb{N}, 1/q)$ be independent and coupled with Π_n by the coupling provided by Corollary 4.2. We set

$$U_{2i,n} := \left|\left\{i\text{-cycles of } \Pi' \circ \Pi \text{ completely contained in } \left[\lfloor n/2 \rfloor\right]\right\}\right|.$$

By (7) in the proof of Theorem 1.2 (setting $\ell = i, a_i = 1, a_1 = \cdots = a_{i-1} = 0$) we have that

$$|C_{2i}(\Pi_n) - U_{2i,n}| = o(n),$$

with probability $1 - o(1)$. Since also $0 \leq U_{2i,n}/n \leq 1$, it follows that

$$\lim_{n \to \infty} \frac{\mathbb{E}U_{2i,n}}{n} = \mu_{2i}. \tag{14}$$

Using Lemma 7.1, we can couple Π, Π' to $\Sigma, \Sigma' \sim \text{Mallows}(\mathbb{Z}, 1/q)$ in such a way that Σ, Σ' are independent and $\mathbb{P}(E) = 1 - o(1)$, where

$$E := \{\Sigma(j) = \Pi(j), \Sigma'(j) = \Pi'(j) \text{ for all } \log^2 n \leq j \leq n - \log^2 n\}.$$

We also have $\mathbb{P}(F) = 1 - o(1)$, where

$$F := \{|\Sigma(j) - j|, |\Sigma'(j) - j| \leq \log^2 n \text{ for all } 1 \leq j \leq n),$$

by an application of Part (vii) of Lemma 2.2 and the union bound. Hence, if A_j denotes the event that j is in an i-cycle of $\Pi' \circ \Pi$ that is completely contained in $\lfloor \lfloor n/2 \rfloor \rfloor$, and B_j denotes the event that j is in an i-cycle of $\Sigma' \circ \Sigma$ then

$$\left| \mathbb{P}(A_j) - \mathbb{P}(B_j) \right| \leq \mathbb{P}(E^c) + \mathbb{P}(F^c) = o(1),$$

for all $2i \cdot \log^2 n \leq j \leq n/2 - 2i\log^2 n$. It follows that

$$\begin{aligned}
\mathbb{E}U_{2i,n} &= \frac{1}{i} \sum_{1 \leq j \leq n/2} \mathbb{P}(A_j) \\
&= \frac{1}{i} \sum_{1 \leq j \leq n/2} \mathbb{P}(B_j) + o(n) \\
&= \frac{1}{i} \cdot \lfloor n/2 \rfloor \cdot \mathbb{P}\left(0 \text{ in an } i\text{-cycle of } \Sigma' \circ \Sigma\right) + o(n),
\end{aligned}$$

using Part (iv) of Lemma 2.2 and the remarks following that lemma for the last identity (applied to both Σ and Σ'). Dividing LHS and RHS by n and sending $n \to \infty$ (and recalling (14)) gives $\mu_{2i} = \mathbb{P}\left(0 \text{ in an } i\text{-cycle of } \Sigma' \circ \Sigma\right)$.

Finally, we briefly clarify how the expression $\mu_{2i} = \sum_{i \in \mathbb{Z}} \mathbb{P}(\Sigma(0) = i)^2$ is obtained. We have

$$\begin{aligned}
\mathbb{P}\left(0 \text{ in a 1-cycle of } \Sigma' \circ \Sigma\right) &= \sum_{i \in \mathbb{Z}} \mathbb{P}(\Sigma(0) = i, \Sigma'(i) = 0) \\
&= \sum_{i \in \mathbb{Z}} \mathbb{P}(\Sigma(0) = i)\mathbb{P}(\Sigma(0) = -i) \\
&= \sum_{i \in \mathbb{Z}} \mathbb{P}(\Sigma(0) = i)^2,
\end{aligned}$$

using that Σ, Σ' are i.i.d. and Part (iv) of Lemma 2.2, and the remarks following that lemma, for the second identity; and in the last identity that

$$-\Sigma(0) = (r \circ \Sigma \circ r)(0) \overset{d}{=} \Sigma(0),$$

also by the remarks following Lemma 2.2.

7.2 │ The proof of Part (ii) of Theorem 1.5

Let $\ell \geq 1$ and define

$$a_i = 2i, \qquad \text{for } i = 1, \dots, \ell.$$

Define $\varphi(\Pi_n) = \sum_{i=1}^{\ell} a_i C_{2i}(\Pi_n)$ and X_1, Y_1 as in the proof of Theorem 1.2. We have $0 \leq \varphi(\Pi_n)/n \leq 1$, so that Theorem 1.2 in fact implies

$$\mathbb{E}\left[\frac{\varphi(\Pi_n)}{n}\right] \to \sum_{i=1}^{\ell} a_i \mu_{2i} = \sum_{i=1}^{\ell} 2i \mu_{2i},$$

while (7) implies

$$\mathbb{E}\left[\frac{\varphi(\Pi_n)}{n}\right] \to \frac{\mathbb{E}Y_1}{2\mathbb{E}X_1}.$$

The function Y_1 counts the number of elements in $[X_1] \cup \{n - X_1 + 1, \ldots, n\}$ that are contained in even cycles of length at most 2ℓ. If $X_1 \leq \ell$ then we have $Y_1 = 2X_1$, as all elements are in cycles of length $\leq 2\ell$. If $X_1 > \ell$ then certainly $Y_1 \leq 2X_1$. Let

$$a := \mathbb{E}[X_1 1_{\{X_1 \leq \ell\}}] \qquad \text{and} \qquad b := \mathbb{E}[X_1 1_{\{X_1 > \ell\}}].$$

We have $\mathbb{E}X_1 = a + b$. We also have the bounds $2a \leq \mathbb{E}X_1 \leq 2a + 2b$. As $\mathbb{E}X_1 < \infty$, for any $\varepsilon > 0$ we can choose $\ell_0 = \ell_0(\varepsilon)$ large enough so that for every $\ell > \ell_0$ we have $b < \varepsilon$. As $1 \leq \mathbb{E}X_1$, having chosen ℓ_0 sufficiently large, we can also ensure $a \geq 1 - \varepsilon$ for all $\ell > \ell_0$. In this case

$$1 \geq \frac{\mathbb{E}Y_1}{2\mathbb{E}X_1} \geq \frac{a}{a + \varepsilon} = 1 - \frac{\varepsilon}{a + \varepsilon} \geq 1 - \varepsilon.$$

So

$$\sum_{i=1}^{\infty} 2i \, \mu_{2i} = \lim_{\ell \to \infty} \left(\frac{\mathbb{E}Y_1}{2\mathbb{E}X_1}\right) = 1.$$

7.3 | The proof of Part (iii) of Theorem 1.5

As in the proof of Part (iii) of Theorem 1.4, it suffices to prove pointwise convergence (convergence for each μ_{2i} separately), and the convergence in L_1 will follow using that $\sum_i 2i\mu_{2i} = 1$.

We first show that $(\mu_2, \mu_4, \ldots) \to (0, 0, \ldots)$ as $q \downarrow 1$. Let $r \geq 1$ and $\Pi_n \sim \text{Mallows}(n, q)$ where $q > 1$. By Part (i) of Theorem 1.5, we have

$$\mu_{2r} = \lim_{n \to \infty} \frac{\mathbb{E}C_{2r}(\Pi_n)}{n}. \tag{15}$$

If $i_1, i_2, \ldots, i_{2r}$ form a $2r$-cycle in $\Pi_n \sim \text{Mallows}(n, q)$ with $q > 1$, and $\Pi_n^* := r_n \circ \Pi_n$ then $\Pi_n^* \overset{d}{=} \text{Mallows}(n, 1/q)$ and we must have

$$\Pi_n^*(i_1) = r_n(i_2), \ldots, \Pi_n^*(i_{2r-1}) = r_n(i_{2r}), \Pi_n^*(i_{2r}) = r_n(i_1).$$

Let us define

$$A_j := \{j \text{ is the largest element of a } 2r\text{-cycle of } \Pi_n\}.$$

(To clarify j being the largest element, we of course just mean $j = \max(i_1, \ldots, i_{2r})$ if $i_1, \ldots, i_{2r}$ is a $2r$-cycle as above.) We consider the iterative procedure for generating Π_n^*. When we generate the

image of $\Pi_n^*(j)$, having already determined $\Pi_n^*(1), \dots, \Pi_n^*(j-1)$ it may not be possible that j is the largest element of some $2r$-cycle. If it is still possible, then we need $\Pi_n^*(j)$ to be some specific value among the still available ones. (To be precise, $(r_n \circ \Pi_n^*)^{(-2r-1)}(j)$.) Since, we sample according to a truncated geometric distributions, the probability is thus at most

$$\mathbb{P}(A_j) \leq \frac{1 - (1/q)}{1 - (1/q)^{n+1-j}} \leq 2 \cdot (1 - 1/q),$$

where the last inequality holds for all $j \leq n - \log^2 n$. To follows that

$$\mathbb{E} C_{2r}(\Pi_n) = \sum_{j=1}^{n} \mathbb{P}(A_j) \leq 2n(1 - 1/q) + o(n),$$

dividing by n, sending $n \to \infty$ and recalling (15), shows $\mu_{2r} \leq 2(1 - 1/q)$. In particular $\lim_{q \downarrow 1} \mu_{2r} = 0$.

We now proceed to show that $(\mu_2, \mu_4, \mu_6 \dots) \to (1/2, 0, 0, \dots)$ as $q \to \infty$. By Part (ii) of Theorem 1.5 it is enough to show that $\mu_2 \to 1/2$ as $q \to \infty$. Let $\Sigma \sim \text{Mallows}(\mathbb{Z}, 1/q)$. We have

$$2\mu_2 = \mathbb{P}[\Sigma(0) = 0]^2 + 2\sum_{i=1}^{\infty} \mathbb{P}[\Sigma(0) = i]^2 = \mathbb{P}[\Sigma(0) = 0]^2 + O\left(\sum_{i \geq 1} q^{-2i}\right).$$

The last equality is due to Part (vi) of Lemma 2.2. By Theorem 1.4 the above is

$$(1 - 2/q + O(1/q^2))^2 + O(1/q^2) = 1 - 4/q + O(1/q^2), \qquad \text{as } q \to \infty.$$

8 | THE PROOF OF THEOREM 1.6

8.1 | The proof Part (i) of Theorem 1.6

Let $\Sigma \sim \text{Mallows}(\mathbb{Z}, 1/q)$ with $0 < 1/q < 1$. By definition of c_e we have

$$c_e := \sum_{i \in \mathbb{Z}} \mathbb{P}[(r \circ \Sigma)(i) = i] = \sum_{i=-\infty}^{\infty} \mathbb{P}[\Sigma(i) = -i].$$

By the remarks following Lemma 2.2, $\Sigma(i) - i \stackrel{d}{=} \Sigma(0)$. Thus

$$c_e = \sum_{i=-\infty}^{\infty} \mathbb{P}[\Sigma(0) = -2i] = \mathbb{P}[\Sigma(0) \text{ even}].$$

Similarly

$$c_o = \sum_{i \in \mathbb{Z}} \mathbb{P}[(\rho \circ \Sigma)(i) = i] = \sum_{i \in \mathbb{Z}} \mathbb{P}[\Sigma(i) = 1 - i] = \sum_{i \in \mathbb{Z}} \mathbb{P}[\Sigma(0) = 1 - 2i] = \mathbb{P}[\Sigma(0) \text{ odd}].$$

8.2 | The proof of Part (ii) of Theorem 1.6

We will make use of the following relatively elementary observation.

Lemma 8.1. *Let $0 < q < 1$ and $\Sigma \sim \text{Mallows}(\mathbb{Z}, q)$. Then*

$$q \, \mathbb{P}\left[\Sigma(0) = j\right] \leq \mathbb{P}\left[\Sigma(0) = j+1\right] \leq \frac{1}{q} \, \mathbb{P}\left[\Sigma(0) = j\right]. \tag{16}$$

Proof. Fix some $j \in \mathbb{Z}$, let $n > j$, $I_n = \{-n, \ldots, n\}$ and $\Sigma_{2n+1} \sim \text{Mallows}(I_n, q)$. Let A be the set of all permutations σ of I_n such that $\sigma(0) = j$, and B be the set of all such permutation σ with $\sigma(0) = j + 1$. Let $\sigma_{j,j+1} = (j \;\; j+1)$ be the permutation of I_n swapping j and $j+1$. Then the map $\phi(\sigma) = \sigma_{j,j+1} \circ \sigma$ is a bijection from A to B. To see this, note that $\phi(\phi(\sigma)) = \sigma$ as $\sigma_{j,j+1}$ is its own inverse. Moreover, if $\sigma(0) = j$, then $\phi(\sigma(0)) = \sigma_{j,j+1}(j) = j + 1$. We claim that $|inv(\sigma) - inv(\phi(\sigma))| = 1$: Let $k = \sigma^{-1}(j+1) \neq 0$. All inversions (a, b) with $\{a, b\} \cap \{0, k\} = \emptyset$ are still inversions in $\phi(\sigma)$ as their images are unchanged in $\phi(\sigma)$. Inversions (a, b) of σ with $|\{a, b\} \cap \{0, k\}| = 1$ are also unchanged, as for all $s \notin \{0, k\}$ we have $\Sigma_{2n+1}(s) < j$ if and only if $\Sigma_{2n+1}(s) < j + 1$. So the only ordered pair that can be an inversion of exactly one of σ and $\phi(\sigma)$ is either $(0, k)$ or $(k, 0)$, depending on whether or not $0 < k$. Thus we have

$$\mathbb{P}\left[\Sigma_{2n+1}(0) = j+1\right] = \sum_{\sigma \in A} \mathbb{P}\left[\phi(\sigma)\right] \leq \frac{1}{q} \sum_{\sigma \in A} \mathbb{P}\left[\sigma\right] = \frac{1}{q} \, \mathbb{P}\left[\Sigma_{2n+1}(0) = j\right].$$

The bound $\mathbb{P}\left[\Sigma_{2n+1}(0) = j+1\right] \geq q \, \mathbb{P}\left[\Sigma_{2n+1}(0) = j\right]$ follows in the same manner. As $\Sigma_{I_n} \to \Sigma$ with probability 1 by Lemma 2.2, Part (ii), the inequality in (16) holds. ∎

The last lemma allows us to give short proof of the following explicit bounds on c_e, c_o, that will immediately imply the value of the $q \downarrow 1$ limits equals $1/2$.

Lemma 8.2. *For all $q > 1$ we have*

$$\frac{1}{1+q} \leq c_e, c_o \leq \frac{q}{1+q}.$$

Proof. The second equality will follow from the first as $c_e + c_o = 1$. Let $q > 1$ and $\Sigma \sim \text{Mallows}(\mathbb{Z}, 1/q)$. By Lemma 8.1 we have $\frac{1}{q}\mathbb{P}\left[\Sigma(0) = j\right] \leq \mathbb{P}\left[\Sigma(0) = j+1\right] \leq q \, \mathbb{P}\left[\Sigma(0) = j\right]$. Thus

$$\frac{1}{q}\mathbb{P}\left[\Sigma(0) \text{ odd}\right] \leq \mathbb{P}\left[\Sigma(0) \text{ even}\right] \leq q \, \mathbb{P}\left[\Sigma(0) \text{ odd}\right].$$

Adding $\mathbb{P}\left[\Sigma(0) \text{ odd}\right]$ and rearranging gives the two inequalities

$$\mathbb{P}\left[\Sigma(0) \text{ odd}\right] \geq \frac{1}{1+q}\mathbb{P}\left[\Sigma(0) \text{ odd or even}\right] = \frac{1}{1+q},$$
$$\mathbb{P}\left[\Sigma(0) \text{ odd}\right] \leq \frac{q}{1+q}\mathbb{P}\left[\Sigma(0) \text{ odd or even}\right] = \frac{q}{1+q}.$$

The same bounds hold for $\mathbb{P}\left[\Sigma(0) \text{ even}\right]$ as $\mathbb{P}\left[\Sigma(0) \text{ even}\right] = 1 - \mathbb{P}\left[\Sigma(0) \text{ odd}\right]$. ∎

As mentioned, from Lemma 8.2 it follows immediately that

$$\lim_{q \downarrow 1} c_e = \lim_{q \downarrow 1} c_o = \frac{1}{2}.$$

It remains to prove the following asymptotic expressions for c_e, c_o.

Lemma 8.3. $c_o = 1 - 2/q + O(1/q^2)$ *and* $c_e = 2/q + O(1/q^2)$ *as* $q \to \infty$.

Proof. The second statement again follows from the first by $c_e + c_o = 1$. Let $q > 1$ and $\Sigma \sim \text{Mallows}(\mathbb{Z}, 1/q)$. We have

$$c_o = \mathbb{P}\left[\Sigma(0) \text{ even}\right] = \mathbb{P}\left[\Sigma(0) = 0\right] + \sum_{i=1}^{\infty} \mathbb{P}\left[|\Sigma(0)| = 2i\right].$$

The quantity $m_1 = m_1(1/q)$ is equal to $\mathbb{P}\left[\Sigma(0) = 0\right]$ by Theorem 1.4. By Part (iii) of Theorem 1.4 it has asymptotic expansion $1 - 2/q + O(1/q^2)$ as $q \to \infty$. By Part (vi) of Lemma 2.2, the sum on the right hand side above is $O(1/q^2)$ as $q \to \infty$. Thus $c_o = 1 - 2/q + O(1/q^2)$. ∎

9 | THE PROOF OF THEOREM 1.7

We will prove the following more detailed result that implies Theorem 1.7.

Proposition 9.1. *For* $0 < q < 1$ *we have, as* $k \to \infty$

(i) $\mathbb{P}(C_1(\rho \circ \Sigma) \geq 2k) = \Omega\left(q^{\binom{2k}{2}} \cdot (1-q)^{2k}\right),$

(ii) $\mathbb{P}(C_1(r \circ \Sigma) \geq 2k) = o\left(q^{\binom{2k}{2}} \cdot (1-q)^{2k}\right).$

(iii) $\mathbb{P}(C_1(r \circ \Sigma) \geq 2k+1) = \Omega\left(q^{\binom{2k+1}{2}} \cdot (1-q)^{2k+1}\right),$

(iv) $\mathbb{P}(C_1(\rho \circ \Sigma) \geq 2k+1) = o\left(q^{\binom{2k+1}{2}} \cdot (1-q)^{2k+1}\right).$

Proof of Proposition 9.1, Parts (i) and (iii). The proofs of (i) and (iii) are very similar. We start with (i). Consider $\Pi_n \sim \text{Mallows}(I_n, q)$ with $I_n := \{-n, \ldots, n\}$. By Part (ii) of 2.2, it suffices to show that $\liminf_{n \to \infty} \mathbb{P}((\rho \circ \Pi_n)(i) = i$ for $i = -k+1, \ldots, k)$ is lower bounded by const $\cdot q^{\binom{2k}{2}} \cdot (1-q)^k$.

For $J \subseteq \mathbb{Z}$ let S_J denote the set of all permutations of J. Consider the set of all permutations $\pi \in S_{I_n}$ constructed as follows. We pick arbitrary permutations $\sigma \in S_{\{-n, \ldots, -k\}}, \sigma' \in S_{\{k+1, \ldots, n\}}$ and set

$$\pi(i) := \begin{cases} 1-i & \text{if } -k+1 \leq i \leq k, \\ \sigma(i) & \text{if } -n \leq i \leq -k, \\ \sigma'(i) & \text{if } k+1 \leq i \leq n. \end{cases}$$

Notice such a permutation satisfies

$$C_1(\rho \circ \pi) = 2k, \quad \text{and} \quad \text{inv}(\pi) = \text{inv}(\sigma) + \text{inv}(\sigma') + \binom{2k}{2}.$$

We have

$$\mathbb{P}(C_1(\rho \circ \Pi_n) \geq 2k) \geq \mathbb{P}(\Pi_n(i) = 1-i \text{ for all } -k+1 \leq i \leq k)$$

$$\geq \frac{1}{Z(2n+1,q)} \cdot \left(\sum_{\sigma \in S_{\{-n,\ldots,-k\}}} \sum_{\sigma\prime \in S_{\{k+1,\ldots,n\}}} q^{\text{inv}(\sigma)+\text{inv}(\sigma\prime)+\binom{2k}{2}} \right)$$

$$= \frac{Z(n-k+1,q) \cdot Z(n-k,q)}{Z(2n+1,q)} \cdot q^{\binom{2k}{2}}.$$

Now we recall that

$$Z(m,q) = \prod_{i=1}^{m} \frac{1-q^i}{1-q} = \Theta\left(\frac{1}{(1-q)^m}\right),$$

as $m \to \infty$. Here we use that $1 > \prod_{i=1}^{m}(1-q^i) > \prod_{i=1}^{\infty}(1-q^i) > 0$. (This last inequality can be easily seen using the Taylor expansion $\log(1-x) = x + O(x^2)$.) So

$$\frac{Z(n-k+1,q) \cdot Z(n-k,q)}{Z(2n+1,q)} = \Theta\left((1-q)^{2k}\right),$$

and Part (i) follows.

The proof of (iii) is essentially the same as the proof of (i). Now we put the elements of $\{-k,\ldots,k\}$ in reverse order, and put arbitrary permutations on $\{-n,\ldots,-k-1\}$ and $\{k+1,\ldots,n\}$ and the proof carries through with only minor adaptations in notation. ∎

We next turn attention to Part (ii) of Proposition 9.1. This proof is a bit more involved, and we break it down into several steps. The first step is the following observation.

Lemma 9.2. *For $0 < q < 1$ we have*

$$\mathbb{P}\left(\text{sign}(\Sigma(i)) \neq \text{sign}(i) \text{ for some } i \text{ with } |i| \geq k^3\right) = o\left(q^{\binom{2k}{2}} \cdot (1-q)^{2k}\right),$$

as $k \to \infty$.

Proof. This immediately follows by the result on displacements $\mathbb{P}(|\Sigma(i)-i| > m) = \Theta(q^m)$ as in Lemma 2.2 part (vi) and the union bound. ∎

By this last lemma, when determining the probability $\mathbb{P}(C_1(r \circ \Sigma) \geq 2k)$, we can restrict attention to the event that there are $2k$ points $x_1,\ldots,x_{2k}$ in the interval $(-k^3,k^3)$ such that $\Sigma(x_1) = -x_1,\ldots,\Sigma(x_{2k}) = -x_{2k}$.

For $0 < i_1 < i_2 < \cdots < i_\ell$ and $0 < j_1 < j_2 < \cdots < j_r$ let us write

$$p_{\underline{i},\underline{j}} := \mathbb{P}\left(\Sigma(-i_1) = i_1,\ldots,\Sigma(-i_\ell) = i_\ell, \Sigma(j_1) = -j_1,\ldots,\Sigma(j_r) = -j_r\right).$$

We point out that, as Σ^{-1} and Σ follow the same distribution, we have

$$p_{\underline{j},\underline{i}} = p_{\underline{i},\underline{j}}. \tag{17}$$

Lemma 9.3. *There is a constant c such that we have*

$$p_{\underline{i},\underline{j}} \leq (1-q)^{\ell+r} \cdot q^{\Psi(\underline{i},\underline{j})},$$

where

$$\Psi(\underline{i}, \underline{j}) := \binom{\ell}{2} + \binom{r}{2} + 2\sum_{a=1}^{\ell} i_a + \frac{1}{2} \sum_{c \leq a \leq r-c} \min(a, r-a) \cdot (j_a - j_{a-1} - 1).$$

Proof. We are again going to consider $\Pi_n \sim \text{Mallows}(I_n, q)$ with $I_n = \{-n, \ldots, n\}$ and n large. For notational convenience we set

$$A_n := \{\Pi_n(j_1) = -j_1, \ldots, \Pi_n(j_r) = -j_r\},$$

$$B_n := \{\Pi_n(-i_1) = i_1, \Pi_n(-i_2) = i_2, \ldots, \Pi_n(-i_\ell) = i_\ell\}.$$

By Mallows' iterative procedure for generating Π_n we have

$$\mathbb{P}(\Pi_n(i) = j | \Pi_n(-n) = x_{-n}, \ldots, \Pi_n(i-1) = x_{i-1}) = \begin{cases} \left(\frac{1-q}{1-q^{n-i+1}}\right) \cdot q^{j-i+k} & \text{if } j \notin \{x_{-n}, \ldots, x_{i-1}\}, \\ 0 & \text{otherwise.} \end{cases},$$

where

$$k = k(x_{-n}, \ldots, x_{i-1}) := |\{a : x_a > j\}|.$$

It follows that

$$\mathbb{P}(\Pi_n(-i_\ell) = i_\ell) \leq \left(\frac{1-q}{1-q^{n+i_\ell+1}}\right) q^{2i_\ell},$$

$$\mathbb{P}(\Pi_n(-i_{\ell-1}) = i_{\ell-1} | \Pi_n(-i_\ell) = i_\ell) \leq \left(\frac{1-q}{1-q^{n+i_{\ell-1}+1}}\right) q^{2i_{\ell-1}+1},$$

$$\mathbb{P}(\Pi_n(-i_{\ell-2}) = i_{\ell-2} | \Pi_n(-i_\ell) = i_\ell, \Pi_n(-i_{\ell-1}) = i_{\ell-1}) \leq \left(\frac{1-q}{1-q^{n+i_{\ell-1}+1}}\right) q^{2i_{\ell-2}+2},$$

$$\vdots$$

$$\mathbb{P}(\Pi_n(-i_1) = i_1 | \Pi_n(-i_\ell) = i_\ell, \ldots, \Pi_n(-i_2) = i_2) \leq \left(\frac{1-q}{1-q^{n+i_1+1}}\right) q^{2i_1+\ell-1}.$$

In other words,

$$\mathbb{P}(B_n) \leq (1-q)^\ell \cdot q^{2\sum_{a=1}^{\ell} i_a + \binom{\ell}{2}} \cdot \prod_{a=1}^{\ell} \left(\frac{1}{1-q^{n+i_a+1}}\right).$$

Note that for A_n to hold, it must be the case that

$$\Pi_n[\{-n, \ldots, j_1-1\}] \cap \{-j_1, \ldots, -j_r\} = \emptyset,$$
$$\Pi_n[\{-n, \ldots, j_2-1\}] \cap \{-j_2, \ldots, -j_r\} = \emptyset,$$
$$\vdots$$
$$\Pi_n[\{-n, \ldots, j_r-1\}] \cap \{-j_r\} = \emptyset.$$

In particular, when choosing the image of j_1 we must "skip" at least the first $r - 1$ available numbers, when choosing the image of j_2 we must "skip" at least $r - 2$ available numbers, and so on.

We let c be a large constant to be specified later on. Suppose that, for some $c \leq a \leq r - c$, there exists a $j_a < j < j_{a+1}$. If both A_n and $\Pi_n(j) > -j_a$ are to hold then when determining $\Pi_n(j)$ we must skip over the (still available) numbers $-j_r, \ldots, -j_{a+1}$. That is, we skip over at least the first $r - a$ of the available numbers. If on the other hand, A_n and $\Pi_n(j) < -j_a$ are to hold then when determining each of $j_1, \ldots, j_a$ must have skipped an additional available number.

It follows that

$$
\begin{aligned}
\mathbb{P}(A_n | B_n) &\leq \left(\frac{1-q}{1-q^{n-j_1+1}} \right) q^{r-1} \cdot \left(\frac{1-q}{1-q^{n-j_2+1}} \right) q^{r-2} \cdots \left(\frac{1-q}{1-q^{n-j_r+1}} \right) q^0 \\
&\quad \cdot \prod_{a=c}^{r-c} \prod_{j_{a-1} < j < j_a} \left(\frac{q^{r-a} + q^a}{1 - q^{n-j+1}} \right) \\
&\leq (1-q)^r q^{\binom{r}{2} + \frac{1}{2} \sum_{a=c}^{r-c} \min(a, r-a) \cdot (j_a - j_{a-1} - 1)} \cdot \prod_{a=1}^{i_r} \left(\frac{1}{1 - q^{n-a+1}} \right),
\end{aligned}
$$

where in the last line we assume without loss of generality that c has been chosen large enough so that $q^{c/2} < 1/2$. (Which implies that

$$
q^{r-a} + q^a \leq 2q^{\min(a, r-a)} \leq q^{\frac{1}{2} \min(a, r-a)},
$$

for all $c \leq a \leq r - c$.)

Combining the bounds on $\mathbb{P}(A_n | B_n)$ and $\mathbb{P}(B_n)$ gives

$$
\mathbb{P}(A_n \cap B_n) \leq (1-q)^{\ell+r} \cdot q^{\Psi(\underline{i}, \underline{j})} \cdot \prod_{a=-i_\ell}^{j_r} \left(\frac{1}{1 - q^{n-a+1}} \right).
$$

Since

$$
\mathbb{P}(\Sigma(-i_1) = i_1, \ldots, \Sigma(-i_\ell) = i_\ell, \Sigma(j_1) = -j_1, \ldots, \Sigma(j_r) = -j_r) \leq \limsup_{n \to \infty} \mathbb{P}(A_n \cap B_n),
$$

the inequality in the lemma follows. ∎

We make an additional definition:

$$
p_{\underline{i}, \underline{j}}^{x \to y} := \mathbb{P}\left(\Sigma(x) = y, \Sigma(-i_1) = i_1, \ldots, \Sigma(-i_\ell) = i_\ell, \Sigma(j_1) = -j_1, \ldots, \Sigma(j_r) = -j_r \right).
$$

We observe that, analogously to (17), we have

$$
p_{\underline{j}, \underline{i}}^{y \to x} = p_{\underline{i}, \underline{j}}^{x \to y}.
$$

Since $r \circ \Sigma \circ r$ (the map $i \mapsto -\Sigma(-i)$) also has the same distribution as Σ, we have in addition that

$$
p_{\underline{j}, \underline{i}}^{-x \to -y} = p_{\underline{i}, \underline{j}}^{x \to y}. \tag{18}
$$

Lemma 9.4. *We have*

(i) If $x < 0$ and $y > 0$ are such that $|x|, y \notin \{i_1, \ldots, i_\ell\}$ then

$$p_{\underline{i},\underline{j}}^{x \to y} \leq (1-q)^{\ell+r} q^{\Psi(\underline{i},\underline{j})+|x|+y}.$$

(ii) If $x > 0$ and $y < 0$ are such that $x, |y| \notin \{j_1, \ldots, j_r\}$ and $x, |y| > j_{\lceil r/2 \rceil}$ then

$$p_{\underline{i},\underline{j}}^{x \to y} \leq (1-q)^{\ell+r} q^{\Psi(\underline{i},\underline{j})+r/4}.$$

Proof. We start with the proof of (i). The proof is nearly identical to the proof of the previous lemma, and we only mention the changes that need to be made. Now, we define

$$C_n := \{\Pi_n(x) = y\} \cap B_n,$$

and let $m \geq 0$ be such that $i_a > x$ for all $a > m$.

We have

$$\mathbb{P}(\Pi_n(-x) = y | \Pi_n(-i_a) = i_a \text{ for all } a > m) \leq \left(\frac{1-q}{1-q^{n+x}}\right) \cdot q^{x+y},$$

while for $a \leq m$

$$\mathbb{P}(\Pi_n(-i_a) = i_a | \Pi_n(-i_\ell) = i_\ell, \ldots, \Pi_n(-i_{a-1}) = i_{a-1}, \Pi_n(-x) = y) \leq \left(\frac{1-q}{1-q^{n+i_a}}\right) \cdot q^{2i_a+\ell-a}.$$

We can conclude that

$$\mathbb{P}(C_n) \leq (1-q)^{\ell+1} \cdot q^{2\sum_{a=1}^{\ell} i_a + \binom{\ell}{2} + x + y} \cdot \prod_{a=1}^{\ell} \left(\frac{1}{1-q^{n+i_a}}\right) \cdot \left(\frac{1}{1-q^{n+x}}\right).$$

The same reasoning as before shows

$$\mathbb{P}(A_n | C_n) \leq (1-q)^r q^{\binom{r}{2} + \frac{1}{2}\sum_{a=c}^{r-c} \min(a,r-a)\cdot(j_a-j_{a-1}-1)} \cdot \prod_{a=1}^{i_r} \left(\frac{1}{1-q^{n-a}}\right).$$

Hence

$$p_{\underline{i},\underline{j}}^{x \to y} \leq \limsup_{n \to \infty} \mathbb{P}(A_n \cap C_n) \leq (1-q)^{\ell+r+1} q^{\Psi(\underline{i},\underline{j})+x+y},$$

as claimed in (i).

For the proof of (ii), we again proceed similarly. This time we define

$$D_n := \{\Pi_n(j_1) = -j_1, \ldots, \Pi_n(j_r) = -j_r, \Pi_n(x) = y\}.$$

Let $m \geq 0.8r$ be such that $j_1, \ldots, j_m < x < j_{m+1}$ (where the upper bound is void if $m = r$). We can compute $\mathbb{P}(D_n | B_n)$ in a manner analogous to the way we determined $\mathbb{P}(A_n | B_n)$ in the proof of Lemma 9.3. When computing $\mathbb{P}(A_n | B_n)$ in the term for the mth

gap, the choice of $\Pi_n(x)$ contributed a factor $\left(\frac{q^{\frac{1}{2}\min(m,r-m)}}{1-q^{n-x}} \right)$ in case $m \leq r-c$ and contributed a factor one otherwise. If we know that $\Pi_n(x) \leq j_{\lfloor r/2 \rfloor}$ then we can replace this by $q^{r/2}$ as all of $j_1, \ldots, j_{\lfloor r/2 \rfloor}$ must have skipped over the additional available number y. This gives

$$\mathbb{P}(D_n|B_n) \leq \mathbb{P}(A_n|B_n) \cdot q^{0.8r - \frac{1}{2}\min(m,r-m)} \geq q^{r/4}.$$

Hence also

$$p_{\underline{i},\underline{j}}^{x \to y} \leq \limsup_{n \to \infty} \mathbb{P}(D_n \cap A_n) \leq q^{r/2} \cdot \limsup_{n \to \infty} \mathbb{P}(A_n \cap B_n),$$

and the result follows. ∎

For notational convenience, we introduce a notation for the gaps between consecutive entries of $\underline{i}$ and $\underline{j}$:

$$g_1 := i_1 - 1, \quad g_a := i_a - (i_{a-1} + 1) \quad (a = 2, \ldots, \ell),$$

and

$$h_1 := j_1 - 1, \quad h_a := j_a - (j_{a-1} + 1) \quad (a = 2, \ldots, r).$$

Notice that $i_a = g_1 + \cdots + g_a + a$ so that

$$2 \sum_{a=1}^{\ell} i_a = 2 \cdot ((g_1 + 1) + (g_1 + g_2 + 2) + (g_1 + g_2 + g_3 + 3) + \cdots + (g_1 + \cdots + g_\ell + \ell))$$

$$= \ell(\ell + 1) + 2 \sum_{a=1}^{\ell} (\ell + 1 - a) \cdot g_a.$$

This gives the following alternative expression for the upper bound on $p_{\underline{i},\underline{j}}$:

$$\Psi(\underline{i}, \underline{j}) = \ell(\ell + 1) + \binom{\ell}{2} + \binom{r}{2} + 2 \sum_{a=1}^{\ell} (\ell + 1 - a) \cdot g_a + \frac{1}{2} \sum_{c \leq a \leq r-c} \min(a, r - a) \cdot h_a$$

$$= \binom{\ell + r}{2} + \ell(\ell + 1 - r) + 2 \sum_{a=1}^{\ell} (\ell + 1 - a) \cdot g_a + \frac{1}{2} \sum_{c \leq a \leq r-c} \min(a, r - a) \cdot h_a.$$

Next, we establish that the probability that $C_1(r \circ \Sigma) \geq 2k$ and $\Sigma(0) \neq 0$ is small compared to the target expression $(1 - q)^{2k} \cdot q^{\binom{2k}{2}}$.

Lemma 9.5. *We have*

$$\sum_{\ell + r = 2k} \sum_{\substack{0 < i_1 < \cdots < i_\ell < k^3, \\ 0 < j_1 < \cdots < j_r < k^3}} p_{\underline{i},\underline{j}} = o\left((1 - q)^{2k} \cdot q^{\binom{2k}{2}}\right),$$

as $k \to \infty$.

Proof. By (17), the sought sum is at most $2S$ where

$$S := \sum_{\substack{\ell+r=2k, \\ \ell \geq r}} \sum_{\substack{0 < i_1 < \cdots < i_\ell < k^3, \\ 0 < j_1 < \cdots < j_r < k^3}} p_{i,j}.$$

We have

$$S \leq \sum_{\substack{\ell+r=2k, \\ \ell \geq r}} \sum_{\substack{0 \leq g_1, \ldots, g_\ell < k^3, \\ 0 \leq h_1, \ldots, h_r < k^3}} (1-q)^{2k} \cdot q^{\binom{2k}{2} + \ell(\ell+1-r) + 2\sum_{a=1}^{\ell}(\ell+1-a) \cdot g_a + \frac{1}{2} \sum_{c \leq a \leq r-c} \min(a, r-a) \cdot h_a}$$

$$= (1-q)^{2k} \cdot q^{\binom{2k}{2}} \cdot \Bigg(\sum_{\substack{\ell+r=2k, \\ \ell \geq r}} q^{\ell(\ell+1-r)} \cdot \prod_{a=1}^{\ell} \Bigg(\sum_{0 \leq g_a < k^3} q^{(\ell+1-a)g_a} \Bigg) \cdot \prod_{a=1}^{c-1} \Bigg(\sum_{0 \leq h_a < k^3} 1 \Bigg)$$

$$\cdot \prod_{c \leq a < r/2} \Bigg(\sum_{0 \leq h_a < k^3} q^{\frac{1}{2}ah_a} \Bigg) \cdot \prod_{r/2 \leq a \leq r-c} \Bigg(\sum_{0 \leq h_a < k^3} q^{\frac{1}{2}(r-a)h_a} \Bigg) \cdot \prod_{a=r-c+1}^{r} \Bigg(\sum_{0 \leq h_a < k^3} 1 \Bigg) \Bigg).$$

Now, we remark that

$$\prod_{a=1}^{\ell} \Bigg(\sum_{0 \leq g_a < k^3} q^{(\ell+1-a)g_a} \Bigg) \leq \prod_{a=1}^{\ell} \Bigg(\frac{1}{1-q^{\ell+1-a}} \Bigg) \leq \prod_{b=1}^{\infty} \Bigg(\frac{1}{1-q^b} \Bigg) < \infty,$$

and analogously

$$\prod_{c \leq a < r/2} \Bigg(\sum_{0 \leq h_a < k^3} q^{\frac{1}{2}ah_a} \Bigg), \quad \prod_{r/2 \leq a \leq r-c} \Bigg(\sum_{0 \leq h_a < k^3} q^{\frac{1}{2}(r-a)h_a} \Bigg) \leq \prod_{b=1}^{\infty} \Bigg(\frac{1}{1-\left(\sqrt{q}\right)^b} \Bigg) < \infty.$$

Of course we also have $\sum_{0 \leq h_a < k^3} 1 = k^3$.

This gives

$$S = O\Bigg((1-q)^{2k} \cdot q^{\binom{2k}{2}} \cdot k^{6c} \cdot \Bigg(\sum_{\substack{\ell+r=2k, \\ \ell \geq r}} q^{\ell(\ell+1-r)} \Bigg) \Bigg).$$

As $\ell + r = 2k, \ell \geq r$ implies that $\ell(\ell + 1 - r) \geq k$ we have

$$\sum_{\substack{\ell+r=2k, \\ \ell \geq r}} q^{\ell(\ell+1-r)} \leq (2k+1)q^k.$$

We find that

$$S = O\Bigg((2k+1) \cdot k^{6c} \cdot (1-q)^{2k} \cdot q^{\binom{2k}{2}+k} \Bigg) = o\Bigg((1-q)^{2k} \cdot q^{\binom{2k}{2}+k} \Bigg),$$

as required. ∎

We proceed by showing that the the probability that **(a)** $C_1(r \circ \Sigma) \geq 2k$, and; **(b)** $\Sigma(0) = 0$, and; **(c)** the number of fixed points of $r \circ \Sigma$ below zero differs by more than one from the number of fixed points of $r \circ \Sigma$ above zero, is small compared to the target expression $(1-q)^{2k} \cdot q^{\binom{2k}{2}}$.

Lemma 9.6. *We have*

$$\sum_{\substack{\ell+r=2k-1, \\ |\ell-r|>1}} \sum_{\substack{0<i_1<\cdots<i_\ell<k^3, \\ 0<j_1<\cdots<j_r<k^3}} p_{\underline{i},\underline{j}} = o\left((1-q)^{2k} \cdot q^{\binom{2k}{2}}\right),$$

as $k \to \infty$.

Proof. Arguing as in the previous lemma, the sought sum is at most

$$k^{O(1)} \cdot (1-q)^{2k-1} \cdot q^{\binom{2k-1}{2}} \cdot \sum_{\substack{\ell+r=2k-1, \\ \ell \geq k+1}} q^{\ell(\ell+1-r)}.$$

We remark that when $\ell + r = 2k - 1, \ell \geq k + 1, r \leq k - 2$ we have

$$\binom{2k-1}{2} + \ell(\ell+1-r) \geq \binom{2k-1}{2} + 3k + 3 = \binom{2k}{2} + k + 4.$$

In particular the sought sum is at most

$$k^{O(1)} \cdot (1-q)^{2k} q^{\binom{2k}{2}+k} = o\left((1-q)^{2k} q^{\binom{2k}{2}}\right).$$

(The number of choices of ℓ, r with $\ell + r = 2k, \ell > r + 1$ is $O(k)$ and is absorbed in the polynomial term $k^{O(1)}$.) ∎

By the results so far, we can restrict attention to the situation where $\Sigma(0) = 0$ and $\ell = k, r = k - 1$. Next, we establish that the contribution from the situation in which $\Sigma(-i) \neq i$ for some $1 \leq i \leq 0.9k$ is negligible.

Lemma 9.7. *We have*

$$\sum_{\substack{0<i_1<\cdots<i_k<k^3, \\ 0<j_1<\cdots<j_{k-1}<k^3, \\ i_a \neq a \;\; \text{for some} \;\; 1 \leq a \leq 0.9k}} p_{\underline{i},\underline{j}} = o\left((1-q)^{2k} \cdot q^{\binom{2k}{2}}\right),$$

as $k \to \infty$.

Proof. If $i_a \neq a$ for some $1 \leq a \leq 0.9k$ then there is also a $1 \leq a \leq 0.9k$ for which $g_a \neq 0$. For this a we have $2(k+1-a) \cdot g_a \geq 0.2k$. Also note that

$$\binom{2k-1}{2} + k(k+1-(k-1)) = \binom{2k-1}{2} + 2k = \binom{2k}{2} + 1.$$

Arguing as in previous lemmas, it follows that the sought sum is at most

$$k^{O(1)} \cdot (1-q)^{2k-1} \cdot q^{\binom{2k}{2}+0.2k} = o\left((1-q)^{2k} q^{\binom{2k}{2}}\right).$$

(The polynomial term $k^{O(1)}$ also absorbs the k ways of choosing an index a for which $g_a \neq 0$ and the k^3 ways of choosing a value for g_a.) ∎

Lemma 9.8. *We have*

$$\sum_{\substack{0 < i_1 < \cdots < i_k < k^3, \\ 0 < j_1 < \cdots < j_{k-1} < k^3, \\ j_a \neq j_{a-1}+1 \ \text{for some} \ 0.1k \leq a \leq 0.9k}} p_{\underline{i},\underline{j}} = o\left((1-q)^{2k} \cdot q^{\binom{2k}{2}}\right),$$

as $k \to \infty$.

Proof. If $j_a \neq j_{a-1} + 1$ for some $0.1k \leq a \leq 0.9k$ then, provided k is sufficiently large, $c \leq a \leq k - 1 - c$ and moreover $\min(a, k - 1 - a) \cdot h_a \geq 0.05k$. Arguing as in previous lemmas, the sought sum is therefore at most

$$k^{O(1)} \cdot (1-q)^{2k-1} \cdot q^{\binom{2k}{2}+0.05k} = o\left((1-q)^{2k} q^{\binom{2k}{2}}\right).$$

(The polynomial term also absorbs the $0.9k$ ways of choosing a and the k^3 ways of choosing h_a.) ∎

Lemma 9.9. *We have*

$$\sum_{\substack{0 < i_1 < \cdots < i_k < k^3, \\ 0 < j_1 < \cdots < j_{k-1} < k^3, \\ h_1 + \cdots + h_{\lfloor 0.1k \rfloor} \geq 2}} p_{\underline{i},\underline{j}} = o\left((1-q)^{2k} \cdot q^{\binom{2k}{2}}\right),$$

as $k \to \infty$.

Proof. For each $p_{\underline{i},\underline{j}}$ occurring in the sought sum, we have

$$p_{\underline{i},\underline{j}} = p_{\underline{j},\underline{i}} \leq (1-q)^{2k} q^{\binom{2k-1}{2}+2\sum_{a=1}^{k-1}(r-a)h_a + \frac{1}{2}\sum_{c \leq a \leq k-c} \min(a,k-a) \cdot g_a},$$

by Lemma 9.3. (Notice the roles of $\ell = k$ and $r = k - 1$ and g_a and h_a are switched. Notice in particular $(k - 1) \cdot ((k - 1) + 1 - k) = 0$.) The sought sum is thus at most

$$S := \sum_{\substack{1 \leq x < y \leq 0.1k \\ h_x+h_y \geq 2}} \sum_{\substack{0 \leq h_1, \ldots, h_{k-1} < k^3, \\ 0 \leq g_1, \ldots, g_k < k^3,}} (1-q)^{2k} q^{\binom{2k-1}{2}+2\sum_{a=1}^{k-1}(r-a)\Delta_a + \frac{1}{2}\sum_{c \leq a \leq k-c} \min(a,k-a) \cdot \Delta'_a}.$$

By computations similar to those in previous proofs:

$$S = k^{O(1)} \cdot (1-q)^{2k} \cdot q^{\binom{2k-1}{2}+4 \cdot 0.9k}$$

$$= k^{O(1)} \cdot (1-q)^{2k} \cdot q^{\binom{2k}{2}+1.6k}$$

$$= o\left((1-q)^{2k} \cdot q^{\binom{2k}{2}}\right).$$

∎

We next observe that if $\Sigma(0) = 0$ and $\Sigma(-i_1) = i_1, \dots, \Sigma(-i_k) = i_k$ and $\Sigma(j_1) = -j_1, \dots, \Sigma(j_{k-1}) = -j_{k-1}$ then there must be some $x \geq 1$ and $y \leq -1$ such that $\Sigma(x) = y$ and $x \notin \{j_1, \dots, j_{k-1}\}$, and $y \notin \{-j_1, \dots, -j_{k-1}\}$. (Since Σ is "balanced.")

This will allow us to improve over our previous bounds on $p_{\underline{i},\underline{j}}$.

Proof of Proposition 9.1, Part (ii). By the previous lemmas and the observation immediately preceding the present proof, it suffices to show that

$$S := \sum p_{\underline{i},\underline{j}}^{x \to y} = o\left((1-q)^{2k} \cdot q^{\binom{2k}{2}}\right),$$

as $k \to \infty$, where the sum is over all $0 < j_1 < \cdots < j_k \leq k^3$ and $0 < i_1 < \cdots < i_{k-1} \leq k^3$ and $0 < x < k^3$ and $-k^3 < y < 0$ such that

- $i_a = a$ for all $1 \leq a \leq 0.9k$, and;
- $j_a = j_{a-1} + 1$ for all $0.1k \leq a \leq 0.9k$, and;
- there is at most one $1 \leq a \leq 0.1k$ such that $j_a \neq j_{a-1} + 1$ and if such an a exists then we have $j_a = j_{a-1} + 2$, and;
- $x, |y| \notin J := \{j_1, \dots, j_{k-1}\}$.

We first notice that if $\underline{i}, \underline{j}, x, y$ are as described then either **(a)** $\{j_1, \dots, j_{\lfloor 0.9k \rfloor}\} = \{1, \dots, \lfloor 0.9k \rfloor\}$ or **(b)** $\{j_1, \dots, j_{\lfloor 0.9k \rfloor}\} = \{1, \dots, \lfloor 0.9k \rfloor + 1\} \setminus \{z\}$ for some $1 \leq z \leq 0.1k$.

In case **(a)** we thus have $x, |y| > j_{\lfloor 0.9k \rfloor}$ and in particular

$$p_{o,\underline{i},\underline{j}}^{x \to y} \leq (1-q)^{2k-1} \cdot q^{\Psi(\underline{i}\underline{j})+r/4}.$$

The corresponding sum $S_{\mathbf{a}}$ thus satisfies

$$S_{\mathbf{a}} := \sum_{\substack{0 < i_1 < \cdots < i_k < k^3, \\ 0 < j_1 < \cdots < j_{k-1} < k^3, \\ i_a = j_a = a \text{ for } 1 \leq a \leq 0.9k}} \sum_{\substack{0.9k \leq x < k^3, \\ -k^3 < y \leq 0.9k, \\ x, |y| \notin J}} p_{\underline{i},\underline{j}}^{x \to y}$$

$$\leq \sum_{\underline{i},\underline{j}} \sum_{\substack{0.9k \leq x < k^3, \\ -k^3 < y \leq 0.9k, \\ x, |y| \notin J}} (1-q)^{2k-1} \cdot q^{\Psi(\underline{i}\underline{j})+r/4}$$

$$\leq k^6 \sum_{\underline{i}\underline{j}} (1-q)^{2k-1} \cdot q^{\Psi(\underline{i}\underline{j})+r/4}$$

$$\leq k^6 \cdot (1-q)^{2k-1} \cdot q^{\binom{2k}{2}+r/4} \sum_{\substack{0 \leq g_1, \dots, g_k < k^3, \\ 0 \leq h_1, \dots, h_{k-1} < k^3}} q^{2\sum_{a=1}^{k}(k+1-a)g_a + \frac{1}{2}\sum_{c \leq a \leq k-1-c} \min(a, k-1-a)h_a}$$

$$\leq k^{O(1)}(1-q)^{2k-1}q^{\binom{2k}{2}+r/4}$$

$$= o\left((1-q)^{2k}q^{\binom{2k}{2}}\right).$$

(Using the familiar observation that $\binom{2k-1}{2} + k(k+1-(k-1)) = \binom{2k}{2} + 1$ in the fourth line, and computations as in previous lemmas.)

In case **(b)** it is possible that either **(b-1)** $x, |y| \geq 0.9k$, or **(b-2)** $x = z$ and $y = -z$, or **(b-3)** $x > 0.9k$ and $y = -z$, or **(b-4)** $x = z$ and $y < -0.9k$.

In the case **(b-1)** the same bound on $p_{o,i,j}^{x \to y}$ applies as in the case **(a)**, and via similar computations we obtain that the corresponding contribution to the sum satisfies

$$S_{\mathbf{b-1}} := \sum_{1 \leq b \leq 0.1k} \sum_{\substack{0 < i_1 < \cdots < i_k < k^3, \\ 0 < j_1 < \cdots < j_{k-1} < k^3, \\ i_a = a \text{ for } 1 \leq a \leq 0.9k, \\ j_a = a \text{ for } a \leq b, \\ j_a = a+1 \text{ for } b < a \leq 0.9k}} \sum_{\substack{0.9k \leq x < k_3, \\ -k^3 < y \leq 0.9k, \\ x,|y| \notin J}} p_{i,j}^{x \to y}$$

$$\leq k^{O(1)} (1-q)^{2k-1} q^{\binom{2k}{2} + r/4}$$

$$= o\left((1-q)^{2k} q^{\binom{2k}{2}} \right).$$

In case **(b-2)** we have $p_{i,j}^{x \to y} = p_{i,j'}$, where $j_1' = 1, \ldots, j_{\lfloor 0.9k \rfloor}' = \lfloor 0.9k \rfloor$ and $j_a' = j_{a-1}$ for $0.9k \leq a \leq k-1$. (i.e., we've filled the "gap at z" and made a vector of length k.) So the sum corresponding to case **(b-2)** is

$$S_{\mathbf{b-2}} \leq \sum_{\substack{0 < i_1 < \cdots < i_k < k^3, \\ 0 < j_1' < \cdots < j_k' < k^3}} p_{i,j'} = k^{O(1)} (1-q)^{2k} q^{\binom{2k}{2} + k} = o\left((1-q)^{2k} q^{\binom{2k}{2}} \right).$$

To deal with case **(b-3)** we note that in this case

$$p_{i,j}^{x \to y} = p_{j,i}^{y \to x}$$

$$\leq (1-q)^{2k-1} q^{\binom{2k-1}{2} + 2(k-1-x) + x + |y| + 2 \sum_{0.9k \leq a \leq k-1} (k-1-a)h_a + \frac{1}{2} \sum_{0.9k \leq a \leq k-c} (k-a)g_a}$$

$$\leq (1-q)^{2k-1} q^{\binom{2k-1}{2} + 3.7k - 2 + 2 \sum_{0.9k \leq a \leq k-1} (k-1-a)h_a + \frac{1}{2} \sum_{0.9k \leq a \leq k-c} (k-a)g_a}$$

$$\leq (1-q)^{2k-1} q^{\binom{2k}{2} + 1.7k + 2 \sum_{0.9k \leq a \leq k-1} (k-1-a)h_a + \frac{1}{2} \sum_{0.9k \leq a \leq k-c} (k-a)g_a}.$$

Hence, the corresponding sum satisfies

$$S_{\mathbf{b-3}} := \sum_{1 \leq x \leq 0.1k} \sum_{\substack{0 < i_1 < \cdots < i_k < k^3, \\ 0 < j_1 < \cdots < j_{k-1} < k^3, \\ i_a = a \text{ for } 1 \leq a \leq 0.9k, \\ j_a = a \text{ for } a < x, \\ j_a = a+1 \text{ for } x \leq a \leq 0.9k}} \sum_{\substack{-k^3 < y \leq 0.9k, \\ |y| \notin J}} p_{i,j}^{x \to y}$$

$$= o\left((1-q)^{2k} q^{\binom{2k}{2}} \right).$$

Finally, we deal with case **(b-4)**. We now use (18) to see that in this case

$$
\begin{aligned}
p_{\underline{i},\underline{j}}^{x\to y} &= p_{\underline{j},\underline{i}}^{-x\to -y} \\
&\leq (1-q)^{2k-1} q^{\binom{2k-1}{2}+2(k-1-x)+x+|y|+2\sum_{0.9k\leq a\leq k-1}(k-1-a)h_a+\frac{1}{2}\sum_{c\leq a\leq k-c}\min(a,k-a)g_a} \\
&\leq (1-q)^{2k-1} q^{\binom{2k-1}{2}+3.7k-2+2\sum_{0.9k\leq a\leq k-1}(k-1-a)h_a+\frac{1}{2}\sum_{c\leq a\leq k-c}\min(a,k-a)g_a},
\end{aligned}
$$

and hence repeating the computations bounding $S_{\mathbf{b-3}}$ we find

$$
S_{\mathbf{b-4}} := \sum_{-0.1k\leq y\leq 1}\ \sum_{\substack{0<i_1<\cdots<i_k<k^3,\\ 0<j_1<\cdots<j_{k-1}<k^3,\\ i_a=a\quad\text{for}\quad 1\leq a\leq 0.9k,\\ j_a=a\quad\text{for}\quad a<|y|,\\ j_a=a+1\quad\text{for}\quad |y|\leq a\leq 0.9k}}\ \sum_{\substack{0.9k\leq x\leq k^3,\\ x\notin J}} p_{\underline{i},\underline{j}}^{x\to y} = o\left((1-q)^{2k}q^{\binom{2k}{2}}\right).
$$

This establishes that $S = o\left((1-q)^{2k}q^{\binom{2k}{2}}\right)$. Proposition 9.1, Part (ii) is proved. ∎

Proof of Proposition 9.1, Part (iv). The proof proceeds in the same manner as the proof of Proposition 9.1 Part (ii), we highlight here only the differences. We will now define for sequences $0 < i_1 < \cdots < i_\ell$ and $1 < j_1 < \cdots < j_r$ the probability

$$
\tilde{p}_{\underline{i},\underline{j}} = \mathbb{P}\left(\Sigma(-i_1) = i_1 + 1, \ldots, \Sigma(-i_\ell) = i_\ell + 1, \Sigma(j_1) = -j_1 + 1, \ldots, \Sigma(j_r) = -j_r + 1\right).
$$

For sequences satisfying the above we have $(\rho \circ \Sigma)(i_s) = i_s$ and $(\rho \circ \Sigma)(j_s) = j_s$. For $\tilde{p}_{\underline{i},\underline{j}}$ the bound in Lemma 9.3 may be replaced by the stronger bound

$$
\tilde{p}_{\underline{i},\underline{j}} \leq (1-q)^{\ell+r} q^{\ell+r+\Psi(\underline{i},\underline{j})}, \tag{19}
$$

the proof is the same, but we now have

$$
\begin{aligned}
&\mathbb{P}(\Pi_n(i) = j+1 \mid \Pi_n(-n) = x_{-n}, \ldots, \Pi_n(i-1) = x_{i-1}) \\
&\quad = \begin{cases} \left(\frac{1-q}{1-q^{n-i+1}}\right) \cdot q^{j-i+k+1} & \text{if } j \notin \{x_{-n}, \ldots, x_{i-1}\}, \\ 0 & \text{otherwise,} \end{cases}
\end{aligned}
$$

which gives the additional $\ell + r$ term in the exponent of q in 19. Also, as in Lemma 9.5, the contribution from such sequences where $\Sigma(1) \neq 0$ or $\Sigma(0) \neq 1$ is $o\left(q^{\binom{2k+1}{2}} \cdot (1-q)^{2k+1}\right)$. Then we need only consider cases with $\Sigma(1) = 0$ and $\Sigma(0) = 1$, so that we consider the case $\ell + r = 2k - 1$. The remaining estimates for the different such sequences $i_1, \ldots, i_\ell$ and $j_1, \ldots, j_r$ follow in the same manner as the estimates for $r \circ \Sigma$, where now the $q^{\ell+r}$ term in (19) contributes q^{2k-1}. In the above proofs we estimate $\mathbb{P}[\Sigma(0) = 0]$ by a a coupling with $\Pi_n \sim \text{Mallows}(n, q)$, which we can sample element by element. If during this sampling 0 does not yet have a preimage when determining the image of 0, then $\mathbb{P}[\Sigma(0) = 1] \leq q\,\mathbb{P}[\Sigma(0) = 0]$. Subsequently, when determining the image of 1, if $\Sigma(0) = 1$, then $\mathbb{P}[\Sigma(1) = 0]$ equals the probability that $\Sigma(0) = 0$ in the previous step, as 1

now needs to skip over exactly all elements below 0 that are not yet selected. In the estimations of the $r \circ \Sigma$ case, the additional contribution of $q^{\ell+r}q = q^{2k}$ is exactly as needed as $q^{\binom{2k}{2}+2k} = q^{\binom{2k+1}{2}}$. ∎

10 | SUGGESTIONS FOR FURTHER WORK

We expect that much more information can be extracted about the constants c_e, c_o, m_i, μ_{2i} and the probability measures of $C_{2i-1}(r \circ \Sigma)$ and $C_{2i-1}(\rho \circ \Sigma)$ from the explicit expressions for them in terms of q-hypergeometric series that can be obtained from Theorem 5.1 in [17]. There might well be some low-hanging fruit available to someone better versed in q-hypergeometric series than the present authors. In particular, the plot in Figure 3 strongly suggests that $c_o > 1/2 > c_e$ for all $q > 1$, but we have not been able to show this rigorously. We leave it as an open problem for other teams. Other obvious open problems are the (strict) monotonicity of c_o and c_e (which together with our results will imply $c_o > 1/2 > c_e$ for all $q > 1$), of m_1 shown in Figure 1 and of μ_2 shown in Figure 2. The plots clearly suggest these to be (strictly) monotonic in q, but as the reader may have noticed our proofs do not show this to be the case and there does not seem to be any obvious way to adapt the proofs to obtain this.

A curious phenomenon we've observed is that if $q \downarrow 1$ then the expected number of 1-cycles in the limiting distribution tends to $1/2$, which is different from the value of 1 that we get when $q = 1$ (which corresponds to sampling a permutation uniformly at random). At the moment we do not even have a reasonable intuitive explanation for this phenomenon. Clearly something interesting must be going on in the phase change regime when $q = q(n)$ is a function of n that approaches one from above. It is also intriguing that the odd and even cycle counts behave so differently for $q > 1$, the odd cycle counts being "tight" and the even ones being linear in n, but at $q = 1$ there appears to be no trace of this difference. We would be very interested to see an analysis of the regime when $q = q(n) \to 1$ that can shed some light on these phenomena.

Finally, given the Gaussian fluctuations for fixed q, and Poisson fluctuations when q is close to 1, it would be interesting to understand what happens in the intermediate regime, and whether there is a phase transition. Let us note that the techniques of the renewal theoretic techniques used in this article are not very helpful, because the lengths between regenerations grow extremely quickly as $q \to 1$.

REFERENCES

1. L. Addario-Berry and B. Corsini, *The height of Mallows trees*, Ann. Probab. **49** (2021), no. 5, 2220–2271.
2. O. Angel, A. E. Holroyd, T. Hutchcroft, and A. Levy, *Mallows permutations as stable matchings*, Canad. J. Math. **73** (2021), no. 6, 1531–1555.
3. R. Basu and N. Bhatnagar, *Limit theorems for longest monotone subsequences in random Mallows permutations*, Ann. Inst. Henri Poincaré Probab. Stat. **53** (2017), no. 4, 1934–1951.
4. 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** (2005), no. 8, 3013–3029.
5. N. Bhatnagar and R. Peled, *Lengths of monotone subsequences in a Mallows permutation*, Probab. Theory Relat. Fields **161** (2015), no. 3-4, 719–780.

6. P. Billingsley, Probability and measure, Wiley series in probability and mathematical statistics, 3rd ed., John Wiley & Sons, Inc., New York, 1995.

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

8. M. Braverman and E. Mossel. Sorting from noisy information, 2009. http://arxiv.org/abs/0910.1191.

9. H. Crane and S. DeSalvo, *Pattern avoidance for random permutations*, Discrete Math. Theor. Comput. Sci **19** (2017), no. 2, 1–24.

10. H. Crane, S. DeSalvo, and S. Elizalde, *The probability of avoiding consecutive patterns in the Mallows distribution*, Random Struct. Algorithms **53** (2018), no. 3, 417–447.

11. P. Diaconis and A. Ram, *Analysis of systematic scan Metropolis algorithms using Iwahori-Hecke algebra techniques*, Michigan Math. J. **48** (2000), 157–190.

12. X. Fang, H. L. Gan, S. Holmes, H. Huang, E. Peköz, A. Röllin, and W. Tang, *Arcsine laws for random walks generated from random permutations with applications to genomics*, J. Appl. Probab. **58** (2021), no. 4, 851–867.

13. R. G. Gallager, Stochastic processes: Theory for applications, Cambridge University Press, Cambridge, 2013.

14. G. Gasper and M. Rahman, Basic hypergeometric series: 96 (Encyclopedia of mathematics and its applications), 2nd ed., Cambridge University Press, Cambridge, 2004 With a foreword by Richard Askey.

15. A. Gladkich and R. Peled, *On the cycle structure of Mallows permutations*, Ann. Probab. **46** (2018), no. 2, 1114–1169.

16. A. Gnedin and G. Olshanski, *q-exchangeability via quasi-invariance*, Ann. Probab. **38** (2010), no. 6, 2103–2135.

17. A. Gnedin and G. Olshanski, *The two-sided infinite extension of the Mallows model for random permutations*, Adv. Appl. Math. **48** (2012), no. 5, 615–639.

18. W. Gontcharoff, *Sur la distribution des cycles dans les permutations*, C. R. (Doklady) Acad. Sci. URSS (N.S.) **35** (1942), 267–269.

19. A. Gut and S. Janson, *The limiting behaviour of certain stopped sums and some applications*, Scand. J. Stat. **10** (1983), no. 4, 281–292.

20. J. He. *A central limit theorem for descents of a mallows permutation and its inverse*, 2020. https://arxiv.org/abs/2005.09802.

21. J. He. *A central limit theorem for cycles of mallows permutations*, 2021. https://arxiv.org/abs/2112.09789.

22. A. E. Holroyd, T. Hutchcroft, and A. Levy, *Mallows permutations and finite dependence*, Ann. Probab. **48** (2020), no. 1, 343–379.

23. S. Janson, T. Łuczak, and A. Rucinski, Random graphs, Wiley-Interscience series in discrete mathematics and optimization, Wiley-Interscience, New York, 2000.

24. K. Jin, *The length of the longest common subsequence of two independent Mallows permutations*, Ann. Appl. Probab. **29** (2019), no. 3, 1311–1355.

25. V. F. Kolchin, *A problem of the allocation of particles in cells and cycles of random permutations*, Theory Probab. Appl. **16** (1971), no. 1, 74–90.

26. J. G. Propp and D. B. Wilson, "*Coupling from the past*," Markov chains and mixing times, 2nd ed., D. A. Levin, Y. Peres, and E. L. Wilmer (eds.), American Mathematical Society, Providence, RI, 2017.

27. C. L. Mallows, *Non-null ranking models. I*, Biometrika **44** (1957), 114–130.

28. 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.

29. R. G. Pinsky, *Permutations avoiding a pattern of length three under Mallows distributions*, Random Struct. Algorithms **58** (2021), no. 4, 676–690.

30. J. Pitman and W. Tang, *Regenerative random permutations of integers*, Ann. Probab. **47** (2019), no. 3, 1378–1416.

31. R. P. Stanley, Enumerative combinatorics, Cambridge Studies in Advanced Mathematics (49), 2nd ed., Vol **1**, Cambridge University Press, Cambridge, 2012.

32. S. Starr, *Thermodynamic limit for the Mallows model on* S_n, J. Math. Phys **50** (2009), no. 9, 095208.

33. S. Starr and M. Walters, *Phase uniqueness for the Mallows measure on permutations*, J. Math. Phys. **59** (2018), no. 6, 063301.

34. W. Tang, "*Mallows ranking models: maximum likelihood estimate and regeneration*," Proceedings of the 36th International Conference on Machine Learning, ICML 2019, Vol **97**, K. Chaudhuri and R. Salakhutdinov (eds.), Proceedings of Machine Learning Research, Long Beach, CA, 2019, pp. 6125–6134.