PATTERN AVOIDANCE FOR RANDOM PERMUTATIONS

HARRY CRANE AND STEPHEN DESALVO

Abstract. We use techniques from Poisson approximation to prove explicit error bounds on the number of permutations that avoid any pattern. Most generally, we bound the total variation distance between the joint distribution of pattern occurrences and a corresponding joint distribution of independent Bernoulli random variables, which as a corollary yields a Poisson approximation for the distribution of the number of occurrences. We also investigate occurrences of consecutive patterns in random Mallows permutations, of which uniform random permutations are a special case. These bounds allow us to estimate the probability that a pattern occurs any number of times and, in particular, the probability that a random permutation avoids a given pattern.

Contents

1. Introduction

A *permutation* of $[n] = \{1, \ldots, n\}$ is a bijection $\sigma : [n] \to [n]$, $i \mapsto \sigma(i) = \sigma_i$, written $\sigma = \sigma_1 \cdots \sigma_n$. For each $n = 1, 2, \ldots$, we write $\mathcal{S}_n$ to denote the set of permutations of $[n]$. Given permutations $\sigma = \sigma_1 \cdots \sigma_n$ and $\tau = \tau_1 \cdots \tau_m$, we say that σ *avoids* τ if there does not exist a subsequence $1 \le i_1 < \cdots < i_m \le n$ such that $\sigma_{i_1} \cdots \sigma_{i_m}$ is order-isomorphic to τ, and we say that σ *avoids* τ *consecutively* if there is no $j = 1, \ldots, n-m+1$ such that $\sigma_j \sigma_{j+1} \cdots \sigma_{j+m-1}$ is order-isomorphic to τ. Here we study pattern avoidance probabilities for a wide class of random permutations from the Mallows distribution, which is of particular interest in the fields of statistics and probability but special cases provide insights into questions in enumerative and extremal combinatorics.

An *inversion* in $\sigma = \sigma_1 \cdots \sigma_n$ is a pair (i, j), $i < j$, such that $\sigma_i > \sigma_j$. For example, $\sigma = 34125$ has four inversions, $(1, 3), (1, 4), (2, 3), (2, 4)$. We write $\mathrm{inv}(\sigma)$ to denote the set of inversions

Date: March 21, 2017.

1991 *Mathematics Subject Classification.* 05A05; 05A15; 05A16; 60C05.

Key words and phrases. pattern avoidance; Mallows distribution; random permutation; Poisson approximation.

of σ. With Σ_n denoting a random permutation of $[n]$, the *Mallows distribution with parameter* $q \in (0, \infty)$ on $\mathcal{S}_n$ assigns probability

$$P\{\Sigma_n = \sigma\} = P_n^q(\sigma) = q^{|\operatorname{inv}(\sigma)|}/I_n(q), \quad \sigma \in \mathcal{S}_n, \tag{1}$$

where $I_n(q) = \prod_{j=1}^n \sum_{i=0}^{j-1} q^i$ is the *inversion polynomial* and $|\operatorname{inv}(\sigma)|$ is the number of inversions in σ. Note that $q = 1$ corresponds to the uniform distribution on $\mathcal{S}_n$, i.e., $P\{\Sigma_n = \sigma\} = 1/n!$ for all $\sigma \in \mathcal{S}_n$, and is the critical point at which the Mallows family switches from penalizing inversions, $q < 1$, to favoring them, $q > 1$. Results for pattern avoidance probabilities under the uniform distribution translate to enumerative results for the number of pattern avoiding permutations. Combinatorial inequalities bounding the number of pattern avoiding permutations obtained from our techniques are contained in Equations (11) and (14).

The Mallows distribution [28] was introduced as a one-parameter model for rankings that occur in statistical analysis. More recently, Mallows permutations have been studied in the context of the longest increasing subsequence problem [6] and quasi-exchangeable random sequences [19, 20]. Our study of pattern avoidance for this class relates to recent work in the combinatorics literature, which considers the Wilf equivalence classes of the inversion polynomials for permutations that avoid certain sets of patterns [10, 16]. For general values of $q > 0$, we consider the problem of consecutive pattern avoidance for random Mallows permutations, but our main theorems go quite a bit further by establishing explicit error bounds on the entire distribution of the number of occurrences of patterns in a random permutation. In the uniform setting ($q = 1$), our method provides an estimate for the number of permutations with a prescribed number of occurrences as well as the number of permutations with a prescribed number of consecutive occurrences of a given pattern. Our main theorems, therefore, complement prior work by Elizalde & Noy [18], Perarnau [32], and the more recent work by the current authors & Elizalde [14] on consecutive pattern avoidance, as well as Nakamura [31], who used functional equations to enumerate sets with a prescribed number of occurrences of a given pattern. Other related works include [33] for patterns τ of length $n-1$, and [34] for patterns τ of length $n-2$. And it is left as an open problem in [38] to perform the same analysis for patterns τ of length $n-3$.

Our approach also differs from previous work in a few key respects. While most prior work seeks either exact or asymptotic enumeration of the sets that avoid a given pattern or collection of patterns, we use the Chen–Stein Poisson approximation method [12], in particular [2], to bound the total variation distance between the collection of all *dependent* indicator random variables indicating pattern occurrence at a prescribed set of indices, and a joint distribution of *independent* Bernoulli random variables with the same marginal distributions. From these bounds, we can approximate the probability that a random permutation avoids a given pattern, i.e., the pattern occurs zero times, or contains any number of occurrences of that pattern.

We summarize our main theorems in Section 3.

2. Motivation

Restricted permutations fall into two broad classes. The first, more tractable type is of the form $\sigma(a) \neq b$ for $a, b \in [n]$, whose study dates to the classical *problèmes des rencontres* in the early 1700s [15]; see also [5, Chapter 4]. A special case counts the number D_n of *derangements* of $[n]$, i.e., permutations of $[n]$ without fixed points, yielding the asymptotic expression

$$D_n = n! \sum_{i=0}^{n} \frac{(-1)^i}{i!} \sim n!/e \quad \text{as } n \to \infty. \tag{2}$$

Equation (2) can be stated in probabilistic terms by letting Σ_n be a uniform random permutation of $[n]$, i.e., $P\{\Sigma_n = \sigma\} = 1/n!$ for each $\sigma \in \mathcal{S}_n$, in which case

$$P\{\Sigma_n \text{ is a derangement}\} = D_n/n! \sim 1/e \quad \text{as } n \to \infty. \tag{3}$$

See [3, 36] for more thorough treatments involving the cycle structure of random permutations.

We can also derive the expression in (2) by Poisson approximation. With W denoting the number of fixed points in a random permutation of $[n]$, we demonstrate in Section 4.2, see also [5, Chapter 4], that the distribution of W converges in total variation distance to the distribution of an independent Poisson random variable with expected value 1. In addition to the asymptotic value for the probability that a random permutation has no fixed points, this approach bounds the absolute error of probabilities that involve any measurable function of the number of fixed points in a random permutation.

The second type of restriction is *pattern avoidance*, which attracts increasing attention in the modern probability [6, 21, 22] and modern combinatorics literature [7]. Any sequence of distinct positive integers $w = w_1 \cdots w_k$ determines a permutation of $[k]$ by *reduction*: with $\{w_{(1)}, \ldots, w_{(k)}\}_<$ denoting the set of elements listed in increasing order, we define the map $w_{(i)} \mapsto i$, under which w maps to a permutation $\mathrm{red}(w)$ of $[k]$, called the *reduction* of w. For example, $w = 826315$ reduces to $\mathrm{red}(w) = 625314$. We call any fixed $\tau \in \mathcal{S}_m$ a *pattern* and say that $\sigma \in \mathcal{S}_n$ *contains* τ if there exists a subsequence $1 \le i_1 < \cdots < i_m \le m$ such that $\mathrm{red}(\sigma_{i_1} \cdots \sigma_{i_m}) = \tau$. We say $\sigma \in \mathcal{S}_n$ *avoids* τ if it does not contain it. We say that σ *contains* τ *consecutively* if there exists an index $j \in [n - m + 1]$ such that $\mathrm{red}(\sigma_j \sigma_{j+1} \cdots \sigma_{j+m-1}) = \tau$; otherwise, we say σ *avoids* τ *consecutively*. For any pattern τ, we define

$$\begin{aligned} \mathcal{S}_n(\tau) &:= \{\sigma \in \mathcal{S}_n : \sigma \text{ avoids } \tau\} \quad \text{and} \\ \overline{\mathcal{S}}_n(\tau) &:= \{\sigma \in \mathcal{S}_n : \sigma \text{ avoids } \tau \text{ consecutively}\}, \end{aligned} \tag{4}$$

which we extend to any subset $A \subset \bigcup_{n\ge 1} \mathcal{S}_n$ by

$$\begin{aligned} \mathcal{S}_n(A) &:= \{\sigma \in \mathcal{S}_n : \sigma \text{ avoids all } \tau \in A\} \quad \text{and} \\ \overline{\mathcal{S}}_n(A) &:= \{\sigma \in \mathcal{S}_n : \sigma \text{ avoids all } \tau \in A \text{ consecutively}\}. \end{aligned} \tag{5}$$

Much effort has been devoted to exact enumeration of $\mathcal{S}_n(A)$ for certain choices of A, see, e.g., [1, 4, 17, 25]. For the most part, we are interested in sets $\mathcal{S}_n(\tau)$ containing all permutations that avoid a given pattern τ, though our approach extends in a straightforward manner for more general sets A.

Attempts to enumerate $\mathcal{S}_n(\tau)$ are notoriously difficult for patterns of fixed length larger than 3. Knuth [26] initiated interest in pattern avoidance in the study of algorithms by identifying the 231-avoiding permutations as exactly those that can be sorted by a single run through a stack; see Bona [7, Chapter 8] for further discussion. In fact, it is now well known that the avoidance sets $\mathcal{S}_n(\tau)$ for every length-3 pattern τ are enumerated by the Catalan numbers [37, A000108]:

$$|\mathcal{S}_n(\tau)| = \binom{2n}{n}/(n+1), \quad \tau \in \{123, 132, 213, 231, 312, 321\}.$$

Just as in the derangement problem above, enumeration of $\mathcal{S}_n(A)$ has an elementary probabilistic interpretation that motivates much of our paper. With Σ_n denoting a uniform random permutation of $[n]$ and A a set of permutations, the probability that Σ_n avoids A is

$$P\{\Sigma_n \text{ avoids every } \tau \in A\} = |\mathcal{S}_n(A)|/n!.$$

The Stanley–Wilf theorem [29] states that $|\mathcal{S}_n(\tau)|$ grows exponentially with n for every fixed τ. For example, the Catalan numbers are known to grow asymptotically like $4^n/\sqrt{\pi n^3}$, yielding the asymptotic avoidance probability

$$P\{\Sigma_n \text{ avoids } 231\} = \binom{2n}{n}/(n+1)! \sim \frac{1}{\pi n^2 \sqrt{2}} \left(\frac{4e}{n}\right)^n \quad \text{as } n \to \infty.$$

Such calculations quickly become intractable. For example, the sets of 1324-avoiding permutations have only been enumerated up to $n = 31$ [23]. Even precise asymptotics for $|\mathcal{S}_n(1324)|$ have not yet been established [8, 9, 13].

3. Main Results

3.1. **Definitions.** Throughout the paper, we write $\mathcal{L}(X)$ to denote the *distribution*, or *law*, of a random variable X and $\mathcal{L}(Y \mid X)$ to denote the *conditional distribution of Y given X*. For random variables X and Y, we write $d_{TV}(\mathcal{L}(X), \mathcal{L}(Y))$ to denote the *total variation distance* between the distributions of X and Y, which in the special case of non-negative integer-valued random variables can be computed as

$$d_{TV}(\mathcal{L}(X), \mathcal{L}(Y)) = \frac{1}{2} \sum_{n=0}^{\infty} |P(X = n) - P(Y = n)|.$$

Define the set of all unordered, distinct j-tuples of elements from $[n]$ by

$$\Gamma_j := \{\{i_1, \ldots, i_j\} : 1 \leq i_1 < \cdots < i_j \leq n\}, \qquad j \in [n].$$

For each $\alpha \in \Gamma_j$, let D_α be the set of all j-element subsets of $[n]$ that overlap with α in at least one element, i.e., $D_\alpha := \{\beta \in \Gamma_j : \beta \cap \alpha \neq \emptyset\}$. For example, if $\alpha = \{1, 5, 8\}$, then $D_\alpha^c = \{\{j_1, j_2, j_3\} : j_i \notin \{1, 5, 8\}, i = 1, 2, 3\}$.

With Σ_n denoting a uniform random permutation of $[n]$, $\alpha \in \Gamma_j$, and τ a fixed pattern of length j, we define $X_\alpha = X_{i_1,\ldots,i_m}$ as the indicator random variable for the event that the reduction of Σ_n at positions $i_1, \ldots, i_m$ form the pattern τ, i.e.,

$$X_{i_1,\ldots,i_m} := \mathbb{I}(\mathrm{red}(\Sigma_n(i_1) \cdots \Sigma_n(i_m)) = \tau). \tag{6}$$

Let $\mathbb{X} \equiv \mathbb{X}_j := (X_\alpha)_{\alpha \in \Gamma_j}$, and let $\mathbb{B} \equiv \mathbb{B}_j = (B_\alpha)_{\alpha \in \Gamma_j}$ denote a joint distribution of independent Bernoulli random variables with marginal distributions satisfying $\mathbb{E}\, B_\alpha = \mathbb{E}\, X_\alpha$ for all $\alpha \in \Gamma_j$. The random variable

$$W = \sum_{1 \leq i_1 < \ldots < i_m \leq n} X_{i_1,\ldots,i_m} \tag{7}$$

counts the total number of occurrences of τ in Σ_n.

For any $\tau \in \mathcal{S}_m$, for each $s = 1, \ldots, m-1$ we define $L_s(\tau)$ as the *overlap of size s*, i.e., the number of permutations $\sigma \in \mathcal{S}_{2m-s}$ such that there are indices $1 \leq i_1 < \cdots < i_m \leq 2m - s$ and $1 \leq j_1 < \cdots < j_m \leq 2m - s$ such that $\{i_1, \ldots, i_m\}$ and $\{j_1, \ldots, j_m\}$ have exactly s elements in common and

$$\mathrm{red}(\sigma_{i_1} \cdots \sigma_{i_m}) = \mathrm{red}(\sigma_{j_1} \cdots \sigma_{j_m}) = \tau.$$

For consecutive pattern avoidance, we similarly define the set of all j-tuples of the form $\{i, i+1, \ldots, i+j-1\}$, $1 \le i \le n-j+1$, as

$$\overline{\Gamma}_j := \{\{i, i+1, \ldots, i+j-1\} : 1 \le i \le n-j+1\}, \qquad j \in [n-j+1].$$

Let $\overline{\mathbb{X}} \equiv \overline{\mathbb{X}}_j := (X_\alpha)_{\alpha \in \overline{\Gamma}_j}$, and let $\overline{\mathbb{B}} \equiv \overline{\mathbb{B}}_j = (B_\alpha)_{\alpha \in \overline{\Gamma}_j}$ denote a joint distribution of independent Bernoulli random variables with marginal distributions which satisfy $\mathbb{E}\, B_\alpha = \mathbb{E}\, X_\alpha$ for all $\alpha \in \overline{\Gamma_j}$. Next, we define analogously the random variable

$$\overline{W} := \sum_{1 \le s \le n-m+1} X_{s,s+1,\ldots,s+m-1}, \tag{8}$$

where $X_{s,s+1,\ldots,s+m-1}$ is defined as in (6) for a fixed pattern $\tau \in \mathcal{S}_m$ and a uniform random permutation Σ_n. We also define $\overline{L}_s(\tau)$ as the *sequential overlap of size s*, i.e., the number of permutations $\sigma \in \mathcal{S}_{2m-s}$ such that

$$\mathrm{red}(\sigma_1 \cdots \sigma_m) = \mathrm{red}(\sigma_{m-s+1} \cdots \sigma_{2m-s}) = \tau.$$

3.2. **Main corollaries.** We begin with several limit theorems which follow from the quantitative bounds given in Section 3.3.

Corollary 3.1. *Fix any $\eta \ge 0$, and let $j \equiv j(n)$ be some increasing, integer-valued function of n such that $j \ge (e\, e^{1/e} + \eta)\sqrt{n}$. For any sequence of patterns $\tau_n \in \mathcal{S}_j$, and for any measurable function $h : \{0,1\}^{\binom{n}{j}} \to \mathbb{R}$ and Borel set $A \subseteq \mathbb{R}$, as n tends to infinity we have*

$$\mathbb{P}(h(\mathbb{X}) \in A) = \mathbb{P}(h(\mathbb{B}) \in A) + o(1).$$

We also have an analogous theorem for consecutive patterns, but with a stronger result.

Corollary 3.2. *Fix any $t > 0$, and define $M(t,n) := \left\lfloor \frac{\log(n/t)}{\log\log(n/t) - \log\log\log(n/t)} - \frac{1}{2} \right\rfloor$. Let $m \equiv m(n)$ be some increasing, integer-valued function of n such that $m \ge M(t,n)$. For any sequence of patterns $\tau_n \in \mathcal{S}_m$, and for any measurable function $h : \{0,1\}^{n-m+1} \to \mathbb{R}$ and Borel set $A \subseteq \mathbb{R}$, as n tends to infinity we have*

$$\mathbb{P}(h(\overline{\mathbb{X}}) \in A) = \mathbb{P}(h(\overline{\mathbb{B}}) \in A) + o(1).$$

In particular, denoting by Y_t an independent Poisson random variable with parameter t, and taking $m(n) = M(t,n)$, we have

$$d_{TV}(\overline{W}, Y_t) = O\left(\frac{1}{m}\right);$$

whence,

$$|\overline{\mathcal{S}}_n(\tau_n)| \sim n!\, e^{-t} \quad \text{as } n \to \infty.$$

Finally, we present an analogous result for permutations chosen according to the Mallows(q) distribution.

Corollary 3.3. *Let $m \equiv m(n)$ be a non-decreasing integer-valued sequence, $\tau_n \in \mathcal{S}_m$ be a sequence of patterns, and $q \equiv q(n)$ be a sequence of parameters. For each $n \ge 1$, let Σ_n be a random permutation from the Mallows distribution* (1) *with parameter $q(n)$, with $\overline{\mathbb{X}}_q$ and $\overline{\mathbb{B}}_q$ defined analogously. For any measurable function $h : \{0,1\}^{n-m+1} \to \mathbb{R}$ and Borel set $A \subseteq \mathbb{R}$, we have*

$$\mathbb{P}(h(\overline{\mathbb{X}}_q) \in A) = \mathbb{P}(h(\overline{\mathbb{B}}_q) \in A) + o(1),$$

provided either

$$q(n) \le n^{-1/\,inv(\tau_{m(n)})} \quad \text{for almost all } n \ge 1,$$
$$q(n) \ge n^{1/(\binom{m(n)}{2} - |\,inv(\tau_{m(n)})|)} \quad \text{for almost all } n \ge 1,$$
$$|\,inv(\tau_{m(n)})| \le -\log(n)/\log(q(n)) \quad \text{for almost all } n \ge 1 \text{ and } q < 1 \quad \text{or}$$
$$|\,inv(\tau_{m(n)})| \ge -\log(n)/\log(q(n)) + m(n)^2/2 \quad \text{for almost all } n \ge 1 \text{ and } q > 1.$$

3.3. **Quantitative bounds for large patterns.** Corollaries 3.1 and 3.2 provide an asymptotic analysis for sequences of patterns which also grow in size. It is too much to expect a general asymptotic formula for any fixed pattern—we have already noted the difficulty of nailing down the asymptotic growth of 1324-avoiding sets—but Poisson approximation, see Section 4, provides a general approach for obtaining *quantitative* bounds on various quantities when all sizes are fixed.

Theorem 3.4. *Assume $n \ge j \ge 3$, and τ is any pattern of length j. Let $\mathbb{X} = (X_\alpha)_{\alpha\in\Gamma_j}$, and let $\mathbb{B} = (B_\alpha)_{\alpha\in\Gamma_j}$ denote an independent Bernoulli process with marginal distributions which satisfy $\mathbb{E}\, B_\alpha = \mathbb{E}\, X_\alpha$ for all $\alpha \in \Gamma_j$. We have*

$$d_{TV}(\mathcal{L}(\mathbb{X}), \mathcal{L}(\mathbb{B})) \le 4D_{n,j} + \frac{2\lambda}{j!}, \tag{9}$$

where $D_{n,j} = \min(1, \lambda^{-1})(d_1 + d_2)$,

$$\lambda = \binom{n}{j}/j!, \qquad d_1 = \binom{n}{j}\left(\binom{n}{j} - \binom{n-j}{j}\right)\frac{1}{j!^2},$$
$$d_2 = \sum_{s=1}^{j-1} \binom{n}{2j-s}\frac{2\,L_s(\tau)}{(2j-s)!}. \tag{10}$$

Furthermore, for Y a Poisson random variable with mean $\lambda = \mathbb{E}\, W$, we have

$$d_{TV}(\mathcal{L}(W), \mathcal{L}(Y)) \le 2D_{n,j},$$

and also

$$n!\, e^{-\lambda}\left(1 - e^{\lambda} D_{n,j}\right) \le |\,\mathcal{S}_n(\tau)| \le n!\, e^{-\lambda}\left(1 + e^{\lambda} D_{n,j}\right). \tag{11}$$

Remark 3.5. *There are several noteworthy aspects to Theorem 3.4.*

1. *The expression for λ is the same for all patterns of length j.*
2. *The expression for d_1 cannot be improved by our approach.*
3. *We are unaware of any efficient means to calculate the values $L_s(\tau)$ in general. For a simple and explicit upper bound, applicable for all patterns τ of length j, we suggest*

$$d_2 \le \binom{n}{j}\sum_{s=1}^{j-1}\binom{n-j}{j-s}\binom{j}{s}\frac{s!}{j!^2}, \tag{12}$$

   but we suspect this bound can be improved.
4. *As Theorem 3.1 suggests, these bounds are most useful when $D_{n,j} < 1$, i.e., when j is large.*

Theorem 3.6. *Assume $n \ge m \ge 3$ and τ is any pattern of length m. We have*

$$d_{TV}\left(\mathcal{L}\left(\overline{\mathbb{X}}\right), \mathcal{L}\left(\overline{\mathbb{B}}\right)\right) \le 4\overline{D}_{n,m} + \frac{2\overline{\lambda}}{m!}, \tag{13}$$

where $\overline{D}_{n,m} = \min(1, 1/\overline{\lambda})(\overline{d}_1 + \overline{d}_2)$,

$$\overline{\lambda} = \frac{n-m}{m!}, \qquad \overline{d}_1 = \frac{n_1}{m!^2}, \qquad \overline{d}_2 = \sum_{s=1}^{m-1} \frac{(n-2m+s)\, 2\, \overline{L}_s(\tau)}{(2m-s)!}, \qquad n_1 \quad = 2mn - 3m^2 + m.$$

Let $\overline{Y}$ *denote a Poisson random variable with parameter* $\lambda = \mathbb{E}\,\overline{W}$. *We have*

$$d_{TV}\left(\mathcal{L}\left(\overline{W}\right), \mathcal{L}\left(\overline{Y}\right)\right) \le 2\overline{D}_{n,m},$$

and also

$$n!\, e^{-\overline{\lambda}}\left(1 - e^{\overline{\lambda}}\, \overline{D}_{n,m}\right) \le |\,\overline{\mathcal{S}}_n(\tau)| \le n!\, e^{-\overline{\lambda}}\left(1 + e^{\overline{\lambda}}\, \overline{D}_{n,m}\right). \tag{14}$$

3.4. **Limitations of Poisson Approximation.** It is tempting to conjecture that Corollary 3.1 holds even when λ tends to some fixed positive constant, but we suspect this is not possible, which we now demonstrate. First, we note the following.

Lemma 3.7 ([30]). *Fix any* $t > 0$ *and let* $\lambda = \binom{n}{j}/j!$. *Then* $\lambda \to t$ *for*

$$j \sim e\sqrt{n} - \frac{1}{4}\log(n) - \frac{1}{2}\log(2\pi t) - \frac{1}{4}e^2 - \frac{1}{2} \quad \text{as } n \to \infty. \tag{15}$$

Lemma 3.8. *Suppose* n, j *and* $n - j$ *tend to infinity, then we have*

$$d_1 \sim \lambda^2\left(1 - e^{-j^2/n}\right).$$

In particular, for $j \sim e\sqrt{n} - \frac{1}{4}\log(n)$, *we have* $d_1 \to c \in (0, \infty)$.

Thus, a necessary condition for d_1 to tend to zero is

$$j \sim e\sqrt{n} - \left(\frac{1}{4} - \epsilon\right)\log(n). \tag{16}$$

It is also well known, see [24, 27], that the typical size of the longest increasing subsequence is asymptotically of order $2\sqrt{n}$, and so one cannot have a Poisson limit theorem which applies to the increasing pattern $12\ldots j$. It would be interesting to investigate the behavior in the gap, i.e., for $j \sim c\sqrt{n}$ with any $2 < c < e$.

3.5. **Consecutive pattern avoidance for Mallows permutations.** In Section 5, we discuss several special properties of the Mallows distribution that are helpful for studying consecutive pattern avoidance. Using these properties, we obtain analogous bounds to those in Theorems 3.4 and 3.6.

Recall the definition of the *restriction* $\Sigma_{n|A}$ of Σ_n to a subset $A \subseteq [n]$, and recall $\overline{\Gamma}_m$ denotes the set of subsets of size m whose elements are consecutive in $\{1, 2, \ldots, n\}$.

Theorem 3.9. *Fix* $q > 0$ *and let* $\Sigma_n \sim$ *Mallows*(q). *For any* $m \ge 2$, *let* τ_m *be any pattern of size* m. *For any* $\alpha \in \overline{\Gamma}_m$, *let*

$$X_\alpha = \mathbb{I}(red(\Sigma_{n|\alpha}) = \tau_m).$$

Let $W = \sum_{\alpha\in\overline{\Gamma}_m} X_\alpha$ *and let* Y *be an independent Poisson random variable with expected value* $\lambda = \mathbb{E}\, W$. *Then*

$$d_{TV}(\mathcal{L}(W), \mathcal{L}(Y)) \le 2(b_1 + b_2),$$

where

$$\lambda = (n-m)\frac{q^{|\,inv(\tau_m)|}}{I_m(q)}, \qquad b_1 = \frac{n_1 q^{2|\,inv(\tau_m)|}}{I_m(q)^2}, \qquad b_2 = \sum_{s=1}^{m-1}(n-2m+s)\sum_{\rho\in\overline{L}_s(\tau_m)}\frac{q^{|\,inv(\rho)|}}{I_{2m-s}(q)},$$

$$n_1 = 2mn - 3m^2 + m.$$

The key to obtaining asymptotic formulas and Poisson limit theorems for general Mallows permutations relies on the interplay between the parameters n, m, $|\mathrm{inv}(\tau)|$, and q. In particular, we need the expected number of occurrences to converge to a constant $\lambda \in (0,\infty)$. In the case of consecutive pattern avoidance, the expected number of occurrences of a pattern $\tau \in \mathcal{S}_n$ in $\Sigma_n \sim \text{Mallows}(q)$ is

$$\lambda = (n-m)q^{|\,\mathrm{inv}(\tau)|}/I_m(q),$$

which, for m fixed, produces non-trivial limiting behavior as long as

$$q \sim n^{-1/|\,\mathrm{inv}(\tau)|} \quad \text{or} \quad q \sim n^{1/\left(\binom{m}{2}-|\,\mathrm{inv}(\tau)|\right)}.$$

We can also allow m to vary and keep q fixed so that

$$|\,\mathrm{inv}(\tau)| \sim -\log(n)/\log(q) \quad \text{or} \quad |\,\mathrm{inv}(\tau)| \sim -\log(n)/\log(q) + m^2/2.$$

Combined with Theorem 3.9, these observations yield Corollary 3.3.

In Section 5.3.2, we demonstrate Theorem 3.9 for all patterns of length 3. In Section 6.2, we compute the bounds in Theorems 3.4 and 3.9 for the specific patterns 2341 and 23451 and we plot the estimated pattern avoidance probabilities in the appropriate asymptotic regime for q from Corollary 3.3.

4. Poisson approximation

4.1. **Chen–Stein method.** Stein's method is an approach to proving the central limit theorem that was adapted by Chen to Poisson convergence [12]. The advantage of this method is that it provides guaranteed error bounds on the total variation distance between the distribution of a sum of possibly dependent random variables and the distribution of an independent Poisson random variable with the same mean.

Theorem 4.1 (Chen [12]). *Suppose $X_1, X_2, \ldots, X_n$ are indicator random variables with expectations $p_1, p_2, \ldots, p_n$, respectively, and let $W = \sum_{i=1}^n X_i$. Let Y denote an independent Poisson random variable with expectation $\lambda = \sum_{i=1}^n p_i$. Suppose, for each $i \geq 1$, a random variable V_i can be constructed on the same probability space as W such that*

$$\mathcal{L}(1+V_i) = \mathcal{L}(W \mid X_i = 1).$$

Then

$$d_{TV}(\mathcal{L}(W), \mathcal{L}(Y)) \leq \frac{1-e^{-\lambda}}{\lambda}\sum_{i=1}^n p_i \mathbb{E}\,|W - V_i|. \tag{17}$$

4.2. **Fixed Points Example.** As a simple example, let $e(n)$ denote the number of fixed-point free permutations of n. With Σ_n a uniform permutation of $[n]$, we define indicator random variables

$$X_i = \mathbb{I}(\ i \text{ is a fixed point of } \Sigma_n\ ), \qquad i = 1, \ldots, n.$$

(Note that these random variables are *not* independent.) We then define the sum

$$W = \sum_{i=1}^{n} X_i$$

so that $P(W = 0) = e(n)/n!$ and $\lambda = \mathbb{E}W = \sum_{i=1}^{n} \frac{1}{n} = 1$, the expected number of fixed points. Even before we proceed with the bound, we obtain the heuristic estimate of $n!\, e^{-1}$ for $e(n)$, just as in (2).

To apply Theorem 4.1, we need to construct W and $1 + V_i$ on the same probability space, and construct an explicit coupling. This is done for more general restrictions in [5], but we shall write out the full calculation on fixed points to demonstrate how one can construct such a coupling.

The random variable $1 + V_i$ is the random sum W conditioned on $X_i = 1$. For a random permutation σ, suppose $\sigma(i) = j$, for some $j \in [n]$. The coupling is: swap elements i and j. The resulting permutation has the same marginal distribution as a random permutation conditioned to have a fixed point at i. In fact, $|W - V_i| \in \{0, 1, 2\}$ for each i since we modify at most 2 elements, and the elements not involved in the swap cancel out (i.e., any fixed points occurring on indices other than these swapping positions remain unchanged). Let us denote the random variables after the coupling by $X_1', X_2', \ldots, X_n'$; that is, $\mathcal{L}(X_j') = \mathcal{L}(X_j \mid X_i = 1)$, so that $1 + V_i = \sum_{j=1}^{n} X_j'$. We have

$$|W - V_i| = \left| X_i + \sum_{k \neq i} (X_k - X_k') \right| = |X_i + X_J \mathbb{I}(J = \sigma(i), J \neq i)| = \begin{cases} 0, & \sigma(i) \neq i, i \text{ not in a 2-cycle}, \\ 1, & \sigma(i) = i, \\ 2, & i \text{ in a 2-cycle}. \end{cases}$$

The probability that two given elements i and j are part of a 2-cycle is precisely $1/(n(n-1))$, and the probability that i is part of a 1-cycle is $1/n$. Thus,

$$\mathbb{E}\,|W - V_i| = \frac{1}{n} + \frac{2}{n} = \frac{3}{n},$$

and Equation (17) becomes

$$d_{TV}(\mathcal{L}(W), \mathcal{L}(Y)) \leq (1 - e^{-1}) \sum_{i=1}^{n} \frac{1}{n} \frac{3}{n} = \frac{3(1 - e^{-1})}{n}.$$

For all $n \geq 1$, we have

$$|P(W = 0) - P(Y = 0)| = \left| \frac{e(n)}{n!} - e^{-1} \right| \leq \sup_i |P(W = i) - P(Y = i)| \leq d_{TV}(W, Y) \leq \frac{3(1 - e^{-1})}{n}.$$

Rearranging yields

$$n!e^{-1} - 3(n-1)!(1 - e^{-1}) \leq e(n) \leq n!e^{-1} + 3(n-1)!(1 - e^{-1}).$$

Note that this is a guaranteed error bound that holds for all $n \geq 1$, and as a corollary we get $e(n) = n!e^{-1}(1 + o(n^{-1}))$.

The error bounds derived from the Chen–Stein method can be improved in special cases, e.g., $e(n)$ above can be obtained exactly by rounding $n!/e$ to the nearest integer for all $n \geq 1$, but the appeal of Poisson approximation is that it applies more generally. Our main theorems (Theorems 3.1 and 3.2) identify cases in which the bounds provide an asymptotically efficient estimate.

4.3. **The Arratia–Goldstein–Gordon Theorem.** Arratia, Goldstein, & Gordon [2] provide another approach that is sometimes more practical for Poisson approximation.

Theorem 4.2 (Arratia, Goldstein, & Gordon [2])**.** *Let I be a countable set of indices and for each $\alpha \in I$ let X_α be an indicator random variable. Let $\mathbb{X} = (X_\alpha)_{\alpha\in I}$ denote a collection of Bernoulli random variables, and let $\mathbb{B} = (B_\alpha)_{\alpha\in I}$ denote a collection of independent Bernoulli random variables with marginal distributions which satisfy $\mathbb{E}\, B_\alpha = \mathbb{E}\, X_\alpha$ for all $\alpha \in I$. Define $p_\alpha := \mathbb{E}\, X_\alpha = P(X_\alpha = 1) > 0$ and $p_{\alpha\beta} := \mathbb{E}\, X_\alpha X_\beta$. Also define $W := \sum_{\alpha\in I} X_\alpha$ and $\lambda := \mathbb{E}\, W = \sum_{\alpha\in I} p_\alpha$. For each $\alpha \in I$, define sets $D_\alpha \subset I$ (typically, this will be the set of all indices $\beta \in I$ for which X_α and X_β are dependent, but this is not necessary) and the quantities*

$$\begin{aligned} b_1 &:= \sum_{\alpha\in I} \sum_{\beta\in D_\alpha} p_\alpha p_\beta, \\ b_2 &:= \sum_{\alpha\in I} \sum_{\alpha\neq\beta\in D_\alpha} p_{\alpha\beta}, \quad \text{and} \\ b_3 &:= \sum_{\alpha\in I} \mathbb{E}\, |\mathbb{E}\, \{X_\alpha - p_\alpha \mid \sigma(X_\beta : \beta \notin D_\alpha)\}|. \end{aligned} \tag{18}$$

We have

$$d_{TV}(\mathcal{L}(\mathbb{X}), \mathcal{L}(\mathbb{B})) \leq 2(2b_1 + 2b_2 + b_3) + 2\sum_{\alpha\in I} p_\alpha^2.$$

Furthermore, let Y denote a Poisson random variable with mean $\lambda = \mathbb{E}\, W$. We have

$$d_{TV}(\mathcal{L}(W), \mathcal{L}(Y)) \leq 2(b_1 + b_2 + b_3),$$

and also

$$|P(W = 0) - P(Y = 0)| \leq (b_1 + b_2 + b_3)\, \frac{1 - e^{-\lambda}}{\lambda}.$$

In our applications, we are able to define sets D_α, $\alpha \in I$, so that $b_3 = 0$ always holds; whence, the calculations of the bounds in our theorems require only calculations involving first and second (unconditioned) moments. For uniform random permutations this is straightforward, but that the analogous properties hold for consecutive patterns under random Mallows permutations is less obvious.

5. Consecutive pattern avoidance of Mallows permutations

To fix notation, we write $\sigma = \sigma_1 \cdots \sigma_n$ to denote a generic permutation, for which we define the *reversal* by $\sigma^r = \sigma_n \cdots \sigma_1$. For any subset $A \subseteq [n]$, we write $\sigma_{|A}$ to denote the *restriction* of σ to a permutation of A obtained by removing those elements among $\sigma_1, \ldots, \sigma_n$ that are not in A. For example, with $\sigma = 867531924$ and $A = \{1, 3, 5, 7, 9\}$, we have $\sigma_{|A} = 75319$. We write Σ_n to denote a random permutation of $[n]$.

From (1) it is apparent that $P_n^q(\sigma) = P_n^{1/q}(\sigma^r)$ for all $\sigma \in \mathcal{S}_n$, and so we can focus on the case $0 < q \leq 1$ in our analysis.

5.1. **Sequential construction.** The Mallows distribution (1) enjoys several nice properties that are amenable to the study of pattern avoidance. These properties are readily observed by the following sequential constructions, both of which are well known and have been leveraged in previous studies of the Mallows distribution; see, for example, [6, 19]. While the properties below are well known, we are not aware of their appearance in relation to pattern avoidance. We provide proofs for completeness.

For $q > 0$, we say that random variable X has the *truncated Geometric(q)* distribution on $[n]$, written as $X \sim \text{Geometric}(n, q)$, when the point probabilities of X are given by

$$P^{n,q}(X = k) = q^{k-1}/(1 + \cdots + q^{n-1}), \quad k = 1, \dots, n. \tag{19}$$

A Mallows permutation can be generated from the truncated Geometric distribution in two ways, which we call the *ordering* and *bumping* constructions.

For the ordering construction, we generate $X_1, X_2, \dots$ independently, with each X_n distributed as $\text{Geometric}(n, 1/q)$. To initialize, we have $\Sigma_1 = 1$, the only permutation of [1]. Given $\Sigma_n = \sigma_1 \cdots \sigma_n$ and $X_{n+1} = k$, we define

$$\Sigma_{n+1} = \sigma_1 \cdots \sigma_{k-1}(n + 1)\sigma_k \cdots \sigma_n.$$

For every $n = 1, 2, \dots$, it is apparent that Σ_n is a Mallows(q) permutation because the probability that element $n + 1$ appears in position k of Σ_{n+1} is

$$P\{\Sigma_{n+1}(k) = n + 1\} = P^{n+1,1/q}(X = k) = P^{n+1,q}(X = n + 1 - k) = q^{n+1-k}/(1 + q + \cdots + q^n).$$

Since $X_1, \dots, X_n$ are chosen independently and each event $\{\Sigma_n = \sigma\}$ corresponds to exactly one sequence $X_1, \dots, X_n$, we observe

$$P\{\Sigma_n = \sigma\} = q^{|\operatorname{inv}(\sigma)|}/I_n(q), \quad \sigma \in \mathcal{S}_n,$$

as in (1).

Definition 5.1 (Mallows process)**.** *A collection* $(\Sigma_n)_{n\ge1}$ *generated by the ordering construction for fixed* $q > 0$ *is called a* Mallows(q) process*.*

For the bumping construction, we generate $X_1, X_2, \dots$ independently with each X_n distributed as $\text{Geometric}(n, 1/q)$ as before, and again we initialize with $\Sigma_1 = 1$. Given $\Sigma_n = \sigma_1 \cdots \sigma_n$ and $X_{n+1} = k$, we obtain Σ_{n+1} by appending k to the end of Σ_n and "bumping" all elements of Σ_n that are greater or equal to k. More formally, $(\Sigma_n, X_{n+1}) \mapsto \Sigma'_1 \cdots \Sigma'_n X_{n+1}$, where

$$\Sigma'_j = \begin{cases} \Sigma_j + 1, & \Sigma_j \ge X_{n+1}, \\ \Sigma_j, & \text{otherwise.} \end{cases}$$

For example, if $\Sigma_5 = 24135$ and $X_6 = 3$, then $\Sigma_6 = 251463$. Again, the resulting distribution of Σ_n is Mallows(q) because $X_{n+1} = k$ introduces exactly $n + 1 - k$ new inversions in Σ_{n+1} and $X_1, X_2, \dots$ are generated independently.

5.2. **Properties of Mallows permutations.** Throughout this section, we let $(\Sigma_n)_{n\ge1}$ be a family of random permutations so that each Σ_n is a permutation of $[n]$. We say that $(\Sigma_n)_{n\ge1}$ is *consistent* if for all $1 \le m \le n$

$$P\{\Sigma_{n|[m]} = \sigma\} = P\{\Sigma_m = \sigma\}, \quad \sigma \in \mathcal{S}_m\,. \tag{20}$$

It is immediate from the ordering construction that the Mallows process $(\Sigma_n)_{n\ge1}$ is consistent for every $q > 0$.

Recall the reduction map described in Section 1. We call $(\Sigma_n)_{n\geq 1}$ *homogeneous* if for all $1 \leq m \leq n$ and every subsequence $1 \leq i_1 < \cdots < i_m \leq n$

$$P\{\mathrm{red}(\Sigma_n(i_1)\cdots\Sigma_n(i_m)) = \sigma\} = P\{\Sigma_m = \sigma\}, \quad \sigma \in \mathcal{S}_m . \tag{21}$$

We call $(\Sigma_n)_{n\geq 1}$ *consecutively homogeneous* if (21) holds only for consecutive subsequences $i_1, i_1+1, \ldots, i_1+m-1$.

Lemma 5.2. *The Mallows(q) process is consecutively homogeneous for all $q > 0$ and homogeneous for $q = 1$.*

Proof. The $q = 1$ case corresponds to the uniform distribution, which is well known to be homogeneous. For arbitrary $q > 0$, consider the event $\{\mathrm{red}(\Sigma_n(j)\cdots\Sigma_n(j+m-1)) = \sigma\}$ for some $\sigma \in \mathcal{S}_n$. By the ordering construction, we can first generate $\Sigma_m = \Sigma_m(1)\cdots\Sigma_m(m)$ from the Mallows($1/q$) distribution on $[m]$. We then obtain Σ_{m+j-1} from Σ_m using the bumping construction for Mallows($1/q$) distribution. Thus, we have $\Sigma_{m+j-1} \sim$ Mallows($1/q$) and its reversal $\Sigma^r_{m+j-1} \sim$ Mallows(q) with $\mathrm{red}(\Sigma^r_{m+j-1}(j)\cdots\Sigma^r_{m+j-1}(m+j-1)) = \Sigma_m(m)\cdots\Sigma_m(1) \sim$ Mallows(q). Finally, we obtain Σ_n by adding to Σ^r_{m+j-1} according to the bumping construction, so that

$$\mathrm{red}(\Sigma_n(j)\cdots\Sigma_n(m+j-1)) = \mathrm{red}(\Sigma^r_{m+j-1}(j)\cdots\Sigma^r_{m+j-1}(m+j-1)) = \Sigma_m(m)\cdots\Sigma_m(1) \sim \text{Mallows}(q).$$

This completes the proof. □

We say that Σ_n is *dissociated* if $\Sigma_{n|A}$ and $\Sigma_{n|B}$ are independent for all non-overlapping subsets $A, B \subseteq [n]$. If, instead, $\Sigma_{n|A}$ and $\Sigma_{n|B}$ are independent only when A and B are disjoint and each consists of consecutive indices, then we call Σ_n *weakly dissociated*.

Lemma 5.3. *For all $n \geq 1$, the Mallows(q) distribution on $\mathcal{S}_n$ is weakly dissociated for all $q > 0$ and dissociated for $q = 1$.*

Proof. For $i' > i \geq 1$ and $m, m' \geq 0$ satisfying $i + m - 1 < i'$ and $i' + m' - 1 \leq n$, let $A = \{i, i+1, \ldots, i+m-1\}$ and $B = \{i', i'+1, \ldots, i'+m'-1\}$. For any $n \geq 1$, we can construct a Mallows(q) permutation of $[n]$ by first generating Σ_{i+m-1}, for which we know that $\mathrm{red}(\Sigma_{i+m-1}(i)\cdots\Sigma_{i+m-1}(i+m-1)) \sim$ Mallows(q) by Lemma 5.2. We then construct Σ_n from Σ_{i+m-1} by the bumping construction. Since bumping does not affect the reduction of any part of $\Sigma_n(1)\cdots\Sigma_n(i+m-1)$, we have

$$\begin{aligned} &P\{\mathrm{red}(\Sigma_n(i)\cdots\Sigma_n(i+m-1)) = \sigma \mid \mathrm{red}(\Sigma_n(i')\cdots\Sigma_n(i'+m'-1)) = \sigma'\} = \\ &= P\{\mathrm{red}(\Sigma_{i+m-1}(i)\cdots\Sigma_{i+m-1}(i+m-1)) = \sigma \mid \mathrm{red}(\Sigma_n(i')\cdots\Sigma_n(i'+m'-1)) = \sigma'\} \\ &= P\{\Sigma_m(1)\cdots\Sigma_m(m) = \sigma\}, \end{aligned}$$

proving that Σ_n is weakly dissociated. Dissociation of the uniform distribution ($q = 1$) is well known and so we omit its proof. The proof is complete. □

Together, the above properties facilitate study of consecutive pattern avoidance for Mallows permutations with arbitrary $q > 0$. For example, the pattern 231 has probability $q^2/(1 + 2q + 2q^2 + q^3)$ to occur in any stretch of three consecutive positions of a Mallows(q) permutation. Since there are $n - 2$ consecutive patterns of length 3 in a permutation of $[n]$, the expected number of occurrences is $(n-2)q^2/(1 + 2q + 2q^2 + q^3)$. For large n and small q, this expected value behaves asymptotically as nq^2, so that taking $q \sim 1/\sqrt{n}$ gives an expected number on the order of a constant. When q is large, the expected number of occurrences behaves as nq^{-1} for large n, and taking $q \sim n$ gives an expected number on the order of a constant.

5.3. **Poisson convergence theorems.** Theorem 3.9 and its corollary follow by combining the above properties of Mallows permutations with Theorem 4.2. The calculations and resulting bounds for the general Mallows measure follow the same program as the uniform case proven in Section 7, with the key distinction that we only consider *consecutive* patterns for the general Mallows distribution. Unlike the uniform setting, the bounds for the Mallows distribution depend non-trivially on the parameter q and the structure of τ. It is more fruitful to illustrate this dependence with specific examples than to regurgitate the same proof for Mallows permutations.

5.3.1. *Monotonic patterns under Mallows distribution.* Consider the set of permutations that avoid the pattern 123. There are no inversions, and the size of the pattern is 3; thus, the probability of seeing this pattern in any given set of three consecutive indices of a Mallows(q) permutation is $1/I_3(q)$. We also need to consider second moments, i.e., the probability of seeing two 123 patterns. By Lemma 5.3 we need only consider overlapping sets of indices. There are two cases, either two indices overlap or one does. If two indices overlap and the first three and last three both reduce to pattern 123, then the segment must reduce to 1234. Similarly, if one index overlaps, then the segment must reduce to 12345.

The results below extend this notion to monotonic patterns.

Lemma 5.4. *Fix $q > 0$ and let $\Sigma_n \sim$ Mallows(q). For each $m \geq 1$, let τ_m denote the pattern $12\cdots m$. For each $\alpha \in J_m$, define*

$$X_\alpha = \mathbb{I}(red(\Sigma_{n|\alpha}) = \tau_m).$$

For a random permutation generated using the Mallows measure, we have

$$\mathbb{E}\, X_\alpha = \frac{1}{I_m(q)}, \qquad \alpha \in J_m,$$

and for $\alpha, \beta \in J_m$, $\alpha \neq \beta$, we have

$$\mathbb{E}\, X_\alpha X_\beta = \begin{cases} \frac{1}{I_m(q)^2} & \alpha, \beta \text{ have no overlapping elements} \\ \frac{1}{I_{2m-s}(q)} & \alpha, \beta \text{ have exactly } s \text{ overlapping elements, } s = 1, 2, \ldots, m-1. \end{cases}$$

Proof. The expression when α and β do not overlap is a consequence of the weak dissociation property of Mallows permutations (Lemma 5.3), whereby

$$\mathbb{E}\, X_\alpha X_\beta = \mathbb{E}\, X_\alpha \mathbb{E}\, X_\beta = P\{\text{red}(\Sigma_{n|\alpha}) = \tau_m\}^2 = (1/I_m(q))^2.$$

When α and β overlap in s elements, the event $\{X_\alpha = X_\beta = 1\}$ requires that both $\Sigma_{n|\alpha}$ and $\Sigma_{n|\beta}$ reduce to the increasing permutation, which can occur only if $\Sigma_{n|\alpha\cup\beta}$ reduces to the increasing permutation of $2m - s$. □

Proposition 5.5. *Fix $q > 0$ and let $\Sigma_n \sim$ Mallows(q). For any $m \geq 2$, let τ_m be the increasing pattern $12\cdots m$. For any $\alpha \in J_m$, let*

$$X_\alpha = \mathbb{I}(red(\Sigma_{n|\alpha}) = \tau_m).$$

Let $W = \sum_{\alpha \in J_m} X_\alpha$ and let Y be an independent Poisson random variable with expected value $\lambda = \mathbb{E}\, W$. Then

$$d_{TV}(\mathcal{L}(W), \mathcal{L}(Y)) \leq 2(b_1 + b_2),$$

where

$$\lambda = \frac{n-m}{I_m(q)}, \qquad b_1 = \frac{n_1}{I_m(q)^2}, \qquad b_2 = n_2 \sum_{s=1}^{m-1} \frac{1}{I_{2m-s}(q)},$$

and n_1 and n_2 are given by

$$n_1 = 2mn - 3m^2 + m, \quad \text{and}$$
$$n_2 = 3m - 3m^2 - 2n + 2mn. \tag{22}$$

Proof. When $2m-1 \le n$, we have $n_1 = 2\sum_{s=1}^{m}(n-2m+s) = 2mn - 3m^2 + m$, and similarly $n_2 = 2\sum_{s=1}^{m-1}(n-2m+s) = 3m - 3m^2 - 2n + 2mn$. The factor of 2 is from exchanging the role of α, β. When $2m-1 > n$, the stated expressions for n_1 and n_2 are still valid upper bounds, but they can be slightly improved. □

In addition, it is easy to state the complementary result about the decreasing pattern $m\cdots 21$.

Proposition 5.6. *Fix $q > 0$ and let $\Sigma_n \sim$ Mallows(q). For any $m \ge 2$, let η_m be the decreasing pattern $m\cdots 21$. For any $\alpha \in J_m$, let*

$$X_\alpha = \mathbb{I}(red(\Sigma_{n|\alpha}) = \eta_m).$$

Let $W = \sum_{\alpha \in J_m} X_\alpha$ and let Y denote an independent Poisson random variable with expected value $\lambda = \mathbb{E}\,W$. Then

$$d_{TV}(\mathcal{L}(W), \mathcal{L}(Y)) \le 2(b_1 + b_2),$$

where

$$\lambda = (n-m)\frac{q^{\binom{m}{2}}}{I_m(q)}, \qquad b_1 = \frac{n_1\, q^{\binom{m}{2}}}{I_m(q)^2}, \qquad b_2 = n_2 \sum_{s=1}^{m-1} \frac{q^{\binom{2m-s}{2}}}{I_{2m-s}(q)},$$

and n_1 and n_2 are given by

$$n_1 = 2mn - 3m^2 + m, \quad \text{and}$$
$$n_2 = 3m - 3m^2 - 2n + 2mn.$$

5.3.2. *Other patterns of length 3.* We now demonstrate the dependence of the total variation bound on q for the small patterns 132, 213, 231, and 312. In this section, we again recall that $\overline{\Gamma}_m$ denotes the set of all m-tuples with consecutive elements in $\{1, 2, \ldots, n\}$, and for a given $\alpha \in \overline{\Gamma}_3$, X_α denotes the indicator random variable defined in (6).

For $\tau = 132$, we have $\mathbb{E}\,X_\alpha = q/I_3(q)$, and there can be no consecutive occurrences of τ that overlap with two indices. The only possible ways in which we can have one overlapping index are the patterns 13254, 15243, and 14253. In these cases, we have

$$\lambda = \frac{(n-3)\,q}{I_3(q)} \quad \text{and}$$

$$\mathbb{E}\,X_\alpha X_\beta = \frac{1}{I_5(q)} \times \begin{cases} q^2, & 13254, \\ q^3, & 14253, \\ q^4, & 15243. \end{cases}$$

Letting $W = \sum_{\alpha \in \overline{\Gamma}_3} \mathbb{E}\,X_\alpha$ and defining Y as an independent Poisson random variable with expectation $\lambda = \mathbb{E}\,W = (n-3)\,q/I_3(q)$, the total variation distance bound is given by

$$d_{TV}(\mathcal{L}(W), \mathcal{L}(Y)) \le 2\left((3n-13)\,\frac{q^2}{I_3(q)^2} + 2(n-5)\,\frac{q^2+q^3+q^4}{I_5(q)}\right).$$

The $2(n-5)$ term comes from the 2 sets of triplets $\{1,2,3\}$ and $\{2,3,4\}$ for which the overlapping pair can only occur to the right of the elements, and similarly from the 2 sets of triplets $\{n-2,n-1,n\}$ and $\{n-3,n-2,n-1\}$ for which the overlapping pair can only occur to the left of the elements, and finally the $(n-4-3)$ triplets in between for which the overlapping pairs are both to the left and the right; hence $2+2(n-7)+2=2(n-5)$. Similarly, the $3n-13$ comes from $2\cdot 2+3(n-7)+2\cdot 2$. Let $t>0$ be fixed. If $q \sim t\,n^{-1}$ or $q \sim t\,n^{1/2}$, we have $\lambda \to t$, and $d_{TV}(\mathcal{L}(W),\mathcal{L}(Y)) = O(n^{-1})$.

For $\tau = 213$, we similarly have

$$\lambda = \frac{(n-3)\,q}{I_3(q)},$$

$$\mathbb{E}\,X_\alpha X_\beta = \frac{1}{I_5(q)} \times \begin{cases} q^2, & 21435, \\ q^3, & 31425, \\ q^4, & 32415, \end{cases} \quad \text{and}$$

$$d_{TV}(\mathcal{L}(W),\mathcal{L}(Y)) \le 2\left((3n-13)\,\frac{q^2}{I_3(q)^2} + 2(n-5)\,\frac{q^2+q^3+q^4}{I_5(q)}\right),$$

and $q \sim t\,n^{-1}$ or $q \sim t\,n^{1/2}$ implies $\lambda \to t$ and $d_{TV}(\mathcal{L}(W),\mathcal{L}(Y)) = O(n^{-1})$.

For $\tau = 231$:

$$\lambda = \frac{(n-3)\,q^2}{I_3(q)},$$

$$\mathbb{E}\,X_\alpha X_\beta = \frac{1}{I_5(q)} \times \begin{cases} q^6, & 34251, \\ q^7, & 35241, \\ q^8, & 45231, \end{cases} \quad \text{and}$$

$$d_{TV}(\mathcal{L}(W),\mathcal{L}(Y)) \le 2\left((3n-13)\,\frac{q^4}{I_3(q)^2} + 2(n-5)\,\frac{q^6+q^7+q^8}{I_5(q)}\right),$$

and $q \sim t^{1/2} n^{-1/2}$ or $q \sim n\,t^{1/2}$ implies $\lambda \to t$ and $d_{TV}(\mathcal{L}(W),\mathcal{L}(Y)) = O(n^{-1})$.

And finally for $\tau = 312$:

$$\lambda = \frac{(n-3)\,q^2}{I_3(q)},$$

$$\mathbb{E}\,X_\alpha X_\beta = \frac{1}{I_5(q)} \times \begin{cases} q^6, & 51423, \\ q^7, & 52413, \\ q^8, & 53412, \end{cases} \quad \text{and}$$

$$d_{TV}(\mathcal{L}(W),\mathcal{L}(Y)) \le 2\left((3n-13)\,\frac{q^4}{I_3(q)^2} + 2(n-5)\,\frac{q^6+q^7+q^8}{I_5(q)}\right),$$

and $q \sim t^{1/2}\,n^{-1/2}$ or $q \sim n\,t^{1/2}$ implies $\lambda \to t$ and $d_{TV}(\mathcal{L}(W),\mathcal{L}(Y)) = O(n^{-1})$.

6. Numerical examples

6.1. **Numerical values.** Using Theorem 3.4, we can estimate $|\mathcal{S}_n(\tau)|$ for various sizes of patterns τ. Table 1 shows the lower bound thresholds for various values of n and patterns

n	j	lower	$n!$
100	36	6.85456×10^{157}	9.3326×10^{157}
1000	133	3.4433×10^{2567}	4.6045×10^{2567}
10000	442	8.3847×10^{35658}	2.8463×10^{35659}
100000	14353	$9.9451 \times 10^{65657058}$	$1.2024 \times 10^{65657059}$

TABLE 1. Bounds on $|\mathcal{S}_n(\tau)|$ for $\tau \in \mathcal{S}_j$, for various values of n and j.

n	j	lower	$n!$
100	6	3.98735×10^{157}	9.33262×10^{157}
1000	7	5.77948×10^{2566}	4.02387×10^{2567}
10000	9	$2.49966 \times 10^{35659}$	$2.84626 \times 10^{35659}$
100000	10	$2.48004 \times 10^{456573}$	$2.82423 \times 10^{456573}$
1000000	11	$7.34802 \times 10^{5565708}$	$8.26393 \times 10^{5565708}$

TABLE 2. Bounds on $|\overline{\mathcal{S}}_n(\tau)|$ for $\tau \in \mathcal{S}_j$, for various values of n and j.

permutation	no. inversions	permutation	no. inversions
3452671	9	3462571	10
3472561	11	3562471	11
3572461	12	4562371	12
4572361	13	3672451	13
4672351	14	5672341	15

TABLE 3. List of all permutations that have pattern 2341 in overlapping positions along with the number of inversions.

of size j. Similarly, using Theorem 3.6, we estimate $|\overline{\mathcal{S}}_n(\tau)|$ in Table 2. In the case of $n = 1000$ and $j = 133$, we have more specifically $3.4433 \times 10^{2567} \le |\mathcal{S}_n(\tau)| \le 4.0239 \times 10^{2567}$.

6.2. **Detailed illustration for the patterns 2341 and 23451.** Propositions 5.5 and 5.6 give an expression for the total variation bound between the number of occurrences of the increasing and decreasing patterns and an independent Poisson random variable. In principle, these bounds can be computed exactly for any pattern by way of the Arratia–Goldstein–Gordon theorem (Theorem 4.2); we need only compute the quantities b_1, b_2, and b_3 as in Theorem 4.2.

By Lemma 5.3, all Mallows(q) permutations are weakly dissociated and, therefore, $b_3 \equiv 0$ for all patterns in the case of consecutive pattern avoidance. For any pattern τ, homogeneity of the Mallows measure implies $p_\alpha = q^{|\,\mathrm{inv}(\tau)|}/I_m(q)$ for all α, and so b_1 is easy to compute. The only complication involves the consideration of overlapping patterns in the calculation of b_2. We cannot provide anything more general than Arratia–Goldstein–Gordon for arbitrary patterns; instead, we compute these bounds in the special cases of $\tau = 2341$ and $\tau = 23451$. Figure 1 shows the performance of these bounds at the critical values $q \sim n^{-1/3}$ and $q \sim n^{1/3}$ for $\tau = 2341$, and $q \sim n^{-1/4}$ and $q \sim n^{1/6}$ for $\tau = 23451$.

6.2.1. *The pattern* 2341. For $\tau = 2341$, we have $p_\alpha = q^3/I_4(q)$ and $b_1 = (n-4)q^3/I_4(q)$. The structure of τ only permits overlap with the first or last position. Table 3 lists all permutations

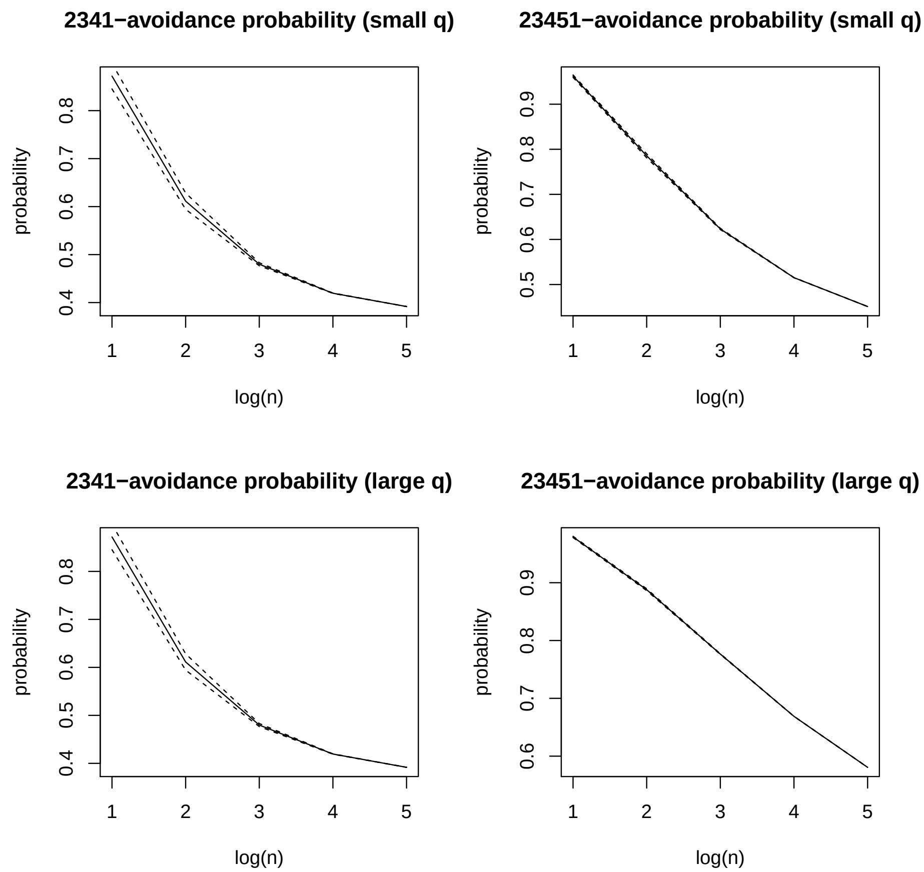

FIGURE 1. Plot of pattern avoidance probabilities of Mallows(q) distribution for: (top left) pattern $\tau = 2341$ with $q = n^{-1/3}$; (top right) pattern $\tau = 23451$ with $q = n^{-1/4}$; (bottom left) pattern $\tau = 2341$ with $q = n^{1/3}$; and (bottom right) pattern $\tau = 23451$ with $q = n^{1/6}$. The dashed lines represent the upper and lower error bounds from the Arratia–Goldstein–Gordon theorem, and the solid line represent their average, i.e., the heuristic approximation. In all panels, the horizontal axis is on the logarithmic scale with base 10.

that have pattern 2341 in the first 4 and last 4 positions. These are the only permutations that contribute to b_2 in the bound of Theorem 4.2. We assume $n \geq 7$. For positions $5, \ldots, n-5$, each of these overlapping patterns can occur twice; otherwise, the patterns occur only once for a total of $2(n-8)+8 = 2n-8$ possibilities. There are $6(n-8)+2(5+4+3) = 6n-24$ overlapping patterns α and β that contribute to b_1. Thus,

$$\begin{aligned}
\lambda &= (n-4)q^3/I_4(q), \\
b_1 &= (6n-24)q^6/I_4(q)^2, \quad \text{and} \\
b_2 &= (2n-8)q^9(1+q+2q^2+2q^3+2q^4+q^5+q^6)/I_7(q),
\end{aligned}$$

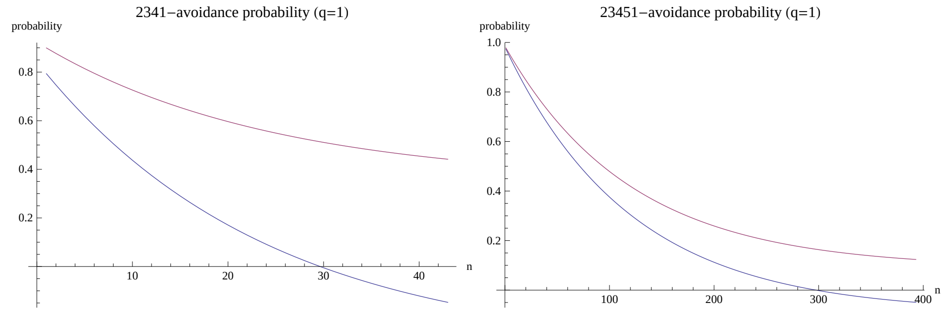

Figure 2. Plot of lower and upper bounds on pattern avoidance probabilities of uniform distribution ($q = 1$) for: (left) pattern $\tau = 2341$; (right) pattern $\tau = 23451$.

permutation	no. inversions	permutation	no. inversions	permutation	no. inversions
345627891	12	347825691	16	456923781	18
345726891	13	347925681	17	467823591	19
345826791	14	348925671	18	467923581	20
345926781	15	357824691	17	468923571	21
346725891	14	357924681	18	567823491	20
346825791	15	358924671	19	567923481	21
346924781	16	367824591	18	568923471	22
356724891	15	367924581	19	378924561	21
356824781	16	368924571	20	478923561	22
356924781	17	457823691	18	578923461	23
456723891	16	457923681	19	678923451	24
456823791	17	458923671	20		

Table 4. List of all permutations that have pattern 23451 in overlapping positions along with the number of inversions.

producing the bounds

$$e^{-\lambda} - (b_1 + b_2)\frac{1 - e^{-\lambda}}{\lambda} \leq P(W = 0) \leq e^{-\lambda} + (b_1 + b_2)\frac{1 - e^{-\lambda}}{\lambda}.$$

6.2.2. *The pattern* 23451. For $\tau = 23451$, we have $p_\alpha = q^4/I_5(q)$ and $b_1 = (n-5)q^4/I_5(q)$. The structure of τ only permits overlap with the first or last position. Table 4 lists all permutations that have pattern 23451 in the first 5 and last 5 positions. These are the only permutations that contribute to b_2 in the bound of Theorem 4.2. We assume $n \geq 9$. For positions $5, \ldots, n-5$, each of these overlapping patterns can occur twice; otherwise, the patterns occur only once for a total of $2(n-10) + 10 = 2n - 10$ possibilities. There are $8(n-10) + 2(7+6+5+4) = 8n - 36$ overlapping patterns α and β that contribute to b_1. Thus,

$$\lambda = (n-5)q^4/I_5(q),$$
$$b_1 = (8n-36)q^8/I_5(q)^2, \quad \text{and}$$
$$b_2 = (2n-10)q^{12}(1 + q + 2q^2 + 3q^3 + 4q^4 + 4q^5 + 5q^6 + 4q^7 + 4q^8 + 3q^9 + 2q^{10} + q^{11} + q^{12})/I_9(q),$$

producing the bounds

$$e^{-\lambda} - (b_1 + b_2)\frac{1 - e^{-\lambda}}{\lambda} \le P(W = 0) \le e^{-\lambda} + (b_1 + b_2)\frac{1 - e^{-\lambda}}{\lambda}.$$

7. Proofs

7.1. **Bounds on p_α and $p_{\alpha\beta}$.** We first prove several lemmas, from which the theorems follow. By the homogeneity property of uniform permutations we have

$$p_\alpha = \mathbb{E}\, X_\alpha = 1/j! \quad \text{for all } \alpha \in \Gamma_j.$$

To calculate the Poisson rate λ, we use linearity of expectation: there are $\binom{n}{j}$ possible j-tuples of elements in $[n]$, and so the expected number of subsets of j elements that reduce to the pattern $k_1 k_2 \cdots k_j$ is

$$\lambda = |\Gamma_j|\, p_\alpha = \binom{n}{j}/j!.$$

Next, we consider the joint expectation $p_{\alpha\beta} = \mathbb{E}\, X_\alpha X_\beta$.

Lemma 7.1. *Fix $\alpha, \beta \in \Gamma_j$ and let $s = 1, \ldots, j-1$ denote the number of elements that α, β have in common. For any such pair α, β, we have*

$$p_{\alpha\beta} \le \frac{s!}{j!^2}. \tag{23}$$

Proof. First we condition on X_α, which contributes a factor of $1/j!$. By conditioning on X_α, we assume that the s common elements are in their proper order with respect to X_β. It may so happen that, conditional on X_α, no such event can occur, which justifies the inequality.

Consider first $s = j - 1$, i.e., condition on $j - 1$ of the entries being in their proper order. Assuming that it is possible to realize both events simultaneously, the remaining element has probability $1/j$ of appearing in its correct order.

Consider, for general s, conditional on s entries being in their proper order, the probability that the remaining $j - s$ elements appear in their proper order is then $((s+1)(s+2)\cdots j)^{-1}$. □

7.2. **Proof of Theorem 3.4.** We have the following lemma.

Lemma 7.2. *For D_α defined as in Section 3.1 and b_3 as in Theorem 4.2, we have $b_3 = 0$ for all patterns τ.*

Proof. This follows from the dissociated property of uniform permutations (Lemma 5.3). We interpret the conditioning on $\sigma(X_\beta : \beta \notin D_\alpha)$ as the σ-algebra containing all possible information about just the order of a particular set of three elements. Since these three elements do not overlap any of the elements in α, knowing only their order does not affect X_α because uniform permutations are dissociated. □

Remark 7.3. *Note that the conditioning in the expression for b_3 is not the σ-algebra containing all information about the elements indexed by each tuple. If it were, then knowing their particular location would have an impact. However, simply knowing their order does not reveal any more information about X_α.*

Next we obtain the size of D_α, which is the same for all $\alpha \in \Gamma_j$.

Lemma 7.4. *For each $\alpha \in \Gamma_j$,*

$$|D_\alpha| = \binom{n}{j} - \binom{n-j}{j}.$$

Proof. Fix any $\alpha, \beta \in \Gamma_j$ and let $s = 1, \ldots, j$ denote the number of elements that α, β have in common. (This includes the case $\alpha = \beta$.) For each $s = 1, 2, \ldots, j$, we select any s elements out of the j for the two sets of indices α, β to have in common, then we select the remaining $j - s$ elements from the $n - j$ remaining elements that are not in α. That is,

$$|D_\alpha| = \sum_{s=1}^{j} \binom{n-j}{j-s}\binom{j}{s} = \binom{n}{j} - \binom{n-j}{j}, \qquad \text{for all } \alpha \in \Gamma_j. \qquad \square$$

Lemma 7.5. *We have*

$$d_1 = \binom{n}{j}\left(\binom{n}{j} - \binom{n-j}{j}\right) / j!^2.$$

Proof. Follows immediately from Lemma 7.4 and Lemma 7.1. $\square$

The expression given for d_2 in Equation (10) in the statement of Theorem 3.4 is straightforward, although it contains the overlap quantities $L_s(\tau)$, which can vary wildly for different patterns τ, and for which we are unaware of any general explicit or asymptotic expression. We calculate explicit upper bounds for d_2 in Lemma 7.7.

7.3. **Proof of Theorem 3.6 and Theorem 3.9.** Theorem 3.6 follows from Theorem 3.9 using $q = 1$ and the fact that $I_n(1) = n!$. Theorem 3.9 is a straightforward generalization of Proposition 5.5.

7.4. **Proof of corollaries.** For the proof of Corollary 3.1 it is sufficient that the bounds for d_1 and d_2 in Theorem 3.4 converge to 0 as $n \to \infty$.

Lemma 7.6. *Suppose $j \geq (e + \epsilon)\sqrt{n}$. Then for d_1 as in Theorem 3.4, we have $d_1 \to 0$.*

Proof. The proof is an elementary application of Stirling's formula. $\square$

For d_2, the asymptotic analysis is not so straightforward, which is why we instead use the inequality in Equation (23).

Lemma 7.7. *We have*

$$d_2 \leq \frac{\binom{n}{j}}{j!^2} \sum_{s=1}^{j-1} \binom{n-j}{j-s}\binom{j}{s} s! \leq \frac{e^{2\left(\frac{1}{12} - \frac{3}{13}\right)}}{(2\pi)^2} \left(\frac{e^2(n-j)}{j^2}\right)^j \frac{1}{j}\, e^{2\sqrt{n-j}} \log(j).$$

Whence, for any $\eta \geq 0$, taking $j \geq (e\, e^{1/e} + \eta)\sqrt{n}$, we have $d_2 \to 0$.

Proof. We count the number of pairs (α, β), $\alpha \in \Gamma_j$ and $\beta \in D_\alpha$ with exactly s shared elements. We may first choose any $2j - s$ locations among the n possible choices for the patterns to occur on. Of those $2j - s$ locations, we can choose any j of them for the elements of α. Then, of those j locations, any s can also be shared with β. Thus, for a given $s \in \{1, 2, \ldots, j-1\}$, there are

$$\binom{n}{2j-s}\binom{2j-s}{j}\binom{j}{s} = \binom{n}{j}\binom{n-j}{j-s}\binom{j}{s}$$

terms in the sum. Using Equation (23), we have

$$d_2 \le \frac{\binom{n}{j}}{j!^2} \sum_{s=1}^{j-1} \binom{n-j}{j-s} \binom{j}{s} s! = \frac{\binom{n}{j}}{j!} \sum_{s=1}^{j-1} \frac{\binom{n-j}{j-s}}{(j-s)!}.$$

In order to handle the sum, first we recall the quantitative bounds of Robbins [35], i.e.,

$$e^{\frac{1}{12n+1}} < \frac{n!}{(n/e)^n \sqrt{2\pi n}} < e^{\frac{1}{12n}}, \qquad \text{for all } n \ge 1.$$

Again we emphasize that this inequality holds for all $n \ge 1$, which allows us to provide the simpler bound of

$$d_2 \le \lambda \exp\left(\left(\frac{1}{12} - \frac{3}{13}\right)\right) \sum_{s=1}^{j-1} \left(\frac{e^2(n-j)}{(j-s)^2}\right)^{j-s} \frac{e^{-\frac{(j-s)^2}{n-j}}}{2\pi(j-s)}.$$

The term $\left(\frac{e^2(n-j)}{(j-s)^2}\right)^{j-s}$ is maximized when $j - s = \sqrt{n-j}$, whence

$$d_2 \le \frac{\lambda}{2\pi} \exp\left(\left(\frac{1}{12} - \frac{3}{13}\right)\right)\left(e^2\right)^{\sqrt{n-j}} \sum_{s=1}^{j-1} \frac{e^{-s^2/(n-j)}}{s} \le \frac{\lambda}{2\pi} \exp\left(\left(\frac{1}{12} - \frac{3}{13}\right)\right) e^{2\sqrt{n-j}} \log(j).$$

Note next that

$$\lambda \le \left(\frac{e^2\, n}{j^2}\right)^j \frac{e^{-\frac{j^2}{n}}}{2\pi\, j},$$

so that for $j \sim (e + \epsilon)\sqrt{n}$, we have

$$\lambda \le \left(1 + \frac{\epsilon}{e}\right)^{2j} \frac{e^{-(e+\epsilon)^2}}{2\pi\, j},$$

and so

$$d_2 \le \left(\left(1 + \frac{\epsilon}{e}\right)\left(e^{\frac{1}{e+\epsilon}}\right)\right)^{2j} \frac{\log j}{2\pi j} \le \left(\frac{e^{1/e}}{1 + \frac{\epsilon}{e}}\right)^{2j} \frac{\log j}{2\pi j}.$$

Letting $\epsilon = e(e^{1/e} - 1) + \eta$ for any $\eta \ge 0$, we conclude that taking $j \ge (e\, e^{1/e} + \eta)\sqrt{n}$ implies $d_2 \to 0$. □

We now compute an upper bound on $\overline{d}_2$ which is explicit and establishes Corollary 3.2.

Lemma 7.8.

$$\overline{d}_2 \le \sum_{s=1}^{m-1} (n - 2m + s) \frac{s!}{(m)!^2} \sim \frac{n}{m!} \frac{1}{m}.$$

Taking $m \ge \Gamma^{(-1)}(n/t) - 1$ and $m \sim \Gamma^{(-1)}(n/t) - 1$, we have $\overline{d}_2 \le \frac{t}{m} \to 0$ as n tends to infinity.

Proof. It is easy to see that

$$\sum_{s=1}^{m-1} \frac{1}{(m+s-1)!} = \frac{1}{m!}\left(1 + O(m^{-1})\right),$$

and also

$$\sum_{s=1}^{m-1} \frac{s}{(m+s-1)!} = \frac{1}{m!}\left(1 + O(m^{-1})\right).$$

Using Equation (23), the result immediately follows. □

For a more explicit form of the growth of m, we define $\psi(x) = \Gamma'(x)/\Gamma(x)$, the digamma function, as the logarithmic derivative of the gamma function, and denote by k_0 the smallest positive root of $\psi(x)$, i.e., $k_0 = 1.46163\ldots$. Also, let $c = e^{-1}\sqrt{2\pi} - \Gamma(k_0) = 0.036534\ldots$, and denote by $W(x)$ the Lambert W function, i.e., the solution to $x = W(x)e^{W(x)}$. Finally, let $L(x) := \log((x+c)/2\pi)$.

Lemma 7.9 ([11]). *As x tends to infinity, we have*

$$\Gamma^{(-1)}(x) \sim \frac{L(x)}{W(L(x)/e)} + \frac{1}{2} \sim \frac{\log(x)}{\log\log(x) - \log\log\log(x)}.$$

Lemma 7.10. *Suppose $t > 0$ is some fixed constant and $m = \lceil \Gamma^{(-1)}(n/t) - 1 \rceil$. Then*

$$\overline{\lambda} = \frac{n-m}{m!} \to t.$$

Remark 7.11. *We must be slightly careful when specifying the length of the pattern m in Lemma 7.10, since in general $\Gamma^{(-1)}(n/t) - 1$ will not be an integer. However, as long as m always exceeds this value, which we have ensured by setting it equal to the smallest integer exceeding it, then the asymptotic expressions still hold.*

8. Acknowledgements

The authors would like to acknowledge helpful comments from Brendan McKay, Dan Romik, Igor Rivin, Jim Pitman, Richard Arratia, Igor Pak, Michael Albert, and Doron Zeilberger.

References

[1] Albert, M. H., Atkinson, M. D., and Brignall, R. The enumeration of permutations avoiding 2143 and 4231. *Pure Math. Appl. (PU.M.A.) 22*, 2 (2011), 87–98.

[2] Arratia, R., Goldstein, L., and Gordon, L. Two moments suffice for poisson approximations: the chen-stein method. *The Annals of Probability* (1989), 9–25.

[3] Arratia, R., and Tavaré, S. The cycle structure of random permutations. *Ann. Probab. 20*, 3 (1992), 1567–1591.

[4] Atkinson, M. D. Restricted permutations. *Discrete Math. 195*, 1-3 (1999), 27–38.

[5] Barbour, A. D., Holst, L., and Janson, S. *Poisson approximation*. Clarendon Press Oxford, 1992.

[6] Bhatnagar, N., and Peled, R. Lengths of Monotone Subsequences in a Mallows Permutation. *Probability Theory and Related Fields* (2014).

[7] Bóna, M. *Combinatorics of Permutations*. Discrete Mathematics and Its Applications. CRC Press, 2012. Second Edition.

[8] Bona, M. A new record for 1324-avoiding permutations. *European Journal of Mathematics 1* (2014), 1–9.

[9] Bona, M. A new upper bound for 1324-avoiding permutations. *Combinatorics, Probability and Computing 23*, 5 (2014), 717–724.

[10] Cameron, N., and Killpatrick, K. Inversion polynomials for permutations avoiding consecutive patterns. *Adv. in Appl. Math. 67* (2015), 20–35.

[11] Cantrell, D. `http://mathforum.org/kb/message.jspa?messageID=342551&tstart=0`.

[12] Chen, L. H. Poisson approximation for dependent trials. *The Annals of Probability* (1975), 534–545.

[13] Claesson, A., Jelinek, V., and Steingrímsson. Upper bounds for the Stanley–Wilf limit of 1324 and other layered patterns. *Journal of Combinatorial Theory, Series A 119*, 8 (2012), 1680–1691.

[14] Crane, H., DeSalvo, S., and Elizalde, S. The probability of avoiding consecutive patterns in the mallows distribution. *arXiv preprint arXiv:1609.01370* (2016).
[15] de Montmort, P. R. Essay d'analyse sur les jeux de hazard. *Revue & augmentée de plusieurs Lettres* (1708).
[16] Dokos, T., Dwyer, T., Johnson, B., Sagan, B., and Selsor, K. Permutation Patterns and Statistics. *Discrete Mathematics 312* (2012), 2760–2775.
[17] Egge, E. S., and Mansour, T. Permutations which avoid 1243 and 2143, continued fractions, and Chebyshev polynomials. *Electron. J. Combin. 9*, 2 (2002/03), Research paper 7, 35 pp. (electronic). Permutation patterns (Otago, 2003).
[18] Elizalde, S., and Noy, M. Consecutive patterns in permutations. *Advances in Applied Mathematics 30* (2003), 110–125.
[19] Gnedin, A., and Olshanski, G. A q-analogue of de Finetti's theorem. *Electronic Communications in Probability 16*, R78 (2009).
[20] Gnedin, A., and Olshanski, G. q-exchangeability via quasi-invariance. *Annals of Probability 38*, 6 (2010), 2103–2135.
[21] Hoffman, C., Rizzolo, D., and Slivken, E. Pattern-avoiding permutations and Brownian excursion, Part I: Shapes and fluctuations. *arXiv:1506.04174* (2016).
[22] Hoffman, C., Rizzolo, D., and Slivken, E. Pattern-avoiding permutations and Brownian excursion, Part II: Fixed points. *arXiv:1406.5156* (2016).
[23] Johansson, F., and Nakamura, B. Using functional equations to enumerate 1324-avoiding permutations. *Advances in Applied Mathematics 56* (2014), 20–34.
[24] Kerov, S. V., and Vershik, A. M. The characters of the infinite symmetric group and probability properties of the Robinson-Schensted-Knuth algorithm. *SIAM J. Algebraic Discrete Methods 7*, 1 (1986), 116–124.
[25] Kitaev, S., and Mansour, T. Simultaneous avoidance of generalized patterns. *Ars Combin. 75* (2005), 267–288.
[26] Knuth, D. E. *The Art of Computer Programming*, vol. Volume 1: Fundamental Algorithms. Addison Wesley, Massachusetts, 1973.
[27] Logan, B. F., and Shepp, L. A. A variational problem for random Young tableaux. *Advances in Math. 26*, 2 (1977), 206–222.
[28] Mallows, C. Non-null ranking models. *Biometrika 44* (1957), 114–130.
[29] Marcus, A., and Tardos, G. Excluded permutation matrices and the Stanley–Wilf conjecture. *Journal of Combinatorial Theory, Series A 107*, 1 (2004), 153–160.
[30] McKay, B. http://mathoverflow.net/questions/216268/asymptotic-expression-for-j-which-satisfies-binomnj-j-sim-k-as-n-to.
[31] Nakamura, B. Approaches for enumerating permutations with a prescribed number of occurrences of patterns.
[32] Perarnau, G. A probabilistic approach to consecutive pattern avoiding permutations. *J. Comb. Th., Series A 120* (2013), 998–1011.
[33] Pratt, V. R. Computing permutations with double-ended queues, parallel stacks and parallel queues. In *Proceedings of the fifth annual ACM symposium on Theory of computing* (1973), ACM, pp. 268–277.
[34] Ray, N., and West, J. Posets of matrices and permutations with forbidden subsequences. *Annals of Combinatorics 7*, 1 (2003), 55–88.
[35] Robbins, H. A remark on Stirling's formula. *Amer. Math. Monthly 62* (1955), 26–29.
[36] Shepp, L. A., and Lloyd, S. P. Ordered cycle lengths in a random permutation. *Trans. Amer. Math. Soc. 121* (1966), 340–357.
[37] Sloane, N. Online Encyclopedia of Integer Sequences. *Published electronically at http://www.oeis.org/*.
[38] Vatter, V. Problems and conjectures presented at the problem session. *Permutation Patterns, LMS Lecture Notes Series 376*, 339–345.

Rutgers University, 110 Frelinghuysen Road, Piscataway, NJ, 08854

UCLA, 520 Portola Plaza, Los Angeles, CA, 90095