

# Pattern avoidance for random permutations

Harry Crane<sup>1\*</sup>Stephen DeSalvo<sup>2</sup><sup>1</sup> Rutgers University, 110 Frelinghuysen Road, Piscataway, NJ, 08854<sup>2</sup> University of California Los Angeles, 520 Portola Plaza, Los Angeles, CA, 90095received 22<sup>nd</sup> Mar. 2017, revised 26<sup>th</sup> Sep. 2017, accepted 3<sup>rd</sup> Aug. 2018.

Using techniques from Poisson approximation, we 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 of any pattern. 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.

**Keywords:** pattern avoidance; Mallows distribution; random permutation; Poisson approximation

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Motivation</b>	<b>3</b>
<b>3</b>	<b>Main Results</b>	<b>4</b>
<b>4</b>	<b>Poisson approximation</b>	<b>9</b>
<b>5</b>	<b>Consecutive pattern avoidance of Mallows permutations</b>	<b>11</b>
<b>6</b>	<b>Numerical examples</b>	<b>16</b>
<b>7</b>	<b>Proofs</b>	<b>18</b>
<b>8</b>	<b>Acknowledgements</b>	<b>21</b>

## 1 Introduction

A permutation of  $[n] = \{1, \dots, n\}$  is a bijection  $\sigma : [n] \rightarrow [n]$ ,  $i \mapsto \sigma(i) = \sigma_i$ , written  $\sigma = \sigma_1 \cdots \sigma_n$ . For each  $n = 1, 2, \dots$ , 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  $\sigma$  *avoids*  $\tau$  if there does not exist a subsequence  $1 \leq i_1 < \dots < i_m \leq n$  such that  $\sigma_{i_1} \cdots \sigma_{i_m}$  is order-isomorphic to  $\tau$ , and we say that  $\sigma$  *avoids*  $\tau$  *consecutively* if there is no  $j = 1, \dots, n-m+1$  such that  $\sigma_j \sigma_{j+1} \cdots \sigma_{j+m-1}$  is order-isomorphic to  $\tau$ . Here we study pattern avoidance probabilities for random permutations from the Mallows distribution, which is of particular interest

\*partially supported by NSF CAREER Award (DMS-1554092)

in the fields of statistics and probability but some special cases are insightful for questions in enumerative and extremal combinatorics.

To demonstrate our approach, let  $n$  be any integer larger than or equal to 6. It was shown in [31] that each fixed pattern  $\tau$  of length  $n - 1$  is avoided by exactly  $n! - ((n - 1)^2 + 1)$  permutations in  $\mathcal{S}_n$ . In [32], it was shown that for each fixed pattern  $\tau$  of length  $N \equiv n - 2$ , there exists  $0 \leq j \leq N - 1$  such that exactly

$$n! - \frac{N^4 + 2N^3 + N^2 + 4N + 4 - 2j}{2}$$

permutations in  $\mathcal{S}_n$  avoid  $\tau$ . It is left as an open problem in [36] to perform the same analysis for patterns of length  $n - 3$ . From this, it is natural to ruminate on whether for each pattern  $\tau$  of length  $n - m$ , one has approximately

$$n! - \frac{n!}{(n - m)!} \binom{n}{n - m} \quad (1)$$

permutations in  $\mathcal{S}_n$  which avoid it, and how large  $m$  can be taken. In this paper, we quantify the quality of approximation of (1) for counting the number of permutations in  $\mathcal{S}_n$  which avoid any given pattern  $\tau$  of length  $n - m$ ,  $1 \leq m \leq n$ , via Poisson approximation. This approach yields general results which specialize to the counting problem just described, as well as the analogous problem for consecutive pattern avoidance.

For classical pattern avoidance, we are able to show a good approximation if for any  $\epsilon > 0$  we have  $m \geq (e e^{1/e} + \epsilon) \sqrt{n}$ , but we suspect the  $e e^{1/e}$  constant can be improved with a natural lower bound of  $e$  demonstrated by Lemma 3.4. For consecutive pattern avoidance, we obtain stronger results, namely, we may take any  $m \geq \lceil \Gamma^{-1}(n) - 1 \rceil$ . (In the above expression,  $\Gamma^{-1}(n)$  is the inverse of the gamma function, i.e.,  $\Gamma(t) = \int_0^\infty x^t e^{-x} dx$ , which gives the usual generalization of the factorial function to all positive real numbers.) The above expression is asymptotically best possible since, in the limit, taking  $m$  equal to the right-hand side yields a Poisson approximation with nonzero rate parameter for the number of occurrences of a consecutive pattern. The range of values for  $m$  follows by a detailed analysis of the bounds contained in Theorem 3.6 and Theorem 3.7, respectively.

The Mallows permutations we study here is a general class of random permutations whose distribution is weighted by the number of inversions. 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  $\text{inv}(\sigma)$  to denote the set of inversions of  $\sigma$ . With  $\Sigma_n$  denoting a random permutation of  $[n]$ , the *Mallows distribution with parameter*  $q \in (0, \infty)$  on  $\mathcal{S}_n$  assigns probability

$$P\{\Sigma_n = \sigma\} = P_n^q(\sigma) = q^{|\text{inv}(\sigma)|} / I_n(q), \quad \sigma \in \mathcal{S}_n, \quad (2)$$

where  $I_n(q) = \prod_{j=1}^n \sum_{i=0}^{j-1} q^i$  is the *inversion polynomial* and  $|\text{inv}(\sigma)|$  is the number of inversions in  $\sigma$ . 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,  $q > 1$ .

The Mallows distribution [26] 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 [17, 18]. For general values of  $q > 0$ , we consider the problem of consecutive pattern avoidance for random Mallows permutations, with Theorem 3.8 establishing explicit error bounds on the entire distribution of the number of occurrences of patterns in a random permutation. Our main theorems, therefore, complement prior work by Elizalde & Noy [16], Perarnau [30], and the more recent work by the current authors & Elizalde [13] on consecutive pattern avoidance, as well as Nakamura [29], who used functional equations to enumerate sets with a prescribed number of occurrences of a given pattern.

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 [11], in particular [2], to bound the total variation distance between the collection of all *dependent* indicator random variables indicating pattern occurrence for a prescribed set of indices, and a joint distribution of *independent* Bernoulli random variables with the same marginal distributions. These bounds allow us to approximate any measurable function of the occurrences, e.g., the number of patterns, locations of patterns, etc., via the corresponding independent random variables. We also reach a natural limit of the usefulness of these approximations corresponding to the strength of interactions between occurrences of patterns.

The next section gives a brief historical context of restricted permutations. Section 3 presents the main results, of which there are two kinds: the first type is a preasymptotic bound given in terms of quantities which are complicated to compute but as accurate as the method allows; the second type is a more detailed analysis of the bounds in a form that is better suited to applications. Section 4 presents the Chen–Stein Poisson approximation approach that we utilize throughout the paper. In Section 5 we apply this method to Mallows permutations. Section 6 contains explicit numerical examples. Finally, the main technical results are proved in Section 7.

## 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 [14]; 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, for which we have the asymptotic expression

$$D_n = n! \sum_{i=0}^n \frac{(-1)^i}{i!} \sim n!/e \quad \text{as } n \rightarrow \infty. \quad (3)$$

Equation (3) can be stated in probabilistic terms by letting  $\Sigma_n$  be a uniform random permutation of  $[n]$ , i.e.,  $P\{\Sigma_n = \sigma\} = 1/n!$  for each  $\sigma \in \mathcal{S}_n$ , for which we compute

$$P\{\Sigma_n \text{ is a derangement}\} = D_n/n! \sim 1/e \quad \text{as } n \rightarrow \infty. \quad (4)$$

See [3, 34] for more thorough treatments involving the cycle structure of random permutations.

We can also derive the expression in (3) 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, 19, 20] 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  $\{w_{(1)}, \dots, w_{(k)}\}_<$  denoting the set of elements listed in increasing order, we define the map  $w_{(i)} \mapsto i$ , under which  $w$  maps to a permutation  $\text{red}(w)$  of  $[k]$ , called the *reduction* of  $w$ . For example,  $w = 826315$  reduces to  $\text{red}(w) = 625314$ . We call any fixed  $\tau \in \mathcal{S}_m$  a *pattern* and say that  $\sigma \in \mathcal{S}_n$  *contains*  $\tau$  if there exists a subsequence  $1 \leq i_1 < \cdots < i_m \leq m$  such that  $\text{red}(\sigma_{i_1} \cdots \sigma_{i_m}) = \tau$ . We say  $\sigma \in \mathcal{S}_n$  *avoids*  $\tau$  if it does not contain it. We say that  $\sigma$  *contains*  $\tau$  *consecutively* if there exists an index  $j \in [n - m + 1]$  such that  $\text{red}(\sigma_j \sigma_{j+1} \cdots \sigma_{j+m-1}) = \tau$ ; otherwise, we say  $\sigma$  *avoids*  $\tau$  *consecutively*. For any pattern  $\tau$ , we define

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

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

$$\begin{aligned} \mathcal{S}_n(A) &:= \{\sigma \in \mathcal{S}_n : \sigma \text{ avoids all } \tau \in A\} \quad \text{and} \\ \overline{\mathcal{S}}_n(A) &:= \{\sigma \in \mathcal{S}_n : \sigma \text{ avoids all } \tau \in A \text{ consecutively}\}. \end{aligned} \tag{6}$$

For the most part, we are interested in sets  $\mathcal{S}_n(\tau)$  containing all permutations that avoid a given pattern  $\tau$ , though our approach extends in a straightforward manner for more general sets  $A$ .

Much effort has been devoted to exact enumeration of  $\mathcal{S}_n(A)$  for certain choices of  $A$ , see, e.g., [1, 4, 15, 23]. But enumerating  $\mathcal{S}_n(\tau)$  is notoriously difficult for patterns of fixed length larger than 3. Knuth [24] 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  $\tau$  are enumerated by the Catalan numbers [35, 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  $\Sigma_n$  denoting a uniform random permutation of  $[n]$  and  $A$  a set of permutations, the probability that  $\Sigma_n$  avoids  $A$  is

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

The Stanley-Wilf conjecture as proved in [27] states that  $|\mathcal{S}_n(\tau)|$  grows exponentially with  $n$  for every fixed  $\tau$ . For example, the Catalan numbers are known to grow asymptotically like  $4^n / \sqrt{\pi n^3}$ , yielding the asymptotic avoidance probability

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

Such calculations quickly become intractable as  $n$  grows large. For example, the sets of 1324-avoiding permutations have only been enumerated up to  $n = 31$  [21]. Even precise asymptotics for  $|\mathcal{S}_n(1324)|$  have not yet been established [8, 9, 12].

### 3 Main Results

To fix notation throughout the text, we write  $\sigma = \sigma_1 \cdots \sigma_n$  to denote a generic permutation. For any subset  $A \subseteq [n]$ , we write  $\sigma|_A$  to denote the *restriction* of  $\sigma$  to a permutation of  $A$  obtained by removing those elements among  $\sigma_1, \dots, \sigma_n$  that are not in  $A$ . For example, with  $\sigma = 867531924$  and  $A = \{1, 3, 5, 7, 9\}$ , we have  $\sigma|_A = 75319$ . We write  $\Sigma_n$  to denote a random permutation of  $[n]$ .

#### 3.1 Definitions

Throughout the paper, we write  $\mathcal{L}(X)$  to denote the *distribution*, or *law*, of a random variable  $X$  and  $\mathcal{L}(Y | X)$  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 is defined as

$$d_{TV}(\mathcal{L}(X), \mathcal{L}(Y)) = \sup_{A \subseteq \mathbb{R}} |P(X \in A) - P(Y \in A)|,$$

where the sup is taken over Borel measurable subsets of  $\mathbb{R}$ . In the special case of non-negative integer-valued random variables, the total variation distance 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  $m$ -tuples of elements from  $[n]$  by

$$J_m := \{\{i_1, \dots, i_m\} : 1 \leq i_1 < \dots < i_m \leq n\}, \quad m \in [n].$$

For each  $\alpha \in J_m$ , let  $D_\alpha$  be the set of all  $m$ -element subsets of  $[n]$  that overlap with  $\alpha$  in at least one element, i.e.,  $D_\alpha := \{\beta \in J_m : \beta \cap \alpha \neq \emptyset\}$ . 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  $\Sigma_n$  denoting a uniform random permutation of  $[n]$ ,  $\alpha \in J_m$ , and  $\tau$  a fixed pattern of length  $m$ , we define  $X_\alpha = X_{i_1, \dots, i_m}$  as the indicator random variable for the event that the reduction of  $\Sigma_n$  at positions  $i_1, \dots, i_m$  gives the pattern  $\tau$ , i.e.,

$$X_{i_1, \dots, i_m} := \mathbb{I}(\text{red}(\Sigma_n(i_1) \cdots \Sigma_n(i_m)) = \tau). \quad (7)$$

Let  $\mathbb{X} \equiv \mathbb{X}_m := (X_\alpha)_{\alpha \in J_m}$  be the collection of all such indicators and let  $\mathbb{B} \equiv \mathbb{B}_m = (B_\alpha)_{\alpha \in J_m}$  denote a collection of independent Bernoulli random variables whose marginal distributions satisfy  $\mathbb{E} B_\alpha = \mathbb{E} X_\alpha$  for each  $\alpha \in J_m$ . The random variable

$$W = \sum_{1 \leq i_1 < \dots < i_m \leq n} X_{i_1, \dots, i_m} \quad (8)$$

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

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

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

$$\bar{J}_m := \{\{i, i+1, \dots, i+m-1\} : 1 \leq i \leq n-m+1\}, \quad m \in [n].$$

Let  $\bar{\mathbb{X}} \equiv \bar{\mathbb{X}}_m := (X_\alpha)_{\alpha \in \bar{J}_m}$ , and let  $\bar{\mathbb{B}} \equiv \bar{\mathbb{B}}_m = (B_\alpha)_{\alpha \in \bar{J}_m}$  denote a collection of independent Bernoulli random variables whose marginal distributions satisfy  $\mathbb{E} B_\alpha = \mathbb{E} X_\alpha$  for each  $\alpha \in \bar{J}_m$ . For fixed  $\tau \in \mathcal{S}_m$  and  $\Sigma_n$  a uniform random permutation of  $n$ , we define the random variable

$$\bar{W} := \sum_{1 \leq s \leq n-m+1} X_{s, s+1, \dots, s+m-1}, \quad (9)$$

which counts the number of consecutive occurrences of  $\tau$  in  $\Sigma_n$ . We also define  $\bar{L}_s(\tau)$  as the *sequential overlap of size  $s$* , i.e., the number of permutations  $\sigma \in \mathcal{S}_{2m-s}$  for which

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

### 3.2 Main results on pattern avoidance

We begin with several theorems specifically about pattern avoidance, which follow from the quantitative bounds given in Section 3.3.

**Theorem 3.1.** *Assume  $n \geq m \geq 3$  and  $\tau$  is any pattern of length  $m$ . Define*

$$\begin{aligned} \lambda &= \binom{n}{m} / m!, & d_1 &= \binom{n}{m} \left( \binom{n}{m} - \binom{n-m}{m} \right) \frac{1}{m!^2}, \\ d_2 &= \sum_{s=1}^{m-1} \binom{n}{2m-s} \frac{2L_s(\tau)}{(2m-s)!} \end{aligned} \quad (10)$$

and

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

Then we have

$$n! e^{-\lambda} (1 - e^{\lambda} D_{n,m}) \leq |\mathcal{S}_n(\tau)| \leq n! e^{-\lambda} (1 + e^{\lambda} D_{n,m}). \quad (11)$$

In addition, for any fixed  $\epsilon \geq 0$ , suppose  $m \equiv m(n)$  is some increasing, integer-valued function of  $n$  such that  $m \geq (e e^{1/e} + \epsilon) \sqrt{n}$ . For any sequence of patterns  $\tau_n \in \mathcal{S}_m$ , we have

$$\left| \frac{|\mathcal{S}_n(\tau)|}{n!} - e^{-\lambda} \right| = o(1) \quad \text{as } n \rightarrow \infty.$$

**Remark 3.2.** There are several noteworthy aspects to Theorem 3.1.

1. The expression for  $\lambda$  equals the expected number of occurrences of  $\tau$  in a uniform random permutation of  $[n]$ , and thus is the same for all patterns of length  $m$ .
2. The expression for  $d_1$  cannot be improved using our approach.
3. We are not aware of any efficient means to calculate the values  $L_s(\tau)$  in general. For a simple and explicit upper bound, applicable for all patterns  $\tau$  of length  $m$ , we have used

$$d_2 \leq \binom{n}{m} \sum_{s=1}^{m-1} \binom{n-m}{m-s} \binom{m}{s} \frac{s!}{m!^2}, \quad (12)$$

but we suspect that this bound can be improved.

It is tempting to conjecture that Theorem 3.1 holds even when  $\lambda$  tends to some fixed positive constant, but we suspect this is not possible, as we now demonstrate.

**Lemma 3.3** ([28]). Fix any  $t > 0$  and let  $\lambda = \binom{n}{m}/m!$ . Then  $\lambda \rightarrow t$  as  $n \rightarrow \infty$  provided

$$m \sim e \sqrt{n} - \frac{1}{4} \log(n) - \frac{1}{2} \log(2\pi t) - \frac{1}{4} e^2 - \frac{1}{2} \quad \text{as } n \rightarrow \infty. \quad (13)$$

**Lemma 3.4.** Suppose  $n$ ,  $m$ , and  $n - m$  tend to infinity and  $d_1$  is as defined in Theorem 3.1. Then we have

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

In particular, for any  $\epsilon > 0$  and  $m \geq (e + \epsilon) \sqrt{n}$ , we have  $d_1 \rightarrow 0$ , and for  $m \sim e \sqrt{n} - \frac{1}{4} \log(n)$ , we have  $d_1 \rightarrow c \in (0, \infty)$ .

It follows that a necessary condition for  $d_1$  to tend to zero is that

$$m \gtrsim e \sqrt{n} - \left( \frac{1}{4} - \epsilon \right) \log(n) \quad \text{for any } \epsilon > 0, \quad (14)$$

where the notation  $a \gtrsim b$  means that both  $a \geq b$  and  $a \sim b$ . It is also well known, see [22, 25], that the typical size of the longest increasing subsequence of a random permutation of size  $n$  is asymptotically of order  $2\sqrt{n}$ , and so one cannot have a Poisson limit theorem which applies to the increasing pattern  $12 \dots m$  with  $m < 2\sqrt{n}$ . It is interesting to investigate the behavior in the gap, i.e., for  $m \sim c\sqrt{n}$  with any  $2 < c < e$ , and we leave this as an open problem.

We also have an analogous theorem for consecutive patterns.

**Theorem 3.5.** *Assume  $n \geq m \geq 3$  and  $\tau$  is any pattern of length  $m$ . Define*

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

and

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

Then we have

$$n! e^{-\bar{\lambda}} (1 - e^{\bar{\lambda}} \bar{D}_{n,m}) \leq |\bar{\mathcal{S}}_n(\tau)| \leq n! e^{-\bar{\lambda}} (1 + e^{\bar{\lambda}} \bar{D}_{n,m}). \quad (15)$$

In addition, fix any  $t > 0$ , and define  $M(t, n) := \lfloor \frac{\log(n/t)}{\log \log(n/t) - \log \log \log(n/t)} - \frac{1}{2} \rfloor$ . Let  $m \equiv m(n)$  be some increasing, integer-valued function of  $n$  such that  $m \geq M(t, n)$ . For any sequence of patterns  $\tau_n \in \mathcal{S}_m$ , we have  $\bar{D}_{n,m} \rightarrow 0$ , and thus

$$\left| \frac{|\bar{\mathcal{S}}_n(\tau)|}{n!} - e^{-\bar{\lambda}} \right| = o(1) \quad \text{as } n \rightarrow \infty.$$

Furthermore, define  $T_n(\tau, k)$  as the number of permutations in  $\mathcal{S}_n$  where pattern  $\tau$  occurs exactly  $k$  times, and take  $m(n) = M(t, n)$ . For any sequence of patterns  $\tau_n \in \mathcal{S}_m$ , we have

$$\sum_{k=0}^{\infty} \left| \frac{T_n(\tau, k)}{n!} - \frac{t^k}{k!} e^t \right| = o(1) \quad \text{as } n \rightarrow \infty.$$

In Section 3.4, we present an analogous result for permutations chosen according to the Mallows( $q$ ) distribution.

### 3.3 Main results on uniform permutations

Theorems 3.1 and 3.5 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 preasymptotic bounds on various quantities when all sizes are fixed.

**Theorem 3.6.** *Assume  $n \geq m \geq 3$  and  $\tau$  is any pattern of length  $m$ . Recalling the definition of  $X_\alpha$  in (7), we let  $\mathbb{X} = (X_\alpha)_{\alpha \in J_m}$  be the collection of all such variables and  $\mathbb{B} = (B_\alpha)_{\alpha \in J_m}$  be an independent Bernoulli process with marginal distributions satisfying  $\mathbb{E} B_\alpha = \mathbb{E} X_\alpha$  for all  $\alpha \in J_m$ . We have*

$$d_{TV}(\mathcal{L}(\mathbb{X}), \mathcal{L}(\mathbb{B})) \leq 4D_{n,m} + \frac{2\lambda}{m!}, \quad (16)$$

where  $D_{n,m}$  and  $\lambda$  are as defined in Theorem 3.1. Furthermore, with  $W$  defined as in (8), and for  $Y$  a Poisson random variable with mean  $\lambda = \mathbb{E} W$ , we have

$$d_{TV}(\mathcal{L}(W), \mathcal{L}(Y)) \leq 2D_{n,m}.$$

**Theorem 3.7.** *Assume  $n \geq m \geq 3$  and  $\tau$  is any pattern of length  $m$ . Recall the definitions of  $\bar{\mathbb{X}}$ ,  $\bar{\mathbb{B}}$ , and  $\bar{W}$  given in Section 3.1. We have*

$$d_{TV}(\mathcal{L}(\bar{\mathbb{X}}), \mathcal{L}(\bar{\mathbb{B}})) \leq 4\bar{D}_{n,m} + \frac{2\bar{\lambda}}{m!}, \quad (17)$$

where  $\bar{D}_{n,m}$  and  $\bar{\lambda}$  are as defined in Theorem 3.5. Let  $\bar{Y}$  denote a Poisson random variable with parameter  $\bar{\lambda}$ . We have

$$d_{TV}(\mathcal{L}(\bar{W}), \mathcal{L}(\bar{Y})) \leq 2\bar{D}_{n,m}.$$

### 3.4 Main results on Mallows permutations

In Section 5, we discuss several special properties of the Mallows distribution that are useful for studying consecutive pattern avoidance. Using these properties, we obtain analogous bounds to those in Theorems 3.6 and 3.7.

Recall the definition of the *restriction*  $\Sigma_{n|A}$  of  $\Sigma_n$  to a subset  $A \subseteq [n]$ , as defined at the beginning of Section 3. Also recall  $\bar{J}_m$  denotes the set of subsets of size  $m$  whose elements are consecutive in  $\{1, 2, \dots, n\}$ .

**Theorem 3.8.** *Fix  $q > 0$  and let  $\Sigma_n \sim \text{Mallows}(q)$  as defined in (2). For any  $m \geq 2$ , fix any pattern  $\tau_m$  of size  $m$ . For any  $\alpha \in \bar{J}_m$ , let*

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

and  $W = \sum_{\alpha \in \bar{J}_m} X_\alpha$ , 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

$$\lambda = (n-m) \frac{q^{|\text{inv}(\tau_m)|}}{I_m(q)}, \quad b_1 = \frac{(2mn - 3m^2 + m) q^{2|\text{inv}(\tau_m)|}}{I_m(q)^2}, \quad b_2 = \sum_{s=1}^{m-1} (n-2m+s) \sum_{\rho \in \bar{I}_s(\tau_m)} \frac{q^{|\text{inv}(\rho)|}}{I_{2m-s}(q)},$$

where  $I_m(q)$  is as defined following (2).

Asymptotic formulas and Poisson limit theorems for general Mallows permutations depend on the interplay between the parameters  $n$ ,  $m$ ,  $|\text{inv}(\tau)|$ , 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

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

which, for  $m$  fixed, produces non-trivial limiting behavior whenever

$$q \sim n^{-1/|\text{inv}(\tau)|} \quad \text{or} \quad q \sim n^{1/(\binom{m}{2} - |\text{inv}(\tau)|)} \quad \text{as } n \rightarrow \infty.$$

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

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

Combined with Theorem 3.8, these observations yield Theorem 3.9 below.

**Theorem 3.9.** *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  $n \geq 1$ , let  $\Sigma_n$  be a random permutation from the Mallows distribution (2) with parameter  $q(n)$ , with  $\bar{\mathbb{X}}_q$  and  $\bar{\mathbb{B}}_q$  defined analogously. For any measurable function  $h : \{0, 1\}^{n-m+1} \rightarrow \mathbb{R}$  and Borel set  $A \subseteq \mathbb{R}$ , we have*

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

provided either

$$q(n) \leq n^{-1/\text{inv}(\tau_{m(n)})} \quad \text{for almost all } n \geq 1,$$

$$q(n) \geq n^{1/(\binom{m(n)}{2} - \text{inv}(\tau_{m(n)}))} \quad \text{for almost all } n \geq 1,$$

$$|\text{inv}(\tau_{m(n)})| \leq -\log(n)/\log(q(n)) \quad \text{for almost all } n \geq 1 \text{ and } q < 1 \quad \text{or}$$

$$|\text{inv}(\tau_{m(n)})| \geq -\log(n)/\log(q(n)) + m(n)^2/2 \quad \text{for almost all } n \geq 1 \text{ and } q > 1.$$

In Section 5.3.2, we demonstrate Theorem 3.8 for all patterns of length 3. In Section 6.2, we compute the bounds in Theorem 3.8 for the specific patterns 2341 and 23451 and we plot the estimated pattern avoidance probabilities in the appropriate asymptotic regime for  $q$  from Theorem 3.9.



### 3.5 Classical pattern avoidance for Mallows permutations

We conclude this section by commenting that our approach is silent about classical pattern avoidance in Mallows permutations. The reason for this limitation is readily seen by noting the interplay between the Arratia–Goldstein–Gordon Theorem (Theorem 4.2) and the homogeneity properties of the Mallows distribution, or lack thereof, highlighted in Section 5.2 below. Briefly, for any pattern  $\tau \in \mathcal{S}_m$  and any  $1 \leq i_1 < \dots < i_m \leq n$ , the probability that  $\tau$  occurs in positions  $i_1, \dots, i_m$  of a uniform random permutation of  $[n]$  is  $1/m!$ , regardless of the choice of indices  $i_1, \dots, i_m$ . This property is crucial to computing the constants  $b_1$  and  $b_2$  in our application of Theorem 4.2, and allows us to obtain quantitative bounds when permutations are assumed to be drawn uniformly at random.

The same property does not hold for Mallows permutations. Instead, Mallows permutations only satisfy the weaker property of *consecutive homogeneity*, by which the probability that a pattern  $\tau$  occurs in consecutive locations  $j, j + 1, \dots, j + m - 1$  is the same regardless of  $j$ ; see Section 5.2. We know of no general formula for computing the probability that a given pattern appears in non-consecutive locations of a Mallows permutation. Without such a formula, we have no systematic way to extend our results in this direction, and so we leave this as an open problem.

## 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 [11]. 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 [11]). *Suppose  $X_1, X_2, \dots, X_n$  are indicator random variables with expectations  $p_1, p_2, \dots, p_n$ , 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

$$d_{\text{TV}}(\mathcal{L}(W), \mathcal{L}(Y)) \leq \frac{1 - e^{-\lambda}}{\lambda} \sum_{i=1}^n p_i \mathbb{E}|W - V_i|. \tag{18}$$

### 4.2 Fixed points example

To see how Theorem 18 can be applied, let  $e(n)$  denote the number of fixed-point free permutations of  $[n]$ . With  $\Sigma_n$  a uniform permutation of  $[n]$ , we define indicator random variables

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

(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 (3).

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

We take  $1 + V_i$  to be the random sum  $W$  conditioned on  $X_i = 1$ . For a random permutation  $\sigma$ , 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  $X'_1, X'_2, \dots, X'_n$ ; that is,  $\mathcal{L}(X'_j) = \mathcal{L}(X_j | X_i = 1)$ , 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 (18) 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.

### 4.3 The Arratia–Goldstein–Gordon Theorem

Arratia, Goldstein, & Gordon [2] provide another approach to Poisson approximation 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  $X_\alpha$  be an indicator random variable. Let  $\mathbf{X} = (X_\alpha)_{\alpha \in I}$  denote a collection of (possibly dependent) 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  $D_\alpha \subset I$  and the quantities*

$$\begin{aligned} b_1 &:= \sum_{\alpha \in I} \sum_{\beta \in D_\alpha} p_\alpha p_\beta, \\ b_2 &:= \sum_{\alpha \in I} \sum_{\alpha \neq \beta \in D_\alpha} p_{\alpha\beta}, \quad \text{and} \\ b_3 &:= \sum_{\alpha \in I} \mathbb{E} |\mathbb{E} \{X_\alpha - p_\alpha \mid \sigma(X_\beta : \beta \notin D_\alpha)\}|, \end{aligned} \tag{19}$$

where  $\sigma(X_\beta : \beta \notin D_\alpha)$  denotes the smallest  $\sigma$ -algebra containing  $X_\beta$ . 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 an independent 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  $D_\alpha$ ,  $\alpha \in I$ , 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 establishing the analogous properties for consecutive patterns under random Mallows permutations is less obvious.

## 5 Consecutive pattern avoidance of Mallows permutations

For any permutation  $\sigma = \sigma_1 \cdots \sigma_n$ , we define its *reversal* by  $\sigma^r = \sigma_n \cdots \sigma_1$ . By the definition of the Mallows distribution in (2), 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 (2) 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, 17]. 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

$$P^{n,q}(X = k) = q^{k-1} / (1 + \cdots + q^{n-1}), \quad k = 1, \dots, n. \quad (20)$$

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  $\Sigma_n$  is a  $\text{Mallows}(q)$  permutation because the probability that element  $n+1$  appears in position  $k$  of  $\Sigma_{n+1}$  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

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

as in (2).

**Definition 5.1** (Mallows process). *A collection  $(\Sigma_n)_{n \geq 1}$  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 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  $\Sigma_{n+1}$  by appending  $k$  to the end of  $\Sigma_n$  and ‘‘bumping’’ all elements of  $\Sigma_n$  that are greater or equal to  $k$ . More formally, the permutation  $\Sigma_n$  and the variable  $X_{n+1}$  give rise to the updated permutation  $\Sigma'_1 \cdots \Sigma'_n X_{n+1}$ , where

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

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

## 5.2 Properties of Mallows permutations

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

$$P(\Sigma_{n| [m]} = \sigma) = P(\Sigma_m = \sigma), \quad \sigma \in \mathcal{S}_m. \quad (21)$$

It is immediate from the ordering construction that the Mallows process  $(\Sigma_n)_{n \geq 1}$  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  $1 \leq i_1 < \cdots < i_m \leq n$

$$P(\text{red}(\Sigma_n(i_1) \cdots \Sigma_n(i_m)) = \sigma) = P(\Sigma_m = \sigma), \quad \sigma \in \mathcal{S}_m. \quad (22)$$

We call  $(\Sigma_n)_{n \geq 1}$  *consecutively homogeneous* if (22) holds only for consecutive subsequences  $i_1, i_1 + 1, \dots, i_1 + m - 1$ .

**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  $\{\text{red}(\Sigma_n(j) \cdots \Sigma_n(j + m - 1)) = \sigma\}$  for some  $\sigma \in \mathcal{S}_m$ . 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  $\Sigma_{m+j-1}$  from  $\Sigma_m$  using the bumping construction for Mallows( $1/q$ ) distribution. Thus, we have  $\Sigma_{m+j-1} \sim \text{Mallows}(1/q)$  and its reversal  $\Sigma_{m+j-1}^r \sim \text{Mallows}(q)$  with  $\text{red}(\Sigma_{m+j-1}^r(j) \cdots \Sigma_{m+j-1}^r(m + j - 1)) = \Sigma_m(m) \cdots \Sigma_m(1) \sim \text{Mallows}(q)$ . Finally, we obtain  $\Sigma_n$  by augmenting  $\Sigma_{m+j-1}^r$  according to the bumping construction, so that

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

This completes the proof.  $\square$

We say that  $\Sigma_n$  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  $\Sigma_n$  *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  $A = \{i, i+1, \dots, i+m-1\}$  and  $B = \{i', i'+1, \dots, i'+m'-1\}$ . For any  $n \geq 1$ , we can construct a Mallows( $q$ ) permutation of  $[n]$  by first generating  $\Sigma_{i+m-1}$ , for which we know that  $\text{red}(\Sigma_{i+m-1}(i) \cdots \Sigma_{i+m-1}(i+m-1)) \sim \text{Mallows}(q)$  by Lemma 5.2. We then construct  $\Sigma_n$  from  $\Sigma_{i+m-1}$  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

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

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

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

Theorems 3.8 and 3.9 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, with the key distinction that we only consider *consecutive* patterns for the general Mallows distribution; see Section 3.5. Unlike the uniform setting, the bounds for the Mallows distribution depend non-trivially on the parameter  $q$  and the structure of  $\tau$ . 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 argument to monotonic patterns.

**Lemma 5.4.** Fix  $q > 0$  and let  $\Sigma_n \sim \text{Mallows}(q)$ . For each  $m \geq 1$ , let  $\tau_m$  denote the pattern  $12 \cdots m$ . For each  $\alpha \in \bar{J}_m$ , define

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

(Recall that  $\Sigma_{m|\alpha}$  denotes the restriction of  $\Sigma_n$  to the subset of indices  $\alpha$ .) For a random permutation generated using the Mallows measure, we have

$$\mathbb{E} X_\alpha = \frac{1}{I_m(q)}, \quad \alpha \in \bar{J}_m,$$

and for  $\alpha, \beta \in \bar{J}_m$ ,  $\alpha \neq \beta$ , we have

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

**Proof:** The expression for  $\mathbb{E}X_\alpha X_\beta$  when  $\alpha$  and  $\beta$  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  $\alpha$  and  $\beta$  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  $2m - s$ .  $\square$

**Proposition 5.5.** Fix  $q > 0$  and let  $\Sigma_n \sim \text{Mallows}(q)$ . For any  $m \geq 2$ , let  $\tau_m$  be the increasing pattern  $12 \cdots m$ . For any  $\alpha \in \bar{J}_m$ , let

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

Let  $W = \sum_{\alpha \in \bar{J}_m} X_\alpha$  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

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

and  $n_1$  and  $n_2$  are given by

$$\begin{aligned} n_1 &= 2mn - 3m^2 + m \quad \text{and} \\ n_2 &= 3m - 3m^2 - 2n + 2mn. \end{aligned} \tag{23}$$

**Proof:** When  $2m - 1 \leq n$ , 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  $\alpha, \beta$ . When  $2m - 1 > n$ , the stated expressions for  $n_1$  and  $n_2$  are still valid upper bounds, but they can be improved.  $\square$

It is straightforward to state the complementary result about the decreasing pattern  $m \cdots 21$ .

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

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

and  $W = \sum_{\alpha \in J_m} X_\alpha$ , and allow  $Y$  to denote 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

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

and  $n_1$  and  $n_2$  are given by

$$\begin{aligned} n_1 &= 2mn - 3m^2 + m \quad \text{and} \\ n_2 &= 3m - 3m^2 - 2n + 2mn. \end{aligned}$$

### 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  $\bar{J}_m$  denotes the set of all  $m$ -tuples with consecutive elements in  $\{1, 2, \dots, n\}$ , and for a given  $\alpha \in \bar{J}_3$ ,  $X_\alpha$  denotes the indicator random variable defined in (7).

For  $\tau = 132$ , we have  $\mathbb{E} X_\alpha = q/I_3(q)$ , and there can be no consecutive occurrences of  $\tau$  that overlap with two indices. The only possible ways to 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  $W = \sum_{\alpha \in \bar{J}_3} \mathbb{E} X_\alpha$  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

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

The  $2(n-5)$  term comes from the two 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 two 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  $2 \cdot 2 + 3(n-7) + 2 \cdot 2$ . For fixed  $t > 0$ , we have  $\lambda \rightarrow t$  and  $d_{TV}(\mathcal{L}(W), \mathcal{L}(Y)) = O(n^{-1})$ , provided  $q \sim tn^{-1}$  or  $q \sim tn^{1/2}$ .

For  $\tau = 213$ , we similarly have

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

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

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

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

For  $\tau = 231$ :

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

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

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

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

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

$n$	$j$	lower	$n!$
100	6	$3.98735 \times 10^{157}$	$9.33262 \times 10^{157}$
1000	7	$5.77948 \times 10^{2566}$	$4.02387 \times 10^{2567}$
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	$7.34802 \times 10^{5565708}$	$8.26393 \times 10^{5565708}$

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

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

And finally for  $\tau = 312$ :

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

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

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

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

## 6 Numerical examples

### 6.1 Numerical values

Using Theorem 3.6, we can estimate  $|\mathcal{S}_n(\tau)|$  for different sizes of patterns  $\tau$ . Table 1 shows the lower bound thresholds for several values of  $n$  and patterns of size  $j$ . Similarly, using Theorem 3.7, we estimate  $|\overline{\mathcal{S}}_n(\tau)|$  in Table 2. In the case of  $n = 1000$  and  $j = 133$ , we have more specifically  $3.4433 \times 10^{2567} \leq |\mathcal{S}_n(\tau)| \leq 4.0239 \times 10^{2567}$ .

### 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), and so 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  $\tau$ , homogeneity of the Mallows measure implies  $p_\alpha = q^{|\text{inv}(\tau)|} / I_m(q)$  for all  $\alpha$ , 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 prove anything more general than Arratia–Goldstein–Gordon for arbitrary patterns; instead, we compute these



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

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

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		

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

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  $\tau$  only permits overlap with the first or last position. Table 3 lists all permutations 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  $5, \dots, n-5$ , 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  $\alpha$  and  $\beta$  that contribute to  $b_1$ . Thus,

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

giving 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  $\tau$  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  $5, \dots, n-5$ , 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  $\alpha$  and  $\beta$  that contribute to  $b_1$ . Thus,

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

producing the bounds

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

## 7 Proofs

### 7.1 Bounds on $p_\alpha$ and $p_{\alpha\beta}$

We first prove several lemmas, from which the theorems follow. By the homogeneity property of uniform permutations we have

$$p_\alpha = \mathbb{E} X_\alpha = 1/m! \quad \text{for all } \alpha \in J_m.$$

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

$$\lambda = |J_m| p_\alpha = \binom{n}{m} / m!.$$

Next, we consider the joint expectation  $p_{\alpha\beta} = \mathbb{E} X_\alpha X_\beta$ . In these calculations, recall the notation for  $d_1, d_2$  and  $\bar{d}_1, \bar{d}_2$  from Theorems 3.1 and 3.5, respectively.

**Lemma 7.1.** *Fix  $\alpha, \beta \in J_m$  and let  $s = 1, \dots, m-1$  denote the number of elements that  $\alpha, \beta$  have in common. For any such pair  $\alpha, \beta$ , we have*

$$p_{\alpha\beta} \leq \frac{s!}{m!^2}. \quad (24)$$

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

Consider first  $s = m-1$ , i.e., condition on  $m-1$  of the entries being in their proper order. Assuming that it is possible to realize both events simultaneously, the remaining element has probability  $1/m$  of appearing in its correct order. For general  $s$ , conditional on  $s$  entries being in their proper order, the probability that the remaining  $m-s$  elements appear in their proper order is then  $((s+1)(s+2) \cdots m)^{-1}$ .  $\square$

### 7.2 Proof of Theorem 3.6

We have the following lemma.

**Lemma 7.2.** *For  $D_\alpha$  defined as in Section 3.1 and  $b_3$  as in Theorem 4.2, we have  $b_3 = 0$  for all patterns  $\tau$ .*

**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  $\sigma$ -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  $\alpha$ , knowing only their order does not affect  $X_\alpha$  because uniform permutations are dissociated.  $\square$

**Remark 7.3.** Note that the conditioning in the expression for  $b_3$  is not the  $\sigma$ -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  $X_\alpha$ .

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

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

(Note that the value of  $D_\alpha$  does not depend on  $\alpha$ .)

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

$$|D_\alpha| = \sum_{s=1}^m \binom{n-m}{m-s} \binom{m}{s} = \binom{n}{m} - \binom{n-m}{m}, \quad \text{for all } \alpha \in J_m.$$

□

**Lemma 7.5.** We have

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

**Proof:** Follows immediately from Lemma 7.4 and Lemma 7.1. □

The expression given for  $d_2$  in Equation (10) in the statement of Theorem 3.6 is straightforward, although it contains the overlap quantities  $L_s(\tau)$ , which can vary wildly for different patterns  $\tau$ , and for which we are unaware of any general explicit or asymptotic expression. We calculate explicit upper bounds for  $d_2$  in Lemma 7.6.

### 7.3 Proof of Theorem 3.7 and Theorem 3.8

Theorem 3.7 follows from Theorem 3.8 using  $q = 1$  and the fact that  $I_n(1) = n!$ . Theorem 3.8 is a straightforward generalization of Proposition 5.5.

### 7.4 Proof of Theorem 3.1 and Theorem 3.5

To prove Theorem 3.1 it is sufficient that the bounds for  $d_1$  and  $d_2$  in Theorem 3.6 converge to 0 as  $n \rightarrow \infty$ . For  $d_2$ , the asymptotic analysis is not so straightforward, which is why we instead use the inequality in Equation (24).

**Lemma 7.6.** We have

$$d_2 \leq \frac{\binom{n}{m}}{m!^2} \sum_{s=1}^{m-1} \binom{n-m}{m-s} \binom{m}{s} s! \leq \frac{e^{2(\frac{1}{2} - \frac{3}{8})}}{(2\pi)^2} \left( \frac{e^2(n-m)}{j^2} \right)^m \frac{1}{m} e^{2\sqrt{n-m}} \log(m).$$

Whence, for any  $\epsilon \geq 0$ , taking  $m \geq (e^{1/e} + \epsilon) \sqrt{n}$  gives  $d_2 \rightarrow 0$ .

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

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

terms in the sum. Using equation (24), we have

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

In order to handle the sum, we first recall the quantitative bounds of Robbins [33], i.e.,

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

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

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

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

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

Note next that

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

so that for  $\eta > 0$  and  $m \sim (e + \eta)\sqrt{n}$ , we have

$$\lambda \leq \left(1 + \frac{\eta}{e}\right)^{2m} \frac{e^{-(e+\eta)^2}}{2\pi m}$$

and

$$d_2 \leq \left(\left(1 + \frac{\eta}{e}\right) \left(e^{\frac{1}{e+\eta}}\right)\right)^{2m} \frac{\log m}{2\pi m} \leq \left(\frac{e^{1/e}}{1 + \frac{\eta}{e}}\right)^{2m} \frac{\log m}{2\pi m}.$$

Letting  $\eta = e(e^{1/e} - 1) + \epsilon$  for any  $\epsilon \geq 0$ , we conclude that taking  $m \geq (e e^{1/e} + \epsilon)\sqrt{n}$  implies  $d_2 \rightarrow 0$ .  $\square$

We now compute an explicit upper bound for  $\bar{d}_2$  which establishes Theorem 3.5.

**Lemma 7.7.**

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

Taking any fixed positive  $t$ ,  $m \geq \Gamma^{(-1)}(n/t) - 1$  and  $m \sim \Gamma^{(-1)}(n/t) - 1$ , we have  $\bar{d}_2 \leq \frac{t}{m} \rightarrow 0$  as  $n$  tends to infinity.

**Proof:** It is easy to see that

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

and also

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

Using equation (24), the result immediately follows.  $\square$

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.,  $k_0 = 1.46163\dots$ . Also, let  $c = e^{-1}\sqrt{2\pi} - \Gamma(k_0) = 0.036534\dots$  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 7.8** ([10]). *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 7.9.** *Suppose  $t > 0$  is some fixed constant and  $m = \lceil \Gamma^{(-1)}(n/t) - 1 \rceil$ . Then*

$$\bar{\lambda} = \frac{n-m}{m!} \rightarrow t.$$

**Remark 7.10.** *We must be slightly careful when specifying the length of the pattern  $m$  in Lemma 7.9, 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

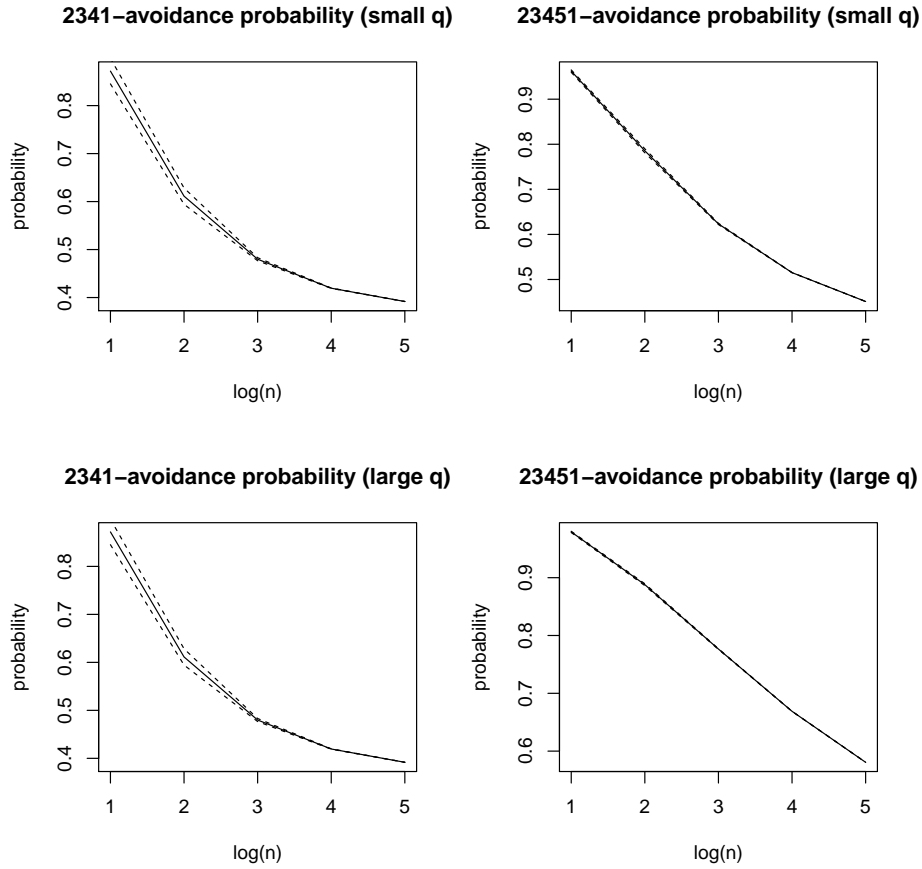
The authors would like to acknowledge helpful comments from Brendan McKay, Dan Romik, Igor Rivin, Jim Pitman, Richard Arratia, Igor Pak, Michael Albert, Vince Vatter, and Doron Zeilberger.

## References

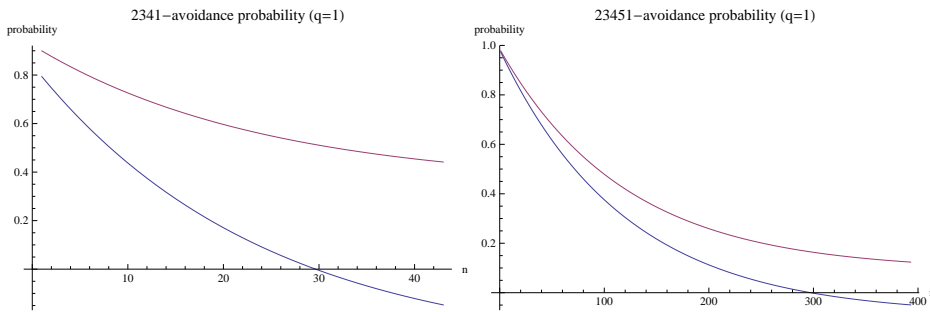
- [1] ALBERT, M. H., ATKINSON, M. D., AND BRIGNALL, R. The enumeration of permutations avoiding 2143 and 4231. *Pure Math. Appl. (P.U.M.A.)* 22, 2 (2011), 87–98.
- [2] ARRATIA, R., GOLDSTEIN, L., AND GORDON, L. Two moments suffice for poisson approximations: the chen-stein method. *The Annals of Probability* (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] CANTRELL, D. <http://mathforum.org/kb/message.jspa?messageID=342551&tstart=0>.
- [11] CHEN, L. H. Poisson approximation for dependent trials. *The Annals of Probability* (1975), 534–545.

- [12] CLAEISSON, 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.
- [13] CRANE, H., DESALVO, S., AND ELIZALDE, S. The probability of avoiding consecutive patterns in the mallows distribution. *Random Structures & Algorithms, arXiv preprint arXiv:1609.01370* (2018).
- [14] DE MONTMORT, P. R. Essay d’analyse sur les jeux de hazard. *Revue & augmentée de plusieurs Lettres* (1708).
- [15] 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).
- [16] ELIZALDE, S., AND NOY, M. Consecutive patterns in permutations. *Advances in Applied Mathematics* 30 (2003), 110–125.
- [17] GNEDIN, A., AND OLSHANSKI, G. A  $q$ -analogue of de Finetti’s theorem. *Electronic Communications in Probability* 16, R78 (2009).
- [18] GNEDIN, A., AND OLSHANSKI, G.  $q$ -exchangeability via quasi-invariance. *Annals of Probability* 38, 6 (2010), 2103–2135.
- [19] HOFFMAN, C., RIZZOLO, D., AND SLIVKEN, E. Pattern-avoiding permutations and Brownian excursion, Part I: Shapes and fluctuations. *arXiv:1506.04174* (2016).
- [20] HOFFMAN, C., RIZZOLO, D., AND SLIVKEN, E. Pattern-avoiding permutations and Brownian excursion, Part II: Fixed points. *arXiv:1406.5156* (2016).
- [21] JOHANSSON, F., AND NAKAMURA, B. Using functional equations to enumerate 1324-avoiding permutations. *Advances in Applied Mathematics* 56 (2014), 20–34.
- [22] 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.
- [23] KITAEV, S., AND MANSOUR, T. Simultaneous avoidance of generalized patterns. *Ars Combin.* 75 (2005), 267–288.
- [24] KNUTH, D. E. *The Art of Computer Programming*, vol. Volume 1: Fundamental Algorithms. Addison Wesley, Massachusetts, 1973.
- [25] LOGAN, B. F., AND SHEPP, L. A. A variational problem for random Young tableaux. *Advances in Math.* 26, 2 (1977), 206–222.
- [26] MALLOW, C. Non-null ranking models. *Biometrika* 44 (1957), 114–130.
- [27] MARCUS, A., AND TARDOS, G. Excluded permutation matrices and the Stanley–Wilf conjecture. *Journal of Combinatorial Theory, Series A* 107, 1 (2004), 153–160.
- [28] MCKAY, B. <http://mathoverflow.net/questions/216268/asymptotic-expression-for-j-which-satisfies-binomnj-j-sim-k-as-n-to>.
- [29] NAKAMURA, B. Approaches for enumerating permutations with a prescribed number of occurrences of patterns.
- [30] PERARNAU, G. A probabilistic approach to consecutive pattern avoiding permutations. *J. Comb. Th., Series A* 120 (2013), 998–1011.

- [31] 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.
- [32] RAY, N., AND WEST, J. Posets of matrices and permutations with forbidden subsequences. *Annals of Combinatorics* 7, 1 (2003), 55–88.
- [33] ROBBINS, H. A remark on Stirling’s formula. *Amer. Math. Monthly* 62 (1955), 26–29.
- [34] SHEPP, L. A., AND LLOYD, S. P. Ordered cycle lengths in a random permutation. *Trans. Amer. Math. Soc.* 121 (1966), 340–357.
- [35] SLOANE, N. Online Encyclopedia of Integer Sequences. *Published electronically at <http://www.oeis.org/>.*
- [36] VATTER, V. Problems and conjectures presented at the problem session. *Permutation Patterns, LMS Lecture Notes Series* 376, 339–345.



**Fig. 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. Note that the dashed lines overlap with the solid line in the two figures on the right, indicating that the error bounds are tightly concentrated about the estimated probability.



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