European Journal of Combinatorics 31 (2010) 241–253

Contents lists available at ScienceDirect

European Journal of Combinatorics

journal homepage: www.elsevier.com/locate/ejc

# Packing sets of patterns

Alexander Burstein [a], Peter Hästö [b]

[a] *Department of Mathematics, Howard University, Washington, DC 20059, USA*

[b] *Department of Mathematical Sciences, P.O. Box 3000, 90014 University of Oulu, Finland*

ARTICLE INFO

*Article history:*
Received 6 June 2008
Received in revised form
16 December 2008
Accepted 17 February 2009
Available online 29 March 2009

ABSTRACT

Packing density is a permutation occurrence statistic which describes the maximal number of permutations of a given type that can occur in another permutation. In this article we focus on containment of sets of permutations. Although this question has been tangentially considered previously, this is the first article focusing exclusively on it. We find the packing density for various special sets of permutations and study permutation and pattern co-occurrence.



## 1. Introduction

The string 413223 contains two subsequences, 133 and 122, each of which is *order-isomorphic* (or simply *isomorphic*) to the string 122, i.e. ordered in the same way as 122. In this situation we call the string 122 a *pattern*. Herb Wilf first proposed the systematic study of pattern containment in his 1992 address to the SIAM meeting on Discrete Mathematics. However, several earlier results on pattern containment exist, for example, those by Knuth [8] and Tarjan [11].

Most results on pattern containment actually deal with *pattern avoidance*, in other words, enumerate or consider properties of strings over a totally ordered alphabet which avoid a given pattern or set of patterns. There is considerably less research on other aspects of pattern containment, specifically, on packing patterns into strings over a totally ordered alphabet, but see [1,6,7,9,10] for the permutation case and [3–5,12–14] for the more general pattern case.

Although several of the above cited papers have defined packing density for sets of patterns, virtually all of them have subsequently restricted the attention to the case when the set contains only one pattern. In this paper we take the next step in studying the set packing question. In Section 2 we study the packing density of so-called layered permutations which have been the focus of much research also in the single permutation case. In Section 3 we calculate the packing density of various

*E-mail addresses:* aburstein@howard.edu (A. Burstein), peter.hasto@helsinki.fi (P. Hästö).

0195-6698/$ – see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ejc.2009.02.004

sets of three-letter patterns. In Section 4 we study the average co-occurrence of patterns and calculate the leading term of the covariance of any two permutation patterns of the same length $n$ (as a polynomial in $n$) and, in general, of any two patterns of the same length over the same alphabet.

### 1.1. Notation

Let $[k] = \{1, 2, \ldots, k\}$ be our canonical totally ordered alphabet on $k$ letters, and consider the set $[k]^n$ of $n$-letter words over $[k]$. We say that a pattern $\pi \in [l]^m$ *occurs* in $\sigma \in [k]^n$, or $\pi$ *hits* $\sigma$, or that $\sigma$ *contains* the pattern $\pi$, if there is a subsequence of $\sigma$ order-isomorphic to $\pi$.

Given a word $\sigma \in [k]^n$ and a set of patterns $\Pi \subseteq [l]^m$, let $\nu(\Pi, \sigma)$ be the total number of occurrences of patterns in $\Pi$ ($\Pi$-patterns, for short) in $\sigma$. Obviously, the largest possible number of $\Pi$-occurrences in $\sigma$ is $\binom{n}{m}$, when each subsequence of length $m$ of $\sigma$ is an occurrence of a $\Pi$-pattern. Define

$$\mu(\Pi, k, n) = \max\{\, \nu(\Pi, \sigma) \mid \sigma \in [k]^n\},$$
$$d(\Pi, \sigma) = \frac{\nu(\Pi, \sigma)}{\binom{n}{m}} \quad \text{and}$$
$$\delta(\Pi, k, n) = \frac{\mu(\Pi, k, n)}{\binom{n}{m}} = \max\{\, d(\Pi, \sigma) \mid \sigma \in [k]^n\},$$

the maximum number of $\Pi$-patterns in a word in $[k]^n$, the probability that a subsequence of $\sigma$ of length $m$ is an occurrence of a $\Pi$-pattern, and the maximum such probability over words in $[k]^n$, respectively. A permutation $\sigma \in [k]^n$ such that $\nu(\Pi, \sigma) = \mu(\Pi, k, n)$ is said to be *$\Pi$-maximal.*

We want to consider the asymptotic behavior of $\delta(\Pi, k, n)$ as $n \to \infty$ and $k \to \infty$. Barton [3] proved that

$$\lim_{n\to\infty} \delta(\Pi, n, n) = \lim_{n\to\infty} \lim_{k\to\infty} \delta(\Pi, n, k) = \lim_{k\to\infty} \lim_{n\to\infty} \delta(\Pi, n, k),$$

so we can amend the definition from [5] and define the common limit $\delta(\Pi)$ to be the packing density of the set of patterns $\Pi$. If $\Pi = \{\pi\}$, then we use also $\delta(\pi)$ for $\delta(\Pi)$.

## 2. Sets of layered permutations

In this section we deal with sets of layered permutations. Recall that a permutation is said to be *layered* if it is a strictly increasing sequence of strictly decreasing substrings. These substrings are called the layers of the permutation. For instance, $\widehat{21}\widehat{543}$ is layered with layers 21 of length 2 and 543 of length 3. Obviously, a layered permutation is determined by the sequence of lengths of its layers, so we can denote $\widehat{21}\widehat{543}$ simply by [2, 3].

It has been shown that if $\Pi$ consists of layered permutations, then there is a $\Pi$-maximal permutation which is layered [1, Theorem 2.2]. For the case of a single layered permutation $\pi = [m_1, \ldots, m_r]$ (i.e. $\pi$ has $r$ layers of lengths $m_1, \ldots, m_r$) Price [9] showed that

$$\delta(\pi) = \lim_{s\to\infty} \max_{\substack{\lambda_1,\ldots,\lambda_s \geq 0 \\ \lambda_1+\cdots+\lambda_s=1}} p_s(\lambda_1, \ldots, \lambda_s), \tag{2.1}$$

where

$$p_s(\lambda_1, \ldots, \lambda_s) = \binom{m}{m_1, \ldots, m_r} \sum_{1\leq i_1<\cdots<i_r\leq s} \lambda_{i_1}^{m_1} \cdots \lambda_{i_r}^{m_r}. \tag{2.2}$$

The sequences $\lambda = (\lambda_1, \ldots, \lambda_s)$ in (2.1) are called *partitions of unity*, and we write $\lambda \vdash 1$. For each $s$, the $\pi$-maximal permutation is approximately $[\lfloor n\lambda_1^* \rfloor, \ldots, \lfloor n\lambda_s^* \rfloor]$ for a maximizing partition of unity $\lambda^* \vdash 1$. If for every $n$, there is a permutation $\sigma_n \in S_n$ with $r$ layers (recall that $r$ is the number of layers of $\pi$) such that

$$\delta(\pi) = \lim_{n\to\infty} d(\pi, \sigma_n),$$

then the permutation $\pi$ is said to be *simple* [6]. For example, $[2, 2, 2, 2, 2]$ is simple, but $[2, 2, 2, 2, 2, 2]$ is not simple (by Examples 6.4 and 6.5 of [9]). A partition $\lambda = (\lambda_1, \ldots, \lambda_s) \vdash 1$ is *optimal* if $p_s(\lambda_1, \ldots, \lambda_s) = \delta(\pi)$ and $s$ is the least integer such that there exists $\lambda \vdash 1$ with this property.

The next result shows that [6, Theorem 3.3] partially generalizes to the case of sets of permutations.

**Proposition 2.3.** *Let $\Pi$ be a set of layered permutations of length $m$ and $r$ layers such that the optimal partition of unity $\lambda = (\lambda_1, \ldots, \lambda_s) \vdash 1$ is increasing (i.e. $\lambda_1 \leq \cdots \leq \lambda_s$). Let $m_-$ be the length of the shortest layer of all permutations in $\Pi$. If $m_- \geq \max\{\log_2(r+1), 2\}$, then $\Pi$ is simple, and the packing density of $\Pi$ equals*

$$\delta(\Pi) = \sup_{(\lambda_1,\ldots,\lambda_r)\vdash 1} \sum_{\pi\in\Pi} \binom{m}{m_1^\pi, \ldots, m_r^\pi} \lambda_1^{m_1^\pi} \cdots \lambda_r^{m_r^\pi},$$

*where $m_i^\pi$ is the length of the* ith *layer of $\pi$ and the supremum is taken over all partitions of unity with $r$ parts.*

**Proof.** As in the proof of [6, Theorem 3.3], we conclude that the maximizing partition of unity of any $\pi \in \Pi$ has $r$ layers. Hence $\Pi$ is simple. The last formula follows directly from this. □

One might think that it is always the case that a set of simple permutations is simple. However, we have not been able to prove it.

In some cases it is easy to show that the condition of the previous proposition holds. A layered permutation is said to be increasing, if its layer sizes are increasing. If $\Pi$ is a set of increasing layered permutations, then the maximizing sequence $(\lambda_i)$ is also increasing. The proof of this fact is the same as in the case of only a single permutation, see [6, Lemma 3.2]. Another obvious case is when the set $\Pi$ is symmetric, in the sense that it contains all the permutations with certain layer sizes, like the set $\{[2, 1, 1], [1, 2, 1], [1, 1, 2]\}$.

Let us next consider some special sets of layered permutations. The prototypical case for the next theorem is permutations $[m, 2]$ and $[m, 1, 1]$. In this case the permutations differ only in that the last two letters are interchanged, but as can be seen below, this is not the reason that we are able to calculate the packing density.

**Theorem 2.4.** *Let $m, n \geq 2$ and let $\Pi(m, n)$ be the set of all permutations whose first layer has length $m$ and whose subsequent layers have total length $n$. Then we have*

$$\delta(\Pi(m,n)) = \binom{m+n-1}{n} \frac{(m-1)^{m-1} n^n}{(m+n-1)^{m+n-1}}.$$

*Note that $\delta(\Pi(m, n)) = \delta([m-1, n])$ if $m \geq 3$.*

**Proof.** Let $P_K$ be the set of sequences $(\lambda_i)_{i=1}^\infty$ of non-negative real numbers with $\sum \lambda_i = 1$ and $\lambda_i = 0$ for $i > K$. Using (2.1) and (2.2) we obtain

$$\delta(\Pi(m,n)) = \lim_{K\to\infty} \sup_{(\lambda_i)\in P_K} F\big((\lambda_i)\big),$$

where

$$\begin{aligned} F\big((\lambda_i)\big) &= \binom{m+n}{n} \sum_{i<j} \lambda_i^m \lambda_j^n + \sum_{p=1}^{n-1} \binom{m+n}{m, n-p, p} \sum_{i<j<k} \lambda_i^m \lambda_j^{n-p} \lambda_k^p + \cdots \\ &= \binom{m+n}{n} \left( \sum_{i<j} \lambda_i^m \lambda_j^n + \sum_{p=1}^{n-1} \binom{n}{p} \sum_{i<j<k} \lambda_i^m \lambda_j^{n-p} \lambda_k^p + \cdots \right) \\ &= \binom{m+n}{n} \sum_i \lambda_i^m \left( \sum_{j>i} \lambda_j \right)^n. \end{aligned}$$

Fix $K \geq 2$ and choose $(\lambda_i) \in P_K$ such that $\sup_{(\bar{\lambda}_i)\in P_K} F\big((\bar{\lambda}_i)\big) = F\big((\lambda_i)\big)$. (This is possible by continuity of $F$ and compactness.) By discarding leading zeros if necessary we may assume that $\lambda_1 > 0$; also $F\big((\lambda_i)\big) = 0$ if $\lambda_1 = 1$, so the maximizing property implies that this is not the case; hence $\lambda_1 \in (0, 1)$. Denote $\Lambda_i = \sum_{i<j} \lambda_j$. We define $\lambda_1' = c\lambda_1$ and $\lambda_i' = d\lambda_i$ for $i > 1$. Moreover, we choose $d(c) = (1 - c\lambda_1)/(1 - \lambda_1)$, in order to have $\sum \lambda_i' = 1$ so that $(\lambda_i') \in P_k$. Since the original sequence $(\lambda_i)_{i=1}^{\infty}$ was maximal in $P_K$, we have

$$\sum_i \lambda_i^m \Lambda_i^n \geq \sum_i \lambda_i'^m \left(\sum_{j>i} \lambda_j'\right)^n = c^m d^n \lambda_1^m \Lambda_1^n + d^{m+n} \sum_{i>1} \lambda_i^m \Lambda_i^n.$$

Let us denote

$$\alpha = \lambda_1^m \Lambda_1^n = \lambda_1^m (1 - \lambda_1)^n$$

and $\beta = \sum_{i>1} \lambda_i^m \Lambda_i^n$. Then our previous conclusion implies that the function

$$G(c) = c^m d(c)^n \alpha + d(c)^{m+n} \beta$$

has a maximum at $c = 1$. Differentiating this function and evaluating at $c = d = 1$ give

$$G'(1) = \Big(m - \frac{n\lambda_1}{1-\lambda_1}\Big)\alpha - (m+n)\frac{\lambda_1}{1-\lambda_1}\beta.$$

Since 1 is a maximum, the derivative equals zero, so

$$\beta = \frac{m(1-\lambda_1) - n\lambda_1}{(m+n)\lambda_1}\,\alpha.$$

Therefore

$$\sum_i \lambda_i^m \Lambda_i^n = \alpha + \beta = \frac{m}{(m+n)\lambda_1}\,\alpha = \frac{m}{m+n}\lambda_1^{m-1}(1-\lambda_1)^n.$$

Clearly, the last expression is maximized by $\lambda_1 = (m-1)/(m+n-1)$. Therefore we have

$$\begin{aligned}\sup_{(\lambda_k)\in P_K} F\big((\lambda_i)\big) &= \binom{m+n}{n}\frac{m}{m+n}\sup_{0<\lambda_1<1}\lambda_1^{m-1}(1-\lambda_1)^n \\ &= \binom{m+n}{n}\frac{m}{m+n}\frac{(m-1)^{m-1}n^n}{(m+n-1)^{m+n-1}}.\end{aligned}$$

Since this expression does not depend on $K$, we see that it is also the limit as $K \to \infty$ of $\sup_{(\lambda_k)\in P_K} F\big((\lambda_i)\big)$, from which we obtain $\delta(\Pi(m, n))$. Taking also into account [6, Theorem 1.2] we see that it equals $\delta([m-1, n])$ when $m \geq 3$. □

**Remark 2.5.** We can make the previous theorem slightly more general by allowing different first terms. Let $M \geq 3$ and let $M > m_1 > \cdots > m_r \geq 2$. Then we can find the packing density of the set

$$\Pi(m_1, M - m_1) \cup \cdots \cup \Pi(m_r, M - m_r)$$

by finding the maximum over $\lambda_1$ of the real valued function

$$\frac{1}{M}\sum_{i=1}^{r} m_i \lambda_1^{m_i - 1}(1-\lambda_1)^{M-m_i}.$$

The proof of this fact is very similar to the proof of Theorem 2.4, and is thus omitted.

Using the method of the previous proof, we get an upper bound for the packing density of much more general types of permutations. In the general case however, the upper bound is not attained.

**Corollary 2.6.** *Let $m_1 \geq 3$ and $m = m_1 + \cdots + m_r \geq 5$. Then*

$$\delta([m_1, \ldots, m_r]) \leq \delta([m_1 - 1, m - m_1])\delta([m_2, \ldots, m_r]).$$

**Proof.** With the notation of the previous proof:

$$\delta([m_1, \ldots, m_r]) = \lim_{K\to\infty} \sup_{(\lambda_k)\in P_K} \binom{m}{m_1, \ldots, m_r} \sum_{i_1<\cdots<i_r} \lambda_{i_1}^{m_1} \cdots \lambda_{i_r}^{m_r}.$$

We split the sum into two parts,

$$\alpha = \lambda_1^{m_1} \sum_{1<i_2<\cdots<i_r} \lambda_{i_2}^{m_2} \cdots \lambda_{i_r}^{m_r},$$

and the rest, denoted by $\beta$. As in the previous proof, we set $\lambda_1' = c\lambda_1$ and $\lambda_i' = \mathrm{d}\lambda_i$ for $i > 1$, construct the function $F(c)$, calculate the derivative, and set it equal to zero. As before, we calculate

$$\alpha + \beta = \frac{m_1}{m\lambda_1}\alpha.$$

Using a rescaling and the definition of packing density we find

$$\frac{\alpha}{\lambda_1^{m_1}} = \sum_{1<i_2<\cdots<i_r} \lambda_{i_2}^{m_2} \cdots \lambda_{i_r}^{m_r} \leq \binom{m-m_1}{m_2, \ldots, m_r}^{-1} (1-\lambda_1)^{m-m_1}\delta([m_2, \ldots, m_r]).$$

We note that $\binom{m}{m_1,\ldots,m_r}\binom{m-m_1}{m_2,\ldots,m_r}^{-1} = \binom{m}{m_1}$. Thus we find that

$$\begin{aligned}\delta([m_1, \ldots, m_r]) &= \binom{m}{m_1, \ldots, m_r} \frac{m_1}{m\lambda_1}\alpha \\ &\leq \binom{m}{m_1} \frac{m_1}{m} \delta([m_2, \ldots, m_r]) \sup_{0<\lambda_1<1} \lambda_1^{m_1-1}(1-\lambda_1)^{m-m_1}.\end{aligned}$$

Clearly the last supremum is reached for $\lambda_1 = (m_1 - 1)/(m - 1)$, from which the claim follows by noting that

$$\delta([m_1 - 1, m - m_1]) = \binom{m}{m_1} \frac{m_1}{m} \frac{(m_1 - 1)^{m_1-1}(m - m_1)^{m-m_1}}{(m - 1)^{m-1}}$$

by [6, Theorem 1.2]. □

## 3. Three-letter patterns

In this section we calculate the packing density for sets of patterns of length three, except the set $\{121, 212\}$, which we did not manage to deal with. For permutations this was done in [1]. In the interest of clarity we first prove some lemmas.

**Lemma 3.1.** *We have $\delta([1, 2], [2, 1]) = 3/4$.*

**Proof.** As in the proof of Theorem 2.4 we have

$$\delta([1, 2], [2, 1]) = \binom{3}{2} \lim_{k\to\infty} \sup_{(\lambda_i)\in P_k} F\big((\lambda_i)\big),$$

where now $P_k$ is the set of sequences $(\lambda_i)_{i=-\infty}^{\infty}$ of positive real numbers with $\sum_i \lambda_i = 1$ and $\lambda_i = 0$ for $|i| > k$; and

$$F\big((\lambda_i)\big) = \sum_{i_1<i_2} \lambda_{i_1}\lambda_{i_2}^2 + \lambda_{i_2}\lambda_{i_1}^2 = \sum_{j=-\infty}^{\infty} \sum_{i\neq j} \lambda_i\lambda_j^2 = \sum_{j=-\infty}^{\infty} \lambda_j^2(1-\lambda_j).$$

Here we used that $\sum_i \lambda_i = 1$ in the last step. We conclude by a simple algebraic identity and the fact that a square is a non-negative number:

$$F\big((\lambda_i)\big) = \sum_{j=-\infty}^{\infty} \lambda_j^2(1-\lambda_j) = \sum_{j=-\infty}^{\infty} \lambda_j\Big(\frac{1}{4} - \Big(\lambda_j - \frac{1}{2}\Big)^2\Big) \le \frac{1}{4}.$$

Hence $\delta([1, 2], [2, 1]) \le 3/4$. For every $k > 0$ we have

$$d(\{[1, 2], [2, 1]\}, [k, k]) = \binom{2k}{3}^{-1} k^2(k-1) = \frac{6k^2(k+1)}{2k(2k-1)(2k-2)} \to \frac{3}{4} \quad \text{as } k \to \infty,$$

which implies that $\delta([1, 2], [2, 1]) \ge 3/4$ and completes the proof. □

**Lemma 3.2.** *We have* $\delta(112, 122) = \delta([2, 1], [1, 2]) = \frac{3}{4}$.

**Proof.** Since both 112 and 122 are non-decreasing, it is clear that the maximizing pattern must be non-decreasing. We may assume that the maximizing pattern is of the form

$$\sigma = 1^{s_1} 2^{s_2} \dots n^{s_n}.$$

Consider then the permutation of type $\sigma' = [s_1, \dots, s_n]$. It is clear that every occurrence of 112 in $\sigma$ corresponds to an occurrence of [2, 1] in $\sigma'$, similarly for 122 and [1, 2]. Thus the claim follows by Lemma 3.1. □

**Lemma 3.3.** *We have* $\delta(112) = \delta(112, 121) = \delta(112, 121, 211) = 2\sqrt{3} - 3$.

**Proof.** The numerical value $\delta(112) = 2\sqrt{3} - 3$ is from [5, Example 2.12]. We next complete the proof by showing that $\delta(112) = \delta(112, 121, 211)$. The remaining equality follows from this, since the density certainly grows if we add more patterns to a set.

Let $\sigma$ be a word and consider adjacent distinct letters at $\sigma_i$ and $\sigma_{i+1}$ and let $\sigma'$ be the pattern with these letters interchanged. Then

$$d(112, 121, 211; \sigma) = d(112, 121, 211; \sigma');$$

to see this notice that the number of occurrences of 112 which hit at most one of the two letters at positions $i$ and $i+1$ is the same in $\sigma$ and $\sigma'$. The same holds for the other two patterns. So it remains to consider occurrences involving both of these positions. Assume $\sigma_i < \sigma_{i+1}$. Then if 112 hits $\sigma$ at positions $j < i < i+1$ it is clear that 121 hits $\sigma'$ at the same positions. Similarly a hit of 121 at $i < i+1 < j$ is turned into a hit of 211 at the same positions. If $\sigma_i > \sigma_{i+1}$, then the situation is reversed. Hence in each case the total number of occurrences is preserved.

We have now shown that we may exchange adjacent letters in $\sigma$. Doing this a sufficient number of times we may assume that $\sigma$ is non-decreasing. But then all the hits are of type 112, hence

$$d(112, 121, 211; \sigma) = d(112; \sigma).$$

Since $\sigma$ was arbitrary, the claim follows. □

**Theorem 3.4.** *Let S be a set of three-letter patterns on* $[2]^3$.

1. *If S includes either of the patterns* 111 *and* 222, *then* $\delta(S) = 1$.
2. *Otherwise, if S includes either of the sets* {112, 122} *or* {211, 221}, *then* $\delta(S) = 3/4$.
3. *Otherwise, if* $S = \{112, 121, 211, 212\}$ *or* $S = \{121, 122, 221, 212\}$, *then* $\delta(S) = \frac{5}{4}(2\sqrt{3} - 3)$.
4. *Otherwise, if S includes any of the patterns* 112, 122, 211 *or* 221, *then* $\delta(S) = 2\sqrt{3} - 3$.
5. *Otherwise, if S equals* {121} *or* {212}, *then* $\delta(S) = \frac{1}{4}(2\sqrt{3} - 3)$.
6. *Otherwise, if* $S = \{121, 212\}$, *then* $\delta(S) \ge 1/4$.

**Proof.** The claim in case 1 is clear. Thus we assume that $111, 222 \notin S$.

For Case 2 we may assume, by taking the complement if necessary, that $\{112, 122\} \subset S \subset \{112, 122, 121, 211, 212, 221\}$. As in the proof of Lemma 3.3 it follows that

$$\delta(112, 122, 121, 211, 212, 221) = \delta(112, 122).$$

Then it follows from Lemma 3.2 that $\delta(S) = 3/4$.

For Case 3 we may assume that $S = \{112, 121, 211, 212\}$. We argue as in Lemma 3.3: let $\sigma$ be a word and consider adjacent distinct letters at $\sigma_i$ and $\sigma_{i+1}$ and let $\sigma'$ be the pattern with these letters interchanged. As we saw before, the set $\{112, 121, 211\}$ hits $\sigma$ and $\sigma'$ the same number of times. So to calculate the hits of $\{112, 121, 211\}$ we may assume that $\sigma$ is increasing and count the hits of 112. Let $\sigma = 1^{\mu_1}2^{\mu_2}\cdots s^{\mu_s}$ where the $\mu_i$ are non-negative integers with $\sum \mu_i = |\sigma|$. Thus the contribution of these hits can be expressed as (a factor times)

$$\mu_1^2 \sum_{i=2}^{s} \mu_i + \mu_2^2 \sum_{i=3}^{s} \mu_i + \cdots + \mu_{s-1}^2 \mu_s.$$

Then we shuffle around the letters to make optimal hits for 212, namely,

$$\left(\frac{1}{2}\mu_1, \ldots, \frac{1}{2}\mu_{s-1}, \mu_s, \frac{1}{2}\mu_{s-1}, \ldots, \frac{1}{2}\mu_1\right),$$

which corresponds to (the same factor times)

$$\left(\frac{\mu_1}{2}\right)^2 \sum_{i=2}^{s} \mu_i + \left(\frac{\mu_2}{2}\right)^2 \sum_{i=3}^{s} \mu_i + \cdots + \left(\frac{\mu_{s-1}}{2}\right)^2 \mu_s.$$

But this is just $1/4$ of the previous sum. Thus, by a "lucky" coincidence, both sums are maximized by the same relative frequencies of letters, so we find that

$$\delta(S) = \delta(\{112, 121, 211\}) + \delta(212) = \frac{5}{4}\delta(\{112, 121, 211\}).$$

Case 4 follows from Lemma 3.3, and combined with the previous equation this implies also Case 3. For Case 5, we found above that $\delta(121) = \delta(212) = \frac{1}{4}(2\sqrt{3} - 3)$.

So it remains to investigate the case $S = \{121, 212\}$. Consider the permutation $\sigma = 121212\ldots$ of length $2n$. The number of hits of 121 with the 2 matching the $k$th 2 in $\sigma$ equals $k(n-k)$. Thus the total number of hits of 121 is

$$\sum_{k=1}^{n-1} k(n-k) = \binom{n+1}{3}.$$

The number of hits of $S$ is twice this, hence $\delta(S) \geq d(S, \sigma) = 2\lim_{n\to\infty} \binom{n+1}{3} / \binom{2n}{3} = 1/4$. □

## 4. Average pattern co-occurrence

In this section we deal with average, rather than maximal, pattern co-occurrence.

Consider $S_n$ as a sample space with uniform distribution. Let $\pi \in S_m$, and let $X_\pi$ be a random variable such that $X_\pi(\tau)$ is the number of occurrences of pattern $\pi$ in a given permutation $\tau \in S_n$.

It is an easy exercise to show that, even though the maximal number of times a pattern can occur in a permutation (or a word, in general) differs with the pattern, the average number of occurrences of any pattern over all permutations of a given length is the same.

**Lemma 4.1.** *$E(X_\pi) = \frac{1}{m!}\binom{n}{m} \sim \frac{1}{(m!)^2} n^m$ for any pattern $\pi \in S_m$.*

**Proof.** Pick an $m$-letter subset $S$ of $[n] = \{1, 2, \dots, n\}$ in $\binom{n}{m}$ ways. There is a unique permutation $\pi(S)$ of $S$ order-isomorphic to $\pi$, out of $m!$ equally likely permutations in which the elements of $S$ can occur in $\tau \in S_n$. Let $Y_{\pi(S)}$ be a random variable such that $Y_{\pi(S)}(\tau)$ is the number of occurrences of $\pi(S)$ in $\tau$. Then

$$P\left(Y_{\pi(S)}(\tau) = 1\right) = \frac{1}{m!} \quad \text{and} \quad P\left(Y_{\pi(S)}(\tau) = 0\right) = 1 - \frac{1}{m!},$$

so $E(Y_{\pi(S)}) = 1/m!$. This is true for any $S \subseteq [n]$ such that $|S| = m$, and we have $X_\pi = \sum_{S\subseteq[n],\ |S|=m} Y_{\pi(S)}$, hence,

$$E(X_\pi) = \sum_{S\subseteq[n],|S|=m} E(Y_{\pi(S)}) = \frac{1}{m!}\binom{n}{m}. \quad \square$$

Therefore, the average pattern occurrence is the same for all $\pi \in S_n$. However, the average pattern co-occurrence, measured by the covariance $\mathrm{Cov}(X_{\pi_1}, X_{\pi_2})$, does depend on the pattern. We will start by considering the average pattern co-occurrence with itself, i.e. $\mathrm{Var}(X_\pi)$. That, via the standard deviation $\sigma(X_\pi)$, will also tell us how tightly the distribution of $X_\pi$ is grouped around the mean of Lemma 4.1.

Let $P_\pi$ be the permutation matrix of $\pi$, in other words, $P_\pi = [\delta(\pi(i), j)]_{m\times m}$, where $\delta$ is the Kronecker symbol. Note that $P_\pi$ is orthogonal, so $P_{\pi^{-1}} = P_\pi^{-1} = P_\pi^T$. Also, for an integer $m > 0$, and integers $1 \le i, j \le m$, define

$$[i, j]_m = \binom{i-1+j-1}{i-1}\binom{m-i+m-j}{m-i}.$$

Let $A_m$ be the $m \times m$ matrix with $(A_m)_{ij} = [i, j]_m$, which have been studied e.g. in [2].

**Theorem 4.2.** $\mathrm{Var}(X_\pi) = c(\pi)n^{2m-1} + O(n^{2m-2})$ *for any pattern* $\pi \in S_m$, $m > 1$, *where*

$$c(\pi) = \frac{1}{((2m-1)!)^2}\left(\mathrm{Tr}(A_m P_\pi A_m P_\pi^{-1}) - \binom{2m-1}{m-1}^2\right) > 0.$$

*The trace in the above formula can be expressed as*

$$\mathrm{Tr}(A_m P_\pi A_m P_\pi^{-1}) = \sum_{i,j=1}^{m} [i, j]_m [\pi(i), \pi(j)]_m.$$

*For the standard deviation this gives* $\sigma(X_\pi) = \sqrt{c(\pi)}\, n^{m-\frac{1}{2}} + O(n^{m-1})$ *for any pattern* $\pi \in S_m$.

**Proof.** Since $\mathrm{Var}(X_\pi) = E(X_\pi^2) - E(X_\pi)^2$, and the value of $E(X_\pi)$ was determined in Lemma 4.1, it remains only to consider $E(X_\pi^2)$. We have

$$E(X_\pi^2) = E\left(\left(\sum_{S\subseteq[n],\ |S|=m} Y_{\pi(S)}\right)^2\right) = \sum_{\substack{S_1,S_2\subseteq[n]\\ |S_1|=|S_2|=m}} E\left(Y_{\pi(S_1)}Y_{\pi(S_2)}\right).$$

Of course, $Y_{\pi(S_1)}Y_{\pi(S_2)} = 1$ if and only if both $\pi(S_1)$ and $\pi(S_2)$ are subsequences of $\tau$, otherwise, $Y_{\pi(S_1)}Y_{\pi(S_2)} = 0$.

Let $S = S_1 \cup S_2$, and $|S_1 \cap S_2| = \ell$, so $|S| = 2m - \ell$. We can pick a subset $S \subseteq [n]$ in $\binom{n}{2m-\ell}$ ways. Note that any such $S$ is order-isomorphic to $[2m-\ell] = \{1, 2, \dots, 2m-\ell\}$. Hence, the number of permutations $\rho(S)$ of $S$ such that $\rho\restriction_{S_1} \cong \pi$ and $\rho\restriction_{S_2} \cong \pi$ is the same for any $S$ of cardinality $2m-\ell$ and depends only on $m$ and $\ell$.

Therefore, $E(X_\pi^2)$ is a linear combination of $\left\{\binom{n}{2m-\ell} \mid 0 \le \ell \le m\right\}$ with coefficients that are rational functions in $m$ and $\ell$. The degrees in $n$ of both $E(X_\pi^2)$ and $E(X_\pi)^2$ are $2m$, and the coefficient of $n^{2m}$ in

$E(X_\pi)^2$ is $1/(m!)^4$. On the other hand, $S = S_1 \cup S_2$, $|S| = 2m$ and $|S_1| = |S_2| = m$ imply that $S_1 \cap S_2 = \emptyset$, so $Y_{\pi(S_1)}$ and $Y_{\pi(S_2)}$ are independent, and hence

$$P\left(Y_{\pi(S_1)}Y_{\pi(S_2)} = 1\right) = P\left(Y_{\pi(S_1)} = 1\right)P\left(Y_{\pi(S_2)} = 1\right) = \left(\frac{1}{m!}\right)^2.$$

Let $[x^d]P(x)$ denote the coefficient of $x^d$ in a given polynomial $P(x)$. Since there are $\binom{2m}{m}$ ways to partition a set $S$ of size $2m$ into two subsets of size $m$, the coefficient of $\binom{n}{2m}$ in $E(X_\pi^2)$ is $\binom{2m}{m}/(m!)^2$. Hence,

$$[n^{2m}]E(X_\pi^2) = \frac{1}{(2m)!}\frac{1}{(m!)^2}\binom{2m}{m} = \frac{1}{(m!)^4}.$$

Thus $[n^{2m}]E(X_\pi^2) = [n^{2m}]E(X_\pi)^2$, so $\deg_n(\mathrm{Var}(X_\tau)) \le 2m - 1$, and hence, $[n^{2m-1}]\,\mathrm{Var}(X_\tau) \ge 0$.

We have

$$[n^{2m-1}]E(X_\pi)^2 = [n^{2m-1}]\left(\frac{1}{m!}\binom{n}{m}\right)^2 = \frac{2}{(m!)^2}\cdot[n^m]\binom{n}{m}\cdot[n^{m-1}]\binom{n}{m}$$
$$= \frac{2}{(m!)^2}\cdot\frac{1}{m!}\cdot\left(-\frac{\binom{m}{2}}{m!}\right) = -\frac{m(m-1)}{(m!)^4}.$$

Similarly, the coefficient of $n^{2m-1}$ in the $\binom{n}{2m}$-term of $E(X_\pi^2)$ is

$$-\frac{\binom{2m}{2}}{(2m)!}\frac{1}{(m!)^2}\binom{2m}{m} = -\frac{m(2m-1)}{(m!)^4},$$

so we only need to find the coefficient of $n^{2m-1}$ in the $\binom{n}{2m-1}$-term of $E(X_\pi^2)$.

As we noted before, all subsets $S \subseteq [n]$ of the same size (in our case, of size $2m-1$) are equivalent, so we may assume $S = [2m-1] = \{1, 2, \ldots, 2m-1\}$. We want to find the number of permutations $\rho$ of $S$ such that there exist subsets $S_1, S_2 \subseteq S$ of size $m$ for which we have $|S_1 \cap S_2| = 1$ (so $S_1 \cup S_2 = S$) and $\rho\restriction_{S_1} \cong \pi$ and $\rho\restriction_{S_2} \cong \pi$.

Suppose that we want to choose $S_1$ and $S_2$ as above, together with their positions in $S$, in such a way that the unique intersection element $e$ is in the $i$th position in $\pi(S_1)$ and the $j$th position in $\pi(S_2)$ (of course, $1 \le i, j \le m$). Then $e$ occupies position $\pi^{-1}(e) = (i-1) + (j-1) + 1 = i + j - 1$ in $S$. Hence, there are $\binom{i-1+j-1}{i-1}$ ways to choose the positions for elements of $\pi(S_1)$ and $\pi(S_2)$ to the left of $e$, and $\binom{m-i+m-j}{m-j}$ ways to choose the positions for elements of $\pi(S_1)$ and $\pi(S_2)$ to the right of $e$. On the other hand, both $\pi(S_1)$ and $\pi(S_2)$ are naturally order-isomorphic to $\pi$, hence, under that isomorphism $e$ maps to $\pi(i)$ as an element of $S_1$ and to $\pi(j)$ as an element of $S_2$. Similarly, since $e$ is the unique intersection element, exactly $\pi(i) - 1$ elements in $S_1$ and exactly $\pi(j) - 1$ elements in $S_2$, all distinct, must be less than $e$, and the rest of the elements of $S$ must be greater than $e$, so we must have $e = (\pi(i) - 1) + (\pi(j) - 1) + 1 = \pi(i) + \pi(j) - 1$. There are $\binom{\pi(i)-1+\pi(j)-1}{\pi(i)-1}$ ways to choose the elements of $S_1$ and $S_2$ which are less than $e$, and $\binom{m-\pi(i)+m-\pi(j)}{m-\pi(j)}$ ways to choose the elements of $S_1$ and $S_2$ which are greater than $e$.

Thus, the positions of $e$ in $\pi(S_1)$ and $\pi(S_2)$ uniquely determine the position $\pi^{-1}(e)$ and value $e$ of the intersection element; there are $[i, j]_m$ ways to choose which other positions are occupied by $\pi(S_1)$ and which ones, by $\pi(S_2)$; and there are $[\pi(i), \pi(j)]_m$ ways to choose which other values are in $\pi(S_1)$ and which ones are in $\pi(S_2)$.

Now that we have chosen both positions and values of elements of $S_1$ and $S_2$, we can produce a unique permutation $\rho(S)$ of $S$ which satisfies our conditions above. Simply fill the positions for $S_1$ and $S_2$ by elements of $\pi(S_1)$ and $\pi(S_2)$, respectively, in the order in which they occur.

Since the total number of permutations of $S$ is $(2m-1)!$, the coefficient of $n^{2m-1}$ in the $\binom{n}{2m-1}$-term of $E(X_\pi^2)$ is

$$\frac{1}{(2m-1)!}\cdot\frac{\sum_{i,j=1}^{m}\binom{i-1+j-1}{i-1}\binom{m-i+m-j}{m-j}\binom{\pi(i)-1+\pi(j)-1}{\pi(i)-1}\binom{m-\pi(i)+m-\pi(j)}{m-\pi(j)}}{(2m-1)!}$$
$$=\frac{\sum_{i,j=1}^{m}[i,j]_m[\pi(i),\pi(j)]_m}{((2m-1)!)^2}.$$

The coefficient of $n^{2m-1}$ in $\mathrm{Var}(X_\pi)$ is, by the previous equations,

$$\begin{aligned}[n^{2m-1}]\,\mathrm{Var}(X_\pi) &= \frac{\sum_{i,j=1}^{m}[i,j]_m[\pi(i),\pi(j)]_m}{((2m-1)!)^2}-\frac{m(2m-1)}{(m!)^4}+\frac{m(m-1)}{(m!)^4}\\ &= \frac{\sum_{i,j=1}^{m}[i,j]_m[\pi(i),\pi(j)]_m}{((2m-1)!)^2}-\frac{1}{(m!(m-1)!)^2}\\ &= \frac{1}{((2m-1)!)^2}\left(\sum_{i,j=1}^{m}[i,j]_m[\pi(i),\pi(j)]_m-\binom{2m-1}{m-1}^2\right).\end{aligned}$$

Since $c(\pi)$ is the leading coefficient of $\mathrm{Var}(X_\pi)$ (a polynomial in $n$), we have $c(\pi)\ge 0$. The following lemma implies that $c(\pi)>0$, which finishes the proof of Theorem 4.2. □

In the next theorem we use the notation $a!! := 0!1!\cdots a!$.

**Lemma 4.3.** *For any* $\pi\in S_m$, $\sum_{i,j=1}^{m}[i,j]_m[\pi(i),\pi(j)]_m > \binom{2m-1}{m-1}^2$.

**Proof.** The matrix $A_m$ is symmetric and hence diagonalizable. Moreover, the identity *(Rabbit)* of [2],

$$\det\left[\binom{i+j+a+b}{i+a}\binom{2n-i-j-a-b}{n-i-a}\right]_{0\le i,j\le m}=\frac{(a+b)!(2n+1)!^{m+1}}{a!b!}$$
$$\times\frac{(2n-m)!!m!!(m+a+b)!!(2n-m-a-b)!!a!!b!!(n-m-a-1)!!(n-m-b-1)!!}{(2n+1)!!(n-a)!!(n-b)!!(m+a)!!(m+b)!!(a+b)!!(2n-2m-a-b-1)!!},$$

implies that all leading principal minors of $A_m$ are positive, so $A_m$ is positive definite, and hence all eigenvalues of $A_m$ are positive. Each row and column of $A_m$ sums to $\frac{1}{2}\binom{2m}{m}=\binom{2m-1}{m-1}$, so $[1,1,\ldots,1]$ is an eigenvector for the eigenvalue $\binom{2m-1}{m-1}$. The same is true of the similar matrix $P_\pi A_m P_\pi^{-1}$. Let $D_m=[d_{ij}]_{m\times m}$ be the $m\times m$ diagonal matrix so that $d_{ii}$'s are the eigenvalues of $A_m$ and $d_{11}=\binom{2m-1}{m-1}$.[1] Then $A_m=BD_mB^{-1}$ for some orthogonal matrix $B$, so recalling that $\mathrm{Tr}(MN)=\mathrm{Tr}(NM)$ for any $M,N$ where $MN$ and $NM$ exist, we have

$$\begin{aligned}\mathrm{Tr}(A_mP_\pi A_mP_\pi^{-1}) &= \mathrm{Tr}(BD_mB^{-1}P_\pi BD_mB^{-1}P_\pi^{-1})=\mathrm{Tr}(D_mB^{-1}P_\pi BD_mB^{-1}P_\pi^{-1}B)\\ &= \mathrm{Tr}(D_m(B^{-1}P_\pi B)D_m(B^{-1}P_\pi B)^{-1})=\mathrm{Tr}(D_mCD_mC^{-1}),\end{aligned}$$

[1] In fact, it is known that the eigenvalues of $A$ are $\binom{2m-1}{i-1}$ for $i=1,\ldots,m$, although we are presently unable to find a reference (or proof) for this claim. It follows that $d_{ii}=\binom{2m-1}{m-i}$. These stronger claims are not needed in the rest of the proof.

and the matrix $C = B^{-1}P_\pi B = [c_{ij}]_{m\times m}$ is orthogonal, i.e. $C^{-1} = C^T$. Let $\mathbf{b}_i$ be the $i$th column of $B = [b_{ij}]_{m\times m}$. Then $c_{ij} = \sum_{k=1}^{m} b_{ki}b_{\pi(k)j}$. In particular, the column $\mathbf{b}_1 = [1, \dots, 1]^T$ remains unchanged for any permutation $\pi \in S_n$, so $c_{1j} = \mathbf{b}_1 \cdot \mathbf{b}_j = \delta(1, j)$ and $c_{i1} = \mathbf{b}_i \cdot \mathbf{b}_1 = \delta(i, 1)$. Now we know that

$$\mathrm{Tr}(D_m C D_m C^T) = \sum_{i,k=1}^{m} d_{ii}d_{kk}c_{ik}^2,$$

$d_{ii} > 0$, $c_{11} = 1$, $c_{i1} = c_{1i} = 0$ for $i > 1$, so $c_{ij}$ are not all zero for $i, j > 1$ (otherwise $C$ is not invertible, let alone orthogonal), and hence

$$\mathrm{Tr}(D_m C D_m C^T) > \binom{2m-1}{m-1}^2.$$

This proves the lemma. □

**Remark 4.4.** Note that the sum $\sum_{i,j=1}^{m}[i, j]_m[\pi(i), \pi(j)]_m$ is invariant under the symmetry operations on $S_m$: reversal $\mathfrak{r} : i \mapsto m - i + 1$, complement $\mathfrak{c} : \pi(i) \mapsto m - \pi(i) + 1$, and inverse $\mathfrak{i} : \pi \mapsto \pi^{-1}$. Invariance under $\mathfrak{r}$ and $\mathfrak{c}$ also extends to permutations of multisets. Thus permutations $\pi$ in the same symmetry class $\bar{\pi}$ have the same $c(\pi)$. The values of

$$\Delta(\pi) = ((2m-1)!)^2 c(\pi) = \sum_{i,j=1}^{m}[i, j]_m[\pi(i), \pi(j)]_m - \binom{2m-1}{m-1}^2$$

for symmetry classes in $S_4$ ($m = 4$) are given in the table below:

| $\bar{\pi}$ | 1234 | 1243 | 1432 | 1342 | 2413 |
|---|---|---|---|---|---|
| $\Delta(\pi)$ | 491 | 359 | 327 | 239 | 91 |

**Remark 4.5.** It is easy to see that, for a given $m$, $\Delta(\pi)$ attains its maximum when $\pi = id_m = 12\cdots m$ since the sequences $\{[i, j]_m\}$ and $\{[\pi(i), \pi(j)]_m\}$ (with multiplicities) are arranged in the same order. It would be interesting to characterize the permutations $\pi_*$ for which $\Delta(\pi_*) = \min_{\pi\in S_m} \Delta(\pi)$. For small values of $m$, these permutations $\pi_*$ are in the symmetry classes of:

| $m$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| $\pi_*(m)$ | 1 | 12 | 132 | 2413 | 25314 | 361452 | 3614725 | 37145826 | 629471583 |

Interestingly, the patterns $\pi_*(m)$ are either involutions (e.g. for all $1 \le m \le 9$ except $m = 4$) or centrally symmetric, i.e. invariant under $\mathfrak{r} \circ \mathfrak{c}$ (e.g. for $m = 1, 2, 4, 5, 7, 8$), or both. They are also less avoided than most patterns of the same length, and in fact, are the least avoided patterns for $m \le 5$. We believe (but cannot prove) that this is not a coincidence.

We can consider the co-occurrence of any two permutation patterns similarly. Since the proof is similar to the variance case, it is omitted.

**Theorem 4.6.** *For any patterns $\pi_1, \pi_2 \in S_m$, $m > 1$, the covariance* $\mathrm{Cov}(X_{\pi_1}, X_{\pi_2})$ *is given by*

$$\mathrm{Cov}(X_{\pi_1}, X_{\pi_2}) = c(\pi_1, \pi_2)n^{2m-1} + O(n^{2m-2}),$$

*where*

$$c(\pi_1, \pi_2) = \frac{1}{((2m-1)!)^2}\left(\mathrm{Tr}(A_m P_{\pi_1} A_m P_{\pi_2}^{-1}) - \binom{2m-1}{m-1}^2\right).$$

*The trace in the above formula is*

$$\mathrm{Tr}(A_m P_{\pi_1} A_m P_{\pi_2}^{-1}) = \sum_{i,j=1}^{m}[i, j]_m[\pi_1(i), \pi_2(j)]_m.$$

Considering symmetry classes of pairs of patterns (see Remark 4.4), we see that there are 7 classes of pairs of 3-letter permutations: {123, 123}, {132, 132}, {123, 132}, {132, 213}, {132, 231}, {123, 231}, {123, 321} (listed in the order of decreasing asymptotical covariance). The first two pairs obviously have a positive covariance, but of the other five pairs, only {123, 132} has a positive covariance.

It would be interesting to characterize the pairs $\{\pi_1, \pi_2\}$ according to the sign or magnitude of their covariance.

We now consider patterns contained in words, where repeated letters are allowed both in the pattern and the ambient string. The additional condition on a pattern $\pi \in [l]^m$ on words, i.e. on a pattern of $m$ letters over an alphabet $[l] = \{1, 2, \ldots, l\}$, is that $\pi$ must contain all letters in $[l]$. We will also assume that the ambient strings are in the set $[k]^n$.

**Theorem 4.7.** *Let $\pi$ be a map of $[m] = \{1, 2, \ldots, m\}$ onto $[l] = \{1, 2, \ldots, l\}$. Then for any positive integers $l \le m$,*

$$\mathrm{Var}(X_\pi) = c(\pi)n^{2m-1}k^{2l-1} + O(n^{2m-2}k^{2l-1} + n^{2m-1}k^{2l-2}),$$

*where*

$$c(\pi) = \frac{1}{(2m-1)!(2l-1)!}\left(\mathrm{Tr}(A_m P_\pi A_l P_\pi^{-1}) - \frac{(2m-1)!(2l-1)!}{((m-1)!)^2(l!)^2}\right).$$

*The trace in the above formula is*

$$\mathrm{Tr}(A_m P_\pi A_l P_\pi^{-1}) = \sum_{i,j=1}^{m} [i,j]_m [\pi(i), \pi(j)]_l.$$

**Remark 4.8.** Note also that, given $1 \le l \le m$, Theorem 4.7 applies to $l!S(m,l)$ patterns $\tau$, where $S(m,l)$ is the Stirling number of the second kind.

The proof of Theorem 4.7 is an obvious extension of the proof of Theorem 4.2. Unfortunately, the same extension to words does not work for Lemma 4.3, but only yields

$$\sum_{i,j=1}^{m} [i,j]_m [\pi(i), \pi(j)]_l > \binom{2m-1}{m}\binom{2l-1}{l} = \frac{l}{m}\left(\frac{(2m-1)!(2l-1)!}{((m-1)!)^2(l!)^2}\right),$$

which is a weaker result than the desired inequality

$$\sum_{i,j=1}^{m} [i,j]_m [\pi(i), \pi(j)]_l > \frac{(2m-1)!(2l-1)!}{((m-1)!)^2(l!)^2}.$$

There is a similar covariance result on words as well.

**Theorem 4.9.** *For any patterns $\pi_1, \pi_2 \in [l]^m$, $1 < l \le m$, the covariance $\mathrm{Cov}(X_{\pi_1}, X_{\pi_2})$ is*

$$\mathrm{Cov}(X_{\pi_1}, X_{\pi_2}) = c(\pi_1, \pi_2)n^{2m-1}k^{2l-1} + O(n^{2m-2}k^{2l-1} + n^{2m-1}k^{2l-2}),$$

*where*

$$c(\pi_1, \pi_2) = \frac{1}{(2m-1)!(2l-1)!}\left(\mathrm{Tr}(A_m P_{\pi_1} A_l P_{\pi_2}^{-1}) - \frac{(2m-1)!(2l-1)!}{((m-1)!)^2(l!)^2}\right).$$

*The trace in the above formula is*

$$\mathrm{Tr}(A_m P_{\pi_1} A_m P_{\pi_2}^{-1}) = \sum_{i,j=1}^{m} [i,j]_m [\pi_1(i), \pi_2(j)]_l.$$

## Acknowledgements

We would like to thank the referee for useful comments. The first author's research was supported in part by the NSA Young Investigator Grant H98230-06-1-0037.

## References

[1] M.H. Albert, M.D. Atkinson, C.C. Handley, D.A. Holton, W. Stromquist, On packing densities of permutations, Electron. J. Combin. 9 (1) (2002) #R5.
[2] T. Amdeberhan, S.B. Ekhad, A condensed condensation proof of a determinant evaluation conjectured by Greg Kuperberg and Jim Propp, J. Combin. Theory Ser. A 78 (1997) 169–170.
[3] R.W. Barton, Packing densities of patterns, Electron. J. Combin. 11 (1) (2004) #R80.
[4] M. Bóna, B.E. Sagan, V.R. Vatter, Pattern frequency sequences and internal zeros, Adv. Appl. Math. 28 (2002) 395–420.
[5] A. Burstein, P. Hästö, T. Mansour, Packing patterns into words, Electron. J. Combin. 9 (2) (2002–2003) #R20.
[6] P.A. Hästö, The packing density of other layered permutations, Electron. J. Combin. 9 (2) (2002) #R1.
[7] M. Hildebrand, B. Sagan, V. Vatter, Bounding quantities related to the packing density of $1(m+1)m\ldots2$, Adv. Appl. Math. 33 (3) (2004) 633–653.
[8] D.E. Knuth, The Art of Computer Programming, Vols.1, 3, Addison-Wesley, NY, 1968, 1973.
[9] A. Price, Packing densities of layered patterns, Ph.D. Thesis, University of Pennsylvania, 1997.
[10] W. Stromquist, Packing layered posets into posets, 1993, unpublished manuscript.
[11] R. Tarjan, Sorting using networks of queues and stacks, J. Assoc. Comput. Mach. 19 (1972) 341–346.
[12] D. Warren, Optimal packing behavior of some two-block patterns, Ann. Comb. 8 (3) (2004) 355–367.
[13] D. Warren, A mild generalization of two results of Hästö, 2005, preprint.
[14] D. Warren, Packing densities of more 2-block patterns, Adv. Appl. Math. 36 (2) (2006) 202–211.