The packing density of other layered permutations

Peter A. Hästö*

Department of Mathematics,
P. O. Box 4, 00014 University of Helsinki, Finland
peter.hasto@helsinki.fi

Submitted: Aug 21, 2002; Accepted: Oct 10, 2002; Published: Oct 31, 2002
MR Subject Classifications: Primary 05A15; Secondary 05A16

Abstract

In this paper the packing density of various layered permutations is calculated, thus solving some problems suggested by Albert, Atkinson, Handley, Holton & Stromquist [Electron. J. Combin. **9** (2002), #R5]. Specifically, the density is found for layered permutations of type $[m_1, \ldots, m_r]$ when $\log(r+1) \le \min\{m_i\}$. It is also shown how to derive good estimates for the packing density of permutations of type $[k, 1, k]$ when $k \ge 3$. Both results are based on establishing the number of layers in near optimal permutations using a layer-merging technique.

1 Introduction

Let $\sigma \in S_n$ (the symmetric group of n letters) and $\pi \in S_m$. The number of occurrences of π in σ is the number of m element subsets E of $[n] := \{1, 2, \ldots, n\}$ such that $\sigma|_E$ and π are isomorphic (as mappings of ordered sets). For instance the permutation 1374625 contains 5 occurrences of the permutation 1423, namely 1746, 1745, 1725, 3746 and 3745. Two quite different problems related to such permutation containment have been studied: the number (or characterization) of permutations not containing a given permutation (e.g. [2, 5, 6, 8]) and the maximum number of containments of a given permutation [1, 4, 7]. It is the latter problem that will concern us in this paper.

Let us denote the number of times that $\pi \in S_m$ is contained in $\sigma \in S_n$ by $\nu(\pi, \sigma)$. If we divide this number by the total number of subsequences of σ of length m (for $m \le n$) we get the density of π in σ:

$$d(\pi, \sigma) := \frac{\nu(\pi, \sigma)}{\binom{n}{m}}.$$

Since we want to determine the maximum number of containments, we further define

$$d_n(\pi) := \max_{\sigma \in S_n} d(\pi, \sigma).$$

*Supported in part by the Academy of Finland

We say that a permutation $\sigma \in S_n$ is π-*maximal* if $d_n(\pi) = d(\pi, \sigma)$. It turns out that $d_n(\pi)$ is decreasing in n and hence it makes sense to define the *packing density of* π by

$$d(\pi) := \lim_{n\to\infty} d_n(\pi)$$

(this is proved in [1, Proposition 1.1], although the authors of that paper consider it a part of combinatorial folklore). In [1] the packing density is also defined for sets of permutations, however, none of the results of this paper pertain to this more general case.

Since the packing density problem seems to be quite difficult in general we restrict our attention to the packing density of layered permutations. We say that the permutation $\pi \in S_m$ is *layered* if there exist numbers $m_1, \ldots, m_r$, the sum of which equals m, such that π starts with the m_1 first positive integers in reverse order, followed by the next m_2 positive integers in reverse order and so on. More specifically, we say that this permutation is of type $[m_1, \ldots, m_r]$. For instance 213654 is layered of type $[2, 1, 3]$. Notice that the type of a layered permutation uniquely determines the permutation. The nice thing about considering layered permutations is that W. Stromquist [7] proved:

Theorem (Theorem 2.2, [1]). *Let π be a layered permutation. Among the π-maximal permutations of each length there will be one that is layered. Furthermore, if all the layers of π have size greater than 1, then every π-maximal permutation is layered.*

In this paper we will only consider the packing density of layered permutations and therefore we introduce the following convention:

Notation 1.1. Throughout this paper we denote by π a layered permutation of type $[m_1, \ldots, m_r]$ and by m the sum $m_1 + \ldots m_r$. All other permutations are also assumed to be layered, unless specified to the contrary.

The central theme of the results in this paper is the number of layers in near π-maximal permutations. It was shown in [4] for some permutations the number of layers in π-maximal $\sigma_n \in S_n$ is bounded as $n \to \infty$ whereas for others it is unbounded (we say that these are of the *bounded* and *unbounded type*, respectively). Albert, Atkinson, Handley, Holton & Stromquist (hereafter referred to as AAHH&S) stated the following conjecture.

Conjecture 2.9 from [1]. *Suppose that π is a layered permutation whose first and last layers have size greater than 1 and which has no adjacent layers of size 1. Then π is of the bounded type.*

These authors showed that the conjecture is true when we consider only layered permutations with at most three layers [1, Proposition 2.8] or permutations with every layer of size two or greater [1, Theorem 2.7]. Also, in Proposition 2.10 they showed that the assumption on the first and last layers is necessary. Knowing that a permutation is of the bounded type has certain implications, in particular it allows us to estimate (and in principle, also to calculate the exact value of) the packing density by finding a maximum of a certain function introduced by Price in [4] (see Section 2 in this paper). Nevertheless, the bounds on the number of layer given by the previous finiteness results are so large that they are virtually useless in determining the packing density. For instance, Theorem 2.7 of [1] implies that the number of layers in a π-maximal permutation for π of type $[2,3,2]$ is less than 30 [1, p. 19] whereas AAHH&S suggested that the correct number in this case should be three. The contributions of this paper are two results which apply only to more limited classes of layered permutations, but conversely give optimal bounds for the number of layers in near π-maximal permutations.

Let us say that the layered permutation π is *simple* if there exists a sequence $\{\sigma_n\}$ with $\sigma_n \in S_n$ such that every σ_n has r layers and $\lim_{n\to\infty} d(\pi,\sigma_n) = d(\pi)$. It turns out that it is very easy to calculate the packing density of a simple permutation, see Lemma 3.1. The next result shows that there are many simple permutations:

Theorem 1.2. *Let $\pi \in S_m$ be a layered permutation of type $[m_1,\dots,m_r]$. If $\log_2(r+1) \le \min\{m_i\}$ then π is simple and*

$$d(\pi) = \frac{m!}{m^m}\prod_{k=1}^{r}\frac{m_k^{m_k}}{m_k!},$$

where $m := m_1 + \ldots + m_r$.

In Lemma 3.5 we show that there exists a permutation with

$$\min\{m_i\} \le \frac{\log(r+1)}{r\log(1+1/r)}$$

which is not simple. This implies that the logarithmic bound in the previous theorem is asymptotically off by at most a factor of $1/\log 2$.

Notice that Theorem 1.2 solves the packing density problem for layered permutations with two or three layers none of which is a *singleton* (i.e. has length 1). Since A. Price [4] has previously solved the packing density problem for permutations of the type $[1,k]$ this means that the we now know how to handle all the two layer cases.

The (non-trivial) layered permutations with three layers not covered by the theorem are of type $[1,k_1,k_2]$, $[1,k_1,1]$, $[1,1,k_1]$ or $[k_1,1,k_2]$ (with $k_1,k_2 \ge 2$). Recall that the first three of these were shown to be of the unbounded type in Proposition 2.10, [1], which suggests that it will be difficult to calculate or estimate their packing density (it might be possible to handle the case $[1,1,k_1]$ as in [1, Proposition 2.4] but the generalization is not straightforward). Section 4 is devoted to a special case of the fourth type, $[k_1,1,k_2]$.

We will show that permutations with a singleton layer are never simple; however, in some cases near π-maximal permutation can be chosen to have exactly one layer more than the packed permutation. More precisely, let us say that a permutation π is almost simple if it is not simple, but there there exists a sequence $\{\sigma_n\}$ with $\sigma_n \in S_n$ such that every σ_n has $r+1$ layers and $\lim_{n\to\infty} d(\pi,\sigma_n) = d(\pi)$.

Theorem 1.3. *Let π be a layered permutation of type $[k,1,k]$ with $k \ge 3$. Then π is almost simple.*

It turns out that this result gives us very good estimates of the packing densities of these permutations, see Theorem 4.3 and Table 1 on page 14. Unfortunately, the case $[2,1,2]$ is not covered, which means that we are not able to answer the question asked in [1, p. 19] regarding the packing density of this permutation. The reason for this disclusion is discussed in Remark 4.2. This result also implies that near optimal permutations for symmetric partitions might be far from symmetric, which answers in the negative a question in [1, p. 13], see Proposition 4.4.

The structure of the rest of this paper is as follows: the tools for attacking packing density problems from [1, 4] will be introduced in the next section. Theorem 1.2 is proved in Section 3 and Theorem 1.3 is proved in Section 4. Both sections are concluded by some open problems.

2 Near π-maximal permutations and optimal partitions

In this section we introduce a tool of Alkes Price's for calculating the packing density of layered partitions. Specifically, the claims in this paragraph are from [1, p. 13] and are proved (according to [1]) in [4]. Price defined

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

where $\lambda_1+\ldots+\lambda_s = 1$ and $\lambda_i \ge 0$ for $1 \le i \le s$ (such a sequence of λ's will be called a *partition of unity*). This is approximately the density of π in the permutation of type $\big[\lfloor n\lambda_1\rfloor,\ldots,\lfloor n\lambda_s\rfloor\big]$ (the approximation getting better as n increases). Price further defined

$$p_s := \max p_s(\lambda_1,\ldots,\lambda_s)$$

where the maximum is also over partitions of unity. It is clear that p_s is increasing as a function of s and moreover $p_s \to d(\pi)$ as $s \to \infty$. It was also shown by Price that π is of the bounded type if and only if $p_s = d(\pi)$ for some s.

For π of the bounded type let us say that the partition of unity $(\lambda_i)_{i=1}^s$ is an *optimal partition* if s is the least integer with $p_s = p_s(\lambda_1,\ldots,\lambda_s) = d(\pi)$. It is clear that that $\lambda_i > 0$ for every $i \in [s]$ in an optimal partition.

Unfortunately it is not known whether the maximal number of layers in π-maximal permutations is the least s such that $p_s = d(\pi)$. However, this question turns out not to be important for us, since the contribution of possible extra layers is negligible anyway. Indeed, there is an obvious connection between the number of layers in an optimal partition and the number of "large" layers in near π-maximal permutations:

Lemma 2.1. *The permutation π is simple* (*almost simple*) *if and only if there exists an optimal partition with exactly r $(r+1)$ layers.*

Remark 2.2. This result follows from [4, Theorem 3.1], but since that paper is not easily available, a simple proof is given here

Proof of Lemma 2.1. We prove only the claim regarding simple permutation, since the case of almost simple permutations is similar.

Suppose first that $(\lambda_i)_{i=0}^r$ is an optimal partition. Let σ_i be of type $\left[\lfloor i\lambda_1\rfloor, \ldots, \lfloor i\lambda_r\rfloor\right]$. Then $d(\pi, \sigma_i) \to d(\pi)$ and it is clear that π is simple.

Assume conversely that π is simple and let $\{\sigma_i\}$ be a sequence of permutations with r layers such that $\sigma_i \in S_i$ and $\lim d(\pi, \sigma_i) = d(\pi)$. We can choose a subsequence of $\{\sigma_i\}$ such that $n_j(i)/i$ converges for $j = 1, \ldots, r$, where $n_j(i)$ is the length of the j^{th} layer of σ_i. It is clear that $\lambda_j := \lim_{i\to\infty} n_j(i)/i$ defines an optimal partition. $\square$

In order to avoid writing out the $\binom{m}{m_1,\ldots,m_r}$ all the time we define an alternative to the p_s function by

$$q_s(\lambda_1, \ldots, \lambda_s) := \sum_{1\le i_1<\ldots<i_r\le s} \lambda_{i_1}^{m_1} \cdots \lambda_{i_r}^{m_r} \tag{1}$$

where all the λ's are positive and $\lambda_1 + \ldots + \lambda_s = 1$. We also define $q_s := p_s/\binom{m}{m_1,\ldots,m_r}$.

3 On simple permutations

An intuitive guiding principle in searching for (near) π-maximal permutations is that they should be structured in a fashion similar to π. In the simplest case this principle suggests that the π-maximal permutation for the simple permutation π will have layers sizes proportional to those of π. The next lemma, which appears in Price's thesis, can easily be proved by fixing all but two variables and considering extreme points of the resulting (one parameter) expression.

Lemma 3.1 (Theorem 4.1, [4]). *Suppose that π is simple with layers $[m_1, \ldots, m_r]$. Then $(m_1/m, \ldots, m_r/m)$ is an optimal partition and hence*

$$d(\pi) = \frac{m!}{m^m} \prod_{k=1}^{r} \frac{m_k^{m_k}}{m_k!}. \tag{2}$$

We now start the proof that certain permutations, namely those with quite long and not so many layers, are simple. We first need a general structure lemma.

Lemma 3.2. *If $2 \le m_1 \le m_2 \le \ldots \le m_r$ then π is of the bounded type and there exists an optimal partition with $\lambda_1 \le \ldots \le \lambda_s$.*

Proof. It follows from [1, Theorem 2.7] that π is of the bounded type. Let $(\lambda_1, \ldots, \lambda_s)$ be an optimal partition. Let $1 \le k < s$ and suppose $\lambda_1 \le \ldots \le \lambda_k$ but $\lambda_k > \lambda_{k+1}$ (if no such k exists then there we are done). Consider what happens if we swap λ_k and λ_{k+1}. Since the original sequence was optimal we have

$$q_s(\lambda_1, \ldots, \lambda_k, \lambda_{k+1}, \ldots, \lambda_s) \ge q_s(\lambda_1, \ldots, \lambda_{k+1}, \lambda_k, \ldots, \lambda_s).$$

Both sides in this inequality contain a large number of terms, but most of them occur on both sides. Specifically, all terms that do not contain both λ_k and λ_{k+1} appear on both sides. After canceling these terms and moving the remaining ones to the left hand side all that remains is

$$\sum_{t=1}^{r-1}\left(\lambda_k^{m_t}\lambda_{k+1}^{m_{t+1}}-\lambda_{k+1}^{m_t}\lambda_k^{m_{t+1}}\right)\sum_{1\le i_1<\ldots<i_{t-1}\le k-1}\lambda_{i_1}^{m_1}\cdots\lambda_{i_{t-1}}^{m_{t-1}}\times$$
$$\sum_{k+2\le i_{t+2}<\ldots<i_r\le s}\lambda_{i_{t+2}}^{m_{t+2}}\cdots\lambda_{i_r}^{m_r}\ge 0. \quad (3)$$

Since $\lambda_k > \lambda_{k+1}$ we have $\lambda_k^{m_t}\lambda_{k+1}^{m_{t+1}}-\lambda_{k+1}^{m_t}\lambda_k^{m_{t+1}} \le 0$. It then follows from (3) that for every t either $\lambda_k^{m_t}\lambda_{k+1}^{m_{t+1}}-\lambda_{k+1}^{m_t}\lambda_k^{m_{t+1}} = 0$ or

$$\sum_{1\le i_1<\ldots<i_{t-1}\le k-1}\lambda_{i_1}^{m_1}\cdots\lambda_{i_{t-1}}^{m_{t-1}}\sum_{k+2\le i_{t+2}<\ldots<i_r\le s}\lambda_{i_{t+2}}^{m_{t+2}}\cdots\lambda_{i_r}^{m_r}=0.$$

Either way we get

$$q_s(\lambda_1,\ldots,\lambda_k,\lambda_{k+1},\ldots,\lambda_s)=q_s(\lambda_1,\ldots,\lambda_{k+1},\lambda_k,\ldots,\lambda_s)$$

and we see that $(\lambda_1,\ldots,\lambda_{k+1},\lambda_k,\ldots,\lambda_s)$ is also an optimal permutation.

This means that from an optimal partition with the layer sizes increasing till layer k we get an optimal partition with the layer sizes increasing till layer $k+1$. Continung like this we get an optimal partition with layer sizes increasing throughout. $\square$

As was seen in the previous proof the sums over the λ's quickly become very complicated with double indices all over. In view of this we introduce the following notation:

$$\bigoplus_{a\ldots b}(u,v) := \sum_{a\le i_u<\ldots<i_v\le b}\lambda_{i_u}^{m_u}\cdots\lambda_{i_v}^{m_v},$$

if $a \le b$ and $u \le v$ and $\bigoplus_{a\ldots b}(u,v) := 1$ if $a \le b$ and $v < u$.

Theorem 3.3. *If $2 \le m_1 \le m_2 \le \ldots \le m_r$ and $m_1 \ge \log_2(r+1)$ then π is simple.*

Proof. As in Lemma 3.2, let $(\lambda_i)_{i=1}^s$ be an increasing optimal partition and assume that $s > r$. Consider what happens when we merge the first two layers. Then

$$q_{s-1}(\lambda_1+\lambda_2,\ldots,\lambda_s) < q_s(\lambda_1,\lambda_2,\ldots,\lambda_s),$$

since s is minimal. Canceling equal terms and rearranging leaves

$$\left((\lambda_1+\lambda_2)^{m_1}-\lambda_1^{m_1}-\lambda_2^{m_1}\right)\bigoplus_{3\ldots s}(2,r) < \lambda_1^{m_1}\lambda_2^{m_2}\bigoplus_{3\ldots s}(3,r). \quad (4)$$

Let us first show that

$$\lambda_2^{m_2}\bigoplus_{3\ldots s}(3,r) \le (r-1)\bigoplus_{3\ldots s}(2,r) \quad (5)$$

for $s > r$. It is easy to check this for $s = r+1$ since the previous inequality then becomes

$$\lambda_2^{m_2} \bigoplus_{3\ldots s}(3,r) \leq (s-2)\lambda_3^{m_2}\lambda_4^{m_3}\cdots\lambda_s^{m_r}. \tag{6}$$

But here the left hand side has $s-2$ terms, each of which is less than the product of the λ's on the right hand side, so this inequality is clear. It remains to handle the case $s > r+1$.

The case $r = 2$ is also easy, since (5) reduces to $\lambda_2^{m_2} \leq \lambda_3^{m_2} + \ldots + \lambda_s^{m_2}$ which certainly holds since $\lambda_2 \leq \lambda_3$. Assume then by induction that (5) holds for $s-1$ and all $r = 2, \ldots, s-2$. Then we have

$$\begin{aligned}
\lambda_2^{m_2} \bigoplus_{3\ldots s}(3,r) &= \lambda_2^{m_2} \bigoplus_{3\ldots(s-1)}(3,r) + \lambda_2^{m_2}\lambda_s^{m_r} \bigoplus_{3\ldots(s-1)}(3,r-1) \\
&\leq (r-1) \bigoplus_{3\ldots(s-1)}(2,r) + (r-2)\lambda_s^{m_r} \bigoplus_{3\ldots(s-1)}(2,r-1) \\
&\leq (r-1)\bigoplus_{3\ldots s}(2,r),
\end{aligned}$$

where the induction assumption is used twice for the second inequality.

We then combine the estimate (5) with (4). This gives

$$\big((\lambda_1+\lambda_2)^{m_1} - \lambda_1^{m_1} - \lambda_2^{m_1}\big) \bigoplus_{3\ldots s}(2,r) < (r-1)\lambda_1^{m_1} \bigoplus_{3\ldots s}(2,r).$$

Since the sum is not zero, we can divide it out, and all that remains is

$$(1+\lambda_2/\lambda_1)^{m_1} - 1 - (\lambda_2/\lambda_1)_1^m < r-1.$$

Since the left hand side has its minimum at $\lambda_1 = \lambda_2$ (because $\lambda_2/\lambda_1 \geq 1$ and $x \mapsto (1+x)^{m_1} - x^{m_1}$ is increasing) this implies that $2^{m_1} - 2 < r-1$, which contradicts the assumption of the theorem. This contradiction shows that the assumption $s > r$ is false, hence $s = r$ and π is simple, as claimed. $\square$

We next show that the assumption that the m_k are increasing is not really essential! We start with a lemma.

Lemma 3.4. *Suppose that* $\min m_i \geq 2$ *and let* π' *be the layered permutation of type* $[m_1', \ldots, m_r']$ *where* (m_k') *is an increasing reordering of the layer sizes* (m_k) *of* π*. Then* $d(\pi) \leq d(\pi')$.

Proof. Let (λ_k) be an optimal partition. We have

$$q_s = \sum_{1\leq i_1<\ldots<i_r\leq s} \lambda_{i_1}^{m_1}\cdots\lambda_{i_r}^{m_r} \leq \sum_{1\leq i_1<\ldots<i_r\leq s} \sup_{\theta\in S_r} \lambda_{i_1}^{m_{\theta(1)}}\cdots\lambda_{i_r}^{m_{\theta(r)}}$$

where θ is a not necessarily layered permutation. It is easily seen that the permutation θ should be chosen so that the least λ_k is raised to the least m_j, the second to least λ_k to the second to least m_j and so on. But this means that if (λ'_k) is the increasing rearrangement of (λ_k) then

$$\sum_{1\le i_1<\ldots<i_r\le s} \sup_{\sigma\in S_r} \lambda_{i_1}^{m_{\sigma(1)}}\cdots\lambda_{i_r}^{m_{\sigma(r)}} = \sum_{1\le i_1<\ldots<i_r\le s} {\lambda'_{i_1}}^{m'_1}\cdots{\lambda'_{i_r}}^{m'_r}.$$

It therefore follows that

$$d(\pi) = \binom{m}{m_1,\ldots,m_r} \sum_{1\le i_1<\ldots<i_r\le s} \lambda_{i_1}^{m_1}\cdots\lambda_{i_r}^{m_r} \le$$
$$\binom{m}{m_1,\ldots,m_r} \sum_{1\le i_1<\ldots<i_r\le s} {\lambda'_{i_1}}^{m'_1}\cdots{\lambda'_{i_r}}^{m'_r} = p_s(\lambda'_1,\ldots,\lambda'_r) \le d(\pi'),$$

which was to be shown. ☐

Proof of Theorem 1.2. The case $\min\{m_i\} = 1$ is trivial, since the condition of the theorem then implies that π has a single layer. Assume next that $\min\{m_i\} > 1$. Let π' be the layered permutation whose layers are the increasing rearrangement of the layers of π, as in Lemma 3.4. It follows from the lemma that $d(\pi) \le d(\pi')$.

On the other hand we know from Theorem 3.3 that π' is simple, which means that it has an optimal partition with exactly r layers. But rearranging these layers gives a partition of unity which is a maximum of p_r. This implies that $d(\pi) \ge p_r = d(\pi')$. It then follows that $d(\pi) = d(\pi') = p_r$ and that π is simple. ☐

Let us next show that the logarithmic lower bound for $\min\{m_i\}$ in the previous theorem is quite good, that is to say, off by only a constant. Indeed, since $r\log(1+1/r) \to 1$ as $r\to\infty$ the ratio between the sufficient bound from Theorem 3.3 and the necessary bound on $\min\{m_i\}$ from the next lemma approaches $1/\log 2 \approx 1.44$ as $r\to\infty$. A more thorough analysis of the packing density of permutations with all layers of equal size was done by Price, see [4, Theorem 6.1].

Lemma 3.5. *Let $m_k = \mu \ge 2$ for $k = 1,\ldots r$. If*

$$\mu < \frac{\log(r+1)}{r\log(1+1/r)}$$

then π is not simple. (*Here* log *denotes the natural logarithm.*)

Proof. We know from Lemma 3.1 that all the layers in the optimal partition should be of the same size if π is simple. If this were the case then we would have

$$q_r = q_r(1/r,\ldots,1/r) \ge q_{r+1}(1/(r+1),\ldots,1/(r+1)).$$

Writing out this inequality gives $r^{-r\mu} \ge (r+1)(r+1)^{-r\mu}$. Taking logarithms and solving for μ gives the inequality

$$\mu \ge \frac{\log(r+1)}{r\log(1+1/r)}.$$

Since this inequality is not satisfied, we see that the assumption that π is simple is false. ☐

We have seen in this section that the packing density can be calculated easily for many permutations. The most obvious and useful extension of these results would be to improve the bound $\log_2(r+1) \leq \min\{m_i\}$ from Theorem 1.2. There are at least two reasons to think that this is possible. First, it seems intuitively clear that the case $m_i = \mu$ for all i is the hardest one and therefore bounds in line with Lemma 3.5 would be more relevant. Second, the technique used in proving Theorem 3.3 was quite primitive since there is nothing particularly smart in just merging the first two layers.

Another interesting and quite simple development would be to improve inequality (5) to the form

$$\lambda_2^{m_2} \bigoplus_{3\ldots s}(3,r) \leq \frac{r-1}{s-r} \bigoplus_{3\ldots s}(2,r)$$

for $s > r$ and (λ_i) increasing. The reason for thinking that this inequality would hold is that the terms on the right hand side seem to be (on average) larger than those on the left hand side and there are $\binom{s-2}{r-1}$ and $\binom{s-2}{r-2}$ terms on the right and left hand side, respectively. If this inequality would indeed hold then we would get quite a good estimate for the difference $s-r$ even when π is not simple, namely $s-r$ would be bounded from above by $(r-1)/(2^{m_1}-2)$. For instance this would imply that permutations with four or five non-singleton layers would be either simple or almost simple.

4 On almost simple permutations

In the previous section we saw how layered permutations without singletons are often simple in which case it is easy to calculate their packing density. Unfortunately, this approach does not work for layered permutations with singleton layers, since such permutations are never simple. Indeed, suppose that $m_i = 1$ and that $(\lambda_1, \ldots, \lambda_r)$ were an optimal partition with r layers. Then splitting the layer $\lambda_i = a + b$ into two layers $a > 0$ and $b > 0$ increases the density of the permutation, since it removes no occurrences of π but adds some new ones, namely those in which the i^{th} layer of π occurs in the layer a and the $i+1^{\text{st}}$ layer of π occurs in the layer b as well as those in which the i^{th} layer of π occurs in the layer b and the $i-1^{\text{st}}$ layer of π occurs in the layer a.

In this section we show that the permutation of type $[k,1,k]$, though not simple, is almost simple if $k \geq 3$ and that this gives us good estimates for its packing density. Let us first note that these permutations are of the bounded type by [1, Proposition 2.8] and hence the partition tools from Section 2 (and [4, 1]) are applicable. As in the previous section we first need a structure lemma.

Lemma 4.1. *If π is of type $[k,1,k]$ for $k \geq 2$ then there exists a t such that every optimal partition is decreasing until t and then increasing, i.e.*

$$\lambda_1 \geq \ldots \geq \lambda_t \leq \ldots \leq \lambda_s.$$

Also, $\lambda_1 \geq k\lambda_2$ and $\lambda_s \geq k\lambda_{s-1}$.

Proof. Let (λ_k) be an optimal partition. We start by proving the second claim. For this, consider what happens if we vary λ_1 and λ_2 while keeping the other λ's and $\lambda_1+\lambda_2$ fixed. To this end we define

$$g(d) := q_s(\lambda_1+d, \lambda_2-d, \ldots, \lambda_s) = \big((\lambda_1+d)^k+(\lambda_2-d)^k\big) \bigoplus_{3\ldots s}(2,3)+(\lambda_1+d)^k(\lambda_2-d) \sum_{i=3}^{s} \lambda_i^k.$$

First note that $\lambda_1 \geq \lambda_2$, since otherwise exchanging them increases the q_s function. Next differentiate g with respect to d and consider the value of this derivative at $d=0$. Since $d=0$ is a maximum we find that

$$g'(0) = k\big(\lambda_1^{k-1} - \lambda_2^{k-1}\big) \bigoplus_{3\ldots s}(2,3) + \lambda_1^{k-1}\big(k\lambda_2 - \lambda_1\big) \sum_{i=3}^{s} \lambda_i^k = 0.$$

Since $\lambda_1 \geq \lambda_2$ the factor in front of the first sum is positive, and therefore the factor in front of the second sum must be negative. But this means that $k\lambda_2 \leq \lambda_1$, which proves the second claim of the lemma since $\lambda_s \geq k\lambda_{s-1}$ follows by symmetry.

We move on to the proof of the first claim. Consider what happens if we swap λ_i and λ_{i+1}:

$$\begin{aligned} q_s(\lambda_1, \ldots, \lambda_i, \lambda_{i+1}, \ldots, \lambda_s) - q_s(\lambda_1, \ldots, \lambda_{i+1}, \lambda_i, \ldots, \lambda_s) = \\ \big(\lambda_i^k\lambda_{i+1} - \lambda_{i+1}^k\lambda_i\big)\left(\sum_{j=i+2}^{s} \lambda_j^k - \sum_{j=1}^{i-1} \lambda_j^k\right) \geq 0. \end{aligned} \tag{7}$$

Suppose that t is such that λ_t is minimal and strictly less than λ_{t+1} (such t exists, by the what was just proved). Then $\lambda_t^k\lambda_{t+1} - \lambda_{t+1}^k\lambda_t < 0$ and hence

$$\sum_{j=t+2}^{s} \lambda_j^k \leq \sum_{j=1}^{t-1} \lambda_j^k.$$

Using this in (7) for larger values of i shows that $\lambda_i \leq \lambda_{i+1}$ for every $i \geq t$. By symmetry this implies the first claim. □

Proof of Theorem 1.3. Let $(\lambda_i)_{i=1}^s$ be an optimal partition and assume that $s \geq 5$. Let t be such that λ_t is minimal and assume without loss of generality (by symmetry) that $\lambda_{t-1} \geq \lambda_{t+1}$. By Lemma 4.1 we have $\lambda_1 \geq \ldots \geq \lambda_t \leq \ldots \leq \lambda_s$. Consider the effect of merging the layers λ_t and λ_{t+1}. We have

$$\begin{aligned} q_s(\lambda_1, \ldots, \lambda_t, \lambda_{t+1}, \ldots, \lambda_s) - q_{s-1}(\lambda_1, \ldots, \lambda_t + \lambda_{t+1}, \ldots, \lambda_s) = \lambda_t^k\lambda_{t+1} \sum_{i\geq t+2} \lambda_i^k + \\ \lambda_t\lambda_{t+1}^k \sum_{i\leq t-1} \lambda_i^k - \Big[(\lambda_t + \lambda_{t+1})^k - \lambda_t^k - \lambda_{t+1}^k\Big]\Bigg[\bigoplus_{1\ldots(t-1)}(1,2) + \bigoplus_{(t+2)\ldots s}(2,3)\Bigg] \geq 0. \end{aligned} \tag{8}$$

Suppose first that $\lambda_t < \lambda_{t+1}$. Then of the two positive terms on the right hand side of the equality sign the second one is larger then the first one, since the sum is larger by the swapping argument as in the proof of Lemma 4.1 (inequality (7)) and $\lambda_t < \lambda_{t+1}$. This means that twice the second term minus the third term is positive, that is

$$2\lambda_t\lambda_{t+1}^k \sum_{i\le t-1} \lambda_i^k \ge \left[(\lambda_t+\lambda_{t+1})^k - \lambda_t^k - \lambda_{t+1}^k\right]\left[\bigoplus_{1\ldots(t-1)}(1,2) + \bigoplus_{(t+2)\ldots s}(2,3)\right]$$

If we ignore the second $\bigoplus$ in the square brackets and use $(\lambda_t+\lambda_{t+1})^k - \lambda_t^k - \lambda_{t+1}^k \ge k\lambda_t\lambda_{t+1}^{k-1}$ this is seen to imply that

$$2\lambda_t\lambda_{t+1}^k \sum_{i\le t-1} \lambda_i^k \ge k\lambda_t\lambda_{t+1}^{k-1} \bigoplus_{1\ldots(t-1)}(1,2).$$

Divide both sides by $\lambda_t\lambda_{t+1}^{k-1}$ and use the definition of $\bigoplus$:

$$2\lambda_{t+1} \sum_{i\le t-1} \lambda_i^k \ge k \sum_{1\le i<j\le t-1} \lambda_i^k\lambda_j.$$

Then we use that every λ_j occurring on the right hand side is at least as large as λ_{t-1} (by monotony, since $j \le t-1$) and hence at least λ_{t+1}, since $\lambda_{t-1} \ge \lambda_{t+1}$ by assumption. Therefore the previous inequality implies that

$$2\sum_{1\le i\le t-1} \lambda_i^k \ge k \sum_{1\le i<t-1} (t-i-1)\lambda_i^k \ge k(t-2)\lambda_1^k.$$

Since $\lambda_1/k \ge \lambda_2 \ge \ldots \ge \lambda_{t-1}$ this implies that

$$2(1+(t-2)/k^k)\lambda_1^k \ge k(t-2)\lambda_1^k.$$

Since $k \ge 3$ we see that this inequality cannot hold unless $t \le 2$. Since $t = 1$ is ruled out by the inequality $\lambda_1 \ge k\lambda_2$ we get $t = 2$.

Let us next establish the conclusion $t = 2$ in the case $\lambda_t = \lambda_{t+1}$, as well. If the second sum on the right hand side of the equality sign in (7) is larger than the first one then the argument goes exactly as above. Otherwise we relabel the λ_i so that $\lambda_i^{\text{new}} = \lambda_{s-i}$ and take $s-(t+1)$ as new t. This swaps the sums, so the second sum is larger for the new λ's, and we get $t = 2$ as before. Note that the relabeling is permissible since the function q_s is symmetric.

We have now shown that the second layer is minimal. It might seem that we could now argue by symmetry that the second to last layer is also minimal, after which it would be an easy task to finish the proof. However, the assumption $\lambda_{t-1} \ge \lambda_{t+1}$ that we made earlier, and that was central in the above argument, breaks the symmetry, and therefore we still have quite a bit of work left.

With the additional information $t = 2$ we return to (8), which now becomes:

$$\lambda_2^k\lambda_3 \sum_{i\ge 4} \lambda_i^k + \lambda_1^k\lambda_2\lambda_3^k \ge \left[(\lambda_2+\lambda_3)^k - \lambda_2^k - \lambda_3^k\right] \bigoplus_{4\ldots s}(2,3). \tag{9}$$

If $\lambda_2 < \lambda_3$ then the swapping argument at layers 2 and 3 (that is, (7) with $i = 2$) implies that

$$\lambda_1^k \geq \sum_{i \geq 4} \lambda_i^k, \tag{10}$$

hence in particular $\lambda_s \leq \lambda_1$. If $\lambda_2 = \lambda_3$ and (10) does not hold then we derive that $s = 4$ from (9) in the same way as we got $t = 2$ before. Therefore we may assume that (10) holds for $\lambda_2 = \lambda_3$, too. Then (9) and (10) together imply that

$$\begin{aligned}
\lambda_1^k\big(\lambda_2^k\lambda_3 + \lambda_2\lambda_3^k\big) &\geq \lambda_2^k\lambda_3 \sum_{i \geq 4} \lambda_i^k + \lambda_1^k\lambda_2\lambda_3^k \\
&\geq \big[(\lambda_2+\lambda_3)^k - \lambda_2^k - \lambda_3^k\big] \bigoplus_{4\ldots s}(2,3) \\
&\geq k\big[\lambda_2^{k-1}\lambda_3 + \lambda_2\lambda_3^{k-1}\big](\lambda_4 + \ldots + \lambda_{s-1})\lambda_s^k \\
&\geq k\big[\lambda_2^k\lambda_3 + \lambda_2\lambda_3^k\big](s-4)\lambda_s^k.
\end{aligned}$$

It follows from this inequality that

$$\lambda_1^k \geq k(s-4)\lambda_s^k. \tag{11}$$

Consider what happens if we merge the last two layers. As usual it follows from the inequality of the q_s and q_{s-1} functions that

$$\lambda_s^k\lambda_{s-1} \sum_{j \leq s-2} \lambda_j^k \geq \big[(\lambda_s + \lambda_{s-1})^k - \lambda_s^k - \lambda_{s-1}^k\big] \sum_{1 \leq i < j \leq s-2} \lambda_i^k \lambda_j. \tag{12}$$

Let us use (10) in the sum on the left hand side for the terms with $4 \leq j \leq s-2$ as well as use $(\lambda_s + \lambda_{s-1})^k - \lambda_s^k - \lambda_{s-1}^k \geq k\lambda_s^{k-1}\lambda_{s-1}$ and discard all the terms not involving λ_1 from the right hand side. This gives the inequality

$$\lambda_s^k\lambda_{s-1}(2\lambda_1^k + \lambda_2^k + \lambda_3^k) \geq k\lambda_s^{k-1}\lambda_{s-1}\lambda_1^k \sum_{1 < j \leq s-2} \lambda_j.$$

From this we continue by using $\sum_{1<j\leq s-2} \lambda_j = 1 - \lambda_1 - \lambda_s - \lambda_{s-1}$ together with $\lambda_2^k + \lambda_3^k \leq 2\lambda_s^k k^{-k} \leq 2\lambda_1^k k^{-k}$, which follows from Lemma 4.1 and $\lambda_s \leq \lambda_1$. This gives us

$$\lambda_s(2 + 2k^{-k})\lambda_1^k \geq k\lambda_1^k(1 - \lambda_1 - \lambda_s - \lambda_{s-1}). \tag{13}$$

Using $\lambda_{s-1} \leq \lambda_s/k$ and dividing by λ_1^k gives

$$\lambda_s(k + 3 + 2k^{-k}) \geq k(1 - \lambda_1). \tag{14}$$

Using this in (11) and taking k^{th}-roots gives

$$\lambda_1 \geq k^{1/k} \frac{k}{k + 3 + 2k^{-k}}(1 - \lambda_1)$$

(we don't need the term $s-4$ now, and since it is larger than 1 it was omitted). Solving for λ_1 gives

$$\lambda_1 \geq \frac{k^{1+1/k}}{k+3+2k^{-k}+k^{1+1/k}}. \tag{15}$$

Next we derive an upper bound for λ_1. Consider $g(a) := q_s(a\lambda_1, b\lambda_2, \ldots, b\lambda_s)$, where $b := (1-a\lambda_1)/(1-\lambda_1)$. Since (λ_j) is an optimal partition g has a maximum at $a=1$. Let us differentiate g:

$$g'(1) = \left(k-(k+1)\frac{\lambda_1}{1-\lambda_1}\right)\lambda_1^k \bigoplus_{2\ldots s}(2,3) - (2k+1)\frac{\lambda_1}{1-\lambda_1}\bigoplus_{2\ldots s}(1,3) = 0.$$

It directly follows that $k-(k+1)\frac{\lambda_1}{1-\lambda_1} \geq 0$, which is to say that $\lambda_1 \leq k/(2k+1)$. Combining this with (15) implies that

$$\frac{k}{2k+1} \geq \frac{k^{1+1/k}}{k+3+2k^{-k}+k^{1+1/k}}.$$

Multiplying by the denominators and dividing by k leads to the inequality

$$k+3+2k^{-k} \geq (k+1)k^{1/k}.$$

Dividing by $k+1$ and raising to the k^{th} power gives

$$k \leq \left(1+\frac{2}{k+1}+\frac{2}{k^k(k+1)}\right)^k \leq \left(1+\frac{2}{k}\right)^k. \tag{16}$$

It is easily seen that $(1+2/k)^k$ is increasing in k and approaches $e^2 < 8$ as $k \to \infty$, hence the inequality does not hold for $k \geq 8$. We also get contradictions for the cases $k = 4, \ldots, 7$ when we check the first inequality by calculator. These contradictions show that the assumption $s \geq 5$ was false for $k \geq 4$, which implies that $s = 4$, since $s \geq r+1 = 4$ by what was proved at the beginning of this section.

For $k = 3$ inequality (16) holds, and we need to be a bit sneakier to get our contradiction. Let us take a second look at inequality (12):

$$\lambda_s^3\lambda_{s-1}\sum_{j\leq s-2}\lambda_j^3 \geq 3\left[\lambda_s^2\lambda_{s-1}+\lambda_s\lambda_{s-1}^2\right]\sum_{1\leq i<j\leq s-2}\lambda_i^3\lambda_j.$$

As before this implies that

$$\lambda_s\sum_{j\leq s-2}\lambda_j^3 \geq 3\lambda_1^3(1-\lambda_1-\lambda_{s-1}-\lambda_s).$$

Let us use a different estimate for the sum on the right hand side for $s=5$, specifically, we use $\lambda_i \leq \lambda_1/3$ for $i = 2, 3$, which gives

$$\lambda_5(1+2/27) \geq 3(1-\lambda_1-\lambda_4-\lambda_5).$$

Since $\lambda_4 \leq \lambda_5/3$ by Lemma 4.1 we find that $\lambda_5(5+2/27) \geq 3(1-\lambda_1)$. Using this in (11) gives

$$\lambda_1 \geq 3^{4/3}(1-\lambda_1)/(5+2/27).$$

Solving λ_1 gives $\lambda_1 \geq 3^{4/3}/(5+2/27+3^{4/3})$. As before $\lambda_1 \leq k/(2k+1) = 3/7$ which contradicts the previous inequality. For larger s we use (14) to estimate λ_s, and using this estimate in (11) gives

$$\lambda_1 \geq \frac{81}{164}(3(s-4))^{1/3}(1-\lambda_1).$$

Solving for λ_1 and using the usual upper bound gives

$$\frac{3}{7} \geq \frac{81(3(s-4))^{1/3}}{164+81(3(s-4))^{1/3}},$$

hence $492 \geq 324(3(s-4))^{1/3}$, which does not hold for $s \geq 6$, a contradiction which shows that $s = 4$ in the case $k = 3$ also. □

Remark 4.2. Notice how the case $k = 3$ in the previous proof is more complicated than the cases with larger k. This corresponds to the intuitive principle that similar size layers in π usually go together with more layers in the π-maximal permutation. Indeed, for $k = 2$ the situation becomes even more difficult: the initial conclusion that $t = 2$ is not easily achieved (along the above line of reasoning), and second, even using much sharper estimates than in the case $k = 3$ all that seems to come out is that $s \leq 5$.

Let us next see how the above result regarding the number of layers in the optimal partition can be used to estimate easily the packing density of $[k, 1, k]$ for $k \geq 3$.

Theorem 4.3. *For $k \geq 3$ denote by π_k the permutation of type $[k, 1, k]$. We have*

$$\binom{2k+1}{k,k,1}\frac{k^{2k}+(k/2)^k}{(2k+1)^{2k+1}} \leq d(\pi_k) \leq \binom{2k+1}{k,k,1}\frac{k^{2k}+2k^k}{(2k+1)^{2k+1}}.$$

Proof. For the lower bound choose $\lambda_1 = \lambda_4 = k/(2k+1)$ and $\lambda_2 = \lambda_3 = 1/(4k+2)$.

In the expression

$$q_s(\lambda_1, \lambda_2, \lambda_3, \lambda_4) = \lambda_1^k(\lambda_2+\lambda_3)\lambda_4^k + \lambda_1^k\lambda_2\lambda_3^k + \lambda_2^k\lambda_3\lambda_4^k$$

we see (as in the proof of Lemma 3.1) that the first term is bounded from above by $k^{2k}/((2k+1)^{2k+1})$. For the second and third terms we use the bounds from the proof of the previous theorem: $\lambda_1 \leq k/(2k+1)$ and $\lambda_2 \leq \lambda_1/k \leq 1/(2k+1)$ and similarly $\lambda_4 \leq k/(2k+1)$ and $\lambda_3 \leq 1/(2k+1)$. Since $\lambda_1^k\lambda_2\lambda_3^k + \lambda_2^k\lambda_3\lambda_4^k$ is increasing in each of the λ's we see that it is bounded from above by $2k^k/(2k+1)^{2k+1}$ (then $\lambda_1+\lambda_2+\lambda_3+\lambda_4$ no longer equals 1, but this is OK for the estimate). □

k	Upper	Lower	Rel. error
3	0.124502	0.133108	$6.912 \cdot 10^{-2}$
4	0.106597	0.107403	$7.567 \cdot 10^{-3}$
5	0.094881	0.094941	$6.300 \cdot 10^{-4}$
6	0.086331	0.086335	$4.253 \cdot 10^{-5}$

Table 1: Estimates of the packing density of the permutation of type $[k, 1, k]$

The relative error in the previous theorem is of magnitude $O(k^{-k})$ as $k \to \infty$, which means that the estimate is very good. The bounds for some small values k are given in Table 1. One can also note that $d(\pi_k) \sim k^{-1/2}/\sqrt{\pi}e$ as $k \to \infty$, which follows from the well known Stirling formula.

The next proposition answers in the negative the question in [1, p. 13] about whether the p_s (or q_s) function has a unique maximum for sets of a single permutation. Whether or not the optimal permutation is symmetric in the case $k = 3$ could not be established, although it seems reasonable to expect that it is.

Proposition 4.4. *For $k \geq 4$ the optimal partition of the permutation of type $[k, 1, k]$ is not symmetric.*

Proof. If $\lambda_1 = \lambda_4$ then we have

$$q_4(\lambda_1, \lambda_2, \lambda_3, \lambda_1) = \lambda_1^{2k}(1 - 2\lambda_1) + \lambda_1^k \lambda_2 \lambda_3 (\lambda_2^{k-1} + \lambda_3^{k-1}).$$

It is not difficult to see that the maximum of $\lambda_2\lambda_3(\lambda_2^{k-1} + \lambda_3^{k-1})$ (for constant $\lambda_2 + \lambda_3$) is not at $\lambda_2 = \lambda_3$. □

As was pointed out earlier, the case of the permutation of type $[2, 1, 2]$, thought by AAHH&S to be the easiest to deal with, remains unhandled. More generally, it seems (intuitively) that permutations of the type $[k_1, 1, k_2]$ would be almost simple when k_1 and k_2 do not differ greatly from each other (say $|k_1 - k_2| \leq c$ or $1/c \leq k_1/k_2 \leq c$ for some suitable c). If this were so then it is clear that the argument from Theorem 4.3 could be used to derive good estimates of the packing density in this case also. How to go about proving this is quite unclear, though, since the arguments used in this section rely heavily on the fact that $[k, 1, k]$ is symmetric.

We have seen in this paper how the layer-merging technique allows us to bound the number of layers in optimal partitions which gives us the packing density or a good estimate for it. Unfortunately, and perhaps surprisingly, attempts to use this technique to attack [1, Conjecture 2.9] have, to this point, been unsuccessful. The main problem here seems to be that it is difficult to get an appropriate structure lemma so that there is nothing to cling to in the merging estimates.

Apart from Conjecture 2.9 it seems that also its converse remains an interesting open problem: namely, if π contains two consecutive layers of size 1 is π then of the unbounded type? This also appears to be the appropriate place to recall that there is also the whole world of the packing density of non-layered permutations to explore, although the reader is referred back to [1] for more details on this, since the present paper has had nothing new to offer in this respect.

Acknowledgement. I would like to thank the referee for several useful comments and suggestions.

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** (2002), #R5.

[2] M. Bóna, Exact enumeration of 1342-avoiding permutations, A close link with labeled trees and planar maps, *J. Combin. Theory Ser. A* **80** (1997), 257–272.

[3] N. G. de Bruijn, *Asymptotic methods in analysis*, Corrected reprint of the third edition, Dover Publications, Inc., New York, 1981.

[4] A. Price, Packing densities of layered patterns, Ph. D. thesis, University of Pennsylvania, Philadelphia, PA, 1997.

[5] R. Simion & F. W. Schmidt, Restricted permutations, *European J. Combin.* **6** (1985), 383–406.

[6] Z. E. Stankova, Classification of forbidden subsequences of length 4, *European J. Combin.* **17** (1996), no. 5, 501–517.

[7] W. Stromquist, Packing layered posets into posets, Unpublished typescript, 1993.

[8] J. West, Generating trees and the Catalan and Schröder numbers, *Discrete Math.* **146** (1995), 247–262.