AUSTRALASIAN JOURNAL OF COMBINATORICS
Volume **87(3)** (2023), Pages 440–451

Multiple twins in permutations

Andrzej Dudek*

Department of Mathematics
Western Michigan University
Kalamazoo, MI, U.S.A.
andrzej.dudek@wmich.edu

Jarosław Grytczuk†

Faculty of Mathematics and Information Science
Warsaw University of Technology
Warsaw, Poland
j.grytczuk@mini.pw.edu.pl

Andrzej Ruciński‡

Department of Discrete Mathematics
Adam Mickiewicz University
Poznań, Poland
rucinski@amu.edu.pl

Abstract

By an r-tuplet in a permutation we mean a family of r pairwise disjoint subsequences with the same relative order. The length of an r-tuplet is defined as the length of any single subsequence in the family. Let $t^{(r)}(n)$ denote the largest k such that every permutation of length n contains an r-tuplet of length k. We prove that $t^{(r)}(n) = O\left(n^{\frac{r}{2r-1}}\right)$ and $t^{(r)}(n) = \Omega\left(n^{\frac{R}{2R-1}}\right)$, where $R = \binom{2r-1}{r}$. We conjecture that the upper bound brings the correct order of magnitude of $t^{(r)}(n)$ and support this conjecture by proving that it holds for almost all permutations. Our work generalizes previous studies of the case $r = 2$.

* The first author was supported in part by Simons Foundation Grant #522400.
† The second author was supported in part by Narodowe Centrum Nauki, grant 2017/26/D/ST6/00264, as well as, the European Regional Development Fund under the grant No. POIR.01.01.01-00-0124/17-00, on the basis of an agreement between FinAi S.A. and the National Center for Research and Development based in Warsaw.
‡ The third author was supported in part by Narodowe Centrum Nauki, grant 2018/29/B/ST1/00426

ISSN: 2202-3518

1 Introduction

By a *permutation* we mean any finite sequence of *distinct* integers. We say that two permutations $(x_1, \ldots, x_k)$ and $(y_1, \ldots, y_k)$ are *similar* if their entries preserve the same relative order, that is, $x_i < x_j$ if and only if $y_i < y_j$ for all pairs $\{i, j\}$ with $1 \leqslant i < j \leqslant k$. This is, clearly, an equivalence relation. Note that given a permutation $(x_1, \ldots, x_k)$ and a k-element set $\{y_1, \ldots, y_k\}$ of positive integers, there is only one permutation of this set similar to $(x_1, \ldots, x_k)$.

Let $[n] = \{1, 2, \ldots, n\}$ and π be a permutation of $[n]$, called also an *n-permutation*, and let $r \geqslant 2$ be an integer. A family of r pairwise similar and disjoint sub-permutations of π is called an *r-tuplet* and the *length* of an r-tuplet is defined as the number of elements in just one of the sub-permutations. It is common to call 2-tuplets *twins* and 3-tuplets *triplets*.

Let $t^{(r)}(\pi)$ denote the largest integer k such that π contains an r-tuplet of length k. Further, let $t^{(r)}(n)$ denote the minimum of $t^{(r)}(\pi)$ over all permutations π of $[n]$. In other words, $t^{(r)}(n)$ is the largest integer k such that every n-permutation contains an r-tuplet of length k. Our aim is to estimate this function.

By the classical result of Erdős and Szekeres [7] concerning monotone subsequences, we get $t^{(r)}(n) = \Omega(\sqrt{n})$ (all implicit constants throughout the paper are allowed to depend on r). Indeed, in a monotone permutation any r disjoint subsequences of the same length yield an r-tuplet. For $r = 2$, using a probabilistic argument, Gawron [9] proved that $t^{(2)}(n) = O(n^{2/3})$. He conjectured that this bound is tight, namely, that we also have $t^{(2)}(n) = \Omega(n^{2/3})$. A currently best result towards this conjecture was obtained by Bukh and Rudenko [3] (see also [6]) who showed that $t^{(2)}(n) = \Omega(n^{3/5})$.

In the first part of this paper, we generalize both these bounds to arbitrary $r \geqslant 2$, adopting the ideas from [3, 6, 9].

Theorem 1.1. *For every $r \geqslant 2$, with $R = \binom{2r-1}{r}$, we have*

$$t^{(r)}(n) = \Omega\left(n^{\frac{R}{2R-1}}\right) \quad \text{and} \quad t^{(r)}(n) = O\left(n^{\frac{r}{2r-1}}\right).$$

Let Π_n be a (uniformly) random permutation of $[n]$ and let $t^{(r)}(\Pi_n)$ be the corresponding random variable equal to the maximum length of an r-tuplet in Π_n. We say that a property of a random object holds *asymptotically almost surely* (a.a.s. for short) if it holds with probability tending to one as the size of the object grows to infinity. In [6] and [3] it was shown that a.a.s., $t^{(2)}(\Pi_n) = \Theta(n^{2/3})$. Here we generalize this result.

Theorem 1.2. *For every $r \geqslant 2$, a.a.s.,*

$$t^{(r)}(\Pi_n) = \Theta\left(n^{\frac{r}{2r-1}}\right).$$

In view of Theorem 1.2, for almost all permutations π the parameter $t^{(r)}(\pi)$ reaches the upper bound from Theorem 1.1.

In the following two sections we give proofs of Theorems 1.1 and 1.2, while the last section contains some related open problems.

2 Proof of Theorem 1.1

The upper bound in Theorem 1.1 follows from the upper bound in Theorem 1.2. Therefore, here we focus exclusively on the proof of the lower bound.

2.1 Preparations

Beame and Huynh-Ngoc [2] proved that amongst any three permutations of $[m]$ there are two with the same sub-permutation of length at least $m^{1/3}$. This simple fact was used in [3] to prove the lower bound in Theorem 1.1 for $r = 2$. Here, we will need the following extension proved by Bukh and Zhou [4] in the context of estimating the length of r-tuplets in words over finite alphabets. Recall that $R = \binom{2r-1}{r}$.

Lemma 2.1 (Theorem 24 in [4]). *For every $r \geqslant 2$, among any $2r-1$ permutations of $[m]$, there are r permutations with the same sub-permutation of length at least $m^{1/R}$.*

The proof of Lemma 2.1 is quite elementary and uses the idea from the classical proof of the Erdős-Szekeres Theorem. It boils down to assigning to each $i \in [m]$ a vector of length R, where the entries represent, for all r-tuples of permutations, the lengths of the longest common sub-permutations that begin at i. Then it suffices to show that this mapping is an injection.

The main technique behind the proof of the lower bound on $t^{(r)}(n)$ is that of *concatenation* of r-tuplets. Here we reveal sufficient conditions under which two r-tuplets can be merged into one. Let $(x_1^{(j)}, \ldots, x_s^{(j)})$, $j = 1, \ldots, r$, and $(y_1^{(j)}, \ldots, y_t^{(j)})$, $j = 1, \ldots, r$, be two disjoint r-tuplets in a permutation π. Their full concatenation, resulting in an r-tuplet of length $s + t$, is possible if

- the rightmost element of the first r-tuplet is to the left of the leftmost element of the second r-tuplet, and
- for each $1 \leqslant i \leqslant s$ and $1 \leqslant i' \leqslant t$,

$$\text{either} \qquad \max_{1 \leqslant j \leqslant r} y_{i'}^{(j)} < \min_{1 \leqslant j \leqslant r} x_i^{(j)} \qquad \text{or} \qquad \min_{1 \leqslant j \leqslant r} y_{i'}^{(j)} > \max_{1 \leqslant j \leqslant r} x_i^{(j)}.$$

Example 2.2. Let $r = 2$, $n = 30$, and

$$\pi = (\underline{26}, 16, 28, 29, \underline{10}, 5, \overline{\underline{24}}, \underline{27}, 1, 22, 11, \overline{8}, 2, 23, 15,$$
$$19, \overline{25}, 21, 20, \underline{\underline{13}}, 9, 30, \underline{\underline{17}}, \overline{\overline{12}}, \overline{\overline{18}}, 7, \underline{\underline{3}}, 14, \overline{\overline{4}}, 6)$$

be a permutation of [30]. Here we marked by blue (single underline) and pink (single overline) the first pair of twins of length 3 (similar to $(2,1,3)$), and by indigo (double underline) and red (double overline) the second one (similar to $(2,3,1)$). Both above conditions hold. Indeed, elements of the first twins are to the left of 13, the leftmost element of the second twins. Moreover, all nine required inequalities hold, e.g., for $i = 2$, $i' = 1$, we have $\min\{13,12\} = 12 > \max\{10,8\} = 10$. Thus, their concatenation forms a (bluish-reddish) pair of twins of length six, namely: $(26,10,27,13,17,3)$ and $(24,8,25,12,18,4)$, both similar to $(5,2,6,3,4,1)$.

If we only knew the first twins and were after the second ones, we could, obviously, help ourselves by searching only to the right of 25, and by eliminating from our search all elements whose values are "squeezed" between the given twins, that is, lie between $\min_{1\leqslant j\leqslant r} x_i^{(j)}$ and $\max_{1\leqslant j\leqslant r} x_i^{(j)}$ for any i. (In our example this step excludes only 9 ($i = 2$), as 25 ($i = 1$) and 26 ($i = 3$) are already excluded.) But most importantly, the second pair of twins should be "narrow" in a sense that would guarantee the second bulleted condition above. (In our example, we chose consecutive values: $13, 12$ and $17, 18$ and $3, 4$.)

To facilitate the idea mentioned in the above example, we introduce the notion of narrow r-tuplets as follows. The *width* of a set of integers A is defined as $\max A - \min A$. For a positive integer w we call an r-tuplet $(x_1^{(1)},\dots,x_k^{(1)}),\dots,(x_1^{(r)},\dots,x_k^{(r)})$ *w-narrow* if for all $i = 1,\dots,k$, the sets $\{x_i^{(j)} : 1 \leqslant j \leqslant r\}$ have width at most w.

Throughout the paper we will sometimes use the notation $[a,b] = \{a,\dots,b\}$, where $a < b$ are integers.

2.2 Lower bound

The proof of the lower bound is similar to that for $r = 2$ in [3, Proof of Theorem 1]. The main idea is to utilize Lemma 2.1 in a clever way.

Set

$$M = n^{(R-1)/(2R-1)} \quad \text{and} \quad N = n/M = n^{R/(2R-1)}.$$

Our goal is, for a given permutation π of $[n]$, to gradually pick from it $O(M)$-narrow r-tuplets of length $\Omega(N/M)$, and remove them together with all "in-between" elements (plus some more), allowing concatenation of obtained pieces into one r-tuplet of length $\Omega(N)$. To achieve this goal, we need to iterate this procedure $\Omega(M)$ times, meaning that the number of elements discarded each time should not exceed $O(N)$. Now come the details.

Let $A \subset \binom{[n]}{(2r-1)N}$, $A = \{a_1 < a_2 \cdots < a_{(2r-1)N}\}$, and let π be a permutation of A. For any $a \in A$ we denote by $\pi^{-1}(a)$ the position of a in π. E.g., if $A = \{1,3,4,6\}$ and $\pi = (3,1,6,4)$, we have $\pi^{-1}(1) = 2$, $\pi^{-1}(3) = 1$, $\pi^{-1}(4) = 4$ and $\pi^{-1}(6) = 3$.

We split the set A into N blocks of $2r-1$ consecutive elements:

$$A = A_0 \cup \cdots \cup A_{N-1}, \quad \text{where} \quad A_i = \{a_{i(2r-1)+1},\dots,a_{(i+1)(2r-1)}\}, \quad i = 0,\dots,N-1.$$

A crucial observation is that if, for some $2 \leqslant k \leqslant N$ and $1 \leqslant j_1 < j_2 \leqslant 2r-1$, there exists a sequence of *distinct* indices $(i_1, \ldots, i_k) \in \{0, \ldots, N-1\}^k$ such that

$$\pi^{-1}(a_{i_1(2r-1)+j_1}) < \cdots < \pi^{-1}(a_{i_k(2r-1)+j_1})$$

and

$$\pi^{-1}(a_{i_1(2r-1)+j_2}) < \cdots < \pi^{-1}(a_{i_k(2r-1)+j_2}),$$

then sub-permutations $(a_{i_1(2r-1)+j_1}, \ldots, a_{i_k(2r-1)+j_1})$ and $(a_{i_1(2r-1)+j_2}, \ldots, a_{i_k(2r-1)+j_2})$ of π form twins of length k. By the same token, if all the above is true for r indices $1 \leqslant j_1 < j_2 < \cdots < j_r \leqslant 2r-1$, then we obtain an r-tuplet of length k in π.

Example 2.3. Let $A = [30]$, $r = 3$, $N = 6$, and, the same as before,

$$\pi = (26, 16, 28, 29, 10, 5, 24, 27, 1, 22, 11, 8, 2, 23, 15,\\ 19, 25, 21, 20, 13, 9, 30, 17, 12, 18, 7, 3, 14, 4, 6).$$

Here $A_0 = [1, 5]$, $A_1 = [6, 10]$, $A_2 = [11, 15]$, $A_3 = [16, 20]$, $A_4 = [21, 25]$ and $A_5 = [26, 30]$. By inspection, one can see that for $k = 3$, $i_1 = 5$, $i_2 = 3$, $i_3 = 0$, and $j_1 = 1$, $j_2 = 3$, $j_3 = 4$, the above conditions are satisfied. Indeed, the first element of A_5 (26 in π) is to the left of the first element of A_3 (16) which, in turn, is to the left of the first element of A_0 (1), and similar order holds for the third elements of A_5, A_3, A_0 (28, 18, 3), as well as, for the fourth elements (29, 19, 4). This reveals a triplet similar to $(3, 2, 1)$ and indicated by colors blue (underline), pink (overline), and red (asterisk) below:

$$\pi = (\underline{26}, \underline{16}, \overline{28}, 29^*, 10, 5, 24, 27, \underline{1}, 22, 11, 8, 2, 23, 15,\\ 19^*, 25, 21, 20, 13, 9, 30, 17, 12, \overline{18}, 7, \overline{3}, 14, 4^*, 6).$$

An equivalent but somewhat easier way to see what is going on here is to define five disjoint sub-permutations $\pi^{(j)}$ of length 6, by including in $\pi^{(j)}$ all elements of π of the form $a_{i(2r-1)+j}$ (in the order they appear in π): $\pi^{(1)} = (26, 16, 1, 11, 21, 6)$, $\pi^{(2)} = (27, 22, 2, 17, 12, 7)$, ..., etc. Next replace them by similar permutations of $\{0, 1, 2, 3, 4, 5\}$: $(5, 3, 0, 2, 4, 1)$, $(5, 4, 0, 3, 2, 1)$, ..., etc. Now, what we are after, are long common sub-permutations in at least three of these permutations. We found $(5, 3, 0)$ as common in the 1st, 3rd, and 4th permutation, but one could have also picked $(5, 3, 1)$ in the 1st, 2nd, and 4th (and there is at least one more alternative). By inspection, one may realize, however, that there are no longer common sub-permutations in at least three of these five permutations.

In the above example, we dealt with consecutive integers in A, so the width of each subset A_i was the same (and equal to 4), but in general this may not be the case. To make sure that the r-tuplets generated by the above approach are $O(M)$-narrow, we focus only on the subsets A_i of width at most, say, $2M$. Clearly, as $A \subset [n]$, at least $N/2$ of these sets have, indeed, the width at most $2n/N = 2M$. Let

$$I = \{0 \leqslant i \leqslant N-1 : \ A_i \text{ has width at most } 2M\}.$$

So, $|I| \geqslant N/2$. Next, define $2r-1$ disjoint sub-permutations $\pi^{(j)}$ of π, $j = 1, \ldots, 2r-1$, of length $|I|$, by including in $\pi^{(j)}$ all elements of π of the form $a_{i(2r-1)+j}$, where $i \in I$ (in the order they appear in π). To place them on a common ground, we replace each element $a_{i(2r-1)+j}$ by i, obtaining a set of new permutations $\bar{\pi}^{(j)}$, $j = 1, \ldots, 2r-1$, of the same set I.

We apply Lemma 2.1 to permutations $\bar{\pi}^{(j)}$, $j = 1, \ldots, 2r-1$, and, as a result, find r of them with a common sub-permutation of length at least

$$|I|^{1/R} \geqslant (N/2)^{1/R} = 2^{-1/R} n^{1/(2R-1)} \geqslant 2^{-1/3} \frac{N}{M},$$

where the latter inequality follows from $R = \binom{2r-1}{r} \geqslant \binom{3}{2} = 3$. Let $\bar{\pi}^{(j_1)}, \ldots, \bar{\pi}^{(j_r)}$, where $1 \leqslant j_1 < \cdots < j_r \leqslant 2r-1$, be these r permutations and let $I_0 \subset I \subset \{0, \ldots, N-1\}$ be the set of the elements of a common sub-permutation in $\bar{\pi}^{(j_1)}, \ldots, \bar{\pi}^{(j_r)}$ of length $k = |I_0| = \lceil 2^{-1/3} \frac{N}{M} \rceil$. Upon returning to permutations $\pi^{(j_1)}, \ldots, \pi^{(j_r)}$, it can be seen, guided by the above example, that I_0 generates in their union, and thus in π, a $2M$-narrow r-tuplet of length k.

It remains to incorporate the above procedure into an iteration loop and make sure that the obtained r-tuplets can be concatenated into a long one. Let π be a permutation of $[n]$. In the first step of the procedure, take A^1 to be the set of the first $(2r-1)N$ elements of π and let π^1 be the sub-permutation of π consisting of the first $|A^1|$ elements. By the above described argument we find a $2M$-narrow r-tuplet $(x_1^{(j)}, \ldots, x_k^{(j)})$, $j = 1, \ldots, r$ of length k in π^1. We then remove from $[n]$ the set

$$A^1 \cup \bigcup_{i=1}^{k} [\min_{1 \leqslant j \leqslant r} x_i^{(j)}, \min_{1 \leqslant j \leqslant r} x_i^{(j)} + 2M - 1].$$

Note that this set has size at most $(2r-1)N + k(2M) \leqslant (2r+1)N$.

Now, we consider the truncation of π to the remaining elements and the set A^2 of its leftmost $(2r-1)N$ elements. Repeating the above procedure mutatis mutandis, we obtain another $2M$-narrow r-tuplet $(y_1^{(1)}, \ldots, y_k^{(r)})$, which, owing to the truncation, satisfies the conditions for proper concatenation with $(x_1^{(j)}, \ldots, x_k^{(j)})$, $j = 1, \ldots, r$, spelled out in Subsection 2.1. This procedure can be continued as long as there are at least $(2r-1)N$ elements left. Thus, it can be repeated at least $\frac{n}{(2r+1)N} = \frac{M}{2r+1}$ times, yielding at the conclusion an r-tuplet in π of length a least

$$\frac{M}{2r+1} \times k \geqslant \frac{M}{2r+1} \times 2^{-1/3} \frac{N}{M} \geqslant \frac{N}{3r+1}.$$

(For $r \geqslant 3$, the last bound can be improved to $\frac{N}{3r}$.)

3 Proof of Theorem 1.2

3.1 Upper bound

For the upper bound we use the first moment method. Let $\Pi = \Pi_n$ be a random permutation chosen uniformly from the set of all $n!$ permutations of $[n]$. Let k be

a fixed positive integer and let X_k be a random variable counting all r-tuplets of length k in Π. Furthermore, for a family of r pairwise disjoint subsets $A_1, \ldots, A_r$ of $[n]$, each of length k, let $X_{A_1,\ldots,A_r}$ be an indicator random variable equal to 1 if there is an r-tuplet in Π on positions determined by the subsets $A_1, \ldots, A_r$. So, $X_k = \sum_{A_1,\ldots,A_r} X_{A_1,\ldots,A_r}$ and by the linearity of expectation

$$\mathbb{E}X_k = \sum_{A_1,\ldots,A_r} \mathbb{E}X_{A_1,\ldots,A_r} = \sum_{A_1,\ldots,A_r} \mathbb{P}(X_{A_1,\ldots,A_r} = 1).$$

Since

$$\mathbb{P}(X_{A_1,\ldots,A_r} = 1) = \frac{\binom{n}{k}\binom{n-k}{k} \cdots \binom{n-(r-2)k}{k} \cdot (n-(r-1)k)! \cdot 1}{n!} = \frac{1}{k!^{r-1}}$$

and the number of unordered r-tuples $\{A_1, \ldots, A_r\}$ in $[n]$ is

$$\binom{n}{rk} \frac{(rk)!}{k!^r r!} = \frac{n!}{k!^r r!(n-rk)!},$$

it follows that

$$\mathbb{E}X_k = \frac{n!}{k!^r r!(n-rk)!} \times \frac{1}{k!^{r-1}} = \frac{n(n-1)\ldots(n-rk+1)}{r!(k!)^{2r-1}}.$$

Using the inequality $k! > \frac{k^k}{e^k}$ we obtain

$$\mathbb{E}X_k < \frac{n^{rk} e^{(2r-1)k}}{r! k^{(2r-1)k}} < \left(\frac{n^r e^{2r-1}}{k^{2r-1}}\right)^k.$$

It follows that for $k \geqslant 2en^{r/(2r-1)}$,

$$\mathbb{P}(X_k \geqslant 1) \leqslant \mathbb{E}X_k < 2^{-(2r-1)k} \to 0$$

as $n \to \infty$. This completes the proof of the upper bound.

3.2 Lower bound

The proof of the lower bound presented here is similar to that in [6] (see also [3] for its continuous version). Set

$$a = r!^{1/(2r-1)} n^{(r-1)/(2r-1)}, \tag{3.1}$$

assume for convenience that a is an integer and divides n, and partition $[n]$ into n/a consecutive blocks of equal size, that is, set

$$[n] = A_1 \cup \cdots \cup A_{n/a},$$

where $A_1 = \{1, \ldots, a\}$, $A_2 = \{a+1, \ldots, 2a\}$, etc.

For fixed $1 \leqslant i, j \leqslant n/a$, let $X = X_{ij}$ be the number of elements from the set A_j which Π puts on the positions belonging to the set A_i. We construct an auxiliary $n/a \times n/a$ bipartite graph B with vertex classes $U = \{1, \ldots, n/a\}$ and $V = \{1, \ldots, n/a\}$, where $ij \in B$ if, and only if, $X_{ij} \geqslant r$.

Let $M = \{i_1 j_1, \ldots, i_m j_m\}$, $i_1 < \cdots < i_m$, be a matching in B of size $|M| = m$. For every $ij \in M$, let $s_1^{(i)}, s_2^{(i)}, \ldots, s_r^{(i)}$ be some r elements of A_i such that $\Pi(s_1^{(i)}), \ldots, \Pi(s_r^{(i)}) \in A_j$. Then, sub-permutations

$$(\Pi(s_1^{(i_1)}), \ldots, \Pi(s_1^{(i_m)})),\ (\Pi(s_2^{(i_1)}), \ldots, \Pi(s_2^{(i_m)})), \ldots,\ (\Pi(s_r^{(i_1)}), \ldots, \Pi(s_r^{(i_m)}))$$

form an r-tuplet. Indeed, if, say, $\Pi(s_1^{(i_1)}) < \Pi(s_1^{(i_2)})$, then $j_1 < j_2$, and so, $\Pi(s_t^{(i_1)}) < \Pi(s_t^{(i_2)})$ for each $t = 2, \ldots, r$, as $\Pi(s_t^{(i_1)}) \in A_{j_1}$, while $\Pi(s_t^{(i_2)}) \in A_{j_2}$.

Hence, it remains to show that a.a.s. there is a matching in B of size $\Omega\left(n^{r/(2r-1)}\right)$. To this end, we are going to use the obvious fact, coming from the greedy algorithm, that in every graph G there is a matching of size at least $|E(G)|/(2\Delta(G))$, where $\Delta(G)$ is the maximum vertex degree in G.

In fact, we apply this bound to a suitably chosen subgraph B' of B. Let $\nu(B)$ be the size of a largest matching in B. Note that $\Delta(B) \leqslant \Delta_0 := \lfloor a/r \rfloor$. Further, let Z_k be the number of vertices of degree k in B, $k = 0, \ldots, \Delta_0$. Then, for the subgraph B' of B obtained by deleting all vertices of degree at least 7, we get

$$|E(B')| \geqslant |E(B)| - \sum_{k=7}^{\Delta_0} k Z_k$$

and

$$\nu(B) \geqslant \nu(B') \geqslant \frac{|E(B')|}{2\Delta(B')} \geqslant \frac{|E(B)| - \sum_{k=7}^{\Delta_0} k Z_k}{12}. \tag{3.2}$$

Further, for each $i \in U \cup V$, let

$$Y_i = \sum_{j=1}^{n/a} \mathbb{I}(X_{ij} \geqslant r)$$

be the degree of vertex i in B. Then,

$$|E(B)| = \frac{1}{2} \sum_{i \in U} \mathbb{E} Y_i$$

and

$$Z_k = \sum_{i \in U \cup V} \mathbb{I}(Y_i = k),$$

and so

$$\mathbb{E}[|E(B)|] = \frac{n}{a} \mathbb{E} Y_1$$

and

$$\mathbb{E} Z_k = \frac{2n}{a} \mathbb{P}(Y_1 = k).$$

Taking the expectation on the outer sides of (3.2), we arrive at

$$\mathbb{E}[\nu(B)] \geqslant \frac{n}{12a}\left(\mathbb{E}Y_1 - 2\sum_{k=7}^{\Delta_0} k\mathbb{P}(Y_1 = k)\right) \tag{3.3}$$

and it remains to estimate $\mathbb{E}Y_1$ and $\mathbb{P}(Y_1 = k)$.

We have

$$\begin{aligned}\mathbb{E}Y_1 &= \sum_{j=1}^{n/a}\mathbb{P}(X_{1j} \geqslant r) \geqslant \frac{n}{a}\mathbb{P}(X_{11} = r)\\ &= \frac{n}{a}\cdot\frac{\binom{n-a}{a-r}\binom{a}{r}^2 r!(a-r)!(n-a)!}{n!}\\ &\sim \frac{a^{2r-1}}{r!n^{r-1}} \geqslant \frac{1}{2},\end{aligned}$$

for sufficiently large n, as $a^2 = o(n)$ and, by (3.1), $\frac{a^{2r-1}}{n^{r-1}} = r!$.

Now we estimate $\mathbb{P}(Y_1 = k)$ for $k \in \{7, \ldots, \Delta_0\}$. Observe that, by using inequalities $\binom{m}{p} \leqslant \frac{m^p}{p!}$ and $\frac{n}{n-rk} \leqslant \exp\{rk/(n-rk)\}$,

$$\mathbb{P}(Y_1 = k) \leqslant \binom{\frac{n}{a}}{k}\binom{a}{r}^k\binom{a}{rk}\frac{(rk)!(n-rk)!}{n!} \leqslant \frac{a^{2rk-k}e^{(rk)^2/(n-rk)}}{r!^k n^{rk-k}k!}.$$

Consequently, since $(rk)^2 \leqslant a^2 = o(n)$, $k-1 \geqslant 6$, $(k-1)! > \left(\frac{k-1}{3}\right)^{k-1}$, and by (3.1),

$$\begin{aligned}k\mathbb{P}(Y_1 = k) &\leqslant \frac{a^{2rk-k}(1+o(1))}{r!^k n^{rk-k}(k-1)!} \leqslant \frac{2a^{2rk-k}}{r!^k n^{rk-k}((k-1)/3)^k}\\ &\leqslant \frac{2a^{2r-1}}{r!n^{r-1}}\left(\frac{3a^{2r-1}}{r!(k-1)n^{r-1}}\right)^{k-1} \leqslant 4\cdot 2^{-k}.\end{aligned}$$

Thus,

$$\sum_{k=7}^{\Delta_0} k\mathbb{P}(Y_1 = k) \leqslant 4\sum_{k=7}^{\infty} 2^{-k} = \frac{1}{16}.$$

Returning to (3.3), we conclude that

$$\mathbb{E}[\nu(B)] \geqslant \frac{n}{12a}\left(\frac{1}{2} - 2\cdot\frac{1}{16}\right) = \frac{1}{32}\cdot\frac{n}{a} = \Theta\left(n^{r/(2r-1)}\right).$$

Since $t^{(r)}(\Pi_n) \geqslant \nu(B)$, to complete the proof of Theorem 1.2 it remains to show that the random variable $h(\Pi) = \nu(B)$ is highly concentrated about its mean. To this end, we are going to use the Azuma-Hoeffding inequality for random permutations (see, e.g., Lemma 11 in [8] or Section 3.2 in [11]):

Theorem 3.1. *Let $h(\pi)$ be a function of n-permutations such that, for some constant $c > 0$, if permutation π_2 is obtained from permutation π_1 by swapping two elements, then $|h(\pi_1) - h(\pi_2)| \leqslant c$. Then, for every $\eta > 0$,*

$$\mathbb{P}(|h(\Pi_n) - \mathbb{E}[h(\Pi_n)]| \geqslant \eta) \leqslant 2\exp(-\eta^2/(2c^2 n)).$$

To apply it to $h(\Pi) = \nu(B)$, first note that if π_2 is obtained from a permutation π_1 by swapping some two of its elements, then at most two edges of B can be destroyed, and at most two edges can be created, so $|h(\pi_1)-h(\pi_2)| \leqslant 2$. Consequently, Theorem 3.1 applied with $c = 2$ and $\eta = n^{2r/(4r-1)}$ implies that

$$\mathbb{P}\left(|h(\Pi_n) - \mathbb{E}[h(\Pi_n)]| \geqslant n^{2r/(4r-1)}\right) = o(1).$$

As, crucially, $n^{2r/(4r-1)} = o(n^{r/(2r-1)})$, this completes the proof of the lower bound in Theorem 1.2.

4 Concluding Remarks

We conclude the paper with two conjectures and two comments related to the contents of this paper. The first one is a natural generalization of Gawron's conjecture [9] concerning the asymptotic order of the function $t^{(2)}(n)$.

Conjecture 4.1. For every $r \geqslant 2$, we have $t^{(r)}(n) = \Theta(n^{r/(2r-1)})$.

This means that we believe that every permutation contains an r-tuplet of length $\Omega(n^{r/(2r-1)})$. While this statement is wide open for every $r \geqslant 2$, notice that, as r tends to ∞, the difference of the exponents in the upper and lower bounds in Theorem 1.1 tends to 0 (since both converge to 1/2).

Throughout the paper we have kept r fixed, that is, independent of n. However, in the next conjecture we allow $r = r(n)$. We call an r-tuplet of length r an *r-square*. The question is how large a square is contained in every permutation. Let $s(\pi)$ denote the largest integer r such that π contains an r-square and let $s(n)$ be the minimum of $s(\pi)$ over all n-permutations.

Conjecture 4.2. We have $n - (s(n))^2 = o(n)$.

In other words, we believe that every permutation is almost entirely filled by a square. Notice that a repeated application of the Erdős-Szekeres Theorem ([7]) implies that every n-permutation contains an $\Omega(\sqrt{n})$-square. Clearly, the components of this square are monotone. It is, therefore, natural to expect even larger squares with components of an arbitrary pattern.

Next, let us mention r-tuplets with forbidden patterns. For a given permutation τ we say that a permutation π is *τ-free* if no subsequence of π is similar to τ. For $r \geqslant 2$, let $t^{(r)}(n, \tau)$ denote the longest length of a τ-free r-tuplet guaranteed in every n-permutation. In [6] we showed that $t^{(2)}(n, \tau) = \Theta(\sqrt{n})$ for any non-monotone pattern τ. It turns out that, somewhat surprisingly, the same threshold remains valid for any $r \geqslant 2$ (with only the constant hidden in Θ possibly depending on r). We omit the details.

Finally, it should be remarked that there has been an extensive research on twins in other discrete structures, like graphs and words over finite alphabets (see [6]

for more details and references). Here we would like to point to one of the most interesting results in this area. In 2012, Axenovich, Person, and Puzynina proved in [1] that every binary sequence of length n contains r disjoint identical subsequences, each of length $n/r - o(n)$. Equally interesting is their method of proof: they proved and then utilized a special version of the regularity lemma for words, an analog of the celebrated Szemerédi's Regularity Lemma for graphs.

In the context of the results presented in this paper we wonder if a similar tool would help to improve bounds on $t^{(r)}(n)$ and ultimately prove Conjecture 4.1. We are only aware of a regularity lemma for permutations developed by Cooper in [5] (see also [10]), but it seems unsuitable for tackling the problem of twins.

Acknowledgements

We are very grateful to both referees, as well as to the editors of the journal, for their careful reading of the manuscript and several comments which have led to an improvement in clarity and readability of our paper.

References

[1] M. Axenovich, Y. Person and S. Puzynina, A regularity lemma and twins in words, *J. Combin. Theory Ser. A.* 120 (2013), 733–743.

[2] P. Beame and D.-T. Huynh-Ngoc, On the value of multiple read/write streams for approximating frequency moments, *Electronic Colloquium on Computational Complexity (ECCC)*, May 2008; Technical Report TR08-024.

[3] B. Bukh and O. Rudenko, Order-isomorphic twins in permutations, *SIAM J. Discrete Math.* 34 (2020), 1620–1622.

[4] B. Bukh and L. Zhou, Twins in words and long common subsequences in permutations, *Israel J. Math.* 213 (2016), 183–209.

[5] J. N. Cooper, A Permutation Regularity Lemma, *Electron. J. Combin.* 13 (1) (2006), #P1.22.

[6] A. Dudek, J. Grytczuk and A. Ruciński, Variations on twins in permutations, *Electron. J. Combin.* 28 (3) (2021), #P3.19.

[7] P. Erdős and G. Szekeres, A combinatorial problem in geometry, *Compositio Mathematica* 2 (1935), 463–470.

[8] A. Frieze and B. Pittel, Perfect matchings in random graphs with prescribed minimal degree, Mathematics and computer science III, pp. 95–132, Trends Math., Birkhäuser, Basel, 2004.

[9] M. Gawron, Izomorficzne podstruktury w słowach i permutacjach, Master Thesis, 2014.

[10] C. Hoppen, Y. Kohayakawa and R. M. Sampaio, A note on permutation regularity, *Discrete Appl. Math.* 160 (2012), 2716–2727.

[11] C. McDiarmid, Concentration, Probabilistic methods for algorithmic discrete mathematics, pp. 195–248, Algorithms Combin., 16, Springer, Berlin, 1998.

(Received 22 Mar 2023; revised 18 Aug 2023, 13 Sep 2023)