

TWINS IN ORDERED HYPER-MATCHINGS

ANDRZEJ DUDEK¹

Department of Mathematics
Western Michigan University
Kalamazoo, MI, USA

e-mail: andrzej.dudek@wmich.edu

JAROSŁAW GRYTCZUK²

Faculty of Mathematics and Information Science
Warsaw University of Technology
Warsaw, Poland

e-mail: jaroslaw.grytczuk@pw.edu.pl

AND

ANDRZEJ RUCIŃSKI³

Department of Discrete Mathematics
Adam Mickiewicz University
Poznań, Poland

e-mail: rucinski@amu.edu.pl

Abstract

An *ordered r -matching* of size n is an r -uniform hypergraph on a linearly ordered set of vertices, consisting of n pairwise disjoint edges. Two ordered r -matchings are *isomorphic* if there is an order-preserving isomorphism between them. A pair of *twins* in an ordered r -matching is formed by two vertex disjoint isomorphic sub-matchings. Let $t^{(r)}(n)$ denote the maximum size of twins one may find in *every* ordered r -matching of size n .

By relating the problem to that of largest twins in permutations and applying some recent Erdős-Szekeres-type results for ordered matchings, we

¹The first author was supported in part by Simons Foundation Grant #522400.

²The second author was supported in part by Narodowe Centrum Nauki, grant 2020/37/B/ST1/03298.

³The third author was supported in part by Narodowe Centrum Nauki, grant 2018/29/B/ST1/00426.

show that $t^{(r)}(n) = \Omega\left(n^{\frac{3}{5 \cdot (2^{r-1}-1)}}\right)$ for every fixed $r \geq 2$. On the other hand, $t^{(r)}(n) = O\left(n^{\frac{2}{r+1}}\right)$, by a simple probabilistic argument. As our main result, we prove that, for *almost all* ordered r -matchings of size n , the size of the largest twins achieves this bound.

Keywords: hypergraphs, ordered matchings, twins.

2020 Mathematics Subject Classification: 05C65.

1. INTRODUCTION

Let $r \geq 2$ be a fixed integer. An *ordered r -matching* of size n is an r -uniform hypergraph M on a linearly ordered vertex set V , with $|V| = rn$, consisting of n pairwise disjoint edges. One may represent such a matching as a *word* with n distinct letters in which every letter occurs exactly r times. For instance, the word $AABCBDBDACCD$ represents an ordered 3-matching of size 4 whose edges correspond to sets of positions occupied by a given letter, i.e., $A = \{1, 2, 9\}$, $B = \{3, 5, 7\}$, $C = \{4, 10, 11\}$, and $D = \{6, 8, 12\}$. We will frequently identify an ordered matching with its representing word.

Two ordered matchings are *isomorphic* if there is an order-preserving bijection between their vertex sets inducing a one-to-one correspondence between their edge sets. In particular, the words representing isomorphic ordered matchings are identical up to renaming the letters. A pair of two vertex disjoint isomorphic sub-matchings of an ordered matching M is called *twins*. By the *size* of twins we mean the size of just one of the two in the pair. The maximum size of twins in a matching M will be denoted by $t(M)$. The corresponding extremal function is defined by

$$t^{(r)}(n) = \min\{t(M) : M \text{ is an ordered } r\text{-matching of size } n\}.$$

For instance, in the matching M represented by the word

$$\textcolor{red}{A} \textcolor{blue}{A} \textcolor{red}{B} \textcolor{blue}{E} \textcolor{red}{C} \textcolor{blue}{B} \textcolor{red}{D} \textcolor{blue}{E} \textcolor{red}{E} \textcolor{blue}{B} \textcolor{red}{D} \textcolor{blue}{A} \textcolor{red}{C} \textcolor{blue}{C} \textcolor{red}{D}$$

one can find a pair of twins of size two formed by the sub-matchings $\textcolor{red}{AACACC}$ and $\textcolor{blue}{BBDBDD}$. So, $t(M) = 2$ as, trivially, we always have $t(M) \leq n/2$ for a matching of size n .

In a recent paper [5] (see also [6]) we demonstrated that the following inequalities hold for all $n \geq 2$ (note that $t^{(2)}(1) = 0$):

$$\frac{1}{16\sqrt[5]{2}} \cdot n^{\frac{3}{5}} \leq t^{(2)}(n) \leq \frac{e}{\sqrt[3]{2}} \cdot n^{\frac{2}{3}}.$$

The upper bound is obtained by a standard probabilistic argument based on the expectation. In fact, we proved in [5] that this upper bound is a.a.s. (asymptotically almost surely) attained by almost all ordered matching of size n . The lower bound is a consequence of a result by Bukh and Rudenko [3], concerning the related problem for permutations.

By *twins* in a permutation π we mean a pair of disjoint order-isomorphic subsequences of π . Let $\tau(n)$ be the maximum length of twins contained in every permutation of length n . The result in [3] states that $\tau(n) \geq \frac{1}{8} \cdot n^{\frac{3}{5}}$, which is so far the best lower bound, while $\tau(n) = O\left(n^{\frac{2}{3}}\right)$ follows by an elementary probabilistic argument. It is conjectured by Gawron [9] that this upper bound yields the correct order of magnitude for $\tau(n)$. If true, this would imply the same for the function $t^{(2)}(n)$, since we proved in [5, 6] that $t^{(2)}(n) = \Theta(\tau(n))$. For other related results on twins in permutations, as well as for the rich background of the problem, we encourage the reader to look at our paper [4].

In the present paper we extend the above results to ordered r -matchings, $r \geq 3$, in both, the deterministic and the random setting. Let $\mathbb{RM}_n := \mathbb{RM}_n^{(r)}$ denote a *random* ordered r -matching of size n (for precise definition, see the next section).

Theorem 1. *For every $r \geq 2$, a.a.s.,*

$$(1) \quad t(\mathbb{RM}_n) = \Theta\left(n^{\frac{2}{r+1}}\right).$$

The proof of the upper bound follows, once again, by a standard application of the first moment method. To get the lower bound we apply a more sophisticated argument, similar to that in [4] and [5, 6], which is based on a concentration inequality of Talagrand for permutations [11] (see Section 2).

Of course, the upper bound in (1) is also true in the deterministic case, so, we have $t^{(r)}(n) = O\left(n^{\frac{2}{r+1}}\right)$, for every $r \geq 2$. Our second result gives a lower bound for $t^{(r)}(n)$.

Theorem 2. *For every $r \geq 2$, we have*

$$(2) \quad t^{(r)}(n) = \Omega\left(n^{\frac{3}{5} \cdot \frac{1}{2^{r-1}-1}}\right).$$

The proof of Theorem 2 in Section 3, is based on a key lemma (Lemma 5), establishing a double recurrence for $t^{(r)}$, and a recent Erdős-Szekeres type result [8] for ordered matchings obtained by Sauerman and Zakharov [13]. This latter result states that every sufficiently large ordered r -matching M contains a large homogenous “clique”, i.e., a sub-matching M' whose all pairs of edges form pairwise isomorphic r -matchings (of size two). More specifically, the size of such

a “clique” M' in any ordered r -matching with n edges is at least $\frac{1}{2} \cdot n^{\frac{1}{2^r-1}}$, as proved in [13].

Observe that if $M' = \{e_1, \dots, e_m\}$, then by arbitrarily splitting M' in half, e.g., $\{e_1, \dots, e_{\lfloor m/2 \rfloor}\}$ and $\{e_{\lfloor m/2 \rfloor+1}, \dots, e_m\}$, we obtain twins. Thus, we immediately get the lower bound $t^{(r)}(n) = \Omega\left(n^{\frac{1}{2^r-1}}\right)$, which is, however, much worse than the bound in (2). For instance, the exponents of n in these two bounds are equal, respectively, to $\frac{1}{5}$ and $\frac{1}{7}$ ($r = 3$), to $\frac{3}{35}$ and $\frac{1}{15}$ ($r = 4$), and to $\frac{1}{25}$ and $\frac{1}{31}$ ($r = 5$). In fact, their ratio converges to $\frac{6}{5}$ as r is growing. Nonetheless, the ratio of the logarithms of the current upper and lower bounds for $t^{(r)}(n)$ grows rather rapidly with r .

We will use the standard notation $[n] := \{1, 2, \dots, n\}$.

2. RANDOM MATCHINGS

In this section we prove Theorem 1. Recall that $\mathbb{RM}_n := \mathbb{RM}_n^{(r)}$ denotes a random ordered r -matching of size n , that is, an ordered r -matching picked uniformly at random out of all

$$\alpha_n^{(r)} := \frac{(rn)!}{(r!)^n n!}$$

such r -matchings on the set $[rn]$.

The formula for $\alpha_n^{(r)}$ indicates that each ordered matching can be coupled with exactly $(r!)^n n!$ permutations. Indeed, one can generate an ordered matching by the following permutational scheme. Let π be a permutation of $[rn]$. Now π can be chopped off into an r -matching consisting of the following collection of edges

$$\{\pi(1), \dots, \pi(r)\}, \{\pi(r+1), \dots, \pi(2r)\}, \dots, \{\pi(rn-r+1), \dots, \pi(rn)\}.$$

Clearly, there are exactly $(r!)^n n!$ permutations π yielding the same matching. Thus, a (uniformly) random permutation Π_{rn} of $[rn]$ generates a (uniformly) random r -matching \mathbb{RM}_n . This scheme allows one to use concentration inequalities (such as the Talagrand inequality [14]) for random permutations in the context of random matchings.

Proof of Theorem 1: upper bound. The upper bound in formula (1) was already mentioned without proof in [7, Section 2]. The proof is based on the standard first moment method. For each k , let X_k be the number of twins of size k in \mathbb{RM}_n . Then

$$\mathbb{E}X_k = \frac{1}{2!} \cdot \binom{rn}{rk, rk, rn-2rk} \cdot \frac{\alpha_k^{(r)} \cdot 1 \cdot \alpha_{n-2k}^{(r)}}{\alpha_n^{(r)}} = \frac{1}{2} \cdot \frac{n!}{(n-2k)!} \cdot \frac{1}{k!} \cdot (r!)^k \cdot \frac{1}{(rk)!},$$

where the factor of 1 represents the second twin which is fully determined by the first one.

Ignoring $\frac{1}{2}$ and using inequalities $n!/(n-2k)! \leq n^{2k}$, $k! \geq (k/e)^k$, $r! \leq r^r$, and $(rk)! \geq (rk/e)^{rk}$, we thus get

$$\mathbb{E}X_k \leq n^{2k} \cdot \left(\frac{e}{k}\right)^k \cdot r^{rk} \cdot \left(\frac{e}{rk}\right)^{rk} = \left(n^2 \cdot \frac{e}{k} \cdot r^r \cdot \left(\frac{e}{rk}\right)^r\right)^k = \left(\frac{n^2}{(k/e)^{r+1}}\right)^k,$$

which converges to 0 as $n \rightarrow \infty$ for $k \geq cn^{2/(r+1)}$, with any constant $c > e$. Hence, for such k , $\mathbb{P}(X_k > 0) \leq \mathbb{E}X_k = o(1)$ and so, a.a.s. $t(\mathbb{RM}_n) \leq cn^{2/(r+1)}$. ■

An important ingredient of the proof of the lower bound in Theorem 1 is a Talagrand's concentration inequality for random permutations from [14]. We quote here a slightly simplified version from [11, Inequality (2) with $l = 2$] (see also [12, Inequality (1.3)]). Let $\Pi_{n'}$ be a random permutation of order n' (we will be applying this theorem with $n' = rn$).

Theorem 3 (Luczak and McDiarmid [11]). *Let $h(\pi)$ be a function defined on the set of all permutations of order n' which, for some positive constants c and d , satisfies*

- (i) *if π_2 is obtained from π_1 by swapping two elements, then $|h(\pi_1) - h(\pi_2)| \leq c$;*
- (ii) *for each π and $s > 0$, if $h(\pi) = s$, then in order to show that $h(\pi) \geq s$, one needs to specify only at most ds values $\pi(i)$.*

Then, for every $\epsilon > 0$,

$$\mathbb{P}(|h(\Pi_{n'}) - m| \geq \epsilon m) \leq 4 \exp(-\epsilon^2 m / (32dc^2)),$$

where m is the median of the random variable $h(\Pi_{n'})$.

As mentioned above, one can use this lemma for random r -matchings, as they can be generated by random permutations.

Proof of Theorem 1: lower bound. Set

$$a := \beta n^{(r-1)/(r+1)} \quad \text{with} \quad \beta := \frac{1}{(20er!)^{1/(r+1)}}$$

and

$$N := \frac{rn}{a} = \frac{r}{\beta} n^{2/(r+1)}.$$

For simplicity we assume that both a and N are integers. Partition $[rn] = A_1 \cup \dots \cup A_N$, where A_i 's are consecutive blocks of a integers each. For every $I = \{i_1, \dots, i_r\}$ with $1 \leq i_1 < \dots < i_r \leq N$, we call an r -element subset $S \subset [rn]$

an I -set if $|S \cap A_{i_j}| = 1$ for each $1 \leq j \leq r$. Further, define a random variable X_I which counts the number of edges of \mathbb{RM}_n which are I -sets.

Consider an auxiliary r -uniform hypergraph $H := H(\mathbb{RM}_n)$ on vertex set $[N]$, where I is an edge in H if and only if $X_I \geq 2$. For better clarity the edges of H will be sometimes called H -edges. Trivially, the maximum degree in H is at most $\binom{N-1}{r-1}$, but also, due to the disjointness of edges in \mathbb{RM}_n , at most $a/2$.

It is easy to see that a matching of size k in H corresponds to twins in \mathbb{RM}_n of size k . Indeed, let $M = \{I_1, \dots, I_k\}$ be a matching in H . For every $1 \leq \ell \leq k$, let $e_1^{(\ell)}, e_2^{(\ell)} \in \mathbb{RM}_n$ be two I_ℓ -edges. Then, the sub-matchings $M_1 = \{e_1^{(1)}, \dots, e_1^{(k)}\}$ and $M_2 = \{e_2^{(1)}, \dots, e_2^{(k)}\}$, owing to the sequential choice of A_i 's, form twins in \mathbb{RM}_n . Thus, our ultimate goal is to show that a.a.s. H contains a matching of size $\Omega(n^{2/(r+1)})$. Let $\nu(H)$ be the size of the largest matching in H . Our ultimate goal is thus to show that a.a.s. $\nu(H) = \Omega(N)$. Anticipating application of Theorem 3 to $\nu(H)$, it will be sufficient to bound $\mathbb{E}(\nu(H))$ from below. We will do it in a most “silly” way by looking just for isolated H -edges.

Let H_1 be a subgraph of H induced by the set V_1 of vertices of degrees at most 1 in H , that is, $E(H_1)$ consists of isolated edges in H which, of course, form a matching. Set $W = |E(H)|$ and $W_1 = |E(H_1)|$ for the random variables counting the edges in H and H_1 , respectively. Then, $\mathbb{E}(\nu(H)) \geq \mathbb{E}(W_1)$ and

$$W_1 = W - |\{e \in E(H) : e \cap (V \setminus V_1) \neq \emptyset\}| \geq W - \sum_{d=2}^{\lfloor a/2 \rfloor} dZ_d,$$

where Z_d counts the number of vertices of degree d in H . Note that $\mathbb{E}(Z_d) = NP(D = d)$, where D is the degree of a fixed vertex, say vertex 1, in H . Thus, we have

$$(3) \quad \mathbb{E}(W_1) \geq \mathbb{E}(W) - \mathbb{E} \left(\sum_{d=2}^{\lfloor a/2 \rfloor} dZ_d \right) = \mathbb{E}(W) - N \sum_{d=2}^{\lfloor a/2 \rfloor} d\mathbb{P}(D = d)$$

and all we need are a lower bound on $\mathbb{E}(W)$ and an upper bound on $\mathbb{P}(D = d)$, so that (3) would yield $\mathbb{E}(W_1) = \Omega(N)$. We begin with the former task. From the definition of H we have

$$(4) \quad \mathbb{E}(W) = \sum_{I \in \binom{[N]}{r}} \mathbb{P}(X_I \geq 2)$$

which, however, unlike in [6], cannot be applied directly. The reason is that to bound $\mathbb{P}(X_I \geq 2)$ from below one has to handle expressions like $n!/(n-a)!$ which are asymptotic to n^a as long as $a = o(\sqrt{n})$. In [6] the case of graphs

($r = 2$) was considered and we had $a = \Theta(n^{1/3})$, however, for $r \geq 3$ we have $a = \Theta(n^{(r-1)/(r+1)}) = \Omega(\sqrt{n})$.

Hence, instead of applying (4) directly, we define a random variable Y that counts the number of pairs of edges $\{e_1, e_2\}$ for which there exists a set of indices $I = \{i_1 < \dots < i_r\}$ such that both e_1 and e_2 are I -edges. By definitions, Y and X_I are related by the identity

$$Y = \sum_{I \in \binom{[N]}{r}, X_I \geq 2} \binom{X_I}{2}.$$

Observe that for any integer $t \geq 2$, using the obvious bound $X_I \leq a$, we get

$$\begin{aligned} Y &= \sum_{I \in \binom{[N]}{r}, 2 \leq X_I \leq t} \binom{X_I}{2} + \sum_{I \in \binom{[N]}{r}, X_I \geq t+1} \binom{X_I}{2} \\ &\leq \binom{t}{2} \sum_{I \in \binom{[N]}{r}, X_I \geq 2} 1 + \binom{a}{2} \sum_{I \in \binom{[N]}{r}, X_I \geq t+1} 1. \end{aligned}$$

Thus,

$$\begin{aligned} \mathbb{E}(Y) &\leq \binom{t}{2} \sum_{I \in \binom{[N]}{r}} \mathbb{P}(X_I \geq 2) + \binom{a}{2} \sum_{I \in \binom{[N]}{r}} \mathbb{P}(X_I \geq t+1) \\ (5) \quad &= \binom{t}{2} \mathbb{E}(W) + \binom{a}{2} \sum_{I \in \binom{[N]}{r}} \mathbb{P}(X_I \geq t+1), \end{aligned}$$

where we used (4).

From here to obtain a lower bound on $\mathbb{E}(W)$ it suffices to bound $\mathbb{E}(Y)$ from below and show that, for some constant t , the latter summation above is $o(\mathbb{E}(Y))$. Observe that $\mathbb{E}(Y)$ can be calculated directly from its definition as

$$\mathbb{E}(Y) = \binom{N}{r} \binom{a}{2}^r (2!)^{r-1} \cdot \frac{\alpha_{n-2}^{(r)}}{\alpha_n^{(r)}} = \binom{N}{r} \binom{a}{2}^r (2!)^{r-1} \cdot (r!)^2 \frac{n!}{(n-2)!} \frac{(rn-2r)!}{(rn)!}.$$

Indeed, there are $\binom{N}{r}$ choices of $I = \{i_1, \dots, i_r\}$ and, given that, there are $\binom{a}{2}^r (2!)^{r-1}$ choices of two I -sets e_1, e_2 . The quotient $\frac{\alpha_{n-2}^{(r)}}{\alpha_n^{(r)}}$ equals the probability that \mathbb{RM}_n contains e_1 and e_2 as edges and the formula follows by the linearity of expectation.

Since r is fixed (and the asymptotic is taken in n), each binomial coefficient can be easily approximated and furthermore $\frac{n!}{(n-2)!} \sim n^2$ and $\frac{(rn-2r)!}{(rn)!} \sim \frac{1}{(rn)^{2r}}$.

This yields,

$$\mathbb{E}(Y) \sim \frac{N^r}{r!} \left(\frac{a^2}{2!} \right)^r (2!)^{r-1} \cdot \left(\frac{r!n}{(rn)^r} \right)^2 = \frac{N^r}{2r!} \left(\frac{(r-1)!a^r}{(rn)^{r-1}} \right)^2.$$

Now expressing N and a as functions of n (leaving just one factor of N intact), we get

$$\mathbb{E}(Y) \sim N \cdot \frac{\left(\frac{r}{\beta} n^{2/(r+1)} \right)^{r-1}}{2r!} \left(\frac{(r-1)! \beta^r n^{(r-1)r/(r+1)}}{(rn)^{r-1}} \right)^2 = N \cdot \frac{(r-1)! \beta^{r+1}}{2r^r}$$

and thus, for large n ,

$$(6) \quad \mathbb{E}(Y) \geq N \cdot \frac{(r-1)! \beta^{r+1}}{3r^r}.$$

Next, we estimate the second summation in (5). Note that for a fixed $I \in \binom{[N]}{r}$,

$$\mathbb{P}(X_I \geq t+1) \leq \binom{a}{t+1}^r ((t+1)!)^{r-1} \frac{\alpha_{n-(t+1)}^{(r)}}{\alpha_n^{(r)}} \sim \frac{a^{r(t+1)}}{(t+1)!} \frac{(r!)^{t+1} n^{t+1}}{(rn)^{r(t+1)}}.$$

Hence,

$$\mathbb{P}(X_I \geq t+1) = O_{r,t} \left(n^{-\frac{(r-1)(t+1)}{r+1}} \right)$$

and consequently,

$$\binom{a}{2} \sum_{I \in \binom{[N]}{r}} \mathbb{P}(X_I \geq t+1) = O_{r,t} \left(a^2 N^r n^{-\frac{(r-1)(t+1)}{r+1}} \right) = O_{r,t} \left(n^{\frac{2(r-1)+2r-(r-1)(t+1)}{r+1}} \right).$$

Since for any $r \geq 2$ and $t \geq 4$,

$$\frac{2(r-1) + 2r - (r-1)(t+1)}{r+1} = \frac{(r-1)(3-t) + 2}{r+1} < \frac{2}{r+1},$$

we obtain, taking $t = 4$,

$$(7) \quad \binom{a}{2} \sum_{I \in \binom{[N]}{r}} \mathbb{P}(X_I \geq 5) = o(N).$$

Using (6) and (7) in (5), this implies the bounds

$$N \cdot \frac{(r-1)! \beta^{r+1}}{3r^r} \leq \mathbb{E}(Y) \leq \binom{4}{2} \mathbb{E}(W) + o(N)$$

from which it follows that

$$\mathbb{E}(W) \geq N \cdot \frac{(r-1)! \beta^{r+1}}{20r^r}.$$

After having estimated $\mathbb{E}(W)$ we move to the second major task which is to bound $\mathbb{P}(D = d)$ from above where, recall, D is the degree of vertex 1 in H . Every edge of H containing vertex 1 corresponds to a set $I \in \binom{[N]}{r}$ with $1 \in I$ and (not uniquely) to a pair of to-be I -edges e_1, e_2 in \mathbb{RM}_n . The number of choices of the triple (I, e_1, e_2) is $\binom{N-1}{r-1} \binom{a}{2}^r (2!)^{r-1}$ and the same bound applies to the $d-1$ remaining triples (I', e_1, e_2') . Of course, the edges selected to the triples should be vertex-disjoint as otherwise they could not all appear in a matching. As we bound from above, we may ignore this requirement. On the other hand, as the H -edges containing vertex 1 are not ordered, we need to divide by $d!$.

Consequently, very crudely,

$$\begin{aligned} \mathbb{P}(D = d) &\leq \mathbb{P}(D \geq d) \leq \frac{1}{d!} \left(\binom{N}{r-1} \binom{a}{2}^r (2!)^{r-1} \right)^d \frac{\alpha_{n-2d}^{(r)}}{\alpha_n^{(r)}} \\ &\sim \frac{1}{d!} \left(\frac{N^{r-1}}{(r-1)!} \frac{a^{2r}}{2} \right)^d \frac{(rn-2rd)!}{(rn)!} \frac{n!}{(n-2d)!} (r!)^{2d}. \end{aligned}$$

Above we turned to estimating $\mathbb{P}(D \geq d)$ instead of $\mathbb{P}(D = d)$ to avoid the issue of producing incidentally more than d edges of H containing vertex 1.

Using the inequality $1 - x \geq e^{-2x}$ valid for $x \leq 1/2$, the fraction $\frac{(rn-2rd)!}{(rn)!}$ can be estimated as

$$\frac{(rn-2rd)!}{(rn)!} = \frac{1}{(rn)^{2rd} \left(1 - \frac{1}{rn}\right) \cdots \left(1 - \frac{2rd-1}{rn}\right)} \leq \frac{e^{4rd^2/n}}{(rn)^{2rd}}$$

and the fraction $\frac{n!}{(n-2d)!}$ is trivially bounded by n^{2d} . Hence,

$$\begin{aligned} \mathbb{P}(D = d) &\leq \frac{1}{d!} \left(\frac{N^{r-1}}{(r-1)!} \frac{a^{2r}}{2} \right)^d \frac{e^{4rd^2/n}}{(rn)^{2rd}} n^{2d} (r!)^{2d} \\ &= \frac{1}{d!} \left(\frac{N^{r-1}}{(r-1)!} \frac{a^{2r}}{2} \frac{e^{4rd/n}}{(rn)^{2r}} n^2 (r!)^2 \right)^d. \end{aligned}$$

Since $N^{r-1} a^{2r} \frac{n^2}{n^{2r}} = r^{r-1} \beta^{r+1}$ and $e^{4rd/n} \leq 2$ as $d \leq a/2 = o(n)$, we get

$$\mathbb{P}(D = d) \leq \frac{1}{d!} \left(\frac{1}{(r-1)!} \cdot \frac{1}{2} \cdot \frac{2}{r^{2r}} (r!)^2 \cdot r^{r-1} \beta^{r+1} \right)^d = \frac{1}{d!} \left(\frac{\beta^{r+1} r!}{r^r} \right)^d.$$

We are now ready to apply (3) and get

$$\mathbb{E}(W_1) = \mathbb{E}(W) - N \sum_{d \geq 2} d \mathbb{P}(D = d) \geq N \left(\frac{(r-1)! \beta^{r+1}}{20r^r} - \sum_{d \geq 2} \frac{1}{(d-1)!} \left(\frac{\beta^{r+1} r!}{r^r} \right)^d \right).$$

Since $\frac{\beta^{r+1} r!}{r^r} = \frac{1}{20er^r} < 1$, we can bound

$$\sum_{d \geq 2} \frac{1}{(d-1)!} \left(\frac{\beta^{r+1} r!}{r^r} \right)^d \leq \left(\frac{1}{20er^r} \right)^2 \sum_{d \geq 2} \frac{1}{(d-1)!} \leq \left(\frac{1}{20er^r} \right)^2 e,$$

and thus,

$$\mathbb{E}(\nu(H)) \geq \mathbb{E}(W_1) \geq N \cdot \frac{1}{20^2 e r^{r+1}} \left(1 - \frac{1}{r^{r-1}} \right) = \Omega_r(N) = \Omega_r \left(n^{2/(r+1)} \right).$$

Finally, owing to the permutational scheme of generating \mathbb{RM}_n , we are in a position to apply Theorem 3 with $h(\pi) = \nu(H)$. Let us check the assumptions. For a permutation π of $[rn]$, let $M(\pi)$ be the corresponding matching. Observe that if π_2 is obtained from a permutation π_1 by swapping some two of its elements, then at most two edges of $M(\pi_1)$ can be destroyed and at most two edges of $M(\pi_1)$ can be created, and thus the same can be said about the H -edges of in $M(\pi_1)$. This, in turn, implies that the size of the largest matching in H has been altered by at most two, that is, $|h(\pi_1) - h(\pi_2)| \leq 2$. Moreover, to exhibit that $h(\pi) \geq s$, it obviously suffices to reveal $2s$ edges of \mathbb{RM}_n , that is, $2rs$ values of π .

Thus, Theorem 3 with $c = 2$, $d = 2r$, and $\epsilon = 1/2$ yields that

$$\mathbb{P}(|\nu(H) - m| \geq m/2) \leq 4 \exp(-m/(1024r)).$$

Moreover, there is a standard passage from the median to the expectation $\mu = \mathbb{E}(\nu(H))$. Indeed, we have (see for example [14], Lemma 4.6, or [12], page 164) that $|m - \mu| = O(\sqrt{m})$. As demonstrated above, $\mu \rightarrow \infty$, so it follows that $m \rightarrow \infty$ and, in particular, $|m - \mu| \leq 0.01\mu$. This implies that $\mathbb{P}(|\nu(H) - m| \geq m/2) = o(1)$ and

$$\begin{aligned} \mathbb{P}(|\nu(H) - \mu| \geq (3/4)\mu) &= \mathbb{P}(|\nu(H) - m + m - \mu| \geq (3/4)\mu) \\ &\leq \mathbb{P}(|\nu(H) - m| + |m - \mu| \geq (3/4)\mu) \\ &\leq \mathbb{P}(|\nu(H) - m| \geq (2/3)\mu) \leq \mathbb{P}(|\nu(H) - m| \geq m/2) \\ &= o(1), \end{aligned}$$

which means that a.a.s. $\nu(H) = \Omega_r(\mu) = \Omega_r(n^{2/(r+1)})$, from which the existence of twins in \mathbb{RM}_n of size $\Omega_r(n^{2/(r+1)})$ follows. \blacksquare

Notice that we could not apply Theorem 3 directly to the random variable W_1 , as in order to exhibit s isolated edges of H , one would need to reveal an unbounded number of values of π — to make sure that none of the r -element subsets of $[N]$ intersecting the given s forms an edge of H . In fact, for $r = 2$, one could use instead the Azuma-Hoeffding inequality (see, e.g., [6, Theorem 3.7] or the references given there), avoiding the above issue with the witness assumption of Talagrand's inequality. Unfortunately, this does not work for $r \geq 3$, as then we only have $\mathbb{E}(W_1) = \Omega(n^{2/(r+1)})$, so we do not know if $(\mathbb{E}(W_1))^2/n$ tends to infinity which makes the Azuma-Hoeffding inequality useless.

3. GENERAL MATCHINGS

In this section we will give the proof of Theorem 2. We start with stating the main lemma together with some explanations of how it leads to the desired lower bound (2).

3.1. The main lemma and its consequence

In [6, Lemma 3.4, $r = 2$] we showed the following lower bound on $t^{(2)}(n)$ in terms of $\tau(n)$. (Note that in [6], unlike here, r meant multiplicity of twins.)

Proposition 4. *For all $3/5 \leq \alpha \leq 2/3$ and $\beta > 0$, if $\tau(n) \geq \beta n^\alpha$ for all $n \geq 2$, then $t^{(2)}(n) \geq \beta(n/4)^\alpha$ for all $n \geq 2$.*

It follows that $t^{(2)}(n) = \Omega(\tau(n))$, but, in fact, we have $t^{(2)}(n) = \Theta(\tau(n))$ (the upper bound is trivial — see [6, Section 3.1] or [5, Section 3]).

Here we generalize Proposition 4 for all values of r . To this end, we first show a (doubly) iterative lower bound on $t^{(r)}(n)$ which also depends on $\tau(n)$. Set $\ell_2 = n^{1/3}$ and, for every $r \geq 3$, set $\ell_r(n) = \frac{1}{2}n^{1/(2^r-1)}$.

Lemma 5. *For all $n \geq 1$ and $r \geq 2$, $t^{(r)}(n) \geq \min\{t_0, t_1, t_2\}$, where*

$$t_0 = 2t^{(r)}(n/3), \quad t_1 = \tau(\ell_{r-1}(n/6r)),$$

and

$$t_2 = \min_{2 \leq p \leq r-2} \max \left\{ t^{(r-p)}(\ell_p(n/6r)), t^{(p)}(\ell_{r-p}(n/6r)) \right\}.$$

For $r \leq 3$, t_2 is not defined, or, to the same effect, we may set it equal to n in such cases, so that it does not affect the minimum.

From this we may derive recursively lower bounds on $t^{(r)}(n)$ in terms of n and r , assuming a lower bound on $\tau(n)$. For $r \geq 1$, let $\eta_r = \frac{1}{2^r-1}$, so that now we can write $\ell_r(n) = \frac{1}{2}n^{\eta_r}$.

Corollary 6. *For all $3/5 \leq \alpha \leq 2/3$ and $\beta > 0$, if $\tau(n) \geq \beta n^\alpha$ for all $n \geq 2$, then for all $r \geq 2$ and all $n \geq 2$,*

$$(8) \quad t^{(r)}(n) \geq \beta_r (n/6r)^{\alpha \eta_{r-1}},$$

where $\beta_2 = \beta$, while for $r \geq 3$,

$$\beta_r = \min \left\{ \min_{2 \leq p \leq r-2} \beta_p (12r)^{-\alpha \eta_{p-1}}, \beta 2^{-\alpha} \right\}.$$

In particular, $t^{(3)}(n) \geq \beta_3 (n/18)^{\alpha/3}$, where $\beta_3 = \beta 2^{-\alpha}$.

For technical reasons we have not pulled out the constant $(6r)^{-\alpha \eta_{r-1}}$ and incorporated it into β_r above. Note also, that just for the sake of unification, for $r = 2$ the bound in Corollary 6 is slightly weaker than the bound in Proposition 4. The case $r = 2$ is special also in that we have $t^{(2)}(n) = \Theta(\tau(n))$ (the upper bound is trivial — see [6, Section 3.1]). On the other hand, for $r \geq 3$ the lower bound in Corollary 6 does not seem to be close to the truth.

Applying the above mentioned bound $\tau(n) = \Omega(n^{3/5})$, we obtain immediately that

$$t^{(r)}(n) = \Omega\left(n^{\frac{3}{5}\eta_{r-1}}\right)$$

holds for all $r \geq 2$. This coincides with (2) and proves Theorem 2. (For $r = 2$ it was already deduced in [6, Corollary 3.5].) So, to complete the proof of Theorem 2, it remains to prove Lemma 5 and Corollary 6.

3.2. Unavoidable patterns

Given $r \geq 2$, there are exactly $\frac{1}{2} \binom{2r}{r}$ ways, called *patterns*, in which a pair of disjoint edges of order r may intertwine on an ordered vertex set. We call them r -patterns if the order r is to be emphasized. For instance, using convenient letter notation, there are just three patterns for $r = 2$, namely $AABB, ABBA, ABAB$ and ten for $r = 3$, $AAABBB, AABABB, AABBBBA, AABBBAB, ABBBBAA, ABBAAB, ABBABA, ABAABB, ABABBA, ABABAB$. For a pattern P , a P -clique is defined as a matching whose all pairs of edges form pattern P . For example, with $P = ABAABB$, a P -clique is a 3-matching with the structure $A_1 \cdots A_n A_1 A_1 \cdots A_n A_n$. Let $L_P(M)$ be the size of the largest P -clique in a matching M , $L(M) = \max_P L_P(M)$, and $L_r(n) = \min_M L(M)$, where the minimum is taken over all r -matchings M of size n .

In [6] (see also [5]) we showed that $L_2(n) = \lfloor n^{1/3} \rfloor$ and used this result (the lower bound) to prove Proposition 4. Very recently, building upon the concepts and results contained in [7] and in [1], Sauerman and Zakharov proved in [13] the following lower bound for every r . Recall our notation from the previous subsection: $\ell_r(n) = \frac{1}{2} n^{1/(2^r-1)}$ and $\eta_r = \frac{1}{2^r-1}$.

Theorem 7 [13]. *For all $r \geq 2$ and all $n \geq 1$, we have $L_r(n) \geq \ell_r(n)$.*

In fact, in [13] there is a better constant than $\frac{1}{2}$ in front of n^{η_r} which depends on r and goes to 1 as $r \rightarrow \infty$. For our purposes this has very little effect and therefore we stick to the weaker but easier to handle $1/2$.

3.3. Proof of Lemma 5

Let $M = M^{(r)}(n)$ be an ordered r -matching on $[rn] = H_1 \cup H_2$, where $H_1 = \{\lfloor rn/2 \rfloor\}$ and $H_2 = [rn] \setminus H_1$ are the first and second “half” of the vertex set. Further, for $p = 0, 1, 2, \dots, r-2, r-1, r$, let

$$n_p = |\{e \in M : |e \cap H_1| = p\}|.$$

Note that $\sum_{p=0}^r n_p = n$, while $\sum_{p=0}^r pn_p = \lfloor rn/2 \rfloor$ and, by symmetry, $\sum_{p=0}^r (r-p)n_p = \lceil rn/2 \rceil$. The latter identities imply that $\max\{n_r, n_0\} \leq n/2$. We consider three cases (more precisely, two cases one of which splits further into two subcases) with respect to the values of n_p . For two edges $e, f \in M$ we write $e < f$ whenever the leftmost vertex of e is to the left of the leftmost vertex of f , i.e., $\min e < \min f$.

Case 1. $\min\{n_0, n_r\} \geq n/3$. Let M_r and M_0 be the sub-matchings of M consisting of the edges contained in, respectively, H_1 and H_2 . Then, for $i = 0, r$, we have $|M_i| \geq n/3$, so $t^{(r)}(M_i) \geq t^{(r)}(n/3)$. Thus, by concatenation, $t^{(r)}(M) \geq 2t^{(r)}(n/3)$.

Case 2. $\min\{n_0, n_r\} \leq n/3$. In this case, $n_0 + n_r \leq n/2 + n/3 = 5n/6$, and so

$$\sum_{p=1}^{r-1} n_p \geq n/6.$$

Thus, we may consider two subcases (for simplicity we compromise $r-1$ to r in all denominators).

Subcase 2a. $\max\{n_1, n_{r-1}\} \geq n/(6r)$. Without loss of generality we assume that $n_{r-1} \geq n/(6r)$. Let $M_{r-1} \subset M$ consist of all n_{r-1} edges of M which intersect H_1 in exactly $r-1$ vertices (and thus they intersect H_2 in just one vertex). For each $e \in M_{r-1}$ we write $e = e^* \cup \{v_e\}$, where $e^* \subset H_1$ and $v_e \in H_2$. Let $M^* = \{e^* : e \in M_{r-1}\}$. Note that $|M^*| = |M_{r-1}| = n_{r-1}$. By applying Theorem 7 to the $(r-1)$ -matching M^* we conclude that for some $(r-1)$ -pattern P there is in M^* a P -clique $M_P^* = \{e_1^* < \dots < e_m^*\}$ of size $m = L_{r-1}(n_{r-1})$. Set $v_i := v_{e_i^*}$, for convenience.

Let $\pi = v_1, \dots, v_m$ be the permutation of the right ends of the edges of M_P^* and let $\pi' = v_{i_1}, \dots, v_{i_t}$ and $\pi'' = v_{j_1}, \dots, v_{j_t}$ be the longest twins in π . Then, we claim that $M' = \{e_{i_1} < \dots < e_{i_t}\}$ and $M'' = \{e_{j_1} < \dots < e_{j_t}\}$ are twins in M . Indeed, for $1 \leq g < h \leq t$, consider two pairs of edges $(e_{i_g}, e_{i_h}) \in M'$

and $(e_{j_g}, e_{j_h}) \in M''$. The first $r - 1$ vertices of e_{i_g} and e_{i_h} , i.e., $e_{i_g}^*$ and $e_{i_h}^*$, as well, as $e_{j_g}^*$ and $e_{j_h}^*$ form the same pattern P (as they all belong to the clique M_P^*). Moreover, the pairs of rightmost vertices, respectively, v_{i_g}, v_{i_h} and v_{j_g}, v_{j_h} , as being at the same positions in the twins π' and π'' , are in the same relation: $(v_{i_g} < v_{i_h} \text{ and } v_{j_g} < v_{j_h})$ or $(v_{i_g} > v_{i_h} \text{ and } v_{j_g} > v_{j_h})$.

As an example, consider a special case when $r = 5$, $p = 4$, and $P = ABBBAABA$. Let e_A, e_B, e_C, e_D be four edges whose first four vertices form mutually pattern P , while the last vertices satisfy $v_A < v_B$ and $v_C < v_D$. Then e_A and e_B form the pattern $ABBBAAABA|AB$, while e_C and e_D — pattern $CCCCCDDC|CD$ which is the very same pattern indeed.

This shows that $t^{(r)}(M) \geq \tau(L_{r-1}(n_{r-1})) \geq \tau(L_{r-1}(n/6r))$.

Subcase 2b. $\max_{2 \leq p \leq r-2} n_p \geq n/(6r)$. Let, for some $2 \leq p \leq r - 2$, $n_p \geq n/(6r)$. Let $M_p \subset M$ consist of all n_p edges of M which intersect H_1 in exactly p vertices (and thus they intersect H_2 in $r - p$ vertices). For each $e \in M_p$ we write $e = e^* \cup e^{**}$, where $e^* = e \cap H_1$ and $e^{**} = e \cap H_2$. Let $M^* = \{e^* : e \in M_p\}$. Note that $|M^*| = |M_p| = n_p$. By applying Theorem 7 to the p -matching M^* we conclude that for some p -pattern P there is in M^* a P -clique $M_P^* = \{e_1^* < \dots < e_m^*\}$ of size $m \geq L_p(n_p)$.

Let $M^{**} = \{e_i^{**} : i = 1, \dots, m\}$ be the $(r - p)$ -matching in H_2 consisting of the remainders of the edges in M_P^* . Further, let $M^{**'} = \{e_{i_1}^{**} < \dots < e_{i_t}^{**}\}$ and $M^{**''} = \{e_{j_1}^{**} < \dots < e_{j_t}^{**}\}$ be the largest twins in M^{**} . Then, we claim that $M' = \{e_{i_1}^* < \dots < e_{i_t}^*\}$ and $M'' = \{e_{j_1}^* < \dots < e_{j_t}^*\}$ are twins in M . Indeed, for $1 \leq g < h \leq t$, consider two pairs of edges $(e_{i_g}, e_{i_h}) \in M'$ and $(e_{j_g}, e_{j_h}) \in M''$. The p -long prefixes of e_{i_g} and e_{i_h} , i.e., $e_{i_g}^*$ and $e_{i_h}^*$ form pattern P , as the prefixes $e_{j_g}^*$ and $e_{j_h}^*$ do. Moreover, the $(r - p)$ -element suffixes $e_{i_g}^{**}$ and $e_{i_h}^{**}$ form the same $(r - p)$ -pattern Q as $e_{j_g}^{**}$ and $e_{j_h}^{**}$ do. So, the pairs (e_{i_g}, e_{i_h}) and (e_{j_g}, e_{j_h}) form the same r -pattern R , which proves that M' and M'' are twins.

To illustrate this part of the proof, consider a special case when $r = 5$ and $p = 3$. Let $P = AABBBBA$ be a collectible 3-pattern and let $Q = ABBA$. If four edges, e_A, e_B, e_C, e_D each have 3 vertices in H_1 which mutually form pattern P and, moreover, the 2-vertex remainders of e_A and e_B form in H_2 pattern Q , and the same holds for the 2-vertex remainders of e_C and e_D , then edges e_A and e_B form overall pattern $AABBBBAABBA$, while edges e_C and e_D form pattern $CCDDDDCCDDC$, which is the very same pattern (we call it R in our proof).

Hence,

$$t^{(r)}(M) \geq t^{(r-p)}(m) \geq t^{(r-p)}(L_p(n_p)).$$

We could repeat the entire argument in Subcase 2b with the roles of p and $r - p$ swapped yielding the max in the definition of t_2 . Finally, we have to minimize the obtained bound over all $2 \leq p \leq r - 2$ and over all three subcases. ■

3.4. Proof of Corollary 6

The proof is by double induction on $r \geq 2$ and $n \geq 2$. The case $r = 2$ was proved in [6], see Proposition 4 above. Indeed, by Proposition 4 we have

$$t^{(2)}(n) \geq \beta(n/4)^\alpha \geq \beta(n/12)^\alpha = \beta_2(n/12)^\alpha.$$

Moreover, for $r \geq 3$, if $2 \leq n \leq 6r(1/\beta_r)^{\frac{1}{\alpha\eta_{r-1}}}$, then the desired bound becomes $t^{(r)}(n) \geq 1$ which is trivially true. Thus, assume that $r \geq 3$, $n \geq 6r(1/\beta_r)^{\frac{1}{\alpha\eta_{r-1}}}$, equivalently,

$$(9) \quad \beta_r(n/6r)^{\alpha\eta_{r-1}} \geq 1$$

and that (8) holds for all $2 \leq r' < r$ and $2 \leq n' < n$.

Note that the assumption $\tau(n) \geq \beta n^\alpha$ for all $n \geq 2$ implies, by taking $n = 2$, that $1 \geq \beta 2^\alpha$, or $\beta \leq 2^{-\alpha}$. Thus, whenever (9) holds, recalling that $\beta_r \leq 2^{-\alpha}\beta$, it follows that

$$\frac{1}{2}(n/6r)^{1/\eta_{r-1}} \geq \frac{1}{2}\beta_r^{-1/\alpha} \geq \beta^{-1/\alpha} \geq 2.$$

This, in turn, implies that $n/6r > 1$ and, consequently, that in Lemma 5 all arguments appearing in the functions defining t_0 , t_1 , and t_2 are larger or equal to 2 (and, obviously, less than n) allowing to apply the induction assumptions.

By Lemma 5, we know that $t^{(r)}(n) \geq \min\{t_0, t_1, t_2\}$. So, consider three cases.

Case 0. $t^{(r)}(n) \geq t_0$. As $n/3 \geq 2$, by induction's assumption we have

$$t^{(r)}(n) \geq t_0 = 2t^{(r)}(n/3) \geq 2\beta_r(n/18r)^{\alpha\eta_{r-1}} \geq \beta_r(n/6r)^{\alpha\eta_{r-1}},$$

the last inequality equivalent to $2^{1/\eta_{r-1}} = 2^{2^{r-1}-1} \geq 3^\alpha$, which is true for $r \geq 3$ (recall that $\alpha \leq 1$).

Case 1. $t^{(r)}(n) \geq t_1$. In this case, by the assumption,

$$t^{(r)}(n) \geq t_1 = \tau\left(\frac{1}{2}(n/6r)^{\eta_{r-1}}\right) \geq \beta 2^{-\alpha}(n/6r)^{\alpha\eta_{r-1}} \geq \beta_r(n/6r)^{\alpha\eta_{r-1}},$$

since $\beta_r \leq 2^{-\alpha}\beta$.

Case 2. $t^{(r)}(n) \geq t_2$. By induction's assumption,

$$\begin{aligned} t^{(r)}(n) &\geq t_2 = \min_{2 \leq p \leq r-2} \max \left\{ t^{(r-p)}\left(\frac{1}{2}(n/6r)^{\eta_p}\right), t^{(p)}\left(\frac{1}{2}(n/6r)^{\eta_{r-p}}\right) \right\} \\ &\geq \min_{2 \leq p \leq r-2} \max \left\{ \beta_{r-p} \left(\frac{(n/6r)^{\eta_p}}{12(r-p)} \right)^{\alpha\eta_{r-p-1}}, \beta_p \left(\frac{(n/6r)^{\eta_{r-p}}}{12p} \right)^{\alpha\eta_{p-1}} \right\} \\ &\geq \min_{2 \leq p \leq r-2} \max \left\{ \beta_{r-p}(12r)^{-\alpha\eta_{r-p-1}}, \beta_p(12r)^{-\alpha\eta_{p-1}} \right\} (n/6r)^{\alpha\eta_{r-1}} \\ &\geq \beta_r(n/6r)^{\alpha\eta_{r-1}}, \end{aligned}$$

by choosing $\beta_r \leq \min_{2 \leq p \leq r-2} \beta_p(12r)^{-\alpha\eta_{p-1}}$. (Above we used the facts that $n/6r > 1$ and $\eta_p\eta_{r-p-1} > \eta_{r-1}$, both with a big margin.) ■

4. CONCLUDING REMARKS

Let us conclude the paper with some problems for future considerations. Firstly, it is natural to speculate on the true asymptotic order of the function $t^{(r)}(n)$. Based on Theorem 1 and the former results around Gawron's conjecture [9] on twins in permutations (see also [4]), we dare to state the following.

Conjecture 8. *For every $r \geq 2$, $t^{(r)}(n) = \Theta\left(n^{\frac{2}{r+1}}\right)$.*

Notice that the case $r = 2$ of this statement is the original Gawron's conjecture. It seems that even increasing the exponent of n in the lower bound (2) to the inverse of any polynomial in r will be quite a challenge.

One could also consider a generalization to t -tuples, that is, t -tuples of vertex disjoint order-isomorphic sub-matchings of a given r -matching. In fact, the proof techniques from Section 2 yield that this generalized parameter is a.a.s. $\Theta(n^{t/(r(t-1)+1)})$ (for $t = 2$ we thus recover Theorem 1). A highly technical analog of Theorem 2 could be proved as well.

It seems also natural to study the problem of the largest twins in more general classes of ordered graphs ($r = 2$) or hypergraphs ($r \geq 3$). By *twins* in an ordered graph G we mean a pair of edge-disjoint order-isomorphic subgraphs of G . Let $t(G)$ denote the maximum size of twins in G , and let $t(m)$ be the minimum of $t(G)$ over all ordered graphs with m edges. What can be said about the function $t(m)$? By the results for ordered matchings we know only that $t(m) = O(m^{\frac{2}{3}})$, but is it optimal? It is worth mentioning that the analogous problem for unordered graphs was solved by Lee, Loh, and Sudakov [10] who proved that the corresponding function is $\Theta\left((m \log m)^{\frac{2}{3}}\right)$.

Finally, one may also investigate the size of twins in general words over finite alphabets. Therein, twins are defined as pairs of identical subsequences occupying disjoint sets of positions. Actually, our motivation to study twins in permutations and ordered matchings has been ignited by a beautiful result of Axenovich, Person, and Puzynina [2], stating that every binary word of length n contains twins of size $\frac{1}{2}n - o(n)$.

Inspired by the word representation of ordered matchings (used also in this paper), one may consider a relaxed variant of twins in words, in which a pair of disjoint subwords forms *permuted twins* if they are identical up to a permutation of their letters. For instance, in this new setting, the subwords 112142212 and 223213323 of a word over alphabet $\{1, 2, 3, 4\}$, would form permuted twins (under permutation 2341). How large permuted twins may one find in every word of length n over a k -element alphabet? Will they be much bigger, especially for large k , than in the classical case?

Acknowledgments

We would like to thank an anonymous referee for a careful reading of the manuscript and suggesting a number of editorial improvements.

REFERENCES

- [1] M. Anastos, Z. Jin, M. Kwan and B. Sudakov, *Extremal, enumerative and probabilistic results on ordered hypergraph matchings* (2023).
arXiv:2308.12268
- [2] M. Axenovich, Y. Person and S. Puzynina, *A regularity lemma and twins in words*, J. Combin. Theory Ser. A **120** (2013) 733–743.
<https://doi.org/10.1016/j.jcta.2013.01.001>
- [3] B. Bukh and O. Rudenko, *Order-isomorphic twins in permutations*, SIAM J. Discrete Math. **34** (2020) 1620–1622.
<https://doi.org/10.1137/20M1323357>
- [4] A. Dudek, J. Grytczuk and A. Ruciński, *Variations on twins in permutations*, Electron. J. Combin. **28(3)** (2021) #P3.19.
<https://doi.org/10.37236/9734>
- [5] A. Dudek and J. Grytczuk and A. Ruciński, *Patterns in ordered (random) matchings*, in: LATIN 2022: Theoretical Informatics, Lecture Notes in Comput. Sci. **13568**, A. Castañeda and F. Rodríguez-Henríquez (Ed(s)), (Springer Cham, 2022) 544–556.
https://doi.org/10.1007/978-3-031-20624-5_33
- [6] A. Dudek, J. Grytczuk and A. Ruciński, *Ordered unavoidable sub-structures in matchings and random matchings* (2022).
arXiv:2210.14042
- [7] A. Dudek, J. Grytczuk and A. Ruciński, *Erdős-Szekeres type theorems for ordered uniform matchings* (2023).
arXiv:2301.02936
- [8] P. Erdős and G. Szekeres, *A combinatorial problem in geometry*, Compos. Math. **2** (1935) 463–470.
- [9] M. Gawron, *Isomorphic Substructures in Words and Permutations*, Master Thesis, Jagiellonian University (2014) in Polish.
- [10] C. Lee, P.-S. Loh and B. Sudakov, *Self-similarity of graphs*, SIAM J. Discrete Math. **27** (2013) 959–972.
<https://doi.org/10.1137/120861436>
- [11] M.J. Luczak and C. McDiarmid, *Concentration for locally acting permutations*, Discrete Math. **265** (2003) 159–171.
[https://doi.org/10.1016/S0012-365X\(02\)00628-3](https://doi.org/10.1016/S0012-365X(02)00628-3)

- [12] C. McDiarmid, *Concentration for independent permutations*, Combin. Probab. Comput. **11** (2002) 163–178.
<https://doi.org/10.1017/S0963548301005089>
- [13] L. Sauermann and D. Zakharov, *A sharp Ramsey theorem for ordered hypergraph matchings* (2023).
arXiv:2309.04813
- [14] M. Talagrand, *Concentration of measure and isoperimetric inequalities in product spaces*, Publ. Math. Inst. Hautes Études Sci. **81** (1995) 73–205.

Received 2 October 2023

Revised 4 January 2024

Accepted 9 January 2024

Available online 26 January 2024