Ordered Unavoidable Sub-Structures in Matchings and Random Matchings

Andrzej Dudek[a] Jarosław Grytczuk[b] Andrzej Ruciński[c]

Submitted: Mar 10, 2023; Accepted: Mar 2, 2024; Published: Apr 19, 2024

Abstract

An *ordered matching of size* n is a graph on a linearly ordered vertex set V, $|V| = 2n$, consisting of n pairwise disjoint edges. There are three different ordered matchings of size two on $V = \{1,2,3,4\}$: an *alignment* $\{1,2\},\{3,4\}$, a *nesting* $\{1,4\},\{2,3\}$, and a *crossing* $\{1,3\},\{2,4\}$. Accordingly, there are three basic homogeneous types of ordered matchings (with all pairs of edges arranged in the same way) which we call, respectively, *lines*, *stacks*, and *waves*.

We prove an Erdős–Szekeres type result guaranteeing in every ordered matching of size n the presence of one of the three basic sub-structures of a given size. In particular, one of them must be of size at least $n^{1/3}$. We also investigate the size of each of the three sub-structures in a *random* ordered matching. Additionally, the former result is generalized to 3-uniform ordered matchings.

Another type of unavoidable patterns we study are *twins*, that is, pairs of order-isomorphic, disjoint sub-matchings. By relating to a similar problem for permutations, we prove that the maximum size of twins that occur in every ordered matching of size n is $O\left(n^{2/3}\right)$ and $\Omega\left(n^{3/5}\right)$. We conjecture that the upper bound is the correct order of magnitude and confirm it for almost all matchings. In fact, our results for twins are proved more generally for r-multiple twins, $r \geqslant 2$. [1]

Mathematics Subject Classifications: 05C80, 05C30, 05C70

[a] Department of Mathematics, Western Michigan University, Kalamazoo, MI, U.S.A. (`andrzej.dudek@wmich.edu`).

[b] Faculty of Mathematics and Information Science, Warsaw University of Technology, Warsaw, Poland (`j.grytczuk@mini.pw.edu.pl`).

[c] Department of Discrete Mathematics, Adam Mickiewicz University, Poznań, Poland, (`rucinski@amu.edu.pl`).

[1] An extended abstract of this paper appears in [11].

1 Introduction

1.1 Background

A graph G is said to be *ordered* if its vertex set is linearly ordered. Let G and H be two ordered graphs with vertex sets $V(G) = \{v_1, \dots, v_m\}$ and $V(H) = \{w_1, \dots, w_m\}$, and the respective linear orders $v_1 < \cdots < v_m$ and $w_1 < \cdots < w_m$, for some integer $m \geqslant 1$. We say that G and H are *order-isomorphic* if for all $1 \leqslant i < j \leqslant m$, $v_iv_j \in E(G)$ if and only if $w_iw_j \in E(H)$. Note that every pair of order-isomorphic graphs is isomorphic, but not vice-versa. Also, if G and H are distinct graphs on the same linearly ordered vertex set V, then G and H are never order-isomorphic, and so all $2^{\binom{|V|}{2}}$ labeled graphs on V are pairwise non-order-isomorphic. It shows that the notion of order-isomorphism makes sense only for pairs of graphs on distinct vertex sets.

One context in which order-isomorphism makes quite a difference is that of subgraph containment. If G is an ordered graph, then any subgraph G' of G can be also treated as an ordered graph with the ordering of $V(G')$ inherited from the ordering of $V(G)$. Given two ordered graphs, (a "large" one) G and (a "small" one) H, we say that a subgraph $G' \subset G$ is *an ordered copy of H in G* if G' and H are order-isomorphic. We will sometimes denote this fact by writing $G' \preceq G$.

All kinds of questions concerning subgraphs in unordered graphs can be posed for ordered graphs as well (see, e.g., [32] and [5]). For example, in [3] and [9] the authors studied Turán and Ramsey type problems for ordered graphs. In particular, they showed independently that there exists an ordered matching on n vertices for which the (ordered) Ramsey number is super-polynomial in n, a sharp contrast with the linearity of the Ramsey number for ordinary (i.e., unordered) matchings. This shows that it makes sense to study even such seemingly simple structures as ordered matchings. In fact, Jelínek [22] counted the number of matchings avoiding (i.e., not containing) a given small ordered matching.

1.2 Topics and organization

In this paper we focus exclusively on *ordered matchings*, that is, ordered graphs which consist of vertex-disjoint edges (and have no isolated vertices). For example, in Figure 1, we depict two ordered matchings, $M = \{\{1,3\},\{2,4\},\{5,6\}\}$ and $N = \{\{1,5\},\{2,3\},\{4,6\}\}$ on vertex set $\{1,2,3,4,5,6\}$ with the natural linear ordering. Unlike in [22], we study what sub-structures are *unavoidable* in ordered matchings. A frequent theme in both fields, the theory of ordered graphs as well as enumerative combinatorics, are unavoidable sub-structures, that is, *patterns* that appear in every member of a prescribed family of structures. A flagship example providing everlasting inspiration is the famous theorem of Erdős and Szekeres [14] on monotone subsequences (see [4, 6, 7, 13, 15, 26, 31] for some recent extensions and generalizations). In its diagonal form it states that any sequence $x_1, x_2, \dots, x_n$ of distinct real numbers contains an increasing or decreasing subsequence of length at least $\sqrt{n}$.

And, indeed, our first goal is to prove its analog for ordered matchings. The reason why

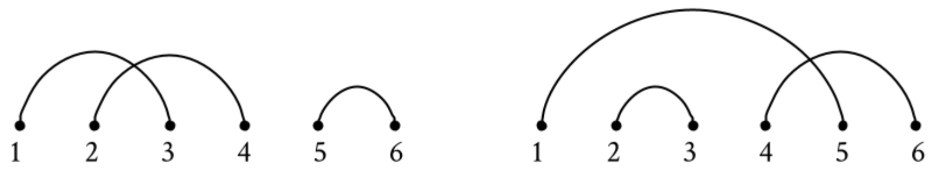

Figure 1: Exemplary matchings M and N.

the original Erdős–Szekeres Theorem lists only two types of subsequences is, obviously, that for any two elements x_i and x_j with $i < j$ there are just two possible relations: $x_i < x_j$ or $x_i > x_j$. For matchings, however, for every two edges $\{x, y\}$ and $\{u, w\}$ with $x < y$, $u < w$, and $x < u$, there are three possibilities: $y < u$, $w < y$, or $u < y < w$ (see Figure 2). In other words, every two edges form either *an alignment*, *a nesting*, or *a crossing* (the first term introduced by Kasraoui and Zeng in [24], the last two terms coined in by Stanley [29]). These three possibilities give rise, respectively, to three "unavoidable" ordered sub-matchings (*lines*, *stacks*, and *waves*) which play an analogous role to the monotone subsequences in the classical Erdős–Szekeres Theorem. (In [29], stacks and waves consisting of k edges were called, respectively, k*-nestings* and k*-crossings*.)

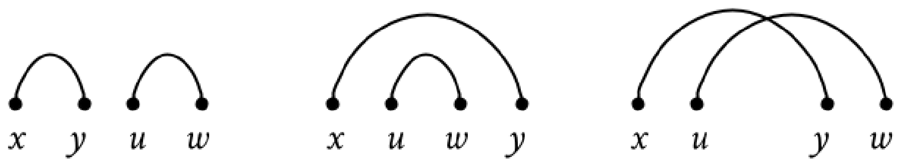

Figure 2: An alignment, a nesting, and a crossing of a pair of edges.

Informally, lines, stacks, and waves are defined by demanding that every pair of edges in a sub-matching forms, respectively, an alignment, a nesting, or a crossing (see Figure 5). In Subsection 2.1 we show, in particular, that every ordered matching of size n contains one of these structures of size at least $n^{1/3}$. This special case was also proved by Huynh, Joos and Wollan (see Lemma 25 in [21]). In the remainder of Section 2, we first extend this result to 3-uniform ordered matchings and then study the size of the largest lines, stacks, and waves in random matchings.

Our second goal is to estimate the size of the largest (ordered) twins in ordered matchings. The problem of twins has been widely studied for other combinatorial structures, including words, permutations, and graphs (see, e.g., [1, 25]). For an integer $r \geqslant 2$, we say that r edge-disjoint (ordered) subgraphs $G_1, G_2, \ldots, G_r$ of an (ordered) graph G form *(ordered) twins in G* if they are pairwise (order-)isomorphic. The size of the (ordered) twins is defined as $|E(G_1)| = \cdots = |E(G_r)|$. For ordinary matchings, the notion of r-twins becomes trivial, as every matching of size n contains twins of size $\lfloor n/r \rfloor$ – just split the matching into r as equal as possible parts. But for ordered matchings the problem becomes interesting. The above mentioned analog of the Erdős–Szekeres Theorem immediately yields (again by splitting into r equal parts) ordered r-twins of length $\lfloor n^{1/3}/r \rfloor$. We provide much better estimates on the size of largest r-twins in ordered matchings (Subsection 3.1) and random matchings (Subsection 3.2) which, not so surprisingly, are of the same order of magnitude as those for r-twins in permutations (see [8, 10]).

1.3 Random matchings

As indicated above, we examine both questions, of unavoidable sub-matchings and of twins, also for *random* matchings. A random (ordered) matching $\mathbb{RM}_n$ is selected uniformly at random from all

$$\alpha_n := \frac{(2n)!}{n!2^n}$$

ordered matchings on vertex set $[2n]$. Among other results, we show that with probability tending to 1, as $n \to \infty$, or *asymptotically almost surely* (a.a.s.), there are in $\mathbb{RM}_n$ lines, stacks, *and* waves of size, roughly, $\sqrt{n}$, as well as (ordered) twins of size $\Theta(n^{2/3})$.

There are two other ways of generating $\mathbb{RM}_n$ which we are going to utilize in the proofs. Besides the above defined *uniform* scheme, we define the *online* scheme as follows. For an arbitrary ordering of the vertices $u_1, \ldots, u_{2n}$ one selects uniformly at random a match, say u_{j_1}, for u_1 (in $2n-1$ ways), then, after crossing out u_1 and u_{j_1} from the list, one selects uniformly at random a match for the first uncrossed vertex (in $2n-3$ ways), and so on. Note that the total number of ways to select a matching this way is $(2n-1)(2n-3)\cdot\ldots\cdot 3\cdot 1 = (2n-1)!!$ which equals α_n. A third equivalent way to generate $\mathbb{RM}_n$ is particularly convenient when one intends to apply concentration inequalities available for random permutations. The *permutation based* scheme boils down to just generating a random permutation $\Pi := \Pi_n$ of $[2n]$ and "chopping it off" into a matching $\{\Pi(1)\Pi(2),\ \Pi(3)\Pi(4), \ldots, \Pi(2n-1)\Pi(2n)\}$. Note that this way each matching corresponds, as it should, to exactly $n!2^n$ permutations. We will stick mostly to the uniform scheme, applying the other two only occasionally.

1.4 Gauss codes

A convenient representation of ordered matchings can be obtained in terms of *double occurrence words* over an n-letter alphabet, in which every letter occurs exactly twice as the label of the ends of the corresponding edge in the matching. For instance, our two exemplary matchings can be written as $M = ABABCC$ and $N = ABBCAC$ (see Figure 3). In fact, this is a special case of an elementary combinatorial bijection between ordered partitions of a set and permutations with repetitions. A minor nuisance here is that ordered matchings correspond to unordered partitions, so every permutation of the letters yields the same matching. Nevertheless, we will sometimes use this representation to better illustrate some notions and ideas.

Interestingly, this type of words was introduced and studied already by Gauss [18] as a way of encoding closed self-intersecting curves on the plane (with points of multiplicity

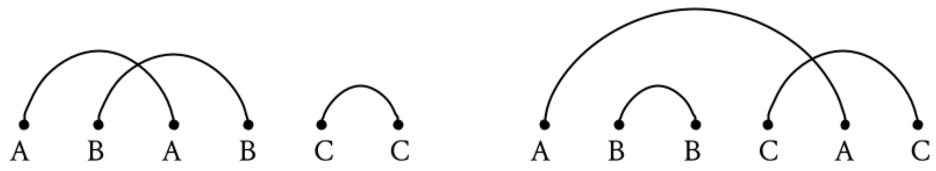

Figure 3: Exemplary matchings $M = ABABCC$ and $N = ABBCAC$.

at most two). Indeed, denoting the self-crossing points by letters and traversing the curve in a fixed direction gives a (cyclic) word in which every letter occurs exactly twice (by the multiplicity assumption) (see Figure 4). A general problem studied by Gauss was to characterize those words that correspond to such curves. He found himself a necessary condition, but a full solution (in terms of some quite involved constraints on an auxiliary graph of crossing pairs) was obtained much later (see [28] for a brief history and further references).

It is, perhaps, also worthwhile to mention that ordered matchings constitute a special case of structures, known as *Puttenham diagrams*, that found an early application in the theory of poetry (see [27, 33]). A basic idea is simple: a rhyme scheme of a poem can be encoded by a word in which same letters correspond to rhyming verses. Of particular interest here are *planar rhyme schemes* which are nothing else but ordered matchings without crossings, or more generally, *noncrossing partitions* (see [30]).

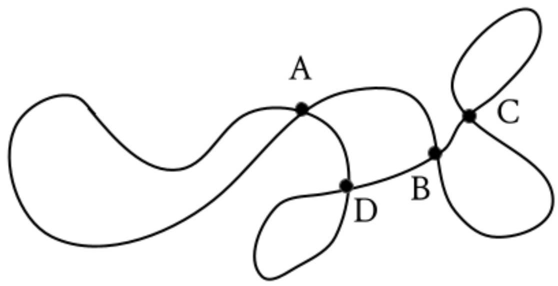

Figure 4: A curve with Gauss code $ABCCBDDA$.

2 Unavoidable sub-matchings

Let us start with formal definitions. Let M be an ordered matching on the vertex set $[2n]$, with edges denoted as $e_i = \{a_i, b_i\}$ so that $a_i < b_i$, for all $i = 1, 2, \ldots, n$, and $a_1 < \cdots < a_n$. We say that an edge e_i is *to the left of* e_j and write $e_i < e_j$ if $a_i < a_j$. That is, in ordering the edges of a matching we ignore the positions of the right endpoints.

We now define the three basic types of ordered matchings:

- Line: $a_1 < b_1 < a_2 < b_2 < \cdots < a_n < b_n$,
- Stack: $a_1 < a_2 < \cdots < a_n < b_n < b_{n-1} < \cdots < b_1$,
- Wave: $a_1 < a_2 < \cdots < a_n < b_1 < b_2 < \cdots < b_n$.

Assigning letter A_i to edge $\{a_i, b_i\}$, their corresponding double occurrence words look as follows:

- Line: $A_1A_1A_2A_2 \cdots A_nA_n$,
- Stack: $A_1A_2 \cdots A_{n-1}A_nA_nA_{n-1} \cdots A_1$.
- Wave: $A_1A_2 \cdots A_nA_1A_2 \cdots A_n$.

Each of these three types of ordered matchings can be equivalently characterized as follows. Let us consider all possible ordered matchings with just two edges. In the double occurrence word notation these are $AABB$ (an *alignment*), $ABBA$ (a *nesting*), and $ABAB$ (a *crossing*). Now a line, a stack, and a wave is an ordered matching in which *every* pair of edges forms, respectively, an alignment, a nesting, and a crossing (see Figure 5).

Note that alignments, crossings and nestings are just special instances (the smallest non-trivial) of, resp., lines, stacks, and waves, and throughout we will use these names interchangeably.

Figure 5: A line, a stack, and a wave of size three.

2.1 In arbitrary matchings

Consider a sub-matching M' of M and an edge $e \in M \setminus M'$, which is to the left of the left-most edge f of M'. Note that if M' is a line and e and f form a line, then $M' \cup \{e\}$ is a line too. Similarly, if M' is a stack and $\{e, f\}$ form a nesting, then $M' \cup \{e\}$ is a stack too. However, an analogous statement fails to be true for waves, as e, though crossing f, may not necessarily cross all other edges of the wave M'. Due to this observation, in the proof of our first result we will need another type of ordered matchings combining lines and waves. We call a matching $M = \{\{a_i, b_i\} : \ i = 1, \dots, n\}$ with $a_i < b_i$, for all $i = 1, 2, \dots, n$, and $a_1 < \cdots < a_n$, *a landscape* if $b_1 < b_2 < \cdots < b_n$, that is, the right-ends of the edges of M are also linearly ordered (a first-come-first-serve pattern). Notice that there are no non-trivial stacks in a landscape. In the double occurrence word notation, a landscape is just a word obtained by a *shuffle* of the two copies of the word $A_1A_2 \cdots A_n$.

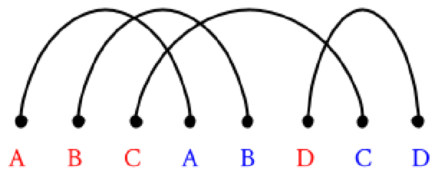

Figure 6: A landscape of size four.

Examples of landscapes for $n = 4$ are, among others, $ABCDABCD$, $AABCBCDD$, $ABCABDCD$ (the last one is depicted in Figure 6).

The following is an Erdős–Szekeres type result for ordered matchings.

Theorem 1. *Let ℓ, s, w be arbitrary positive integers and let $n = \ell sw + 1$. Then, every ordered matching on $2n$ vertices contains a line of size $\ell + 1$, or a stack of size $s + 1$, or a wave of size $w + 1$.*

Proof. Let M be any ordered matching with edges $\{a_i, b_i\}$, $i = 1, 2, \ldots, n$. Let s_i denote the size of a largest stack whose left-most edge is $\{a_i, b_i\}$. Similarly, let λ_i be the largest size of a landscape whose left-most edge is $\{a_i, b_i\}$. Consider the sequence of pairs (s_i, λ_i), $i = 1, 2, \ldots, n$. We argue that no two pairs of this sequence may be equal. Indeed, let $i < j$ and consider the two edges $\{a_i, b_i\}$ and $\{a_j, b_j\}$. These two edges may form a nesting, an alignment, or a crossing. In the first case we get $s_i > s_j$, since the edge $\{a_i, b_i\}$ enlarges the largest stack starting at $\{a_j, b_j\}$. In the two other cases, we have $\lambda_i > \lambda_j$ by the same argument. Since the number of pairs (s_i, λ_i) is $n > s \cdot \ell w$, it follows that either $s_i > s$ for some i, or $\lambda_j > \ell w$ for some j. In the first case we are done, as there is a stack of size $s + 1$ in M.

In the second case, assume that L is a landscape in M of size at least $p = \ell w + 1$. Let us order the edges of L as $e_1 < e_2 < \cdots < e_p$, accordingly to the linear order of their left ends. Decompose L into edge-disjoint waves, $W_1, W_2, \ldots, W_k$, in the following way. For the first wave W_1, pick e_1 and all edges whose left ends are between the two ends of e_1, say, $W_1 = \{e_1 < e_2 < \ldots < e_{i_1}\}$, for some $i_1 \geqslant 1$. Clearly, W_1 is a true wave since there are no nesting pairs in L. Also notice that the edges e_1 and e_{i_1+1} are non-crossing since otherwise the latter edge would be included in W_1. Now, we may remove the wave W_1 from L and repeat this step for $L - W_1$ to get the next wave $W_2 = \{e_{i_1+1} < e_{i_1+2} < \ldots < e_{i_2}\}$, for some $i_2 \geqslant i_1 + 1$. And so on, until exhausting all edges of L, while forming the last wave $W_k = \{e_{i_{k-1}+1} < e_{i_{k-1}+2} < \ldots < e_{i_k}\}$, with $i_k \geqslant i_{k-1} + 1$. Clearly, the sequence $e_1 < e_{i_1+1} < \ldots < e_{i_{k-1}+1}$ of the leftmost edges of the waves W_i must form a line (see Figure 7). So, if $k \geqslant \ell+1$, we are done. Otherwise, we have $k \leqslant \ell$, and because $p = \ell w+1$, some wave W_i must have at least $w + 1$ edges. This completes the proof. □

It is not hard to see that the above result is optimal. For example, consider the case $\ell = 5$, $s = 3$, $w = 4$. Take 3 copies of the wave of size $w = 4$: $ABCDABCD$,

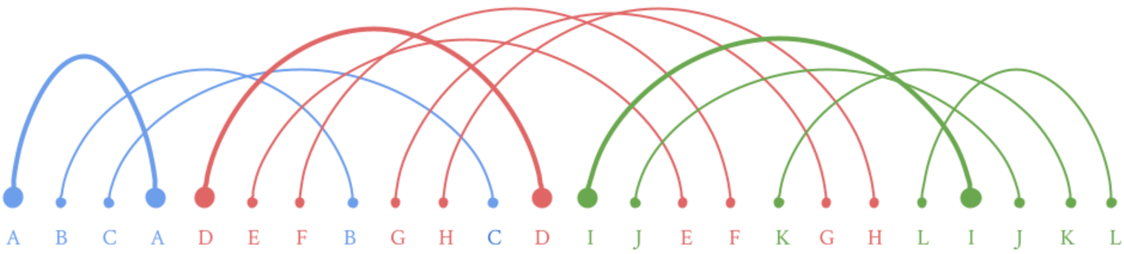

Figure 7: Greedy decomposition of a landscape into waves. The left-most edges of the waves (in bold) form a line.

$PQRSPQRS$, $XYZTXYZT$. Arrange them into a stack-like structure (see Figure 8):

$$ABCDPQRSXYZTXYZTPQRSABCD.$$

Now, concatenate $\ell = 5$ copies of this structure. Clearly, we obtain a matching of size

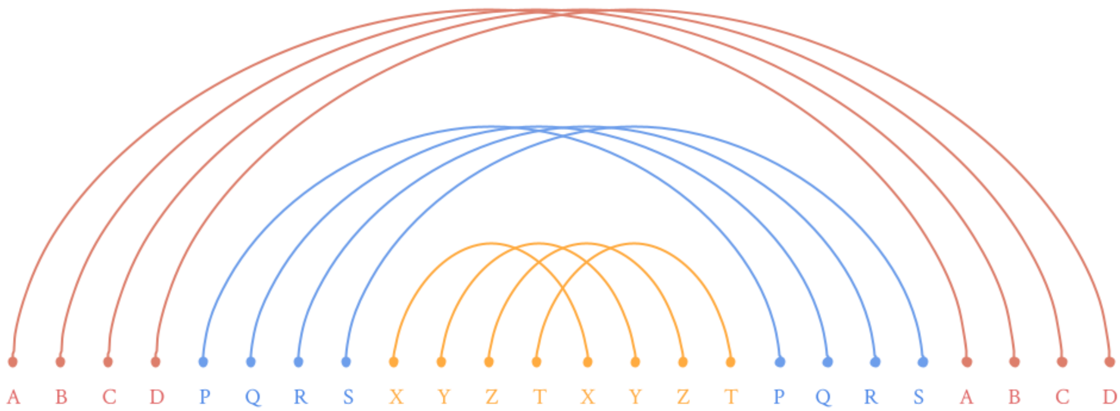

Figure 8: A stack of waves.

$\ell sw = 5 \cdot 3 \cdot 4$ with no line of size 6, no stack of size 4, and no wave of size 5. This example can be easily generalized to get the following fact.

Proposition 2. *For every positive integers ℓ, s and w there exists a matching of size $n = \ell sw$ with all lines, waves, and stacks of size at most ℓ, s and w, respectively.*

By forbidding one of the three basic structures to be present in M and setting the corresponding parameter, ℓ, s, or w to 1, we immediately deduce from Theorem 1 that one of the other two structures of an appropriately large size must be present in M. For example, in a landscape (i.e., no nestings) of size $n \geqslant \ell w + 1$ one can find either a line of size $\ell + 1$ or a wave of size $w + 1$. More interestingly, forbidding an alignment we obtain what we define in Section 3.1 as permutational matchings which are in a one-to-one correspondence with permutations of order n. Moreover, under this bijection waves and stacks in a permutational matching M correspond to, respectively, increasing and decreasing subsequences of the permutation which is the image of M. Thus, we recover the original Erdős–Szekeres Theorem as a special case of our Theorem 1.

Finally, we formulate separately the diagonal case of Theorem 1.

Corollary 3. *Every ordered matching on $2n$ vertices contains a line, a stack, or a wave of size at least $n^{1/3}$.*

Proof. Given n. Clearly, $\lceil n^{1/3}\rceil - 1 < n^{1/3}$. Hence, $(\lceil n^{1/3}\rceil - 1)^3 < n$ and so $(\lceil n^{1/3}\rceil - 1)^3 + 1 \leqslant n$. Applying Theorem 1 with $\ell = s = w = \lceil n^{1/3}\rceil - 1$ implies one of the three structures of size $(\lceil n^{1/3}\rceil - 1) + 1 \geqslant n^{1/3}$. □

2.2 3-uniform matchings

It is perhaps natural and interesting to try to generalize Theorem 1 to *r-uniform ordered matchings*, that is, families of n disjoint r-element subsets of the linearly ordered set $[rn]$. Here we make the first step by showing that the case $r = 3$ follows relatively easily from Theorem 1 itself.

At the start the problem seems a bit overwhelming, as there are $\frac{1}{2}\binom{6}{3} = 10$ different ways in which two triples may intertwine. Using the *triple occurrence words*, they are $AAABBB$, $AABABB$, $AABBBA$, $AABBAB$, $ABBBAA$, $ABBAAB$, $ABBABA$, $ABAABB$, $ABABBA$, $ABABAB$. We will call them *patterns* or *relations* since these words describe a way in which any two edges are interwoven with each other. What is worse, one of them, $AABABB$, stands out as a culprit who spoils the otherwise nice picture. To see it through, call an ordered pair of triples (e, f) *collectable* if for each $k \geqslant 1$ there exists a collection of k triples such that every pair of them is order-isomorphic to the pair (e, f).

For instance, the relation $AAABBB$, which we may call, as before, *an alignment*, is collectable, as for any k one can take $A_1A_1A_1A_2A_2A_2 \dots A_kA_kA_k$. Similarly, for $AABBBA$, say, one may take $(A_1A_1A_2A_2 \dots A_kA_k)(A_k \dots A_2A_1)$. In fact, all nine relations but $AABABB$ are collectable. However, for $AABABB$, which we may call *an engagement*, one cannot even add a third triple CCC. Indeed, if it were possible, then there should be a C between the second and third A, but after the first B, which makes the relation between B's and C's not an engagement, as it would begin with BC. (Here we assumed w.l.o.g. that the first A precedes the first B which precedes the first C.)

Due to this annoying exception, the Erdős–Szekeres type result we are going to prove is not as clean as its predecessors for singletons and pairs. For reasons, which will become clear once we reveal our proof strategy, we "give names to all the animals" as presented in the first two columns of Table 1.

To explain this encoding, let us denote the three basic graph configurations as `L` $= AABB$, `S` $= ABBA$, and `W` $= ABAB$, accordingly to their (alternative) names, that is, line, stack, and wave. Now, each of the ten pairwise relations of triples can be uniquely decomposed into a pair of pairs consisting, resp., of the first two vertices in both triples and the last two vertices in both triples. For example, the relation $AABBAB$ decomposes into a line $AABB$ and a wave $ABAB$. We express this fact by writing $AABBAB = AABB \oplus ABAB$ and denoting relation $AABBAB$ by $\mathcal{R}_{\texttt{LW}}$.

Obviously, for a given 3-uniform relation this decomposition is unique. However, this mapping is not one-to-one as there are only nine different ordered pairs made of three elements. And, indeed, the alignment $AAABBB$ and the engagement $AABABB$ both decompose into the same the pair $AABB$, $AABB$. To distinguish between them, we add asterisk for the latter, that is, we denote $AAABBB$ by $\mathcal{R}_{\texttt{LL}}$ while $AABABB$ by $\mathcal{R}^*_{\texttt{LL}}$.

The last column of Table 1 displays this mapping using, again, the words with underlined and overlined letters. The three relations with the first subscript S differ in that the letters in the second component of the decomposition $\oplus$ are reversed; indeed, if the pairs of the first two elements of two triples form the sequence $ABBA$, then the pairs of the last two elements can only form sequences $BBAA, BABA$, or $BAAB$. But, let us emphasize that as Gauss words $BBAA = AABB$, etc.

RELATION	IN WORDS	DECOMPOSITION $\oplus$
$\mathcal{R}_{\texttt{LL}}$	$AAABBB$	$AABB \oplus AABB = \underline{A}\underline{\overline{A}}\overline{A}\underline{B}\underline{\overline{B}}\overline{B}$
$\mathcal{R}^*_{\texttt{LL}}$	$AABABB$	$AABB \oplus AABB = \underline{A}\underline{\overline{A}}\underline{B}\overline{A}\underline{\overline{B}}\overline{B}$
$\mathcal{R}_{\texttt{LS}}$	$AABBBA$	$AABB \oplus ABBA = \underline{A}\underline{\overline{A}}\underline{B}\underline{\overline{B}}\overline{B}\overline{A}$
$\mathcal{R}_{\texttt{LW}}$	$AABBAB$	$AABB \oplus ABAB = \underline{A}\underline{\overline{A}}\underline{B}\underline{\overline{B}}\overline{A}\overline{B}$
$\mathcal{R}_{\texttt{SL}}$	$ABBBAA$	$ABBA \oplus BBAA = \underline{A}\underline{B}\underline{\overline{B}}\overline{B}\underline{\overline{A}}\overline{A}$
$\mathcal{R}_{\texttt{SS}}$	$ABBAAB$	$ABBA \oplus BAAB = \underline{A}\underline{B}\underline{\overline{B}}\underline{\overline{A}}\overline{A}\overline{B}$
$\mathcal{R}_{\texttt{SW}}$	$ABBABA$	$ABBA \oplus BABA = \underline{A}\underline{B}\underline{\overline{B}}\underline{\overline{A}}\overline{B}\overline{A}$
$\mathcal{R}_{\texttt{WL}}$	$ABAABB$	$ABAB \oplus AABB = \underline{A}\underline{B}\underline{\overline{A}}\overline{A}\underline{\overline{B}}\overline{B}$
$\mathcal{R}_{\texttt{WS}}$	$ABABBA$	$ABAB \oplus ABBA = \underline{A}\underline{B}\underline{\overline{A}}\underline{\overline{B}}\overline{B}\overline{A}$
$\mathcal{R}_{\texttt{WW}}$	$ABABAB$	$ABAB \oplus ABAB = \underline{A}\underline{B}\underline{\overline{A}}\underline{\overline{B}}\overline{A}\overline{B}$

Table 1: Possible relations of two triples and their corresponding decompositions $\oplus$.

To cope with engagements, we will "marry" them with alignments by combining relations $\mathcal{R}_{\texttt{LL}}$ and $\mathcal{R}^*_{\texttt{LL}}$ together. First, as in the case of pairs, define *a line* as an ordered 3-uniform matching M such that all pairs of triples of M are in relation $\mathcal{R}_{\texttt{LL}}$, that is, each pair forms an alignment. Call M *a semi-line* if all pairs of triples are in either relation $\mathcal{R}_{\texttt{LL}}$ or $\mathcal{R}^*_{\texttt{LL}}$, that is, they form an alignment or an engagement.

Proposition 4. *Every semi-line of size k contains a line of size at least $k/2$.*

Proof. Let $e_1, \ldots, e_k$ be a semi-line. Define an auxiliary graph G on vertex set $[k]$ where $ij \in E(G)$ if e_i and e_j form an engagement. It turns out that G is a linear forest. To prove it, assume w.l.o.g. that for all $i < j$ the left-most vertex of e_i is to the left of the left-most vertex of e_j. We claim that every vertex i has at most one neighbor $j > i$. Indeed, if there were edges ij and ih, $i < j < h$, then, identifying e_i, e_j and e_h, respectively, with letters A, B and C, we would have a sequence $AABCA\ldots$, where the last 4 positions are occupied by 2 letters B and 2 letters C in an arbitrary order. This means, however, that B's and C's, or equivalently e_j and e_h, form neither an alignment nor an engagement, a contradiction with the definition of a semi-line. By symmetry, any vertex i has also at most one neighbor $j < i$. So, G is, indeed, a linear forest and as such has an independent number at least $k/2$. This completes the proof. □

We are now ready to formulate an Erdős–Szekeres type result for 3-uniform matchings.

Theorem 5. *Let a_{XY}, where $X, Y \in \{\mathtt{L}, \mathtt{S}, \mathtt{W}\}$, be arbitrary positive integers and let*

$$n = \prod_{X,Y} a_{XY} + 1.$$

Then, every ordered 3-uniform matching M on $3n$ vertices either contains a semi-line of size $a_{\mathtt{LL}} + 1$, or there exists $(X, Y) \neq (\mathtt{L}, \mathtt{L})$ such that M contains a sub-matching of $a_{XY} + 1$ triples every two of which are in relation $\mathcal{R}_{XY}$ (as defined in Table 1).

Remark 6. As in Corollary 3 one can show that every 3-uniform ordered matching of size n contains one of the nine sub-structures listed in Theorem 5 of size at least $n^{1/9}$.

Remark 7. At the moment we are not able to find a construction showing optimality of Theorem 5. It was relatively easy in the graph case, as both, stacks and waves, had the *interval* chromatic number equal to 2, which enabled one to superimpose one into another by blowing up the vertices of one of them and filling its edges with copies of the other (see Figure 8 for the superimposition of a wave of size 4 upon a stack of size 3). Following, e.g., [17], we say that an ordered r-uniform hypergraph has *interval chromatic number r* if it is r-partite with the partition sets forming consecutive blocks of the linearly ordered vertex set.

Now, out of the eight 3-uniform relations, alignments and engagements aside, only four have interval chromatic number 3, namely those not having L within their indices: $\mathcal{R}_{\mathtt{SS}} = ABBAAB$, $\mathcal{R}_{\mathtt{SW}} = ABBABA$, $\mathcal{R}_{\mathtt{WS}} = ABABBA$, and $\mathcal{R}_{\mathtt{WW}} = ABABAB$. Thus, we can provide a counterexample showing optimality of Theorem 5 only in the special case when $a_{\mathtt{LS}} = a_{\mathtt{LW}} = a_{\mathtt{SL}} = a_{\mathtt{WL}} = 1$. We do so by superimposing the four "no-L" relations and then taking $a_{\mathtt{LL}}$ disjoint copies of the obtained construction. The superimposition is done by carefully replacing each triple of the current matching with a matching obeying mutually the new relation. E.g., replacing $ABABAB$ (relation $\mathcal{R}_{\mathtt{WW}}$) with $CD\ EF\ CD\ EF\ DC\ FE$ results in a matching where four pairs of triples retain relation $\mathcal{R}_{\mathtt{WW}}$, while two pairs ($C-D$ and $E-F$) enjoy the new relation $\mathcal{R}_{\mathtt{WS}}$. It is crucial here that consecutive blocks of size 2 in $CD\ CD\ EF$ each contain both letters W and Z, which is equivalent to having interval chromatic number 3.

For the case when $a_{\mathtt{SS}} = a_{\mathtt{SW}} = a_{\mathtt{WS}} = a_{\mathtt{WW}} = 2$ (and all other parameters set to 1) see Figure 9, where the pairwise relations between all 16 triples can be described as follows. Let us first focus on the 8 triples drawn with solid colored lines $(A, \ldots, H)$. Among them, all 4 monochromatic pairs of triples (like $A - B$) are in relation $\mathcal{R}_{\mathtt{SS}}$, the 8 red-blue and orange-green pairs of triples (like $A - C$) obey relation $\mathcal{R}_{\mathtt{SW}}$, while the remaining 16 pairs (like $A-E$) satisfy relation $\mathcal{R}_{\mathtt{WS}}$. The very same is true for the 8 triples drawn with dashed colored lines $(I, \ldots, P)$. Finally, each solid-dashed pair (like $A - I$) satisfies relation $\mathcal{R}_{\mathtt{WW}}$. Note that, viewing this construction as a complete graph K_{16}, each relation corresponds to a bipartite subgraph, and so, crucially, no three triples are mutually in the same relation.

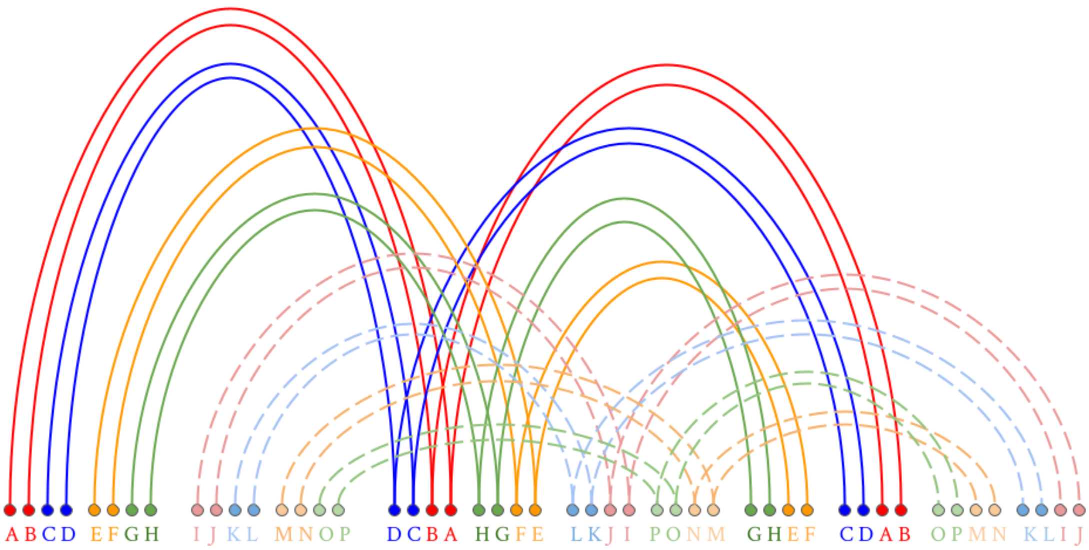

Figure 9: A matching of $2^4 = 16$ triples illustrating optimality of Theorem 5 with $a_{\mathtt{SS}} = a_{\mathtt{SW}} = a_{\mathtt{WS}} = a_{\mathtt{WW}} = 2$ (and all other parameters set to 1).

Remark 8. As in the case of graph matchings it is impossible to employ directly the original Erdős–Szekeres proof idea. To see this, call a relation $\mathcal{R}$ *extendable* if, given a collection M of disjoint triples which are mutually in relation $\mathcal{R}$, every triple to the left of M which is in relation $\mathcal{R}$ with the left-most triple of M, is in relation $\mathcal{R}$ with all triples in M. For graphs, lines and stacks were extendable, while waves were not, which forced us to come up with the notion of a landscape. Now, only $\mathcal{R}_{\mathtt{LL}}$, $\mathcal{R}_{\mathtt{LS}}$, $\mathcal{R}_{\mathtt{SW}}$, and $\mathcal{R}_{SL}$ are extendable, so any direct proof of Theorem 5 would require more sophisticated analogs of landscapes. Fortunately, we deduce Theorem 5 directly from Theorem 1.

Proof of Theorem 5. Let M be a matching (of triples) as in the assumptions of the theorem. Our proof strategy is to apply Theorem 1 twice, first to the (graph) matching M_1 composed of pairs consisting of the first two elements of the triples in M, then to a suitably chosen sub-matching of the (graph) matching M_1 composed of pairs of the last two elements of the triples in M.

Set $b_{\mathtt{L}} = a_{\mathtt{LL}}a_{\mathtt{LS}}a_{\mathtt{LW}}$, $b_{\mathtt{S}} = a_{\mathtt{SL}}a_{\mathtt{SS}}a_{\mathtt{SW}}$, and $b_{\mathtt{W}} = a_{\mathtt{WL}}a_{\mathtt{WS}}a_{\mathtt{WW}}$ and apply Theorem 1 to M_1 with $\ell := b_{\mathtt{L}}$, $s := b_{\mathtt{S}}$, and $w := b_{\mathtt{W}}$. We will now examine the three alternatives of the conclusion.

Case (L). M_1 contains a (graph) line L of size $b_{\mathtt{L}} + 1$. Let M_L be the sub-matching of M_2 composed of pairs of the last two elements of those triples in M whose first two elements form a pair belonging to L. Formally,

$$M_L = \{\{j,k\} : \; i < j < k, \; \{i,j,k\} \in M, \quad \text{and} \quad \{i,j\} \in L\}.$$

We apply Theorem 1 to M_L with $\ell := a_{\mathtt{LL}}$, $s := a_{\mathtt{LS}}$, and $w := a_{\mathtt{LW}}$ and examine the three possible outcomes.

Subcase (LL). M_L contains a line $Ł$ of size $a_{\mathtt{LL}} + 1$. This is the troublesome case. Let $M_{\mathtt{LL}}$ be the sub-matching of M consisting of all triples in M whose first pair of elements belongs to L, while the last pair belongs to $Ł$. Consider any two triples in $M_{\mathtt{LL}}$. There are two possible relations they may be in: the alignment $\underline{A\overline{A}}\overline{A}\underline{B\overline{B}}\overline{B}$ and the engagement $\underline{A\overline{A}B}\overline{A}\underline{\overline{B}}\overline{B}$. Indeed, in both cases the first pairs and the last pairs, underlined and overlined, resp., form (graph) alignments. This shows that $M_{\mathtt{LL}}$ is a semi-line.

Subcase (LS). M_L contains a stack S of size $a_{\mathtt{LS}} + 1$. Let $M_{\mathtt{LS}}$ be the sub-matching of M consisting of all triples in M whose first pair of elements belongs to L, while the last pair belongs to S. Consider any two triples in $M_{\mathtt{LS}}$. Then they necessarily are in relation $\mathcal{R}_{\mathtt{LS}}$, that is, they form the word $\underline{A\overline{A}B\overline{B}}\overline{BA}$.

Subcase (LW). M_L contains a wave W of size $a_{\mathtt{LW}} + 1$. Let $M_{\mathtt{LW}}$ be the sub-matching of M consisting of all triples in M whose first pair of elements belongs to L, while the last pair belongs to W. Consider any two triples in $M_{\mathtt{LW}}$. Then they necessarily are in relation $\mathcal{R}_{\mathtt{LW}}$, that is, they form the word $\underline{A\overline{A}B\overline{B}}\overline{AB}$.

Cases (S) and (W). Each case splits further, as before, into three subcases. It is straightforward to check that these altogether six subcases lead to the remaining six relations $\mathcal{R}_{\mathtt{SL}}$, $\mathcal{R}_{\mathtt{SS}}$, $\mathcal{R}_{\mathtt{SW}}$, $\mathcal{R}_{\mathtt{WL}}$, $\mathcal{R}_{\mathtt{WS}}$, $\mathcal{R}_{\mathtt{WW}}$, yielding each time the required size of the collection. $\square$

2.3 In random matchings

In this section we shall investigate the size of unavoidable structures one can find in *random* ordered matchings with the emphasis on the three canonical patterns: lines, stacks, and waves. Recall that $\mathbb{RM}_n$ is a random (ordered) matching of size n, that is, a matching picked uniformly at random out of the set of all $\alpha_n := (2n)!/(n!2^n)$ matchings on the set $[2n]$.

Baik and Rains in [2] (see also [29, Theorem 17]) determined the asymptotic distribution of the maximum size of two of the three canonical patterns contained in a random ordered matching. As a consequence, their values can be pinpointed very precisely.

Theorem 9 ([2]). *The sizes of the largest stack and the largest wave contained in $\mathbb{RM}_n$ are a.a.s. equal to $(1 + o(1))\sqrt{2n}$.*

A similar result for lines was proved by Justicz, Scheinerman, and Winkler in [23]. Note, however, that the constant is different.

Theorem 10 ([23]). *The size of the largest line contained in $\mathbb{RM}_n$ is a.a.s. equal to $(2 + o(1))\sqrt{n/\pi}$.*

In this section we provide simpler, purely combinatorial proofs of weaker versions of Theorems 9 and 10, with the asymptotic coefficient, resp., $\sqrt{2}$ and $2/\sqrt{\pi}$ replaced by pairs of constants setting lower and upper bounds only. The proof of the upper bounds is quite straightforward and provides a more general result.

Proposition 11. *Let* $(M_k)_1^\infty$ *be a sequence of ordered matchings of size* k*,* $k = 1, 2, \dots$*. Then, a.a.s.*

$$\max\{k : \ M_k \preceq \mathbb{RM}_n\} \leqslant (1 + o(1))e\sqrt{n/2}.$$

Proof. Set $k_0 = \lfloor \left(1 + n^{-1/3}\right) e\sqrt{n/2} \rfloor$, and let X_k be a random variable counting the number of ordered copies of M_k in $\mathbb{RM}_n$. Our goal is to show, via the first moment method, that a.a.s. $X_k = 0$ for all $k \geqslant k_0$. Note that it is *not* enough to show that a.a.s. $X_{k_0} = 0$, as the sequence $(M_k)_1^\infty$ may not be ascending.

To compute the expectation of X_k, one has to first choose the $2k$ vertices of a copy of M_k (the copy itself is placed in just one way), then count the number of extensions of that copy to an entire matching of size n, and, finally, divide by the total number of matchings. Thus,

$$\mathbb{E}X_k = \binom{2n}{2k} \cdot 1 \cdot \frac{\alpha_{n-k}}{\alpha_n} = \frac{2^k}{(2k)!} \cdot \frac{n!}{(n-k)!} \leqslant \frac{2^k}{(2k)!} \cdot n^k \leqslant \frac{2^k}{(2k/e)^{2k}} \cdot n^k = \left(\frac{e^2 n}{2k^2}\right)^k,$$

and, by the union bound applied together with Markov's inequality,

$$\begin{aligned}\Pr(\exists k \geqslant k_0 : \ X_k > 0) &\leqslant \sum_{k=k_0}^{n} \Pr(X_k \geqslant 1) \\ &\leqslant \sum_{k=k_0}^{n} \mathbb{E}X_k \leqslant \sum_{k=k_0}^{n} \left(\frac{e^2 n}{2k^2}\right)^k \leqslant n(1 + n^{-1/3})^{-2k_0} = o(1). \qquad \square\end{aligned}$$

Owing to the very homogeneous structure of lines, stacks, and waves, we are able to establish corresponding lower bounds for their maximum sizes in $\mathbb{RM}_n$. It is, perhaps, interesting to note that, unlike for permutations, the size of the sub-structures guaranteed by the of Erdős–Szekeres-type result (cf. Corollary 3) grows substantially in the random setting.

Theorem 12. *A random matching* $\mathbb{RM}_n$ *contains a.a.s. stacks and waves of size at least* $\frac{1-o(1)}{e\sqrt{2}}\sqrt{n}$ *each, as well as lines of size at least* $\frac{1}{8}\sqrt{n}$*.*

The proof for stacks and waves is very simple and relies mostly on Theorem 1. For lines we will make use of the following lemma which might be of some independent interest. For that reason we state it in a more general setting than what we actually need. The *length* of an edge $\{i, j\}$ in a matching on $[2n]$ is defined as $|j - i|$.

Lemma 13. *Let a sequence* $f(n)$ *be such that* $f(n) \to \infty$ *and* $f(n) = o(n)$*. Then, a.a.s. the number of edges of length at most* $f(n)$ *in* $\mathbb{RM}_n$ *is* $(1 + o(1))f(n)$*.*

Proof. For each pair of vertices $1 \leqslant u < v \leqslant 2n$, let X_{uv} be the indicator random variable equal to one if $\{u, v\} \in \mathbb{RM}_n$ and 0 otherwise. Clearly, $\Pr(X_{uv} = 1) = 1/(2n - 1)$. Let us write for simplicity $m = f(n)$. The sum $X = \sum_{1 \leqslant v-u \leqslant m} X_{uv}$ counts all edges in $\mathbb{RM}_n$ of length at most m. As the number of summands is $(2n-1)+(2n-2)+\dots+(2n-m) = 2nm - \binom{m+1}{2}$, we have

$$\mathbb{E}X = \frac{2nm - \binom{m+1}{2}}{2n-1} = m(1 - O(m/n)).$$

In particular, by the assumptions on $f(n)$, we get $\mathbb{E}X = m(1-o(1)) \to \infty$. To estimate the second moment, observe that

$$\Pr(X_{u_1v_1} = X_{u_2v_2} = 1) = \begin{cases} \frac{1}{(2n-1)(2n-3)}, & \text{if } \{u_1, v_1\} \cap \{u_2, v_2\} = \emptyset, \\ 0, & \text{otherwise.} \end{cases}$$

Hence, estimating quite crudely,

$$\mathbb{E}(X(X-1)) \leqslant \frac{\left(2nm - \binom{m+1}{2}\right)^2}{(2n-1)(2n-3)},$$

and consequently, by Chebyshev's inequality

$$\begin{aligned} \Pr(|X - \mathbb{E}X| \geqslant \varepsilon \mathbb{E}X) &\leqslant \frac{\mathbb{E}(X(X-1)) + \mathbb{E}X - (\mathbb{E}X)^2}{\varepsilon^2(\mathbb{E}X)^2} \\ &\leqslant \frac{1}{\varepsilon^2}\left(\frac{(2n-1)^2}{(2n-1)(2n-3)} + \frac{1}{\mathbb{E}X} - 1\right) = \frac{1}{\varepsilon^2}\left(\frac{2}{2n-3} + \frac{1}{\mathbb{E}X}\right) \to 0, \end{aligned}$$

provided that $\varepsilon^2 m \to \infty$. This implies that a.a.s. $1 - \varepsilon - m/(2n) < X/m \leqslant 1 + \varepsilon$. □

Proof of Theorem 12. We split the proof into two uneven parts.

Stacks and waves: Fix $\varepsilon > 0$ and choose $\varepsilon' > 0$ such that $\frac{1-2\varepsilon'}{1+\varepsilon'} = 1 - \varepsilon$. Let M be the sub-matching of $\mathbb{RM}_n$ consisting of all edges with one end in $[n]$ and the other in $[2n] \setminus [n]$. Set $X := |M|$. We have $\mathbb{E}X = n^2/(2n-1) \sim n/2$ and

$$\mathbb{E}(X(X-1)) = \frac{n^2(n-1)^2}{(2n-1)(2n-3)} \sim (\mathbb{E}X)^2,$$

so, using Chebyshev's inequality, it follows that a.a.s. $|X - n/2| = o(n)$.

Observe that there are no lines in M longer than 1. By Proposition 11, a.a.s. there are no stacks or waves in $\mathbb{RM}_n$ of size $k_0 = \lfloor (1+o(1))e\sqrt{n/2} \rfloor$. Hence, by applying Theorem 1 to M with $\ell = 1$, $s = k_0 - 1$, and $w = \lfloor \frac{X-1}{k_0-1} \rfloor$, we conclude that a.a.s. there is a wave in M, and thus in $\mathbb{RM}_n$, of size at least

$$w + 1 = \frac{1 - o(1)}{e\sqrt{2}}\sqrt{n}.$$

By swapping the roles of waves and stacks in the above argument, we deduce that a.a.s. there is also a stack in M, and thus in $\mathbb{RM}_n$, of size $\frac{1-o(1)}{e\sqrt{2}}\sqrt{n}$.

Lines: Let $m = \lfloor \sqrt{n}/2 \rfloor$. By Lemma 13, a.a.s. the number of edges of length at most m in $\mathbb{RM}_n$ is at least $\sqrt{n}/4$. We will show that among the edges of length at most k, there are a.a.s. at most $\sqrt{n}/8$ pairs forming crossings or nestings. After removing one edge from each crossing and nesting we obtain a line of size at least $\sqrt{n}/4 - \sqrt{n}/8 = \sqrt{n}/8$.

For a 4-element subset $S = \{u_1, u_2, v_1, v_2\} \subset [2n]$ with $u_1 < v_1 < u_2 < v_2$, let X_S be an indicator random variable equal to 1 if both $\{u_1, u_2\} \in \mathbb{RM}_n$ and $\{v_1, v_2\} \in \mathbb{RM}_n$, that is, if S spans a crossing in $\mathbb{RM}_n$. Clearly,

$$\Pr(X_S = 1) = \frac{1}{(2n-1)(2n-3)}.$$

Let $X = \sum X_S$, where the summation is taken over all sets S as above and such that $u_2 - u_1 \leqslant m$ and $v_2 - v_1 \leqslant m$. Note that this implies that $v_1 - u_1 \leqslant m - 1$. Let $f(n,m)$ denote the number of terms in this sum. We have

$$f(n,m) \leqslant \left(2n(m-1) - \binom{m}{2}\right)\binom{m}{2} \leqslant \left(nm - \frac{1}{2}\binom{m}{2}\right) m^2,$$

as we have at most $2n(m-1) - \binom{m}{2}$ choices for u_1 and v_1 (see the proof of Lemma 13 with m replaced by $m-1$) and, once u_1, v_1 have been selected, at most $\binom{m}{2}$ choices of u_2, and v_2. It is easy to see that $f(n,m) = \Omega(nm^3)$. (In fact, one could show that $f(n,m) \sim \frac{2}{3}nm^3$, but we do not care about optimal constants here.) Hence, $\mathbb{E}X = \Omega(m^3/n) \to \infty$, while

$$\mathbb{E}X = \sum_S \mathbb{E}X_S = \frac{f(n,m)}{(2n-1)(2n-3)} \leqslant \frac{m^3}{4n} = \frac{1}{32}\sqrt{n}.$$

To apply Chebyshev's inequality, we need to estimate $\mathbb{E}(X(X-1))$, which can be written as

$$\mathbb{E}(X(X-1)) = \sum_{S,S'} \Pr\left(\{\{u_1, v_1\}, \{u_2, v_2\}, \{u_1', v_1'\}, \{u_2', v_2'\}\} \subset \mathbb{RM}_n\right),$$

where the summation is taken over all (ordered) pairs of sets $S = \{u_1, u_2, v_1, v_2\} \subset [2n]$ with $u_1 < v_1 < u_2 < v_2$ and $S' = \{u_1', u_2', v_1', v_2'\} \subset [2n]$ with $u_1' < v_1' < u_2' < v_2'$ such that $u_2 - u_1 \leqslant m$, $v_2 - v_1 \leqslant m$, $u_2' - u_1' \leqslant m$, and $v_2' - v_1' \leqslant m$. We split the above sum into two sub-sums Σ_1 and Σ_2 according to whether $S \cap S' = \emptyset$ or $|S \cap S'| = 2$ (for all other options the above probability is zero). In the former case,

$$\Sigma_1 \leqslant \frac{f(n,m)^2}{(2n-1)(2n-3)(2n-5)(2n-7)} = (\mathbb{E}X)^2(1 + O(1/n)).$$

In the latter case, the number of such pairs (S, S') is at most $f(n,m) \cdot 4m^2$, as given S, there are four ways to select the common pair and at most m^2 ways to select the remaining two vertices of S'. Thus,

$$\Sigma_2 \leqslant \frac{f(n,m) \cdot 4m^2}{(2n-1)(2n-3)(2n-5)} = O(m^5/n^2) = O(\sqrt{n})$$

and, altogether,

$$\mathbb{E}(X(X-1)) \leqslant (\mathbb{E}X)^2(1+O(1/n)) + O(\sqrt{n}) = (\mathbb{E}X)^2 + O(\sqrt{n}).$$

By Chebyshev's inequality,

$$\begin{aligned}\Pr(|X-\mathbb{E}X| \geqslant \mathbb{E}X) &\leqslant \frac{\mathbb{E}(X(X-1)) + \mathbb{E}X - (\mathbb{E}X)^2}{(\mathbb{E}X)^2} \\ &\leqslant 1 + O(1/\sqrt{n}) + \frac{1}{\mathbb{E}X} - 1 = O(1/\sqrt{n}) \to 0.\end{aligned}$$

Thus, a.a.s. $X \leqslant 2\mathbb{E}X \leqslant \sqrt{n}/16$.

We deal with nestings in a similar way. For a 4-element subset $S = \{u_1, u_2, v_1, v_2\} \subset [2n]$ with $u_1 < v_1 < v_2 < u_2$, let Y_S be an indicator random variable equal to 1 if both $\{u_1, u_2\} \in \mathbb{RM}_n$ and $\{v_1, v_2\} \in \mathbb{RM}_n$, that is, if S spans a nesting in $\mathbb{RM}_n$. Further, let $Y = \sum Y_S$, where the summation is taken over all sets S as above and such that $u_2 - u_1 \leqslant m$ and (consequently) $v_2 - v_1 \leqslant m-2$. It is crucial to observe that, again, $\mathbb{E}Y \leqslant m^3/n = \sqrt{n}/32$. Indeed, this time there are at most $2nm - \binom{m+1}{2}$ choices for u_1 and u_2 and, once u_1, u_2 have been selected, at most $\binom{m-2}{2}$ choices of v_1, and v_2, while the probability of both pairs appearing in $\mathbb{RM}_n$ remains the same as before. The remainder of the proof goes mutatis mutandis.

We conclude that a.a.s. the number of crossings and nestings of length at most m in $\mathbb{RM}_n$ is at most $\sqrt{n}/8$ as was required. □

We close this section with a straightforward generalization of Proposition 11 to random r-uniform ordered matchings, $r \geqslant 2$. Let $\mathbb{RM}_n^{(r)}$ be a random (ordered) r-matching of size n, that is, a matching picked uniformly at random out of the set of all $\alpha_n^{(r)} := (rn)!/(n!(r!)^n)$ matchings on the set $[rn]$.

Proposition 14. *Let $(M_k^{(r)})_1^\infty$ be a sequence of ordered r-matchings of size k, $k = 1, 2, \dots$. Then, a.a.s.*

$$\max\{k : \; M_k^{(r)} \preceq \mathbb{RM}_n^{(r)}\} \leqslant (1+o(1))\frac{e}{r}(r!n)^{1/r}.$$

Proof. Let $k_0 = \lfloor (1+n^{-1/(r+1)})\frac{e}{r}(r!n)^{1/r} \rfloor$ and, for each $k \geqslant k_0$, let $X_k^{(r)}$ be the number of ordered copies of $M_k^{(r)}$ of size k in $\mathbb{RM}_n^{(r)}$. Then

$$\mathbb{E}X_k^{(r)} = \binom{rn}{rk} \cdot 1 \cdot \frac{\alpha_{n-k}^{(r)}}{\alpha_n^{(r)}} = \frac{(r!)^k}{(rk)!} \cdot \frac{n!}{(n-k)!} \leqslant \frac{(r!)^k}{(rk)!} \cdot n^k \leqslant \frac{(r!)^k}{(rk/e)^{rk}} \cdot n^k = \left(\frac{e^r r! n}{(rk)^r}\right)^k,$$

which implies that

$$\Pr(\exists k \geqslant k_0 : \; X_k^{(r)} > 0) \leqslant \sum_{k=k_0}^{n} \left(\frac{e^r r! n}{(rk)^r}\right)^k \leqslant n(1+n^{-1/(r+1)})^{-rk_0} = o(1). \qquad \square$$

In particular, the above statement yields that in $\mathbb{RM}_n^{(3)}$ a.a.s. none of the nine collectable sub-matchings (as defined in Subsection 2.2) has size bigger than $2n^{1/3}$. It is, at the moment, an open problem to match it with a fair lower bound (see discussion in Section 4).

3 Twins

Recall that for $r \geqslant 2$, by *r-twins* in a matching M we mean any collection of disjoint, pairwise order-isomorphic sub-matchings $M_1, M_2, \ldots, M_r$. For instance, the matching $M = AABCDDEBCFGHIHEGFI$ contains 3-twins formed by the triple $M_1 = BDDB$, $M_2 = EHHE$, and $M_3 = FGGF$ (see Figure 10).

Recall also that by the *size* of r-twins we mean the size (the number of edges) in just one of them. Let $t_r(M)$ denote the maximum size of r-twins in a matching M and $t_r^{\text{match}}(n)$ – the minimum of $t_r(M)$ over all matchings M on $[2n]$.

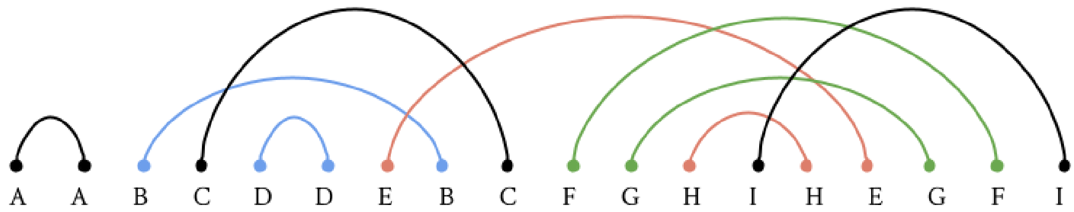

Figure 10: 3-twins of size two.

3.1 In arbitrary matchings

Let us first point to a direct connection between twins in permutations and ordered twins in a certain kind of matchings. By a *permutation* we mean any finite sequence of pairwise distinct positive 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 $1 \leqslant i < j \leqslant k$. For $r \geqslant 2$, any r pairwise similar and disjoint sub-permutations of a permutation π are called *r-twins*. For example, in permutation

$$(6, 1, 4, 7, 3, 9, 8, 2, 5),$$

the red and blue subsequences form a pair of twins, or 2-twins, of length 3, both similar to the permutation $(1, 3, 2)$.

Let $t_r(\pi)$ denote the maximum length of r-twins in a permutation π and $t_r^{\text{perm}}(n)$ – the minimum of $t_r(\pi)$ over all permutations π of $[n]$. Gawron [19] proved that $t_2^{\text{perm}}(n) \leqslant cn^{2/3}$ for some constant $c > 0$. This was easily generalized to $t_r^{\text{perm}}(n) \leqslant c_r n^{r/(2r-1)}$ (see [10]).

As for a lower bound, notice that by the Erdős–Szekeres Theorem, we have $t_r^{\text{perm}}(n) \geqslant \left\lfloor \frac{1}{r} n^{1/2} \right\rfloor$. For $r = 2$, this bound was substantially improved by Bukh and Rudenko [8].

Theorem 15 (Bukh and Rudenko [8])**.** *For all n, $t_2^{\text{perm}}(n) \geqslant \frac{1}{8} n^{3/5}$.*

Using their approach, in [12], we generalized this bound to arbitrary $r \geqslant 3$.

Theorem 16 ([12])**.** *For all n and $r \geqslant 3$, $t_r^{\text{perm}}(n) \geqslant \frac{1}{3r} n^{\frac{R}{2R-1}}$, where $R = \binom{2r-1}{r}$.*

We call an ordered matching M on the set $[2n]$ *permutational* if the left end of each edge of M lies in the set $[n]$ (and so, the right end of each edge lies in $[n+1, 2n]$. In the double occurrence word notation such a matching can be written as $M = A_1A_2 \dots A_nA_{i_1}A_{i_2} \dots A_{i_n}$, where $\pi_M = (i_1, i_2, \dots, i_n)$ is a permutation of $[n]$. There are only $n!$ permutational matchings, nevertheless this connection to permutations turned out to be quite fruitful. Indeed, it is not hard to see that ordered r-twins in a permutational matching M are in one-to-one correspondence with r-twins in π_M. In particular, we have $t_r(M) = t_r(\pi_M)$ for such matchings and, consequently, $t_r^{\text{match}}(n) \leqslant t_r^{\text{perm}}(n)$. In particular, by the above mentioned result of Gawron, $t_2^{\text{match}}(n) = O(n^{2/3})$.

More subtle is the opposite relation.

Proposition 17. *For all $r \geqslant 2$ and $1 \leqslant m \leqslant n$, where $n - m$ is even,*

$$t_r^{\text{match}}(n) \geqslant \min\left\{t_r^{\text{perm}}(m),\ 2t_r^{\text{match}}\left(\frac{n-m+2}{2}\right)\right\}.$$

Proof. Let M be a matching on $[2n]$. Split the set of vertices of M into two halves, $A = [n]$ and $B = [n+1, 2n]$ and let $M' := M(A, B)$ denote the set of edges of M with one end in A and the other end in B. Note that M' is a permutational matching. We distinguish two cases. If $|M'| \geqslant m$, then

$$t_r(M) \geqslant t_r(M') = t_r(\pi_{M'}) \geqslant t_r^{\text{perm}}(|M'|) \geqslant t_r^{\text{perm}}(m).$$

If, on the other hand, $|M'| < m$, that means, $|M'| \leqslant m-2$, due to the assumption of $n-m$ being even, then we have sub-matchings M_A and M_B of M of size at least $(n-m+2)/2$ in sets, respectively, A and B. Thus, in this case, by concatenation,

$$t_r(M) \geqslant t_r(M_A) + t_r(M_B) \geqslant 2t_r^{\text{match}}\left(\frac{n-m+2}{2}\right). \qquad \square$$

Proposition 17 allows, under some mild conditions, to "carry over" any lower bound on $t_r^{\text{perm}}(n)$ to one on $t_r^{\text{match}}(n)$. In view of Theorems 15 and 16, as well as the upper bound on $t_r^{\text{perm}}(n)$ from [10] mentioned above, we may assume that the parameter α introduced in the lemma below falls between $R/(2R-1)$ and $r/(2r-1)$.

Lemma 18. *For all $r \geqslant 2$, $R/(2R-1) \leqslant \alpha \leqslant r/(2r-1)$, and $\beta > 0$, if $t_r^{\text{perm}}(n) \geqslant \beta n^\alpha$ for all $n \geqslant r$, then $t_r^{\text{match}}(n) \geqslant \beta(n/4)^\alpha$ for all $n \geqslant r$.*

Proof. Assume that for some $r \geqslant 2$ and α, β as above, $t_r^{\text{perm}}(n) \geqslant \beta n^\alpha$, for all $n \geqslant r$. We will prove that $t_r^{\text{match}}(n) \geqslant \beta(n/4)^\alpha$ by induction on n. For $n \leqslant 4\left(\frac{1}{\beta}\right)^{1/\alpha}$ the claimed bound is at most 1, so it is trivially true. Assume then that $n \geqslant \left(\frac{1}{\beta}\right)^{1/\alpha}$ and that $t_r^{\text{match}}(n') \geqslant \beta(\gamma n')^\alpha$ for all $r \leqslant n' < n$. Let $n' \in \{\lceil n/4\rceil, \lceil n/4\rceil + 1\}$ have the same parity as n. Then, by Proposition 17 with $m = n'$,

$$t_r^{\text{match}}(n) \geqslant \min\left\{t_r^{\text{perm}}(n'), 2t_r^{\text{match}}\left(\frac{n-n'+2}{2}\right)\right\}.$$

Now, by the assumption of the lemma applied to $n = r$, noticing that $t_r^{\text{match}}(r) = 1$, we have that $1 \geqslant \beta r^\alpha$, or $\beta \leqslant r^{-\alpha}$, so $n' \geqslant n/4 \geqslant r$. Thus, by the assumption of the lemma applied to n', $t_r^{\text{perm}}(n') \geqslant \beta(n/4)^\alpha$. Further, noticing that $n - n' + 2 \geqslant 3n/4$, again by the induction assumption, we also have

$$2t_r^{\text{match}}\left(\frac{n-n'+2}{2}\right) \geqslant 2\beta\left(\frac{1}{4}\cdot\frac{3}{8}n\right)^\alpha \geqslant \beta(n/4)^\alpha,$$

where the last inequality is equivalent to $3 \geqslant 2^{3-1/\alpha}$ which, in turn, follows by estimating the R-H-S by $2^{3-(2r-1)/r} \leqslant 2^{3-3/2} = 2^{3/2}$. □

In particular, Theorem 15 and Lemma 18 with $\beta = 1/8$, $\alpha = 3/5$, and $\gamma = 1/4$ imply immediately the following result.

Corollary 19. *For every n, $t_2^{\text{match}}(n) \geqslant \frac{1}{8}\left(\frac{n}{4}\right)^{3/5}$.* □

Moreover, any future improvement of the bound in Theorem 15 would automatically yield a corresponding improvement of the lower bound on $t_2^{\text{match}}(n)$.

3.2 In random matchings

In this section we study the size of the largest r-twins in a random (ordered) matching $\mathbb{RM}_n$, which is, recall, selected uniformly at random from all $\alpha_n := (2n)!/(n!2^n)$ matchings on vertex set $[2n]$. The first moment method yields that

$$\text{a.a.s.} \quad t_r(\mathbb{RM}_n) < cn^{r/(2r-1)} \quad \text{for any} \quad c > e2^{-(r-1)/(2r-1)}. \tag{1}$$

Indeed, the expected number of r-twins of size k in $\mathbb{RM}_n$ is

$$\frac{1}{r!}\binom{2n}{\underbrace{2k,\dots,2k}_{r\text{ times}},2n-2kr}\frac{\alpha_k\cdot 1^{r-1}\cdot\alpha_{n-rk}}{\alpha_n} = \frac{2^{(r-1)k}n!}{r!(2k)!^{r-1}k!(n-rk)!} < \left(\frac{e^{2r-1}n^r}{2^{r-1}k^{2r-1}}\right)^k,$$

where we used inequalities $n!/(n-rk)! \leqslant n^{rk}$, $(2k)! \geqslant (2k/e)^{2k}$, $k! \geqslant (k/e)^k$, and $r! \geqslant 1$. Thus, it converges to 0, as $n \to \infty$, provided $k \geqslant cn^{r/(2r-1)}$, for any c as in (1).

It turns out that the a.a.s. the lower bound on $t_r(\mathbb{RM}_n)$ is of the same order.

Theorem 20. *For every $r \geqslant 2$, a.a.s.,*

$$t_r(\mathbb{RM}_n) = \Theta\left(n^{r/(2r-1)}\right).$$

In the proof of the lower bound we are going to use the Azuma-Hoeffding inequality for random permutations (see, e.g., Lemma 11 in [16] or Section 3.2 in [20]). Let us recall that Π_n stands for a *random* permutation selected uniformly from all $(2n)!$ permutations of the set $[2n]$.

Theorem 21. *Let $h(\pi)$ be a function defined on the set of all permutations of order $2n$ such that, for some constant $c > 0$, if a permutation π_2 is obtained from a permutation π_1 by swapping two elements, then $|h(\pi_1) - h(\pi_2)| \leqslant c$. Then, for every $\eta > 0$,*

$$\Pr(|h(\Pi_n) - \mathbb{E}[h(\Pi_n)]| \geqslant \eta) \leqslant 2\exp(-\eta^2/(4c^2 n)).$$

Proof of Theorem 20. In view of (1), it suffices to prove a lower bound, that is, to show that a.a.s. $\mathbb{RM}_n$ contains r-twins of size $\Omega(n^{r/(2r-1)})$. In doing so we are following the proof scheme applied in [12, Theorem 1.2]. Set

$$a := r!^{1/(2r-1)}(2n)^{(r-1)/(2r-1)}$$

and assume for simplicity that both, a and $N := 2n/a = r!^{-1/(2r-1)}(2n)^{r/(2r-1)}$, are integers. Partition $[2n] = A_1 \cup \cdots \cup A_N$, where A_i's are consecutive blocks of a integers each, and define, for every $1 \leqslant i < j \leqslant N$, a random variable X_{ij} which counts the number of edges of $\mathbb{RM}_n$ with one endpoint in A_i and the other in A_j. Consider an auxiliary graph $G := G(\mathbb{RM}_n)$ on vertex set $[N]$ where $\{i,j\} \in G$ if and only if $X_{ij} \geqslant r$.

A crucial observation is that a matching of size t in G corresponds to r-twins in $\mathbb{RM}_n$ of size t. Indeed, let $M = \{i_1 j_1, \ldots, i_t j_t\}$, $i_1 < \cdots < i_t$, be a matching in G. For every $1 \leqslant s \leqslant t$, let $u_h^s \in A_{i_s}$ and $v_h^s \in A_{j_s}$, $h = 1, \ldots, r$, be such that $e_h^s := \{u_h^s, v_h^s\} \in \mathbb{RM}_n$. Then, the sub-matchings $\{e_1^1, \ldots, e_1^t\}, \ldots, \{e_r^1, \ldots, e_r^t\}$, owing to the sequential choice of A_i's form r-twins in $\mathbb{RM}_n$. Thus, our ultimate goal is to show that a.a.s. G contains a matching of size $\Theta(n^{r/(2r-1)})$.

Let $\nu(G)$ be the largest size (as the number of edges) of a matching in G. We will first show that $\nu(G)$ is sharply concentrated around its expectation. For this we appeal to the permutation scheme of generating $\mathbb{RM}_n$ and apply Theorem 21. For a permutation π of $[2n]$, let $M(\pi) = \{\pi(1)\pi(2), \ldots, \pi(2n-1)\pi(2n)\}$ be the corresponding matching. Further, let $h(\pi) = \nu(G(M(\pi)))$. 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 edges of $G(M(\pi_1))$. This, in turn, implies that the size of the largest matching has been altered by at most two, that is, $|h(\pi_1) - h(\pi_2)| \leqslant 2$. Hence, Theorem 21 applied to $h(\pi)$, with $c = 2$, and, say, $\eta = n^{2r/(4r-1)}$ implies that

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

Note that, crucially, $n^{2r/(4r-1)} = o(n^{r/(2r-1)})$, and it thus remains to show that $\mathbb{E}(\nu(G)) = \Omega(n^{r/(2r-1)})$.

Let us first estimate from below the probability of an edge in G, that is, $\Pr(X_{ij} \geqslant r)$. Trivially, $\Pr(X_{ij} \geqslant r) \geqslant \Pr(X_{ij} = r)$. We are going to further bound $\Pr(X_{ij} = r)$ from below by counting matchings on $[2n]$ with precisely r edges between the sets A_i and A_j, but with no edges within A_i or A_j (the latter is a simplifying restriction). To build such a matching one has to first select subsets S_i, S_j, T_i, T_j such that $S_i \subset A_i$, $S_j \subset A_j$, $|S_i| = |S_j| = r$, while $T_i \cap (A_i \cup A_j) = \emptyset$, $T_j \cap (A_i \cup A_j) = \emptyset$, $T_i \cap T_j = \emptyset$, and

$|T_i| = |T_j| = a - r$. The total number of these selections is

$$\binom{a}{r}^2 \binom{2n-2a}{a-r} \binom{2n-3a+r}{a-r}.$$

Then one has to match S_i with S_j, $A_i \setminus S_i$ with T_i, and $A_j \setminus S_j$ with T_j, and find a perfect matching on the set $[2n] \setminus (A_i \cup A_j \cup T_i \cup T_j)$ of the remaining $n - 4a + 2r$ elements. There are

$$r!(a-r)!(a-r)!\alpha_{n-2a+r}$$

ways of doing so.

Multiplying the two products together and dividing by α_n, we obtain the inequality

$$\Pr(X_{ij} = r) \geqslant \frac{a!^2 4^a (2n-2a)!n!}{r!2^r(a-r)!^2(2n)!(n-2a+r)!}.$$

Since $a = o(\sqrt{n})$, we have

$$(2n)!/(2n-2a)! = (1+o(1))(2n)^{2a} \quad \text{and} \quad n!/(n-2a+r)! = (1+o(1))n^{2a-r},$$

and so the R-H-S above equals

$$\frac{(1+o(1))}{r!}\left(\frac{a^2}{2n}\right)^r.$$

Consequently, for n large enough,

$$\Pr(X_{ij} \geqslant r) \geqslant \Pr(X_{ij} = r) \geqslant \frac{1}{2r!}\left(\frac{a^2}{2n}\right)^r.$$

Having estimated the probability of an edge in G, we are now in position to estimate the degree of a vertex. For each $i \in [N]$, let $Y_i = \deg_G(i)$ be the degree of vertex i in G. Then

$$\mathbb{E}(Y_i) = (N-1)\Pr(X_{ij} \geqslant r) \geqslant \frac{N}{4r!}\left(\frac{a^2}{2n}\right)^r = \frac{2n}{4r!a}\left(\frac{a^2}{2n}\right)^r = \frac{1}{4}. \tag{3}$$

There is an obvious bound on the size of the largest matching $\nu(G)$ in $G = (V, E)$ in terms of the vertex degrees, namely

$$\nu(G) \geqslant \frac{|E(G)|}{2\Delta_G} = \frac{\sum_{i=1}^N Y_i}{4\Delta_G},$$

where Δ_G is the maximum degree in G. Note that, trivially, $\Delta_G \leqslant \min\{a/r, N-1\} = a/r$. Unfortunately, since the expected degrees $\mathbb{E}(Y_i)$ are (bounded by) constants – cf. (3), in this form the bound on $\nu(G)$ is of no use, as one cannot show concentration of all degrees Y_i simultaneously. Instead, we resort to an even weaker, but more manageable bound.

For an integer $D > 0$, let G_D be a subgraph of G induced by the set V_D of all vertices of degrees at most D in G, that is, $G_D = G[V_D]$. Then, clearly, $\Delta(G_D) \leqslant D$ and

$$|E(G_D)| = |E(G)| - |\{e \in E(G) :\ e \cap (V \setminus V_D) \neq \emptyset\}| \geqslant \frac{1}{2}\sum_{i=1}^{N} Y_i - \sum_{k=D+1}^{\lfloor a/r \rfloor} kZ_k,$$

where $Z_k = |\{i \in [N] : Y_i = k\}|$, and thus,

$$\nu(G) \geqslant \nu(G_D) \geqslant \frac{|E(G_D)|}{2D} \geqslant \frac{1}{2D}\left(\frac{1}{2}\sum_{i=1}^{N} Y_i - \sum_{k>D} kZ_k\right).$$

Hence, recalling (3) and noticing that $\mathbb{E}Z_k = N\Pr(Y_1 = k)$, we have

$$\mathbb{E}(\nu(G)) \geqslant \frac{N}{2D}\left(\frac{1}{8} - \sum_{k>D} k\Pr(Y_1 = k)\right).$$

It remains to estimate $\Pr(Y_1 = k)$ from above. Very crudely, to create a matching satisfying $Y_1 = k$, one has to select k other sets A_i, choose r vertices from each of them, and match them with some kr vertices of A_1. In the estimates below, we ignore the demand that Y_1 is precisely k, so, in fact, we estimate from above $\Pr(Y_1 \geqslant k)$. We thus have

$$\Pr(Y_1 = k) \leqslant \Pr(Y_1 \geqslant k) \leqslant \binom{N}{k}\binom{a}{r}^k\binom{a}{rk}(rk)!\frac{\alpha_{n-rk}}{\alpha_n} \leqslant \frac{N^k a^{2rk} 2^{rk} n^{rk}}{k! r!^k} \cdot \frac{(2n-2rk)!}{(2n)!}.$$

Using the inequality $1 - x \geqslant e^{-2x}$ valid for $x \leqslant 1/2$, the last fraction can be estimated as

$$\frac{(2n-2rk)!}{(2n)!} = \frac{1}{(2n)^{2rk}\left(1-\frac{1}{2n}\right)\cdot\ldots\cdot\left(1-\frac{2rk-1}{2n}\right)} \leqslant \frac{e^{2r^2k^2/n}}{(2n)^{2rk}}.$$

Since $a^{2r-1} = r!n^{r-1}$ and $k \leqslant a/r$, we have $k^2 = o(n)$ and can bound, roughly, $e^{2r^2k^2/n} \leqslant 2$. Consequently, recalling that $N = 2n/a$ and using the bound $(k-1)! \geqslant ((k-1)/3)^{k-1}$, we infer that

$$k\Pr(Y_1 = k) \leqslant \frac{2a^{2rk-k}}{(k-1)!r!^k(2n)^{rk-k}} = \frac{2}{(k-1)!} \leqslant 2\left(\frac{3}{k-1}\right)^{k-1} \leqslant 2\left(\frac{1}{2}\right)^{k-1}$$

for $k \geqslant 7$. Setting $D = 6$ and summing over all $k > D$, we thus obtain the bound

$$\sum_{k>D}^{\lfloor a/r \rfloor} k\Pr(Y_1 = k) \leqslant 2\sum_{k>6}^{\infty}\left(\frac{1}{2}\right)^{k-1} = 2\sum_{k\geqslant 6}^{\infty}\left(\frac{1}{2}\right)^{k} = 4 \cdot 2^{-6} = \frac{1}{16}.$$

Finally,

$$\mathbb{E}(\nu(G)) \geqslant \frac{N}{12}\left(\frac{1}{8} - \frac{1}{16}\right) \geqslant \frac{N}{200} = \Theta\left(n^{r/(2r-1)}\right)$$

which, together with (2), completes the proof. □

4 Final remarks

Let us conclude the paper with some suggestions for future studies. The first natural goal is to extend our Erdős–Szekers-type results to r-uniform ordered matchings for arbitrary $r \geqslant 4$. We have inspected more carefully the case of $r = 4$ observing some similar phenomena as in the two smaller cases, $r = 2$ and $r = 3$. In particular, among the 35 possible relations between pairs of ordered quadruples there are exactly 27 that are collectable. This boosts some hope that by cleverly handling the remaining 8 relations, one can, indeed, generalize Theorem 5 with the exponents $1/9$ replaced by $1/27$.

Returning to the case $r = 3$, we are not yet completely done, as we were unable to construct a general counterexample showing the optimality of Theorem 5 for arbitrary values of *all* 9 parameters (c.f. Remark 7).

Problem 22. For all positive integers a_{XY}, where $X, Y \in \{\mathtt{L}, \mathtt{S}, \mathtt{W}\}$, $(X, Y) \neq (\mathtt{L}, \mathtt{L})$, construct a matching M of size $n = \prod_{X,Y} a_{xy}$ such that neither M contains a semi-line of size 2, nor for any pair (X, Y) does it contain a sub-matching of $a_{XY} + 1$ triples every two of which are in relation $\mathcal{R}_{XY}$.

As mentioned earlier, setting $a_{\mathtt{LL}} = 1$ is not a restriction at all, as in the general case one may simply concatenate $a_{\mathtt{LL}}$ disjoint copies of the matching described in Problem 22

A related problem is to estimate the size of unavoidable patterns in random r-matchings $\mathbb{RM}_n^{(r)}$ with $r \geqslant 3$. For $r = 2$, we have established (c.f. Theorem 12) that it is a.a.s. of the order $\Theta(\sqrt{n})$. As it was already mentioned at the end of Section 2.3, as a consequence of Proposition 14, for $r = 3$ a.a.s. none of the nine collectable sub-matchings (as defined in Subsection 2.2) has size bigger than $2n^{1/3}$. It would be nice to prove a complementary lower bound. Although for 3-uniform lines this does not seem to be difficult (as, likely, Lemma 13 can be extended for hyperedges), showing that a.a.s. $\mathbb{RM}_n^{(3)}$ contains *every* collectable sub-matching of size $\Omega(n^{1/3})$ will require some new ideas.

For arbitrary r, it too seems natural to expect that, as in the case $r = 2$, all homogeneous substructures (corresponding to collectable relations) of $\mathbb{RM}_n^{(r)}$ should a.a.s. have size $\Theta(n^{c_r})$, for some constant $0 < c_r < 1$. By Theorem 12 we know that $c_2 = 1/2$, while guided by Proposition 14, we suspect that $c_r = 1/r$.

Other open problems can be formulated for twins in ordered matchings. Based on what we proved here we state the following conjecture.

Conjecture 23. For every fixed $r \geqslant 2$, we have $t_r^{\mathrm{match}}(n) = \Theta\left(n^{\frac{r}{2r-1}}\right)$.

The same statement is conjectured for twins in permutations (see [12]), and, by our results, we know that both conjectures are actually equivalent.

One could also study the size of twins in r-uniform ordered matchings or, more generally, in arbitrary ordered graphs or hypergraphs. For ordinary *unordered* graphs there is a result of Lee, Loh, and Sudakov [25] giving an asymptotically exact answer of order $\Theta(n \log n)^{2/3}$. It would be nice to have an analogue of this result for ordered graphs.

Acknowledgements

The first author was supported in part by Simons Foundation Grant #522400. The second author was supported in part by Narodowe Centrum Nauki, grant 2020/37/B/ST1/03298. The third author was supported in part by Narodowe Centrum Nauki, grant 2018/29/B/ST1/00426.

References

[1] M. Axenovich, Y. Person, and S. Puzynina, *A regularity lemma and twins in words*, J. Combin. Theory Ser. A **120** (2013), no. 4, 733–743.

[2] J. Baik and E. M. Rains, *The asymptotics of monotone subsequences of involutions*, Duke Math. J. **109** (2001), no. 2, 205–281.

[3] M. Balko, J. Cibulka, K. Král, and J. Kynčl, *Ramsey numbers of ordered graphs*, Electronic J. Combin. **27** (2020), no. 1, #P1.16, 32.

[4] I. Bárány, G. Kalai, and A. Pór, *Erdős–Szekeres theorem for k-flats*, Discrete Comput. Geom. **69** (2023), 1232–1240.

[5] J. Barát, A. Gyárfás, and G. Tóth, *Monochromatic spanning trees and matchings in ordered complete graphs*, J. Graph Theory **105** (2024), no. 4, 523–541.

[6] M. Bucić, B. Sudakov, and T. Tran, *Erdős-Szekeres theorem for multidimensional arrays*, J. Eur. Math. Soc. (JEMS) **25** (2023), no. 8, 2927–2947.

[7] B. Bukh and J. Matoušek, *Erdős–Szekeres–type statements: Ramsey function and decidability in dimension 1*, Duke Math. J. **163** (2014), 2243–2270.

[8] B. Bukh and O. Rudenko, *Order-isomorphic twins in permutations*, SIAM J. Discrete Math. **34** (2020), no. 3, 1620–1622.

[9] D. Conlon, J. Fox, C. Lee, and B. Sudakov, *Ordered Ramsey numbers*, J. Combin. Theory Ser. B **122** (2017), 353–383.

[10] A. Dudek, J. Grytczuk, and A. Ruciński, *Variations on twins in permutations*, Electron. J. Combin. **28** (2021), no. 3, #P3.19, 18.

[11] ———, *Patterns in ordered (random) matchings*, LATIN 2022: theoretical informatics, Lecture Notes in Comput. Sci., vol. 13568, Springer, Cham, 2022, pp. 544–556.

[12] ———, *Multiple twins in permutations*, Austral. J. Combin. **87** (2023), no. 3, 440–451.

[13] M. Eliáš and J. Matoušek, *Higher–order Erdős–Szekeres theorems*, Adv. Math. **244** (2013), 1–15.

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

[15] J. Fox, J. Pach, B. Sudakov, and A. Suk, *Erdős–Szekeres–type theorems for monotone paths and convex bodies*, Proc. Lond. Math. Soc. **105** (2012), 953–982.

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

[17] Z. Füredi, T. Jiang, A. Kostochka, D. Mubayi, and J. Verstraëte, *Extremal problems for convex geometric hypergraphs and ordered hypergraphs*, Canad. J. Math. **73** (2021), no. 6, 1648–1666.

[18] C. F. Gauss, *Werke*, Georg Olms Verlag, Hildesheim-New York, 1976, Ergänzungsreihe. IV. Briefwechsel C. F. Gauss–H. W. M. Olbers, I, Nachdruck der 1900 Auflage, herausgegeben von C. Schilling.

[19] M. Gawron, *Izomorficzne podstruktury w słowach i permutacjach*, Master Thesis (2014).

[20] M. Habib, C. McDiarmid, J. Ramirez-Alfonsin, and B. Reed (eds.), *Probabilistic methods for algorithmic discrete mathematics*, Algorithms and Combinatorics, vol. 16, Springer-Verlag, Berlin, 1998.

[21] T. Huynh, F. Joos, and P. Wollan, *A unified Erdős-Pósa theorem for constrained cycles*, Combinatorica **39** (2019), no. 1, 91–133.

[22] V. Jelínek, *Dyck paths and pattern-avoiding matchings*, European J. Combin. **28** (2007), no. 1, 202–213.

[23] J. Justicz, E. R. Scheinerman, and P. M. Winkler, *Random intervals*, Amer. Math. Monthly **97** (1990), no. 10, 881–889.

[24] A. Kasraoui and J. Zeng, *Distribution of crossings, nestings and alignments of two edges in matchings and partitions*, Electron. J. Combin. **13** (2006), no. 1, #R33, 12.

[25] C. Lee, P.-S. Loh, and B. Sudakov, *Self-similarity of graphs*, SIAM J. Discrete Math. **27** (2013), no. 2, 959–972.

[26] G. Moshkovits and A. Shapira, *Ramsey theory, integer partitions and a new proof of the Erdős–Szekeres Theorem*, Adv. Math. **262** (2014), 1107–1129.

[27] G. Puttenham, *The arte of english poesie*, London, 1589.

[28] B. Shtylla, L. Traldi, and L. Zulli, *On the realization of double occurrence words*, Discrete Math. **309** (2009), no. 6, 1769–1773.

[29] R. P. Stanley, *Increasing and decreasing subsequences and their variants*, International Congress of Mathematicians. Vol. I, Eur. Math. Soc., Zürich, 2007, pp. 545–579.

[30] ———, *Catalan numbers*, Cambridge University Press, New York, 2015.

[31] T. Szabó and G. Tardos, *A multidimensional generalization of the Erdős–Szekeres lemma on monotone subsequences*, Combin. Prob. Comput. **10** (2001), 557–565.

[32] G. Tardos, *Extremal theory of ordered graphs*, Proceedings of the International Congress of Mathematicians—Rio de Janeiro 2018. Vol. IV. Invited lectures, World Sci. Publ., Hackensack, NJ, 2018, pp. 3235–3243.

[33] F. Whigham and W. A. Rebhorn, *The art of english poesy by George Puttenham*, Cornell University Press, 2007.