

# EXTREMAL, ENUMERATIVE AND PROBABILISTIC RESULTS ON ORDERED HYPERGRAPH MATCHINGS

MICHAEL ANASTOS, ZHIHAN JIN, MATTHEW KWAN, AND BENNY SUDAKOV

ABSTRACT. An *ordered  $r$ -matching* is an  $r$ -uniform hypergraph matching equipped with an ordering on its vertices. These objects can be viewed as natural generalisations of  *$r$ -dimensional orders*. The theory of ordered 2-matchings is well-developed and has connections and applications to extremal and enumerative combinatorics, probability, and geometry. On the other hand, in the case  $r \geq 3$  much less is known, largely due to a lack of powerful bijective tools. Recently, Dudek, Grytczuk and Ruciński made some first steps towards a general theory of ordered  $r$ -matchings, and in this paper we substantially improve several of their results and introduce some new directions of study. Many intriguing open questions remain.

## 1. INTRODUCTION

A *matching* is a graph with the property that every vertex is incident to exactly one edge (equivalently, a matching is a partition of the vertex set into pairs). Given a partition of the vertex set into two equal-size parts  $V_1, V_2$ , we say that a matching is *bipartite* if every edge is between the two parts.

Matchings are fundamental objects in graph theory (and beyond), where one is usually interested in the existence and identification of matchings inside larger graphs (see for example the monograph of Lovász and Plummer [48]). However, matchings can be very interesting objects in their own right, if one puts an *ordering* on the set of vertices (for example, we could restrict our attention to matchings on vertex sets of the form  $\{1, \dots, 2n\}$ , and consider the natural ordering of the integers).

In particular, in an ordered matching there are three different ways that a pair of distinct edges  $e, e'$  can “interact with each other”, as follows (write  $e[1] < e[2]$  for the two vertices of  $e$ , write  $e'[1] < e'[2]$  for the two vertices of  $e'$ , and assume without loss of generality that  $e[1] < e'[1]$ ).

- We could have  $e[1] < e[2] < e'[1] < e'[2]$  (that is to say, one edge fully comes before the other one). This configuration is called an *alignment*.
- We could have  $e[1] < e'[1] < e[2] < e'[2]$  (that is to say, the two edges are interleaved with each other). This configuration is called a *crossing*.
- We could have  $e[1] < e'[1] < e'[2] < e[2]$  (that is to say, one of the two edges is “within” the other). This configuration is called a *nesting*.

It is a classical result (first proved by Errera [26]) that the crossing-free matchings on  $\{1, \dots, 2n\}$  are enumerated by the Catalan numbers  $C_n$ , and an ingenious bijection (see [57]) shows that there are also exactly  $C_n$  nesting-free matchings on  $\{1, \dots, 2n\}$ . Note that a matching  $M$  on  $\{1, \dots, 2n\}$  is alignment-free if and only if every edge is between  $\{1, \dots, n\}$  and  $\{n+1, \dots, 2n\}$  (i.e., if  $M$  is a bipartite matching with these two parts). Such matchings are in correspondence with permutations  $\sigma \in \mathcal{S}_n$ , and there are therefore  $n!$  of them.

Two of the most important parameters of a permutation  $\sigma \in \mathcal{S}_n$  are the length  $L_{\nearrow}(\sigma)$  of its longest increasing subsequence and the length  $L_{\searrow}(\sigma)$  of its longest decreasing subsequence (see for example [53, 56] for surveys on the study of these parameters). Two of the highlights in this area are the *Erdős–Szekeres theorem* [25], which says that we always have  $L_{\nearrow}(\sigma) \geq \sqrt{n}$  or  $L_{\searrow}(\sigma) \geq \sqrt{n}$  (i.e., it is not possible to simultaneously avoid long decreasing subsequences and long increasing subsequences), and the *Robinson–Schensted–Knuth correspondence* [19, 44, 54] between permutations and Young tableaux, which can be used to enumerate permutations  $\sigma \in \mathcal{S}_n$  by their values of  $L_{\nearrow}(\sigma)$  and  $L_{\searrow}(\sigma)$  (and in particular, to study the behaviour of  $L_{\nearrow}(\sigma)$  and  $L_{\searrow}(\sigma)$  for a *random* permutation  $\sigma \in \mathcal{S}_n$ ).

---

Michael Anastos is supported by the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 101034413. Matthew Kwan is supported by ERC Starting Grant “RANDSTRUCT” No. 101076777, also funded by the European Union . Zhihan Jin and Benny Sudakov are supported by SNSF grant 200021\_1969.

Recalling that permutations  $\sigma \in \mathcal{S}_n$  are in correspondence with bipartite matchings between  $\{1, \dots, n\}$  and  $\{n+1, \dots, 2n\}$ , it turns out that increasing and decreasing subsequences can be described in the language of configurations in matchings: an increasing subsequence corresponds to a set of edges which are pairwise crossing (which we call a *crossing-clique*), and a decreasing subsequence corresponds to a set of edges which are pairwise nesting (which we call a *nesting-clique*).

Extending the huge body of work on increasing and decreasing subsequences, there has been quite some work (see for example [5, 6, 17, 39, 41, 43, 45, 56]), studying nesting-cliques and crossing-cliques (and alignment-cliques, which have the obvious definition) in general matchings. It turns out that many of the techniques that are effective for permutations have natural analogues for general matchings (in particular, there is a variant of the Robinson–Schensted–Knuth correspondence relating matchings to *oscillating tableaux*; see [56]).

Very recently, Dudek, Grytczuk and Ruciński [21, 22] made some first steps towards extending the theory of ordered matchings to ordered *hypergraph* matchings (in an  $r$ -uniform hypergraph matching, every edge contains exactly  $r$  vertices, and every vertex is incident to exactly one edge). The jump from graphs to hypergraphs seems to introduce a number of serious difficulties, which should perhaps not be surprising: in much the same way that an ordered matchings generalise permutations, ordered  $r$ -uniform hypergraph matchings generalise  $(r-1)$ -tuples of permutations (which are sometimes called “ $r$ -dimensional orders”, as they can be described by sets of points in  $r$ -dimensional space). When  $r \geq 3$ , there is no known analogue of the Robinson–Schensted–Knuth correspondence for  $r$ -dimensional orders, and there are a number of longstanding open problems (see for example the survey [12]).

In this paper we substantially extend and improve the results in [21, 22], discovering some surprisingly intricate phenomena and moving towards a more complete theory of ordered hypergraph matchings. We also draw attention to a large number of compelling open problems.

**1.1. Ordered hypergraph matchings: basic notions.** We say that a (hyper)graph  $H$  is *ordered* if its vertex set  $V(H)$  is equipped with a total order (one may think of graphs and hypergraphs which have vertex set  $\{1, \dots, N\}$  for some  $N$ ).

**Definition 1.1.** An ordered hypergraph is said to be an  $r$ -*matching* if every edge has exactly  $r$  vertices, and every vertex is contained in exactly one edge. An  $r$ -matching is said to be  $r$ -*partite* if, when we divide the vertex set into  $r$  contiguous intervals of equal length, every edge of the matching has exactly one vertex in each interval. The *size* of a matching  $M$  is its number of edges.

Note that there are

$$\frac{(rn)!}{(r!)^n n!} \tag{1.1}$$

different  $r$ -matchings on the vertex set  $\{1, \dots, rn\}$ . The  $r$ -partite  $r$ -matchings on  $\{1, \dots, rn\}$  are in correspondence with  $(r-1)$ -tuples of permutations  $(\sigma_1, \dots, \sigma_{r-1}) \in \mathcal{S}_n^{r-1}$ , and there are therefore exactly  $(n!)^{r-1}$  of them.

**Definition 1.2.** An  $r$ -*pattern* is an  $r$ -matching of size 2 (on the vertex set  $\{1, \dots, 2r\}$ , say). We can represent an  $r$ -pattern by a string of “A”s and “B”s starting with “A” (where the vertices from one edge are represented with “A”, and the vertices of the other edge are represented with “B”).

Note that there are exactly  $(2 \cdot 2)! / (2^2 \cdot 2) = 3$  different 2-patterns: the alignment (represented by AABB), the crossing (represented by ABAB) and the nesting (represented by ABBA). The crossing and nesting are 2-partite, but the alignment is not (in fact, as we discussed earlier in the introduction, avoidance of the alignment pattern is equivalent to 2-partiteness).

**Definition 1.3.** For an  $r$ -pattern  $P$ , an  $r$ -matching  $M$  is said to be a  $P$ -*clique* if every pair of edges of  $M$  are order-isomorphic to  $P$ . For any  $r$ -pattern  $P$  and  $r$ -matching  $M$ , let  $L_P(M)$  be the size of the largest  $P$ -clique in  $M$ .

In total, there are  $(2r)! / (2(r!)^2)$  different  $r$ -patterns, but there is a subtlety that occurs only for uniformities  $r \geq 3$ : for some  $r$ -patterns  $P$  it is simply not possible to have a large  $P$ -clique.

**Definition 1.4.** We say that an  $r$ -pattern  $P$  is *collectable* if there are arbitrarily large  $P$ -cliques.

For example, it is easy to see that the 3-pattern AABABB is not collectable. In fact, Dudek, Grytczuk and Ruciński [22] showed for every non-collectable pattern, the largest possible clique size is 2. To this end, they gave a characterisation of the collectable patterns: they are precisely those patterns that are *splittable*, as follows.

**Definition 1.5.** A *run* in a pattern  $P$  is a sequence of consecutive vertices in the same edge of  $P$ . A pattern  $P$  is *splittable* if it can be partitioned into *blocks* of consecutive vertices, each consisting of two runs of the same length coming from different edges of  $P$ . If  $P$  is splittable, then this partition is uniquely determined; we call it the *block partition* of  $P$ .

For example, the pattern AABBABBA is splittable (the divisions between the blocks are described by |AABB|AB|BA| and the block partition has parts  $\{1, 2, 3, 4\}$ ,  $\{5, 6\}$  and  $\{7, 8\}$ ). It is easy to check that there are exactly  $3^{r-1}$  different collectable  $r$ -patterns, exactly  $2^{r-1}$  of which are  $r$ -partite (note that the  $r$ -partite  $r$ -patterns are precisely the collectable  $r$ -patterns whose block partition has  $r$  blocks).

For every collectable  $r$ -pattern  $P$ , there is only one way to form a  $P$ -clique on a given set of vertices. For example, in the case  $r = 2$ , the only way to form an alignment-clique on  $\{1, \dots, 2n\}$  is with the edges

$$\{1, 2\}, \{3, 4\}, \dots, \{2n - 1, 2n\},$$

the only way to form a crossing-clique is with the edges

$$\{1, n + 1\}, \{2, n + 2\}, \dots, \{n, 2n\},$$

and the only way to form a nesting-clique is with the edges

$$\{1, 2n\}, \{2, 2n - 1\}, \dots, \{n, n + 1\}.$$

We are now ready to present the main results of the paper. To briefly summarise: in [Section 1.2](#) we consider Ramsey-type questions (generalising the Erdős–Szekeres theorem), in [Section 1.3](#) we study the size of the largest  $P$ -clique in a *random* ordered matching, and in [Section 1.4](#) we give estimates for the number of ordered hypergraph matchings avoiding  $P$ -cliques of a given size. Throughout, we discuss a number of auxiliary results we prove along the way; in particular, in [Section 1.5](#) we discuss some contributions to the extremal theory of ordered hypergraphs, which we believe to be of independent interest. In [Section 1.6](#) we present a large number of open problems and directions for further study.

**1.2. Ramsey-type questions.** Recall that the Erdős–Szekeres theorem says that every permutation  $\sigma \in \mathcal{S}_n$  has an increasing or decreasing subsequence of length at least  $\sqrt{n}$  (and it is easy to construct a permutation showing that this is best-possible, by taking a certain type of “product” of an increasing sequence of length  $\sqrt{n}$  with a decreasing sequence of length  $\sqrt{n}$ ). The Erdős–Szekeres theorem falls under the umbrella of *Ramsey theory*: no matter how “disordered” a permutation is, it must have a long subsequence which is “completely homogeneous”.

In [21], Dudek, Grytczuk and Ruciński adapted the Erdős–Szekeres theorem to ordered (2-uniform) matchings. They showed that in every 2-matching  $M$  of size  $n$ , one can find an alignment-clique, a crossing-clique or a nesting-clique of size at least  $n^{1/3}$ . They also showed that this result is best-possible, with the same type of product construction used to demonstrate optimality of the Erdős–Szekeres theorem.

To discuss the situation for higher-uniformity hypergraphs, we introduce some notation.

**Definition 1.6.** Let  $L(M) = \max_P L_P(M)$  be the size of the largest clique (of any pattern) in  $M$ , and let  $L_r(n)$  be the minimum value of  $L(M)$  among all ordered  $r$ -matchings of size  $n$ .

In this notation, the aforementioned result of Dudek, Grytczuk and Ruciński [21] is  $L_2(n) = \lceil n^{1/3} \rceil$ .

As first observed by Burkill and Mirsky [15] and Kalmanson [40], it is actually very easy to iterate the Erdős–Szekeres theorem to obtain an optimal bound for  $r$ -partite  $r$ -matchings: every  $r$ -partite  $r$ -matching  $M$  of size  $n$  has  $L(M) \geq n^{1/2^{r-1}}$ , and a product construction shows that this is best possible<sup>1</sup>. It is not quite so obvious how to prove a bound on  $L_r(n)$  by iterating the bound  $L_2(n) \geq n^{1/3}$ , but in a follow-up paper [22], Dudek, Grytczuk and Ruciński showed that this is indeed possible if one is willing

<sup>1</sup>In the language of permutations, this can be equivalently formulated as the fact that in any  $(r - 1)$ -tuple of length- $n$  permutations there is a set of at least  $n^{1/2^{r-1}}$  indices in which each of our permutations is monotone. This is one of many different ways to generalise the Erdős–Szekeres theorem to “higher dimensions”; for example, see also [13, 14, 28, 33, 46, 59].

to give up a constant factor: they proved the general lower bound  $L_r(n) \geq c_r n^{1/3^{r-1}}$  (for some  $c_r > 0$  depending on  $r$ ).

Naïvely, this seems to suggest that  $L_r(n)$  is of order  $n^{1/3^{r-1}}$  (viewing  $r$  as a constant, while  $n$  is large). However, it is not clear how to prove a corresponding upper bound using the usual “product-type” constructions, because cliques can interact with each other in surprisingly intricate ways. Via product-type constructions, Dudek, Grytczuk and Ruciński were only able to prove the upper bound  $L_r(n) \leq n^{1/(2^{r-1}+2)}$ .

As our first result, we significantly improve both the lower and upper bounds for  $L_r(n)$ , showing that the exponent scales roughly like  $1/2^r$ .

**Theorem 1.7.** *For  $r \geq 2$ , we have*

$$\frac{1}{r-1} \cdot n^{1/((r+1)2^{r-2})} \leq L_r(n) \leq \lceil n^{1/(2^r-1)} \rceil.$$

The upper bound in [Theorem 1.7](#) comes from a product-like construction, and the lower bound in our proof of [Theorem 1.7](#) is proved via an argument that partitions the possible  $r$ -patterns into roughly  $2^r$  subsets with “poset-like” structure. It is possible to push this method further, obtaining slightly stronger lower bounds by considering more intricate partitions of the  $r$ -patterns, and we are actually not sure what the limit of the method is. As an illustration, with some intricate combinatorial analysis we managed to use our method to obtain the correct order of magnitude of  $L_3(n)$  and  $L_4(n)$ .

**Theorem 1.8.** *We have*

$$\frac{1}{2} \cdot n^{1/7} \leq L_3(n) \leq n^{1/7}, \quad \frac{1}{4} \cdot n^{1/15} \leq L_4(n) \leq n^{1/15}.$$

In [Section 3](#), we describe our strategy to prove lower bounds, and show how to use it to prove the general lower bound on  $L_r(n)$  in [Theorem 1.7](#) and the sharp lower bound on  $L_3(n)$  in [Theorem 1.8](#). The proof of the sharp lower bound on  $L_4(n)$  follows a similar approach but has much more complicated casework; we defer it to [Appendix A](#). In [Section 4](#) we prove the upper bound in [Theorem 1.7](#) (which also implies the upper bounds in [Theorem 1.8](#)).

**1.3. Random matchings.** One of the most notorious problems in the theory of permutations is the *Ulam–Hammersley problem*, to describe the distribution of the longest increasing permutation  $L_{\succ}(\sigma)$  in a random permutation  $\sigma \in \mathcal{S}_n$ . This problem was famously resolved by Baik, Deift and Johansson [4], but one of the most important milestones along the way was a theorem of Logan and Shepp [47] and Vershik and Kerov [61] (see also the alternative proofs [1, 35, 37, 55]), establishing that the expected value of  $L_{\succ}(\sigma)$  is asymptotically  $2\sqrt{n}$ .

These results were extended to random matchings by Baik and Rains [6] (see also [5, 56]). Specifically, they proved that in a random matching  $M$  on  $\{1, \dots, 2n\}$ , the expected values of  $L_{\text{crossing}}(M)$  and  $L_{\text{nesting}}(M)$  are both  $(\sqrt{2} + o(1))\sqrt{n}$  (as part of a tour-de-force where they found the asymptotic distribution of these quantities).

All this work uses variations on the Robinson–Schensted–Knuth correspondence (or related bijective tools), and does not easily generalise to hypergraphs (or to higher dimensions). For example, if  $M$  is a random  $r$ -partite  $r$ -matching, then for any  $r$ -partite  $r$ -pattern  $P$ , the random variable  $L_P(M)$  has the same distribution as the length of the longest common increasing subsequence in  $r-1$  independent random permutations  $\sigma_1, \dots, \sigma_{r-1} \in \mathcal{S}_n$ . This random variable is notoriously hard to study: Bollobás and Winkler [7] proved that its expected value is asymptotic to  $c_r n^{1/r}$  for some constant  $c_r > 0$ , but the value of  $c_r$  is unknown for all  $r \geq 3$ .

In [22] (improving their results in [21]), Dudek, Grytczuk and Ruciński proved that there are constants  $c'_r, c''_r \geq 0$  such that for any collectable  $r$ -pattern  $P$ , and a random  $r$ -matching  $M$ , we have  $c'_r n^{1/r} \leq L_P(M) \leq c''_r n^{1/r}$  whp<sup>2</sup>. They also conjectured that  $L_P(M)/n^{1/r}$  converges in probability, to a constant that only depends on  $r$ . However, this conjecture is already false for  $r = 2$ : via analysis of a simple renewal process, Justicz, Scheinerman and Winkler [39] proved that for a random matching  $M$  on  $\{1, \dots, 2n\}$ , the size of the largest alignment-clique is  $(2/\sqrt{\pi} + o(1))\sqrt{n}$  whp, while the work of Baik and Rains described

<sup>2</sup>We say that an event holds *with high probability*, or *whp* for short, if it holds with probability  $1 - o(1)$ . Here and for the rest of the paper, asymptotics are as  $n \rightarrow \infty$ , unless explicitly stated otherwise.

above implies that the largest crossing-clique and nesting-clique both have size  $(\sqrt{2} + o(1))\sqrt{n}$ . (This was likely missed by Dudek, Grytczuk and Ruciński on account of the result in [39] being stated in the language of *random interval graphs*, but the equivalence to random matchings is straightforward).

Combining a subadditivity argument with Talagrand’s concentration inequality, we are able to show that for a collectable pattern  $P$  and a random matching  $M$ , the random variable  $L_P(M)/n^{1/r}$  does converge to a limit. This limit may depend on  $P$ , but only through the *type* of  $P$ , as follows.

**Definition 1.9.** Let  $P$  be an  $r$ -pattern with block partition  $J_1 \cup \dots \cup J_\ell$ . The *type* of  $P$  is the partition  $|J_1|/2 + \dots + |J_\ell|/2$  of  $r$  (in the number-theoretic sense).

For example,  $|AABB|AB|BA|$  and  $|AB|AABB|AB|$  both have type  $2 + 1 + 1$ .

**Theorem 1.10.** Fix an  $r$ -pattern  $P$ , and let  $M$  be a random  $r$ -matching of size  $n$ . Then we have

$$\frac{L_P(M)}{n^{1/r}} \xrightarrow{P} b_P$$

for some  $b_P > 0$  depending only on the type of  $P$ .

Actually determining the constants  $b_P$  seems to be a hard problem (in particular, if  $P$  is an  $r$ -partite pattern then our proof shows that  $b_P$  can be expressed in terms of the constant  $c_r$  in the Bollobás–Winkler theorem described above). However, we are able to prove some bounds on the  $b_P$  in some special cases. In particular, we show that the conjecture of Dudek, Grytczuk and Ruciński is false for all  $r \geq 2$ . (Recall that the  $\Gamma$  function is the analytic continuation of the factorial function).

**Proposition 1.11.** Let the constants  $b_P$  be as in [Theorem 1.10](#).

- (1) If  $P$  has type  $r$  (i.e., if the block partition of  $P$  has a single block), then  $b_P = 1/\Gamma((r+1)/r)$ .
- (2) If  $P$  has type  $1 + \dots + 1$  (i.e., if  $P$  is  $r$ -partite), then  $b_P = c_r(r!)^{1/r}/r > 1/\Gamma((r+1)/r)$ , where  $c_r$  is the constant from the Bollobás–Winkler theorem mentioned earlier in this section.

We prove [Theorem 1.10](#) and [Proposition 1.11](#) in [Section 6](#), after proving a certain necessary generalisation of the Bollobás–Winkler theorem in [Section 5](#).

**1.4. Enumeration.** For an  $r$ -pattern  $P$ , let  $N_P(n)$  denote the number of ordered  $r$ -matchings on the vertex set  $\{1, \dots, rn\}$  which are  $P$ -free (i.e., no two edges form  $P$ ). First notice that if  $P$  is not  $r$ -partite then every  $r$ -partite matching is  $P$ -free. In particular,  $N_P(n)$  is at least the number of  $r$ -partite matchings of size  $n$ , which is exactly  $(n!)^{r-1}$ . In combination with [\(1.1\)](#) (and Stirling’s approximation) this is already enough to approximate  $N_P(n)$  up to exponential factors (for constant  $r$ ).

**Proposition 1.12.** Fix a constant  $r \in \mathbb{N}$ . If  $P$  is not  $r$ -partite then

$$N_P(n) = e^{O_r(n)} n^{(r-1)n}.$$

(In this paper, subscripts on asymptotic notation indicate quantities that should be viewed as constants: the constant factor implicit in “ $O_r(n)$ ” is allowed to depend on  $r$ ).

The case where  $P$  is  $r$ -partite is more delicate (and the value of  $N_P(n)$  is quite different), but we are able to obtain estimates of similar quality, as follows.

**Theorem 1.13.** Fix a constant  $r \in \mathbb{N}$ . If  $P$  is  $r$ -partite then

$$N_P(n) = e^{O_r(n)} n^{(r-1-1/(r-1))n}.$$

*Remark 1.14.* In the case  $r = 2$  there are only two possibilities for an  $r$ -partite pattern  $P$ : a crossing  $|AB|AB|$  or a nesting  $|AB|BA|$ . As discussed in the introduction,  $N_P(n)$  is *exactly* the same in both cases (in particular, there is a well-known bijection between crossing-free and nesting-free matchings of a given size). However, this does *not* generalise straightforwardly to higher dimensions: a computer search shows that if  $P$  is the 3-uniform pattern represented by  $|AB|AB|AB|$ , then  $N_P(4) = 8626$ , whereas if  $P$  is the 3-uniform pattern represented by  $|AB|BA|BA|$  we have  $N_P(4) = 8630$ .

The lower bound in [Theorem 1.13](#) is a direct consequence of a result due to Brightwell [11] on linear extensions of  $r$ -dimensional posets, but the upper bound is new (it is obtained by a general connection to extremal ordered hypergraph theory).

We are also interested in enumerating ordered matchings by the size of their largest  $P$ -clique. We write  $N_{P,m}(n)$  for the number of ordered  $r$ -matchings  $M$  on the vertex set  $\{1, \dots, rn\}$  which satisfy

$L_P(M) < m$  (i.e., they do not contain a  $P$ -clique of size  $m$ ). Note that if  $P$  is not  $r$ -partite, then the same considerations as for [Proposition 1.12](#) show that  $N_{P,m}(n) = e^{O_r(n)} n^{(r-1)n}$  for all  $m$  (i.e., varying  $m$  can only affect  $N_{P,m}(n)$  by a factor of  $e^{O_r(n)}$ ). However, if  $P$  is  $r$ -partite, we get a significant dependence on  $m$ , as per the following generalisation of [Theorem 1.13](#) (note that [Theorem 1.13](#) corresponds to the case  $m = 2$ ).

**Theorem 1.15.** *Fix a constant  $r \in \mathbb{N}$ . If  $P$  is  $r$ -partite then for any  $2 \leq m \leq n^{1/r}$  we have*

$$N_{P,m}(n) = e^{O_r(n)} (m-1)^{(r/(r-1))n} n^{(r-1-1/(r-1))n}.$$

We remark that the condition  $m \leq n^{1/r}$  in [Theorem 1.15](#) is not just an artifact of the proof: recall from [Section 1.3](#) that for almost every size- $n$  ordered matching  $M$  we have  $L_P = O(n^{1/r})$ .

We prove the upper bound in [Theorem 1.15](#) via a general lemma ([Theorem 7.1](#)) which estimates  $N_{P,m}(n)$  in terms of a certain *extremal* parameter (namely, the maximum number of edges in an ordered  $r$ -uniform hypergraph on  $\lceil n/2 \rceil$  vertices with  $L_P(M) < m$ ). This type of reduction goes back to Alon and Friedgut [2] (see also [27, 42]); in particular, we adapt a proof in a similar high-dimensional situation due to Cibulka and Kyncl [18]. We believe that the relevant extremal parameter is of independent interest, and we discuss it further in [Section 1.5](#).

For the lower bound in [Theorem 1.15](#), we obtain a new estimate on the number of  $(r-1)$ -tuples of length- $n$  permutations which have no common increasing subsequence of length  $m$  (see [Remark 7.6](#)), by “reverse-engineering” the techniques in the upper bound. This lower bound, and the ideas in its proof, may be of independent interest (in particular, our approach is very different to the probabilistic approach of Brightwell [11] for the case  $m = 2$ ).

We prove [Theorem 1.15](#) (which implies [Theorem 1.13](#)) in [Section 7](#).

**1.5. Extremal results.** As briefly mentioned in [Section 1.4](#), in our study of enumerative questions for ordered matchings we encounter some extremal problems for ordered hypergraphs. We believe our extremal results to be of independent interest so we take a moment to discuss them here. First, we define *ordered extremal numbers*, which are variants of the classical *extremal numbers* (or *Turán numbers*) for unordered graphs.

**Definition 1.16.** Let  $G, F$  be ordered  $r$ -uniform hypergraphs (ordered  $r$ -graphs, for short). We say  $G$  is  $F$ -free if it contains no subgraph isomorphic to  $F$  (where the isomorphism must preserve the order of the vertices). Let  $\text{ex}_{<}(n, F)$  denote the maximum number of edges in an  $F$ -free  $n$ -vertex ordered  $r$ -graph.

In the case  $r = 2$  (i.e., the case of graphs), ordered extremal numbers have been extensively studied (see the survey by Tardos [60] and the references within). For general  $r \geq 2$ , much less is known, though there is literature on similar problems for *cyclically* ordered hypergraphs, motivated by geometric considerations (in particular, due to applications to *convex geometric hypergraphs*; see for example [3, 8–10, 16, 29, 30, 32, 34, 51]), and much of this work has implications for ordered extremal numbers  $\text{ex}_{<}(n, F)$ .

Our first contribution in this direction is that we are able to nail down the exact value  $\text{ex}_{<}(n, P)$  for any  $r$ -partite  $r$ -pattern  $P$  (this extremal parameter is relevant for [Theorem 1.13](#)). For convenience, we define  $(x)_+ := \max(x, 0)$ .

**Theorem 1.17.** *Let  $r, n \geq 1$  and  $P$  be any  $r$ -partite  $r$ -pattern. Then,*

$$\text{ex}_{<}(n, P) = \binom{n}{r} - \binom{(n-r)_+}{r}.$$

We remark that in the case where  $P$  is the  $r$ -uniform “generalised crossing” pattern represented by  $|AB|AB| \cdots |AB|$ , the result of [Theorem 1.17](#) was already known, thanks to recent work of Füredi, Jiang, Kostochka, Mubayi and Verstraëte [30]. (It was also known in the case where  $P$  is a 2-uniform nesting, as we will discuss after the next theorem).

For an  $r$ -pattern  $P$  and a positive integer  $m$ , we use  $P^{(m)}$  to denote the  $P$ -clique of size  $m$  on the vertex set  $[rm]$  (so  $P^{(2)} = P$ ). We are not quite able to pin down the values of the ordered extremal numbers  $\text{ex}_{<}(n, P^{(m)})$ , but we are able to obtain some quite strong bounds (which are an ingredient in our proof of [Theorem 1.15](#)).

**Theorem 1.18.** *Let  $r, n, m \geq 1$  and let  $P$  be an  $r$ -partite  $r$ -pattern.*

(1) In general, we have

$$\text{ex}_{<}(n, P^{(m)}) \geq \binom{n}{r} - \binom{(n - r(m - 1))_+}{r}.$$

(2) If  $n$  is sufficiently large (in terms of  $r$ ) then

$$\text{ex}_{<}(n, P^{(m)}) \leq O\left(r^2(m - 1) \binom{n}{r - 1}\right).$$

(3) If  $P$  is the "alternating"  $r$ -pattern represented by  $|AB|BA|AB|BA|\dots|$ , then

$$\text{ex}_{<}(n, P^{(m)}) = \binom{n}{r} - \binom{(n - r(m - 1))_+}{r}.$$

*Remark 1.19.* If we view  $r$  as a constant, and assume  $m = o_r(n)$ , then one can check that

$$\binom{n}{r} - \binom{(n - r(m - 1))_+}{r} = (1 + o(1))r(m - 1) \binom{n}{r - 1}.$$

So, our upper and lower bounds in (1) and (2) differ by a factor of  $O(r)$ .

The  $r = 2$  case of [Theorem 1.18\(3\)](#) (i.e., an exact result in the case where  $P$  is the 2-uniform nesting  $|AB|BA|$ ) can be derived from the known results on graphs with bounded *queue number* due to Pemmaraju [\[52\]](#) and Dujmović and Wood [\[23\]](#). We also remark that Capoleas and Pach [\[16\]](#) previously studied the case where  $P = |AB|AB|$  is a 2-uniform crossing; in this case they obtained the exact result

$$\text{ex}_{<}(n, P^{(m)}) = \binom{n}{2} - \binom{(n - 2(m - 1))_+}{2}.$$

For general  $r \geq 2$ , Füredi, Jiang, Kostochka, Mubayi and Verstraëte [\[30\]](#) studied the case where  $P$  is the  $r$ -uniform "generalised crossing" pattern represented by  $|AB|AB|\dots|AB|$ . We have already mentioned their exact result in the case  $m = 2$ ; for all  $m$  they proved

$$(1 - o_{r,m}(1))r(m - 1) \binom{n}{r - 1} \leq \text{ex}_{<}(n, P^{(m)}) \leq 2(r - 1)(m - 1) \binom{n}{r - 1}. \quad (1.2)$$

(Recall that subscripts on asymptotic notation indicate quantities that should be viewed as constants, so the " $o_{r,m}(1)$ " term goes to zero as  $n \rightarrow \infty$ , holding  $r, m$  fixed). Their lower bound construction is the same as our construction for (1) (in the case where  $P$  is a generalised crossing), but they only analysed its number of edges asymptotically. Their upper bound for this specific case is stronger than our general bound in (2).

A variety of techniques are involved in the proof of [Theorem 1.17](#) and the various parts of [Theorem 1.18](#). In particular, for our proof of [Theorem 1.18\(2\)](#) we refine a partitioning lemma for ordered hypergraphs due to Füredi, Jiang, Kostochka, Mubayi and Verstraëte [\[31\]](#). We present this refinement, and discuss it further, in [Section 8](#).

The proofs of [Theorems 1.17](#) and [1.18](#) appear in [Section 9](#).

**1.6. Further directions.** There are a great number of compelling further directions for study.

**1.6.1. Ramsey-type questions.** The most obvious open question in this direction is to close the gap between the lower and upper bound in [Theorem 1.7](#). It is tempting to guess that the upper bound is sharp, and that  $L_r(n)$  has order of magnitude  $n^{1/(2^r - 1)}$  for any constant  $r$  (this is true for  $r \in \{2, 3, 4\}$ ). It is even possible that such a bound can be proved via our "poset partitioning" method, with judicious choice of the relevant partitions (perhaps for small  $r$ , appropriate partitions can be found via computer search). However, our investigations do not suggest any general structural reason why a bound of this form should hold.

Even in the cases  $r \in \{2, 3, 4\}$  where we know the order of magnitude of  $L_r(n)$ , it would be interesting to obtain the exact value (or at least an asymptotic estimate). Also, it is possible to consider "off-diagonal" versions of the problem, where one treats different patterns differently. For example, fixing a positive integer  $k_P$  for each  $r$ -pattern  $P$ , what is the maximum possible size of an  $r$ -matching which has no  $P$ -clique of size  $k_P$  for any  $P$ ? Actually, this general setting was considered in [\[21, 22\]](#), and optimal bounds were obtained in the case  $r = 2$ . Our methods seem to be suitable for attacking the larger- $r$  case

as well, but the relevant casework seems likely to be even more fiddly than for [Theorem 1.8](#) (one would need to completely characterise how all the different subsets of patterns can interact with each other).

**1.6.2. Random ordered matchings.** The obvious open question in this direction is to determine the values of the constants  $b_P$  in [Theorem 1.10](#). As we have discussed, this is a difficult problem in the case where  $P$  is  $r$ -partite (in which case it amounts to determination of the Bollobás–Winkler constant  $c_r$ ), but it may be tractable for certain special  $P$ .

For example, in the case where  $P$  is a 3-pattern which has type  $2 + 1$ , the problem of determining  $b_{2+1} := b_P$  seems like it might be of “intermediate difficulty” between the Ulam–Hammersley problem (of determining  $c_2 = 2$ ) and the problem of determining the Bollobás–Winkler constant  $c_3$ . For the Ulam–Hammersley problem, there is a well-known interacting particle process (*Hammersley’s interacting particle process*) which captures the limiting behaviour of the longest increasing subsequence of a random permutation, and Aldous and Diaconis [1] managed to give a fairly “soft” proof that  $c_2 = 2$  by studying this process. It seems that one can design a variant of Hammersley’s process which is suitable for studying  $b_{2+1}$ , though this process lacks certain symmetries that Aldous and Diaconis used in their analysis.

**1.6.3. Enumeration.** The results in [Section 1.4](#) have exponential error terms, and of course it would be interesting to sharpen these. However, there may be some limitations to what is possible to accomplish without determining the limits  $b_P$  in [Theorem 1.10](#): in the regime where  $m$  is of order  $n^{1/r}$ , studying  $N_{P,m}(n)$  amounts to studying the large deviations of  $L_P(M)$  in a random  $r$ -matching  $M$ .

Even if exponential error terms cannot be eliminated, it might be interesting to determine their dependence on  $r$  (recall that in [Theorems 1.13](#) and [1.15](#) we treat  $r$  as a constant).

**1.6.4. Extremal problems.** Given the discussion in [Section 1.5](#), it is natural to conjecture that the lower bound in [Theorem 1.18\(1\)](#) is always sharp, as follows.

**Conjecture 1.20.** *Let  $r, n, m \geq 1$  and  $P$  be an  $r$ -partite  $r$ -pattern. Then,*

$$\text{ex}_{<}(n, P^{(m)}) = \binom{n}{r} - \binom{(n - r(m - 1))_+}{r}.$$

In particular, given [Theorem 1.18\(3\)](#), an indirect route to prove [Conjecture 1.20](#) would be to prove that  $\text{ex}_{<}(n, P^{(m)})$  is the same for every  $r$ -partite  $r$ -pattern  $P$ . In the case  $r = 2$ , there is a bijective proof of (a vast generalisation of) this fact, due to Jonsson and Welker [38] (see also [20, Corollary 2.5]). However, the considerations in [Remark 1.14](#) suggest some difficulties in generalising such bijective proofs to higher uniformities.

Also, although there is no obvious connection to enumeration, it may still be of interest to determine  $\text{ex}_{<}(n, P^{(m)})$  in the case where  $P$  is not  $r$ -partite. It is easy to show that  $\text{ex}_{<}(n, P^{(m)}) \geq (n/r)^r$  when  $n$  is a multiple of  $r$ , but this does not seem to be best-possible.

**1.7. Notation.** We use standard asymptotic notation throughout, as follows. For functions  $f = f(n)$  and  $g = g(n)$ , we write  $f = O(g)$  to mean that there is a constant  $C$  such that  $|f(n)| \leq C|g(n)|$  for sufficiently large  $n$ . Similarly, we write  $f = \Omega(g)$  to mean that there is a constant  $c > 0$  such that  $f(n) \geq c|g(n)|$  for sufficiently large  $n$ . We write  $f = \Theta(g)$  to mean that  $f = O(g)$  and  $f = \Omega(g)$ , and we write  $f = o(g)$  or  $g = \omega(f)$  to mean that  $f(n)/g(n) \rightarrow 0$  as  $n \rightarrow \infty$ . Subscripts on asymptotic notation indicate quantities that should be treated as constants.

For  $n, r \geq 1$ , we use  $K_n^{(r)}$  to indicate the ordered  $r$ -uniform clique on the vertex set  $\{1, \dots, n\}$ . For an ordered hypergraph  $H$ , we write  $V(H)$  for its vertex set,  $E(H)$  for its edge set and  $e(H) = |E(H)|$  for its number of edges. For  $e \in E(H)$ , we write  $e[1], \dots, e[r]$  for the vertices of  $e$  (ordered according to the vertex ordering of  $H$ ). When we say that an ordered hypergraph  $H'$  is a *subgraph* of another ordered hypergraph  $H$ , we mean that the order of the vertices is maintained. (We use the notation  $H' \subseteq H$ ).

We also remind the reader that  $(x)_+$  is defined to be  $\max(x, 0)$ .

## 2. PRELIMINARIES

We first recall the following corollary of Mirsky’s theorem [50], for our proof of [Theorem 1.7](#).

**Theorem 2.1.** *Suppose  $P$  is a partially ordered set (poset) with  $n$  elements, that contains no chain of length  $x$  for some  $x \in \mathbb{R}_+$ . Then,  $P$  contains an antichain of size at least  $n/x$ .*



At several points in the paper (related to the directions introduced in [Section 1.5](#)) we will need the following enumerative identity.

**Lemma 2.2.** *Let  $n \geq r \geq 1$  and  $\delta_1, \dots, \delta_{r-1} \in \mathbb{Z}_{\geq 0}$ . Then, the number of  $r$ -tuples  $(a_1, \dots, a_r) \in \{1, \dots, n\}^r$  satisfying  $a_{i+1} - a_i > \delta_i$  for all  $1 \leq i \leq r-1$  is exactly*

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

*Proof.* We may assume  $n \geq \sum_{k=1}^{r-1} \delta_k$ , or else the quantity under consideration is zero. For an  $r$ -tuple  $(a_1, \dots, a_r)$  as in the lemma statement, let  $f(a_1, \dots, a_r) = (b_1, \dots, b_r)$ , where  $b_1 = a_1$  and  $b_i = a_i - \delta_{i-1}$  for  $2 \leq i \leq r$ . The desired result follows by noting that  $f$  is a bijection between the set of tuples under consideration and the set of unordered  $r$ -tuples in  $\{1, 2, \dots, n - \sum_{k=1}^{r-1} \delta_k\}$ .  $\square$

For [Theorem 1.10](#) we need a number of probabilistic tools. First, we need a bounded-differences concentration inequality for permutations.

**Theorem 2.3.** *Let  $f : \mathcal{S}_n \rightarrow \mathbb{R}$  be a function which has “bounded differences” in the sense that for any  $\sigma \in \mathcal{S}_n$  and any transposition  $\tau \in \mathcal{S}_n$ , we have  $|f(\tau \circ \sigma) - f(\sigma)| \leq c$ . Then, if  $\sigma \in \mathcal{S}_n$  be a uniformly random permutation of length  $n$ , we have*

$$\Pr[|f(\sigma) - \mathbb{E}f(\sigma)| \geq t] \leq \exp\left(-\frac{2t^2}{c^2n}\right).$$

This inequality seems to have first been proved by McDiarmid in the case where  $c = 1$  (see [\[49, p. 18\]](#)<sup>3</sup>).

*Remark 2.4.* Note that one can obtain a random ordered  $r$ -matching on the vertex set  $1, \dots, rn$  via a random permutation  $\sigma \in \mathcal{S}_{rn}$ : simply take the matching whose edges are

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

We also need a version of Talagrand’s concentration inequality (see for example [\[36, Theorem 2.29\]](#) and [Equation \(2.43\)](#))

**Theorem 2.5.** *Consider a function  $X = f(Z_1, \dots, Z_n)$  of independent random objects  $Z_1 \in \Lambda_1, \dots, Z_n \in \Lambda_n$ . Suppose that the following conditions are satisfied.*

- *If  $z, z' \in \prod_{i=1}^n \Lambda_i$  differ only in the  $i$ th coordinate, then  $|f(z) - f(z')| \leq 1$ .*
- *For any  $z \in \prod_{i=1}^n \Lambda_i$ , there is a subset of indices  $J \subseteq \{1, \dots, n\}$  with  $|J| \leq f(z)$ , such that for any  $y \in \prod_{i=1}^n \Lambda_i$  which agrees with  $z$  on the coordinates indexed by  $J$ , we have  $f(y) \geq f(z)$ .*

*Then, for some universal constant  $\gamma > 0$ , we have*

$$\Pr[|X - \mathbb{E}X| \geq t] \leq 4 \exp\left(-\frac{\gamma t^2}{\mathbb{E}X + t}\right)$$

*for any  $t \geq 0$ .*

We will also need a version of Kingman’s subadditive ergodic theorem (see for example [\[53, Theorem A.3\]](#)). We have “flipped” the conditions to make it apply under a *superadditivity* condition rather than a subadditivity condition.

**Theorem 2.6.** *Let  $(X_{m,n})_{0 \leq m < n}$  be a family of nonnegative random variables, defined on some common probability space, such that the following conditions are satisfied.*

- *$X_{0,n} \geq X_{0,m} + X_{m,n}$  for all  $m < n$ .*
- *For any  $k \geq 1$ ,  $(X_{nk,(n+1)k})_{n=0}^{\infty}$  is a sequence of i.i.d. random variables.*
- *For any  $m \geq 1$ , we have  $(X_{0,k})_{k=0}^{\infty} \stackrel{d}{=} (X_{m,m+k})_{k=0}^{\infty}$ .*
- *There exists a constant  $M > 0$  such that  $\mathbb{E}X_{0,n} \leq Mn$  for all  $n$ .*

*Then,  $\mathbb{E}X_{0,n}/n$  converges to a limit  $\gamma$  as  $n \rightarrow \infty$ . Also,  $X_{0,n}/n$  converges almost surely (therefore in probability) to  $\gamma$ .*

<sup>3</sup>There is a typo in this book; in the inequality we are citing, the exponent should really be  $2t^2/n$ , not  $2t^2/n^2$ .

### 3. LOWER BOUNDS ON RAMSEY PARAMETERS

In this section we will outline our general framework for proving lower bounds on  $L_r(n)$ , and show how to use it to prove the lower bounds in [Theorem 1.7](#) and in the  $r = 3$  case of [Theorem 1.8](#). The  $r = 4$  case of [Theorem 1.8](#) follows a similar strategy but has some very complicated casework, so we defer it to [Appendix A](#).

The starting point is that patterns sometimes give rise to posets. For an  $r$ -matching  $M$  and an  $r$ -pattern  $P$ , we define the relation  $\preceq_P$  on the edges of  $M$  by taking  $e \preceq f$  if  $e[1] < f[1]$  and  $e, f$  form pattern  $P$  (or if  $e = f$ ). It is sometimes (but not always!) the case that  $\preceq_P$  is a partial order. The interesting property here is *transitivity*: we need to know whether  $e \preceq_P f$  and  $f \preceq_P g$  implies  $e \preceq_P g$ . For example, in the case  $r = 2$ , the only  $r$ -patterns are alignments, crossings and nestings. It is easy to see that the alignment and nesting pattern always give rise to a poset, but the crossing pattern may not (it is possible to have edges  $e, f, g$  with  $e[1] < f[1] < g[1]$  such that  $e$  and  $f$  form a crossing, and  $f$  and  $g$  form a crossing, but  $e$  and  $g$  form an alignment). However, if we restrict ourselves to matchings that are alignment-free, then the crossing pattern *does* give rise to a poset.

Now, if  $\preceq_P$  is a poset, then Mirsky's theorem ([Theorem 2.1](#)) implies that there is a long chain (which corresponds to a large  $P$ -clique) or a large antichain (which corresponds to a  $P$ -free sub-matching). One can iterate this, in several different ways, to give a simple proof of the optimal bound  $L_2(n) \geq n^{1/3}$  (previously proved with a different method in [\[21\]](#)). For example, in a size- $n$  matching  $M$ , first, we look for an alignment-clique of size  $n^{1/3}$  or an alignment-free sub-matching  $M'$  of size  $n^{2/3}$ , and in the latter case, inside  $M'$  we look for a nesting-clique of size  $n^{1/3}$  or a nesting-free sub-matching (therefore a crossing-clique) of size  $n^{1/3}$ .

For general  $r$ , if one is careful about the order of operations, one can show  $L_r(n) \geq \Omega_r(n^{1/3^{r-1}})$  (as in [\[22, Theorem 1.3\]](#)) by generalising the above proof, applying Mirsky's theorem once for each of the  $3^{r-1}$  collectable  $r$ -patterns (one also needs some separate arguments to handle the non-collectable patterns).

The authors of [\[22\]](#) did not manage to improve on the above bound, and they may have been tempted to conjecture that the “ $1/3^{r-1}$ ” exponent is best-possible. However, they found some clues that the situation is more intricate than it may first appear. For example, they observed that some patterns “cannot interact with each other”: if  $P$  and  $P'$  are the collectable  $r$ -patterns represented by  $|AB|AABB|$  and  $|AABB|AB|$ , and we have a matching  $M$  in which every pair of edges forms pattern  $P$  or  $P'$ , then in fact  $M$  must be a  $P$ -clique or a  $P'$ -clique.

To leverage this type of observation, instead of using Mirsky's theorem to process patterns one-by-one, we partition the set of all  $r$ -patterns into subsets, and process patterns one subset at a time (defining a poset in terms of the entire subset, instead of a single pattern). We need to choose the subsets in our partition judiciously, such that the patterns in each subset “cannot interact much with each other”, and such that each subset gives rise to a poset (after eliminating the patterns from previous subsets).

Both these properties are encapsulated in the following lemma, which is the key ingredient for the lower bound in [Theorem 1.7](#).

**Definition 3.1.** For a set of  $r$ -patterns  $\mathcal{P}$ , say that a matching  $M$  is  $\mathcal{P}$ -free if it is  $P$ -free for all  $P \in \mathcal{P}$ . Say that  $M$  is a  $\mathcal{P}$ -clique if every pair of edges forms a pattern in  $\mathcal{P}$ .

Also, for a matching  $M$  and a set of  $r$ -patterns  $\mathcal{P}$ , define the relation  $\preceq_{\mathcal{P}}$  on the edges of  $M$  by taking  $e \preceq f$  if  $e[1] < f[1]$  and  $e, f$  form a pattern in  $\mathcal{P}$  (or if  $e = f$ ).

**Lemma 3.2.** *Let  $b = (r + 1)2^{r-2}$ . There is a partition of the  $r$ -patterns into subsets  $\mathcal{P}_1, \dots, \mathcal{P}_b$  such that the following properties hold for any  $i \in \{1, \dots, b\}$ .*

- (A) *For any  $(\mathcal{P}_1 \cup \dots \cup \mathcal{P}_{i-1})$ -free matching  $M$ , the relation  $\preceq_{\mathcal{P}_i}$  is a partial order on  $M$ .*
- (B) *For any  $\mathcal{P}_i$ -clique  $M$  of size  $n$ , we have  $L(M) \geq n/(r - 1)$ .*

The proofs of the sharper lower bounds in [Theorem 1.8](#) are a bit more delicate; they use the same idea, with some extra twists. We will discuss them at the end of this section, after the proof of [Lemma 3.2](#).

For [Lemma 3.2](#), our partition into subsets will be defined in terms of *weak patterns*, which we define and investigate next.

**3.1. Weak patterns.** Weak patterns measure the relationship between two edges  $e, f$  in a slightly coarser way than ordinary patterns. Specifically, for weak patterns, we are only concerned with the behaviour of pairs of consecutive vertices in  $e$  and in  $f$ .

**Definition 3.3.** For an  $r$ -pattern  $P$  and some  $1 \leq i < j \leq r$  let  $P_{i:j}$  be the  $(j - i + 1)$ -pattern formed by  $\{e[i], e[i + 1], \dots, e[j]\}$  and  $\{f[i], f[i + 1], \dots, f[j]\}$ .

Then, define the *weak  $r$ -pattern*  $\phi(P) = (\phi_1(P), \dots, \phi_{r-1}(P)) \in \{\alpha, \kappa, \nu\}^{r-1}$  corresponding to  $P$ , by taking

- $\phi_i(P) = \alpha$  if  $P_{i:i+1}$  is the alignment 2-pattern,
- $\phi_i(P) = \kappa$  if  $P_{i:i+1}$  is the crossing 2-pattern,
- $\phi_i(P) = \nu$  if  $P_{i:i+1}$  is the nesting 2-pattern.

In general, a sequence  $W = (W_1, \dots, W_{r-1}) \in \{\alpha, \kappa, \nu\}^{r-1}$  is called a *weak  $r$ -pattern*.

In [Table 1](#), we list all the 3-patterns and their corresponding weak 3-patterns.

Pattern	Weak pattern	Pattern	Weak pattern
AAA BBB	$\alpha\alpha$	AB BA AB	$\nu\nu$
AABABB	$\alpha\alpha$	AB BA BA	$\nu\kappa$
AABB BA	$\alpha\nu$	AB AABB	$\kappa\alpha$
AABB AB	$\alpha\kappa$	AB AB BA	$\kappa\nu$
AB BBAA	$\nu\alpha$	AB AB AB	$\kappa\kappa$

**Table 1.** All  $(3 \cdot 2)! / (2 \cdot (3!)^2) = 10$  different 3-patterns, together with their corresponding weak pattern. Note that there are two different patterns corresponding to the weak pattern  $\alpha\alpha$  (one collectable, and one not).

**Definition 3.4.** For a weak pattern  $W$ , say a pair of edges “form  $W$ ” if they form a pattern  $P$  with  $\phi(P) = W$ . An ordered  $r$ -matching  $M$  is said to be a  $W$ -*clique* if every pair of edges form  $W$ .

In the proof of [Lemma 3.2](#), we will define each of our subsets  $\mathcal{P}_i$  to be a set of all patterns  $P$  such that  $\phi(P)$  has a prescribed number of “ $\alpha$ ”s, and has its “ $\nu$ ”s in prescribed positions. To understand why this works, we need quite a thorough study of weak patterns.

First, we collect a few basic properties of weak patterns. A “generalised alignment” is a pattern with block representation  $AA \cdots ABB \cdots B$  (i.e., a collectable pattern with a single block).

**Lemma 3.5.** Fix an  $r$ -pattern  $P$  with edges  $e, f$  (and assume  $e[1] < f[1]$ ).

- (i) For each  $i \in [r]$ , we have  $e[i] < f[i]$  if and only if there are an even number of “ $\nu$ ”s among  $\phi_1(P), \dots, \phi_{i-1}(P)$ .
- (ii)  $P$  is collectable if and only if:
  - (\*) for all  $1 \leq i < j \leq r$  where

$$\phi(P_{i:j}) = \phi_i(P) \phi_{i+1}(P) \dots \phi_{j-1}(P) = \alpha\alpha \dots \alpha,$$

$P_{i:j}$  is a generalised alignment.

- (iii) For each weak  $r$ -pattern  $W$ , there is exactly one collectable  $r$ -pattern  $P$  satisfying  $\phi(P) = W$ .

*Proof sketch.* For (i), it is a simple observation (immediate from the definitions of alignments, crossings and nestings) that the relative order between  $e[i]$  and  $f[i]$  is different from the relative order between  $e[i - 1]$  and  $f[i - 1]$  if and only if  $\phi(P)_{i-1} = \nu$ . The statement in (i) then follows by a straightforward induction on  $i$ .

(ii) is a bit more involved. Recall from [Definition 1.5](#) that  $P$  is collectable if and only if it is splittable (i.e., it can be partitioned into blocks that have an A-run followed by a B-run of the same length, or vice versa).

For the “only if” direction, it is easy to see that if  $P$  is collectable (splittable) then it satisfies (\*). Indeed, note that if  $P$  is splittable then  $P_{i:j}$  is splittable for all  $1 \leq i < j \leq r$  (this follows immediately from the definition of splittability). In the block representation of  $P_{i:j}$ , each division between consecutive blocks gives rise to a “ $\nu$ ” or “ $\kappa$ ”, so if  $\phi(P_{i:j}) = \alpha \dots \alpha$  then  $P_{i:j}$  must have a single block and is therefore a generalised alignment, as desired.

For the “if” direction, suppose  $P$  is not collectable (splittable). We need to prove that (\*) is violated. Let  $j$  be minimal such that  $P_{1:j}$  is not splittable (note that every 2-pattern is splittable, so we have  $j \geq 3$ ). Let  $|\mathfrak{B}_1| \dots |\mathfrak{B}_k|$  be the block representation of  $P_{1:j-1}$ . If we had  $\phi_{j-1}(P) \in \{\kappa, \nu\}$ , then the

block representation of  $P_{1:j}$  would be  $|\mathfrak{B}_1| \dots |\mathfrak{B}_k|AB|$  or  $|\mathfrak{B}_1| \dots |\mathfrak{B}_k|BA|$  (depending on the parity of the number of “ $\nu$ ”s among  $\phi_1(P), \dots, \phi_{j-1}(P)$ ). In either case,  $P_{1:j}$  would then be splittable (collectable), a contradiction. So, we must have  $\phi_{j-1}(P) = \alpha$ .

Now, let  $i$  be minimal such that  $\phi(P_{i:j}) = \alpha \dots \alpha$ . We have  $i \leq j - 1$ , since  $\phi_{j-1}(P) = \alpha$ . In the case  $i = 1$ , we know  $P_{1:j} = P_{i:j}$  cannot be a generalised alignment because (by the choice of  $j$ ) it is not splittable. This means that  $P$  violates the condition in (\*). So, it remains to consider the case  $i \geq 2$ . By the definition of  $i$ , we know  $\phi_{i-1}(P) \neq \alpha$ , which implies that both  $e[i - 1]$  and  $f[i - 1]$  precede  $e[i]$  and  $f[i]$ . If  $P_{i:j}$  were to form a generalised alignment, the block representation of  $P_{1:j}$  would be  $|\mathfrak{B}'_1| \dots |\mathfrak{B}'_k|AA \dots ABB \dots B|$  or  $|\mathfrak{B}'_1| \dots |\mathfrak{B}'_k|BB \dots BAA \dots A|$  (depending on the parity of the number of “ $\nu$ ”s among  $\phi_1(P), \dots, \phi_{i-1}(P)$ ), where  $|\mathfrak{B}'_1| \dots |\mathfrak{B}'_k|$  is the block representation of  $P_{1:i}$ . In either case,  $P_{1:j}$  would be splittable, a contradiction. So,  $P_{i:j}$  cannot be a generalised alignment, meaning that (\*) is violated. This completes the proof for (ii).

For (iii), fix a weak  $r$ -pattern  $W = (W_1, \dots, W_{r-1}) \in \{\alpha, \kappa, \nu\}^{r-1}$ . We can use properties (i) and (ii) to construct and force the structure of the collectable  $r$ -pattern  $P$  with  $\phi(P) = W$ . Specifically, let  $1 \leq i_1 < \dots < i_m < r$  be the indices  $i$  with  $W_i \in \{\nu, \kappa\}$ . Set  $i_0 = 0$  and  $i_{m+1} = r$  for convenience. For  $1 \leq \ell \leq m + 1$ , let  $\mathfrak{B}_\ell$  be a sequence of  $(i_\ell - i_{\ell-1})$  “A”s followed by a sequence of  $(i_\ell - i_{\ell-1})$  “B”s if the number of “ $\nu$ ”s among  $W_{i_1}, W_{i_2}, \dots, W_{i_{\ell-1}}$  is even; let  $\mathfrak{B}_\ell$  be a sequence of  $(i_\ell - i_{\ell-1})$  “B”s followed by a sequence of  $(i_\ell - i_{\ell-1})$  “A”s otherwise. Then, it is not hard to deduce from (i) and (ii) that  $Q = \mathfrak{B}_1 \mathfrak{B}_2 \dots \mathfrak{B}_{m+1}$  is a collectable  $r$ -pattern with  $\phi(Q) = W$ , and the block representation of  $P$  must be precisely the concatenation  $|\mathfrak{B}_1| |\mathfrak{B}_2| \dots |\mathfrak{B}_{m+1}|$ .  $\square$

Given Lemma 3.5(iii), it makes sense to introduce some notation for the unique collectable pattern corresponding to a particular weak pattern.

**Definition 3.6.** For each weak  $r$ -pattern  $W$ , let  $\psi(W)$  be the unique collectable  $r$ -pattern satisfying  $\phi(\psi(W)) = W$ .

Now, one advantage of considering weak patterns is that we can get a handle on how they can interact with each other by considering how their constituent alignments, crossings and nestings can interact with each other.

**Lemma 3.7.** Suppose  $e, f, g$  are three edges in some ordered  $r$ -matching such that  $e[1] < f[1] < g[1]$ . Let  $W^{e,f}, W^{f,g}, W^{e,g}$  be the weak  $r$ -patterns formed by pairs  $\{e, f\}, \{f, g\}, \{e, g\}$ , respectively. If the “ $\nu$ ”s occur in the same positions in  $W^{e,f}$  and in  $W^{f,g}$ , then, for  $1 \leq i \leq r - 1$ ,

- (i) if  $W_i^{e,f} = W_i^{f,g} = \nu$ , then  $W_i^{e,g} = \nu$ ;
- (ii) if  $W_i^{e,f} = \alpha$  or  $W_i^{f,g} = \alpha$ , then  $W_i^{e,g} = \alpha$ ;
- (iii) if  $W_i^{e,f} = W_i^{f,g} = \kappa$ , then  $W_i^{e,g} \in \{\alpha, \kappa\}$ .

Also, combining (ii) and (iii), we get:

- (iv) if  $W_i^{e,f} \neq \nu$  and  $W_i^{f,g} \neq \nu$ , then  $W_i^{e,g} \neq \nu$ .

*Proof.* By assumption, the number of “ $\nu$ ”s among  $W_1^{e,f}, W_2^{e,f}, \dots, W_{i-1}^{e,f}$  and among  $W_1^{f,g}, W_2^{f,g}, \dots, W_{i-1}^{f,g}$  are the same. We assume that this number is even (the odd case is similar). By Lemma 3.5(i), we have  $e[i] < f[i] < g[i]$ .

If  $W_i^{e,f} = W_i^{f,g} = \nu$ , then  $e[i] < f[i] < f[i + 1] < e[i + 1]$  and  $f[i] < g[i] < g[i + 1] < f[i + 1]$ , implying that  $e[i] < g[i] < g[i + 1] < e[i + 1]$ . This means  $\{e[i], e[i + 1]\}$  and  $\{g[i], g[i + 1]\}$  form a nesting, i.e.,  $W_i^{e,g} = \nu$ . This proves (i).

If  $W_i^{e,f}, W_i^{f,g} \neq \nu$ , we know that  $W_i^{e,f}, W_i^{f,g} \in \{\alpha, \kappa\}$ . This implies

$$\begin{aligned} e[i] < \min(e[i + 1], f[i]) < \max(e[i + 1], f[i]) < f[i + 1] \text{ and} \\ f[i] < \min(f[i + 1], g[i]) < \max(f[i + 1], g[i]) < g[i + 1], \end{aligned}$$

which in turn imply

$$e[i] < \min(e[i + 1], g[i]) < \max(e[i + 1], g[i]) < g[i + 1].$$

That is to say,  $W_i^{e,g} \in \{\kappa, \alpha\}$ , proving (iii). If we furthermore have  $W_i^{e,f} = \alpha$  or  $W_i^{f,g} = \alpha$ , then we know  $e[i] < e[i + 1] < g[i] < g[i + 1]$ , i.e.,  $W_i^{e,g} = \alpha$ , proving (ii).  $\square$

Now, the following lemma will be used to handle non-collectable patterns: if we manage to find a large  $W$ -clique (for some weak pattern  $W$ ), we can very efficiently drop to a  $\psi(W)$ -clique. A similar

fact (with a slightly worse constant) can actually be deduced from the main result of [21], but here we provide a self-contained proof.

**Lemma 3.8.** *Let  $W$  be a weak  $r$ -pattern, and let  $M$  be a  $W$ -clique of size  $n$ . Then,  $M$  contains a  $\psi(W)$ -clique of size at least  $n/\max(1, \delta(W))$ , where  $\delta(W)$  is the length of the longest run of “ $\alpha$ ”s in  $W$ .*

*Proof.* By Lemma 3.5(ii), if  $\delta(W) = 0$  then there is no non-collectable  $r$ -pattern  $P$  with  $\phi(P) = W$ . Thus,  $M$  itself is a  $\psi(W)$ -clique. So, we may assume  $\delta := \delta(W) \geq 1$ .

Let  $e_1, e_2, \dots, e_n$  be the edges in  $M$ , ordered such that  $e_1[1] < e_2[1] < \dots < e_n[1]$ . We claim that for any  $s, t$  with  $t - s \geq \delta$ , the edges  $e_s, e_t$  form  $\psi(W)$ . This suffices to prove the lemma: we can simply take every  $\delta$ -th edge as our  $\psi(M)$ -clique.

For  $1 \leq a < b \leq n$  let  $P^{a,b}$  be the pattern formed by  $e_a, e_b$ . Fix  $s, t$  with  $t - s \geq \delta$ . We wish to show that  $P^{s,t}$  is collectable (which will imply  $P^{s,t} = \psi(W)$  by Lemma 3.5(iii)). By Lemma 3.5(ii), it suffices to show that for any  $1 \leq i < j \leq r$  with  $W_i = W_{i+1} = \dots = W_{j-1} = \alpha$  (i.e.,  $\phi(P_{i;j}^{s,t}) = \alpha \dots \alpha$ ), the  $(j - i + 1)$ -pattern  $P_{i;j}^{s,t}$  is a generalised alignment.

Suppose there are an even number of “ $\nu$ ”s among  $W_1, W_2, \dots, W_{i-1}$  (the odd case is similar). For  $q \in \{s, s+1, \dots, t-1\}$  and  $k \in \{i, i+1, \dots, j-1\}$ , we have  $e_q[k] < e_q[k+1] < e_{q+1}[k] < e_{q+1}[k+1]$  (since  $W_k = \alpha$ ). This means we have

$$e_s[j] < e_{s+1}[j-1] < \dots < e_{s+j-i}[i] \leq e_t[i],$$

where the last inequality holds because  $s+j-i \leq s+\delta \leq t$  and there are an even number of “ $\nu$ ”s among  $W_1, \dots, W_{i-1}$ . In other words,  $P_{i;j}^{s,t}$  is a generalised alignment, as desired.  $\square$

*Remark 3.9.* It is not hard to see that the constant factor  $1/\max(1, \delta(M))$  cannot be improved.

**3.2. Proof of the key lemma.** Now, we have all the preparations in place to prove Lemma 3.2. The subsets of patterns  $\mathcal{P}_i$  will be defined in terms of *signatures*, as follows.

**Definition 3.10.** For each weak  $r$ -pattern  $W \in \{\alpha, \kappa, \nu\}^{r-1}$ , the *signature* of  $W$  is defined to be

$$\sigma(W) := (|\{i : W_i = \alpha\}|, \{i : W_i = \nu\}).$$

That is to say, the signature specifies the number of “ $\alpha$ ”s and the positions of the “ $\nu$ ”s. We define the *weight* of  $\sigma(W)$  to be the number of “ $\alpha$ ”s in  $W$ .

The total number of signatures is

$$\sum_{S \subseteq \{1, \dots, r-1\}} (r - |S|) = \sum_{i=0}^{r-1} \binom{r-1}{i} (r-i) = (r+1)2^{r-2} =: b.$$

Let  $\sigma_1, \dots, \sigma_b$  be an ordering of the signatures in descending weight (breaking ties arbitrarily). Let  $\mathcal{P}_i$  be the set of all patterns  $P$  such that  $\sigma(\phi(P)) = \sigma_i$ .

*Proof of Lemma 3.2(A).* With  $\mathcal{P}_1, \dots, \mathcal{P}_b$  as defined above (in terms of signatures  $\sigma_1, \dots, \sigma_b$ ), we wish to show that for each  $i$ , the relation  $\preceq_{\mathcal{P}_i}$  is a partial order on any  $(\mathcal{P}_1 \cup \dots \cup \mathcal{P}_{i-1})$ -free matching  $M$ .

It suffices to show the transitivity of  $\preceq_{\mathcal{P}_i}$ . Fix a  $(\mathcal{P}_1 \cup \dots \cup \mathcal{P}_{i-1})$ -free matching  $M$  and consider any edges  $e, f, g$  with  $e[1] < f[1] < g[1]$ . Let  $P^{e,f}, P^{f,g}, P^{e,g}$  be the patterns formed by the pairs  $\{e, f\}, \{f, g\}$  and  $\{e, g\}$  respectively, and suppose that  $P^{e,f}, P^{f,g} \in \mathcal{P}_i$ . Our objective is to show that  $P^{e,g} \in \mathcal{P}_i$ . We know that  $\phi(P^{e,f})$  and  $\phi(P^{f,g})$  have their “ $\nu$ ”s in the same positions (as recorded by the signature  $\sigma_i$ ), so by Lemma 3.7(i,iv), the “ $\nu$ ”s in  $\phi(P^{e,g})$  must be in exactly these same positions.

Then, let  $w$  be the weight of  $\sigma_i$  (i.e., the number of “ $\alpha$ ”s in the weak patterns associated with  $\sigma_i$ ). We know that  $\phi(P^{e,f})$  and  $\phi(P^{f,g})$  have exactly  $w$  “ $\alpha$ ”s. Also, since  $M$  is  $(\mathcal{P}_1 \cup \dots \cup \mathcal{P}_{i-1})$ -free (and higher-weight signatures come earlier in our ordering), we know that  $\phi(P^{e,f})$  has *at most*  $w$  “ $\alpha$ ”s. However, by Lemma 3.7(ii), in every position where  $\phi(P^{e,f})$  or  $\phi(P^{f,g})$  have a “ $\alpha$ ”, there is also a “ $\alpha$ ” in  $\phi(P^{e,g})$ . The only way this can happen is if  $\phi(P^{e,f}), \phi(P^{f,g})$  and  $\phi(P^{e,g})$  each have exactly  $w$  “ $\alpha$ ”s, in exactly the same positions. We have proved that  $P^{e,g} \in \mathcal{P}_i$ , as desired.  $\square$

*Proof of Lemma 3.2(B).* Suppose  $M$  is a  $\mathcal{P}_i$ -clique of size  $n$ . We wish to show that  $L(M) \geq n/(r-1)$ .

Consider any three edges  $e, f, g$  with  $e[1] < f[1] < g[1]$ . Let  $P^{e,f}, P^{f,g}, P^{e,g}$  be the patterns formed by the pairs  $\{e, f\}, \{f, g\}$  and  $\{e, g\}$  respectively, so  $P^{e,f}, P^{f,g}, P^{e,g} \in \mathcal{P}_i$ . By the considerations in

the above proof of [Lemma 3.2\(A\)](#), each of  $\phi(P^{e,f}), \phi(P^{f,g}), \phi(P^{e,g})$  must have their “ $\nu$ ”s and “ $\alpha$ ”s (and therefore their “ $\kappa$ ”s) in exactly the same positions, which means that  $e, f, g$  actually form a  $W$ -clique for some weak pattern  $W$ . But if every three edges form a  $W$ -clique, then the whole of  $M$  must be a  $W$ -clique for some weak  $r$ -pattern  $W$  with signature  $\sigma_i$ . By [Lemma 3.8](#), we have  $L_{\psi(W)}(M) \geq n/(r-1)$ .  $\square$

**3.3. Proof of the lower bound in [Theorem 1.7](#).** Now, it is easy to derive the lower bound in [Theorem 1.7](#) by iteratively applying [Lemma 3.2](#) and Mirsky’s theorem.

*Proof of the lower bound in [Theorem 1.7](#).* Let  $M$  be a size- $n$  matching, and recall the subsets of patterns  $\mathcal{P}_1, \dots, \mathcal{P}_b$  in the proof of [Lemma 3.2](#). First, combining Mirsky’s theorem ([Theorem 2.1](#)) and [Lemma 3.2\(A\)](#), it is easy to prove by induction that for each  $i < b$ :

- (\*) Either for some  $j \leq i$  we can find a  $\mathcal{P}_j$ -clique of size at least  $n^{1/b}$ , or we can find a  $(\mathcal{P}_1 \cup \dots \cup \mathcal{P}_i)$ -free sub-matching of size at least  $n^{1-i/b}$ .

A  $(\mathcal{P}_1 \cup \dots \cup \mathcal{P}_{b-1})$ -free matching is nothing more than a  $\mathcal{P}_b$ -clique, so (\*) with  $i = b-1$  actually implies that we can always find a  $\mathcal{P}_j$ -clique  $M'$  of size at least  $n^{1/b}$ , for some  $j \leq b$ . Then, by [Lemma 3.2\(B\)](#), we have  $L(M) \geq L(M') \geq n^{1/b}/(r-1)$ , as desired.  $\square$

**3.4. Further improvements.** Weak  $r$ -patterns are very convenient to work with, due to the fact that we can separately study how their constituent alignments, crossings and matchings interact (with [Lemma 3.7](#)). However, by directly considering how patterns can interact with each other, one can prove stronger bounds (specifically, one can prove an analogue of [Lemma 3.2](#) with a smaller value of  $b$ ). In this subsection, we show how to do this to prove the essentially optimal lower bound  $L_3(n) \geq n^{1/7}/2$  featuring in [Theorem 1.8](#). Note that [Theorem 1.7](#) only gives the non-optimal bound  $L_3(n) \geq n^{1/8}/2$ .

The considerations in this  $r = 3$  proof can be generalised to larger  $r$ . For example, with some more sophisticated case analysis we can prove the optimal lower bound  $L_4(n) \geq \Omega(n^{1/15})$ . As the proof is quite technical, we refer the interested reader to [Appendix A](#).

*Proof of the lower bound on  $L_3(n)$  in [Theorem 1.8](#).* We use the notions of weak patterns and signatures introduced in [Sections 3.1](#) and [3.2](#). Referring to [Table 1](#), we see that there are ten different 3-patterns, including one non-collectable pattern (which we call  $P^*$ ; note that  $\phi(P^*) = \alpha\alpha$ ). Define

$$\mathcal{P}_1 = \{\psi(\alpha\alpha), P^*\}, \quad \mathcal{P}_2 = \{\psi(\alpha\kappa), \psi(\kappa\alpha)\}, \quad \mathcal{P}_3 = \{\psi(\alpha\nu), \psi(\nu\alpha)\}.$$

We observe that these subsets can be used to define posets (we write  $\preceq_i$  instead of  $\preceq_{\mathcal{P}_i}$ ):

- By the proof of [Lemma 3.2\(A\)](#), for any matching  $M$ , the relation  $\preceq_1$  is a partial order. (Note that  $\mathcal{P}_1$  corresponds to the signature  $(2, \emptyset)$ ).
- Similarly, by the proof of [Lemma 3.2\(A\)](#), for any  $\mathcal{P}_1$ -free matching  $M$ , the relation  $\preceq_2$  is a partial order. (Note that  $\mathcal{P}_2$  corresponds to the signature  $(1, \emptyset)$ ).
- One can also see that the relation  $\preceq_3$  is always a partial order (despite  $\mathcal{P}_3$  containing the patterns for two different signatures  $(1, \{1\})$  and  $(1, \{2\})$ ). Observe that the two patterns  $\psi(\alpha\nu), \psi(\nu\alpha) \in \mathcal{P}_3$  have block representations  $|AABB|BA$  and  $|AB|BBAA$ ; informally speaking, these patterns are “generalised nestings”, where all vertices of one edge are fully contained between two consecutive vertices of the other edge, and it is not hard to see that  $\preceq_3$  is therefore a partial order. In detail: note that it suffices to show transitivity. Suppose  $e, f, g$  are three edges in a matching  $M$  with  $e[1] < f[1] < g[1]$  such that  $e, f$  form pattern  $\psi(\alpha\nu)$  and  $f, g$  form pattern  $\psi(\alpha\nu)$  or  $\psi(\nu\alpha)$ . Then,  $e[2] < f[1] < f[3] < e[3]$  (as  $e, f$  form pattern  $\psi(\alpha\nu)$ ) and  $f[1] < g[1] < g[3] < f[3]$  (as  $f, g$  form pattern  $\psi(\alpha\nu)$  or  $\psi(\nu\alpha)$ ). So  $e[2] < g[1] < g[3] < e[3]$ , i.e.,  $e, g$  form pattern  $\psi(\alpha\nu)$ . Similarly, if  $e, f$  form pattern  $\psi(\nu\alpha)$  and  $f, g$  form pattern  $\psi(\alpha\nu)$  or  $\psi(\nu\alpha)$ , then  $e, g$  form pattern  $\psi(\nu\alpha)$ .
- By the proof of [Lemma 3.2\(A\)](#), for any  $(\mathcal{P}_1 \cup \mathcal{P}_2 \cup \mathcal{P}_3)$ -free matching  $M$ , the relation  $\preceq_P$  is a poset when  $P$  is any of the four 3-partite patterns  $\psi(\nu\nu), \psi(\nu\kappa), \psi(\kappa\nu), \psi(\kappa\kappa)$ . (These correspond to the signatures  $(0, S)$ , for  $S \subseteq \{1, 2\}$ ).

Proceeding in the same way as the proof of the lower bound in [Theorem 1.7](#), we can therefore find a  $\mathcal{P}_i$ -clique of size at least  $n^{1/7}$ , for some  $i \in \{1, 2, 3\}$ , or we can find a  $P$ -clique of size at least  $n^{1/7}$  for  $P$  being one of the four 3-partite patterns  $\psi(\nu\nu), \psi(\nu\kappa), \psi(\kappa\nu), \psi(\kappa\kappa)$ .

By the proof of [Lemma 3.2\(B\)](#), if we found a  $\mathcal{P}_1$ -clique or a  $\mathcal{P}_2$ -clique  $M'$  of size at least  $n^{1/7}$ , then  $L(M') \geq n^{1/7}/2$ .

Finally, it suffices to consider the case where we found a  $\mathcal{P}_3$ -clique  $M'$  with at least  $n^{1/7}$  edges. Let  $e_1, \dots, e_m$  be the edges in  $M'$  (with  $m \geq n^{1/7}$ , and  $e_1[1] < \dots < e_m[1]$ ). As discussed in the third bullet point above, we see that for any  $1 \leq i < j < k \leq m$ , the edges  $e_i, e_k$  always form the same 3-pattern as the edges  $e_i, e_j$ . So, every index  $i$  is of one of two types: we say it is of “type  $\psi(\alpha\nu)$ ” if  $e_i, e_j$  form  $\psi(\alpha\nu)$  for each  $i < j$ , and we say it is of “type  $\psi(\nu\alpha)$ ” if  $e_i, e_j$  form  $\psi(\nu\alpha)$  for each  $i < j$ . By the pigeonhole principle, at least half of the indices have the same type, and the corresponding edges give us a  $\psi(\alpha\nu)$ -clique or a  $\psi(\nu\alpha)$ -clique of size  $n^{1/7}/2$ .  $\square$

#### 4. UPPER BOUNDS ON RAMSEY PARAMETERS

In this section we prove the upper bound in [Theorem 1.7](#) (which also implies the upper bounds in [Theorem 1.8](#)). We consider the following notion of “blow-up” introduced by Dudek, Grytczuk and Ruciński [\[22\]](#).

**Definition 4.1.** Suppose  $M, M'$  are ordered  $r$ -matchings, and let  $t$  be the size of  $M'$ . The  $M'$ -blow-up of  $M$ , denoted by  $M[M']$ , is an ordered  $r$ -matching obtained from  $M$  as follows. We replace every vertex  $i \in V(M)$  by an ordered set  $U_i$  of  $t$  contiguous vertices. Then, for each edge  $e \in E(M)$  we place on the vertex set  $U_{e[1]} \cup \dots \cup U_{e[r]}$  a copy  $M'_e$  of the matching  $M'$ . Note that there are two kinds of pairs of edges  $(f_1, f_2)$  in  $M[M']$ :

- If  $f_1, f_2$  both lie in the same  $M'_e$  (for some  $e \in E(M)$ ), then we say  $f_1$  and  $f_2$  comprise an  $M'$ -pair.
- If  $f_1, f_2$  lie in different  $M'_e$ , then we say  $f_1$  and  $f_2$  comprise an  $M$ -pair.

The above definition is useful typically when  $M'$  is  $r$ -partite. In this case, one can check that if two edges  $f_1, f_2$  comprise an  $M$ -pair (say  $f_1$  lie in  $M'_{e_1}$  and  $f_2$  lie in  $M'_{e_2}$ ), then  $f_1, f_2$  form the same  $r$ -pattern as  $e_1, e_2$  do (in  $M$ ). Therefore, if  $M$  contains no  $r$ -partite  $r$ -pattern and  $M'$  is  $r$ -partite, then  $L_P(M[M']) = L_P(M)$  if  $P$  is  $r$ -partite and  $L_P(M[M']) = L_P(M')$  otherwise, and thus  $L(M[M']) = \max(L(M), L(M'))$ . This fact will be used in the proof of the upper bound in [Theorem 1.7](#) momentarily.

For  $r$ -partite  $r$ -patterns, Dudek, Grytczuk and Ruciński [\[22\]](#) used a blow-up construction to prove the following tight upper bound (which is actually equivalent to the result of Burkill and Mirsky [\[15\]](#) and Kalmanson [\[40\]](#) mentioned in the introduction, on  $(r-1)$ -tuples of permutations).

**Theorem 4.2** (Dudek, Grytczuk and Ruciński [\[22\]](#)). *For  $r \geq 2$ , and  $n \geq 1$ , there exists an  $r$ -partite ordered  $r$ -matching  $M$  of size  $n^{2^{r-1}}$  such that for all ( $r$ -partite)  $r$ -patterns  $P$ , the largest  $P$ -clique is of size  $n$ .*

We are now ready to prove that  $L_r(n) \leq \lceil n^{1/(2^r-1)} \rceil$ .

*Proof of the upper bound in [Theorem 1.7](#).* Fix  $r \geq 2$  and  $n \geq 1$ . We are going to construct an ordered  $r$ -matching  $M$  of size  $n^{2^{r-1}}$  such that any pair of edges form a collectable  $r$ -pattern, and such that the largest clique has size  $n$ . We will do this by induction on  $r$ .

The base case  $r = 2$  already appears in [\[21, 22\]](#). Indeed, consider  $M := M_1[M_2]$ , where  $M_1$  is an alignment-clique of size  $n$  and  $M_2$  is the 2-partite ordered 2-matching from [Theorem 4.2](#). Then,  $M$  has size  $n^3$ . A pair of distinct edges  $f_1, f_2 \in E(M)$  form an alignment if  $(f_1, f_2)$  is an  $M_1$ -pair, and they form a nesting or a crossing if  $(f_1, f_2)$  is an  $M_2$ -pair. It is easy to check that there is no clique of size larger than  $n$ .

For the inductive step, suppose we have constructed an ordered  $(r-1)$ -matching  $M_1$  of size  $n^{2^{r-1}-1}$  in which every pair of edges forms a collectable  $(r-1)$ -pattern, such that the largest clique has size  $n$ . For each edge  $e \in E(M_1)$ , we add a new vertex  $v_e$  just to the right of  $e[r-1]$ , and extend  $e$  to an edge  $e'$  of uniformity  $r$  by adding  $v_e$  to  $e$ . Let the resulting ordered  $r$ -matching be  $M_2$ .

Now, we claim that  $M_2$  is free of  $r$ -partite  $r$ -patterns, and its largest clique has size  $n$ . To see this, note that if  $e, f \in E(M_1)$  form a pattern with block representation  $|\mathfrak{B}_1| \dots |\mathfrak{B}_k|$ , then  $e', f' \in E(M_2)$  form the pattern whose block representation is  $|\mathfrak{B}_1| \dots |\mathfrak{B}_{k-1}| \mathfrak{B}'_k|$ , where  $\mathfrak{B}'_k$  is obtained by extending the “A-run” and the “B-run” in  $\mathfrak{B}_k$  by one. (For example, if  $e, f$  form a pattern with block representation  $|AABB|AB|BA|$ , then  $e', f'$  form a pattern with block representation  $|AABB|AB|BBAA|$ ). This latter pattern is never  $r$ -partite, because it has at most  $r-1$  blocks.

Now, let  $M_3$  be the  $r$ -partite ordered  $r$ -matching given by [Theorem 4.2](#) (with size  $n^{2^{r-1}}$ , and whose largest clique has size  $n$ ). Every pair of edges in  $M_3$  form an  $r$ -partite  $r$ -pattern. Then, we take  $M$  to be

the blow-up  $M_2[M_3]$ , which has size  $n^{2^{r-1}-1} \cdot n^{2^{r-1}} = n^{2^r-1}$ . For any distinct  $f_1, f_2 \in E(M)$ , the edges  $f_1, f_2$  form an  $r$ -partite  $r$ -pattern if and only if  $(f_1, f_2)$  is an  $M_3$ -pair. This means a clique in  $M$  comes either from  $M_2$  or from some copy of  $M_3$ . So, the largest clique in  $M$  has size  $n$ , proving the induction step.  $\square$

## 5. VARIANTS ON THE LONGEST INCREASING SUBSEQUENCE PROBLEM

The famous Ulam–Hammersley problem (see [53] for a book-length treatment) asks for the expected length of the longest monotone subsequence in a random set of  $n$  points in a box  $[0, 1]^2$ . This was generalised to higher dimensions by Steele [58] (and studied further by Bollobás and Winkler [7]). In order to prove [Theorem 1.10](#), we will need a further generalisation.

**Definition 5.1.** Fix a partition  $\mathcal{A}$  of  $\{1, \dots, r\}$  into disjoint parts  $I_1, \dots, I_\ell$ . Then, for any vectors  $(x_1, \dots, x_r), (y_1, \dots, y_r) \in \mathbb{R}^r$ , write  $(x_1, \dots, x_r) \preceq_{\mathcal{A}} (y_1, \dots, y_r)$  if  $\max\{x_i : i \in I_j\} \leq \min\{y_i : i \in I_j\}$  for all  $j \in \{1, \dots, \ell\}$ . Note that  $\preceq_{\mathcal{A}}$  is a partial order on  $\mathbb{R}^r$ .

For a set of points  $\mathcal{S} \subseteq \mathbb{R}^d$ , write  $L_{\mathcal{A}}(\mathcal{S})$  for the longest chain in  $\mathcal{S}$  with respect to  $\preceq_{\mathcal{A}}$ .

We will be interested in  $L_{\mathcal{A}}(\mathcal{S}_n)$ , for a set  $\mathcal{S}_n$  of  $n$  independent uniformly random points in  $[0, 1]^r$ .

**Theorem 5.2.** Fix a partition  $\mathcal{A}$  of  $\{1, \dots, r\}$ . For  $m \in \mathbb{N}$ , let  $\mathcal{T}_m \subseteq [0, m]^r$  be a set of points obtained by a Poisson process of rate 1 in  $[0, m]^r$ , and let  $L = L_{\mathcal{A}}(\mathcal{T}_m)$ . Then, as  $m \rightarrow \infty$ ,

- (1)  $\frac{\mathbb{E}L}{m} \rightarrow a_{\mathcal{A}}$ , and
- (2)  $\frac{L}{m} \xrightarrow{p} a_{\mathcal{A}}$ ,

for some  $a_{\mathcal{A}} > 0$  only depending on  $\mathcal{A}$ . (By symmetry, in fact  $a_{\mathcal{A}}$  only depends on the multiset of sizes of the parts in  $\mathcal{A}$ ).

*Proof.* Let  $\mathcal{T} \subseteq \mathbb{R}^r$  be a Poisson process with rate 1 in  $\mathbb{R}^r$ . Let  $L_{m,n} = L(\mathcal{T} \cap [m, n]^r)$  be the longest chain in  $\mathcal{T} \cap [m, n]^r$  with respect to  $\mathcal{P}(\mathcal{A})$ . Noting that  $\mathcal{T} \cap [0, m]^r \stackrel{d}{=} \mathcal{T}_m$ , it suffices to prove that the conditions of Kingman’s subadditive ergodic theorem ([Theorem 2.6](#)) are satisfied, as follows.

- For  $0 \leq m \leq n$  we have  $L_{0,n} \geq L_{0,m} + L_{m,n}$ , because if we have a chain in  $\mathcal{T} \cap [0, m]^r$  and a chain in  $\mathcal{T} \cap [m, n]^r$ , their union is always a chain in  $\mathcal{T} \cap [0, n]^r$ .
- For any  $k \geq 1$ ,  $(L_{nk, (n+1)k})_{n=0}^{\infty}$  is a sequence of i.i.d. random variables.
- For any  $m \geq 1$ , we have  $(L_{0,k})_{k=0}^{\infty} \stackrel{d}{=} (L_{m, m+k})_{k=0}^{\infty}$ .
- There exists a constant  $M > 0$  such that  $\mathbb{E}L_{0,m} \leq Mm$  for all  $m$ . To see this, we first note that if  $\mathcal{S}_N$  is a set of  $N$  independent uniformly random points in  $[0, m]^r$ , then we have  $\mathbb{E}L_{\mathcal{A}}(\mathcal{S}_N) \leq \mathbb{E}L_{\mathcal{B}\mathcal{W}}(\mathcal{S}_N)$ , where  $\mathcal{B}\mathcal{W}$  is the partition of  $\{1, \dots, r\}$  into  $r$  singleton sets. Bollobás and Winkler [7] proved that  $\mathbb{E}L_{\mathcal{B}\mathcal{W}}(\mathcal{S}_N) \leq (1 + o(1))eN^{1/r}$ , so

$$\begin{aligned} \mathbb{E}L_{0,m} &= \mathbb{E}L_{\mathcal{A}}(\mathcal{T}_m) = \mathbb{E}\left[\mathbb{E}\left[L_{\mathcal{A}}(\mathcal{T}_m) \mid |\mathcal{T}_m|\right]\right] = \mathbb{E}\left[\mathbb{E}\left[L_{\mathcal{A}}(\mathcal{S}_{|\mathcal{T}_m|}) \mid |\mathcal{T}_m|\right]\right] \\ &\leq \mathbb{E}\left[(1 + o(1))e|\mathcal{T}_m|^{1/r}\right] = (1 + o(1))em, \end{aligned}$$

where the last equation holds as  $|\mathcal{T}_m|$  has distribution  $\text{Poisson}(m^r)$ .  $\square$

**Theorem 5.3.** Fix a partition  $\mathcal{A}$  of  $\{1, \dots, r\}$ . Consider a set  $\mathcal{S}_n$  of  $n$  independent uniformly random points in  $[0, 1]^r$ , and let  $L = L_{\mathcal{A}}(\mathcal{S}_n)$ . Let  $a_{\mathcal{A}}$  be as in [Theorem 5.2](#). Then, the following hold.

- (1)  $\Pr[|L - \mathbb{E}L| \geq t] \leq 4 \exp\left(-\frac{\gamma t^2}{\mathbb{E}L + t}\right)$  for some universal constant  $\gamma > 0$ .
- (2)  $\frac{L}{n^{1/r}} \xrightarrow{p} a_{\mathcal{A}}$ .
- (3)  $\frac{\mathbb{E}L}{n^{1/r}} \rightarrow a_{\mathcal{A}}$ .

*Proof.* First, (1) follows directly from Talagrand’s inequality ([Theorem 2.5](#)):

- If we change a single point of  $\mathcal{S}_n$ , we change  $L$  by at most 1.
- Whenever  $L \geq r$  then there is a set of  $r$  points which certifies that  $L \geq r$ .



Then, for (2), recall the random set  $\mathcal{T}_m$  from [Theorem 5.2](#). Note that  $|\mathcal{T}_m|$  has a  $\text{Poisson}(m^r)$  distribution, so by Chebyshev's inequality we can choose  $m_1, m_2 \in \mathbb{N}$ , both of the form  $(1 + o(1))n^{1/r}$ , such that  $|\mathcal{T}_{m_1}| \leq n$  whp and  $n \leq |\mathcal{T}_{m_2}|$  whp.

Note that if we condition on  $|\mathcal{T}_m|$ , then  $\mathcal{T}_m$  is conditionally a set of that many independent uniformly random points in  $[0, m]^r$  (which is equivalent to taking random points in  $[0, 1]^r$  and rescaling them by a factor of  $m$ ). So, there is a coupling of  $\mathcal{S}_n, \mathcal{T}_{m_1}, \mathcal{T}_{m_2}$  for which  $L(\mathcal{T}_{m_1}) \leq L \leq L(\mathcal{T}_{m_2})$  whp. Then, (2) follows from [Theorem 5.2\(2\)](#).

Finally, (3) follows from (1) and (2).  $\square$

**Proposition 5.4.** *Fix a partition  $\mathcal{A}$  of  $\{1, \dots, r\}$ , and define the constant  $a_{\mathcal{A}}$  as in [Theorem 5.2](#).*

(1) *If  $\mathcal{A}$  has a single part of size  $r$ , then we have  $a_{\mathcal{A}} = \frac{r}{\Gamma(1/r)}$*

(2) *If  $\mathcal{A}$  has  $r$  parts of size 1, then we have  $a_{\mathcal{A}} > \frac{r}{(r!)^{1/r}\Gamma(1/r)}$ .*

*Proof.* For (1), we use the algorithmic approach of Justicz, Scheinerman and Winkler [\[39\]](#): we can build a longest chain  $\vec{x}^{(1)}, \vec{x}^{(2)}, \dots, \vec{x}^{(L)}$  by first taking the point  $\vec{x}^{(1)} = (x_1^{(1)}, \dots, x_r^{(1)}) \in \mathcal{T}_m$  with the smallest value of  $\max_i x_i^{(1)}$ , then throwing out all other points  $\vec{y} \in \mathcal{T}_m$  which do not satisfy  $\vec{x} \preceq \vec{y}$ , and taking the point  $\vec{x}^{(2)} = (x_1^{(2)}, \dots, x_r^{(2)})$  with the next smallest value of  $\max_i x_i^{(2)}$ , and so on.

If we let  $\mathcal{T}$  be the set of points corresponding to a Poisson process of rate 1 in the semi-infinite box  $[0, \infty)^r$  (instead of a finite box  $[0, m]^r$ ) then we obtain an infinite sequence of points  $(\vec{x}^{(k)})_{k=1}^{\infty}$  as above. Then,  $(\vec{x}^{(k)})_{k=1}^{\infty}$  has the following two properties. First, for every  $j \geq 1$  the sequence of points  $\vec{x}^{(1)}, \vec{x}^{(2)}, \dots, \vec{x}^{(j)}$  is a longest chain of points in  $\mathcal{T} \cap [0, \max_i x_i^{(j)}]^r$ , and second, the increments  $\max_i x_i^{(k+1)} - \max_i x_i^{(k)}$  are independent and identically distributed (let  $Z$  be a random variable with this common distribution). We have  $\Pr[Z > z] = \Pr[\mathcal{T} \cap [0, z]^r = \emptyset] = \exp(-z^r)$  so the density of  $Z$  is

$$\frac{d}{dz}(1 - \exp(-z^r)) = rz^{r-1} \exp(-z^r)$$

and

$$\mathbb{E}Z = \int_0^{\infty} z \cdot rz^{r-1} \exp(-z^r) dz = \frac{\Gamma(1/r)}{r}.$$

So, for any  $\varepsilon > 0$ , taking  $L_1 \leq (r/\Gamma(1/r) - \varepsilon)m$  and  $L_2 \geq (r/\Gamma(1/r) + \varepsilon)m$ , we have  $\max_i x_i^{(L_1)} \leq m \leq \max_i x_i^{(L_2)}$  whp, by the law of large numbers. Since  $\varepsilon$  was arbitrary, it follows that whp  $L = (r/\Gamma(1/r) - o(1))m$ ; that is,  $a = r/\Gamma(1/r)$ .

For (2), we recall the algorithmic lower bound of Bollobás and Winkler [\[7\]](#). They build a chain  $\vec{x}^{(1)}, \vec{x}^{(2)}, \dots, \vec{x}^{(Q)}$  by first taking the point  $\vec{x}^{(1)} = (x_1^{(1)}, \dots, x_r^{(1)}) \in \mathcal{T}_m$  such that  $\sum_i x_i^{(1)}$  is minimal, then throwing out all other points  $\vec{y} \in \mathcal{T}_m$  which do not satisfy  $\vec{x} \preceq \vec{y}$ , and taking the point  $\vec{x}^{(2)} = (x_1^{(2)}, \dots, x_r^{(2)})$  with the next smallest value of  $\sum_i x_i^{(2)}$ , and so on. It is proved in [\[7\]](#), by a very similar method as above, that this algorithm produces a chain of length  $r^2/(r!)^{1/r}\Gamma(1/r)$ .

We then just need to observe that this algorithm is not optimal: for example, instead of finding our chain point-by-point, we can build our chain two points at a time, at step  $k$  choosing two points  $\vec{x}^{(2k-1)}, \vec{x}^{(2k)}$  such that  $\vec{x}^{(2k-1)} \preceq \vec{x}^{(2k)}$  and such that  $\max\left(\sum_i x_i^{(2k-1)}, \sum_i x_i^{(2k)}\right)$  is minimized. It is not hard to see that the expected “distance travelled” in one such step is strictly less than two times the distance travelled in a step of the Bollobás–Winkler algorithm.  $\square$

## 6. CLIQUES IN RANDOM ORDERED MATCHINGS

In this section we prove [Theorem 1.10](#) and [Proposition 1.11](#), using the results in [Section 5](#). To give a brief overview, the key insight is that every  $P$ -clique in an ordered matching  $M$  can be interpreted as a chain in a certain point set (defined in terms of a partition of the vertices of  $M$ ), with respect to a certain partial order  $\preceq_{\mathcal{A}}$  (as in [Definition 5.1](#)). Taking advantage of the strong concentration in [Theorem 5.3\(1\)](#), we can take the union bound over *all* appropriate vertex partitions of a random matching  $M$ , giving us control over the longest chains in the corresponding point sets.

*Proof of [Theorem 1.10](#).* Suppose  $P$  has block partition given by  $J_1 \cup \dots \cup J_{\ell} = \{1, \dots, 2r\}$ . Let  $\mathcal{A}$  be the partition of  $\{1, \dots, r\}$  with parts  $I_i = \{j : 2j \in J_i\}$  for  $1 \leq i \leq \ell$ . Note that one can uniquely recover  $I_1, \dots, I_{\ell}$  from  $J_1, \dots, J_{\ell}$ ; we have  $J_i = \bigcup_{j \in I_i} \{2j - 1, 2j\}$  for all  $i \leq \ell$ .

Observe that for every  $P$ -clique  $C$  in  $M$ , there is some partition  $Q_1 \cup \dots \cup Q_\ell$  of  $V(M) = \{1, \dots, rn\}$  into contiguous intervals, such that for every edge  $e \in C$  and every  $i \in \{1, \dots, \ell\}$ , we have  $|e \cap Q_i| = |I_i|$ . Say that  $C$  is *consistent* with  $Q_1, \dots, Q_\ell$  if this is the case.

Fix a partition  $\mathcal{B}$  of  $\{1, \dots, rn\}$  into contiguous intervals  $Q_1, \dots, Q_\ell$ ; we will only consider cliques consistent with  $\mathcal{B}$  (at the end we will take a union bound over all partitions).

Let  $M_{\mathcal{B}}$  be the sub-matching of  $M$  containing the edges of  $M$  which are consistent with  $\mathcal{B}$ . Condition on some outcome  $N$  for the number of edges in  $M_{\mathcal{B}}$ .

We will realise the conditional distribution of  $M_{\mathcal{B}}$  via an alternative construction, as follows.

- (1) Let  $V = [0, \ell] \subseteq \mathbb{R}$  and for each  $i \in \{1, \dots, \ell\}$ , let  $V_i = [i-1, i]$ , so  $V_1 \cup \dots \cup V_\ell = [0, \ell]$ .
- (2) For each  $j \in \{1, \dots, r\}$ , let  $i(j)$  be such that  $j \in I_{i(j)}$ .
- (3) Let  $\mathcal{R}$  be a set of  $N$  independent uniformly random points in  $\prod_{j=1}^r V_{i(j)}$ .
- (4) For each  $\vec{v} \in \mathcal{R}$ , let  $e(\vec{v})$  be the set of coordinates of  $\vec{v}$  (so  $|e(\vec{v}) \cap V_i| = |I_i|$  for each  $i$ ).
- (5) Let  $M^*$  be the hypergraph on the vertex set  $\bigcup_{\vec{v} \in \mathcal{R}} e(\vec{v})$  whose edges are the sets  $e(\vec{v})$  for  $\vec{v} \in \mathcal{R}$ .

Note that with probability 1,  $M^*$  is a matching, and the distribution of (the order-isomorphism class of)  $M^*$  is the same as the conditional distribution of (the order-isomorphism class of)  $M_{\mathcal{B}}$ , by symmetry.

Our next goal is to realise the distribution of our random point set  $\mathcal{R} \subseteq \prod_{j=1}^r V_{i(j)}$  in terms of a set  $\mathcal{S}_N$  of  $N$  independent uniformly random points in  $[0, 1]^r$ , in such a way that the size of the largest  $P$ -clique  $L_P(M^*)$  in  $M^*$  corresponds precisely to the length of the longest chain  $L_{\mathcal{A}}(\mathcal{S}_N)$  in  $\mathcal{S}_N$  with respect to the partial order  $\preceq_{P_{\mathcal{A}}}$  defined in [Definition 5.1](#).

Recall that in the block representation ([Definition 1.5](#)) of  $P$ , the incident vertices of one edge are assigned with label ‘‘A’’ while those of the other edge are assigned with label ‘‘B’’. Now, consider an isometry  $\phi : [0, 1]^r \rightarrow \prod_{j=1}^r V_{i(j)}$  defined as follows.

- (1) For each  $i \in \{1, \dots, r\}$ , define the function  $\sigma_i : \mathbb{R} \rightarrow \mathbb{R}$  by taking  $\sigma_i(x) = x$  if the block  $I_i$  consists of an A-run followed by a B-run, and taking  $\sigma_i(x) = 1 - x$  if the block  $I_i$  consists of a B-run followed by an A-run.
- (2) Let  $\phi_0 : [0, 1]^r \rightarrow [0, 1]^r$  be the isometry  $(x_1, \dots, x_r) \mapsto (\sigma_1(x_1), \dots, \sigma_r(x_r))$ .
- (3) Let  $\vec{t}$  be such that  $[0, 1]^r + \vec{t} = \prod_{j=1}^r V_{i(j)}$ , and let  $\phi_1 : [0, 1]^r \rightarrow \prod_{j=1}^r V_{i(j)}$  be the translation  $\vec{x} \mapsto \vec{x} + \vec{t}$ .
- (4) Let  $\phi = \phi_1 \circ \phi_0$ .

It is easy to check that  $L_P(M^*) = L_{I_1, \dots, I_\ell}(\mathcal{S}_N)$ . So, still conditioning on an outcome  $|M_{\mathcal{B}}| = N$ , by [Theorem 5.3](#), with conditional probability  $1 - N^{-\omega(1)}$  we have that the random variable  $L_P(M_{\mathcal{B}})$  is of the form  $(a_{\mathcal{A}} + o(1))N^{1/r}$ .

Also, by the concentration inequality [Theorem 2.3](#) (see [Remark 2.4](#)), with probability at least  $1 - n^{-\omega(1)}$  we have

$$|M_{\mathcal{B}}| = \left( f_P \left( \frac{|Q_1|}{rn}, \dots, \frac{|Q_\ell|}{rn} \right) + o(1) \right) n,$$

where  $f_P(q_1, \dots, q_\ell)$  is the probability that for a set of  $r$  independent uniformly random points in  $[0, 1]$ , exactly  $I_i$  of them fall between  $q_{i-1}$  and  $q_{i-1} + q_i$  (here we take  $q_0 = 0$  for convenience). Note that  $f_P(q_1, \dots, q_\ell)$  is a homogeneous polynomial of degree  $r$  in the variables  $q_1, \dots, q_\ell$ .

Let  $\alpha_P$  be the maximum value of  $f_P(q_1, \dots, q_\ell)$  over all  $q_1, \dots, q_\ell \geq 0$  summing to 1. Then, taking a union bound over all possible partitions  $\mathcal{B}$  (of which there are at most  $n^{\ell-1}$ ), we see that

$$\frac{L(Q)}{n^{1/r}} \xrightarrow{P} a_{\mathcal{A}} \alpha_P^{1/r}.$$

So, the desired result follows with  $b_P = a_{\mathcal{A}} \alpha_P^{1/r}$ .  $\square$

*Proof of [Proposition 1.11](#).* Recall the definitions of  $f_P$  and  $\alpha_P$  from the proof of [Theorem 1.10](#), and recall the estimates in that proof.

- (1) Suppose  $P$  has type  $r$ . As in the proof of [Theorem 1.10](#), let  $\mathcal{A}$  be the trivial partition of  $\{1, \dots, r\}$  into one part. Note that  $\alpha_P = f_P(1) = 1$ , and by [Proposition 5.4](#),  $a_{\mathcal{A}} = r/\Gamma(1/r) = 1/\Gamma((r+1)/r)$ . So,  $b_P = a_{\mathcal{A}} \alpha_P^{1/r} = 1/\Gamma((r+1)/r)$ .
- (2) Suppose  $P$  has type  $1 + \dots + 1$ . As in the proof of [Theorem 1.10](#), let  $\mathcal{A}$  be the partition of  $\{1, \dots, r\}$  into  $r$  singleton parts. Note that  $\alpha_P = f_P(1/r, \dots, 1/r) = r!/r^r$ , and note that  $a_{\mathcal{A}}$  is precisely the Bollobás–Winkler constant  $c_r$ , which by [Proposition 5.4](#) is strictly larger than  $\frac{r^2}{(r!)^{1/r} \Gamma(1/r)}$ . Therefore we have  $b_P = c_r (r!/r^r)^{1/r} = c_r (r!)^{1/r} / r > r/\Gamma(1/r) = 1/\Gamma((r+1)/r)$ .  $\square$

## 7. ENUMERATION

In this section we estimate  $N_{P,m}(n)$  up to an exponential factor, proving [Theorem 1.15](#) and therefore [Theorem 1.13](#). The upper bound and the lower bound in [Theorem 1.15](#) are proved separately. Despite [Theorem 1.15](#) being about ordered matchings, both the upper and lower bound require the consideration of general ordered  $r$ -graphs; in particular, the upper bound requires the extremal results introduced in [Section 1.5](#) (proved in [Section 9](#)).

For the upper bound, we prove the following general estimate on the number of ordered matchings  $M$  avoiding a particular sub-matching  $Q$ .

**Theorem 7.1.** *Let  $r \geq 2$  and  $Q$  be an ordered  $r$ -matching with at least two edges. Suppose there exists some  $n_Q > 0$  such that  $\text{ex}_<(n, Q) \leq \alpha n^{r-1}$  for  $n \geq n_Q$ . Then, if  $n$  is sufficiently large in terms of  $r$  and  $n_Q$ , the number of  $Q$ -free ordered  $r$ -matchings of size  $n$  is at most*

$$e^{40n} (\alpha r^r n^{r-2})^{(r/(r-1))n}.$$

In the above statement, we allow  $\alpha$  to depend on  $n$ . To obtain the upper bound in [Theorem 1.15](#), we can simply apply [Theorem 7.1](#) with  $Q = P^{(m)}$  and the estimate in [Theorem 1.18\(2\)](#).

In order to prove the general statement in [Theorem 7.1](#), we need a few basic facts about extremal numbers of general ordered  $r$ -matchings. First, we need the fact that extremal numbers always have order of magnitude at least  $n^{r-1}$ . (This is not hard to prove; in particular it follows from the  $m = 2$  case of [Theorem 9.1](#), which we prove in [Section 9](#)).

**Proposition 7.2.** *Let  $r \geq 2$ , and let  $Q$  be any  $r$ -matching with at least two edges. Then*

$$\text{ex}_<(n, Q) \geq \binom{n}{r} - \binom{(n-r)_+}{r}.$$

We also need a basic monotonicity property of extremal numbers of ordered matchings.

**Proposition 7.3.** *Let  $r, n \geq 2$  and let  $Q$  be an ordered  $r$ -matching with more than one edge. Then,*

$$\text{ex}_<(n, Q) \geq 2 \text{ex}_<(\lceil n/2 \rceil, Q).$$

*Proof.* Write the edges of  $Q$  as  $e_1, \dots, e_m$ , with  $e_1[1] < \dots < e_m[1]$ . Say that  $Q$  splits into sub-matchings if there is an index  $\ell \in \{1, \dots, m-1\}$  such that  $e_i[r] < e_j[1]$  for any  $1 \leq i \leq \ell < j \leq m$  (i.e.,  $Q$  consists of two matchings placed side-by-side). We proceed differently depending on whether  $Q$  splits into sub-matchings.

If  $Q$  does not split into sub-matchings, then given any  $Q$ -free ordered  $r$ -graph  $G$ , we can “glue together two copies of  $G$ ” to make a larger  $Q$ -free ordered  $r$ -graph  $G'$ . Specifically, we can take two copies of  $G$ , and identify the last vertex of the first copy with the first vertex of the second copy (so except for the single identified vertex, all vertices of the first copy come before all vertices of the second copy). Clearly, if the resulting graph  $G'$  were to contain some copy of  $Q$ , then this copy must completely lie in one of the two copies of  $G$ . But  $G$  is  $Q$ -free, meaning that  $G'$  is also  $Q$ -free. In addition, if  $G$  had exactly  $\lceil n/2 \rceil$  vertices and  $\text{ex}_<(\lceil n/2 \rceil, Q)$  edges, then  $G'$  has  $2\lceil n/2 \rceil - 1 \leq n$  vertices and  $2 \text{ex}_<(\lceil n/2 \rceil, Q)$  edges, proving the desired inequality.

On the other hand, if  $Q$  does split into sub-matchings, then we can consider the  $r$ -graph  $G$  on the vertex set  $\{1, \dots, n\}$  containing all possible edges  $e \in E(K_n^{(r)})$  with  $e[1] \leq \lceil n/2 \rceil < e[r]$ . Note that  $G$  is  $Q$ -free. We deduce

$$\text{ex}_<(n, Q) \geq e(G) \geq \binom{n}{r} - \binom{\lceil n/2 \rceil}{r} - \binom{\lfloor n/2 \rfloor}{r} \geq 2 \binom{\lceil n/2 \rceil}{r} \geq 2 \text{ex}_<(\lceil n/2 \rceil, Q),$$

as desired. □

Now, our proof of [Theorem 7.1](#) proceeds by a “contraction” argument, similar to an argument for  $r$ -dimensional orders by Cibulka and Kyncl [18]. We execute this argument in two steps: first we use an inductive contraction argument to estimate the number of  $Q$ -free graphs on  $n$  vertices, then we use this bound and a different contraction argument to upper bound the number of  $Q$ -free ordered matchings of size  $n$ . The first of these steps is encapsulated in the following lemma.

**Lemma 7.4.** Fix a constant  $r \geq 2$ , and suppose  $n$  is sufficiently large in terms of  $r$ . For any ordered  $r$ -matching  $Q$  with more than one edge, the number of ordered  $Q$ -free  $r$ -graphs on the vertex set  $\{1, \dots, n\}$  is at most

$$\exp\left(2^{r+2} \text{ex}_<(\lceil n/2 \rceil, Q)\right).$$

Note that if  $\text{ex}_<(n, Q) \leq \alpha n^{r-1}$  for  $n \geq n_Q$  (as in the statement of [Theorem 7.1](#)), the bound in [Lemma 7.4](#) is at most  $\exp(O(\alpha n^{r-1}))$  for  $n \geq 2n_Q$ .

*Proof.* For any  $n \in \mathbb{N}$ , let  $\mathcal{G}_n$  be the set of all  $r$ -graphs on the vertex set  $\{1, \dots, n\}$  which are  $Q$ -free.

First, we describe how to “contract pairs of vertices” to transform an  $n$ -vertex ordered graph  $G \in \mathcal{G}_n$  into an  $\lceil n/2 \rceil$ -vertex ordered graph  $\phi(G) \in \mathcal{G}_{\lceil n/2 \rceil}$ . For  $G \in \mathcal{G}_n$ , partition its vertex set  $\{1, \dots, n\}$  into  $\lceil n/2 \rceil$  contiguous intervals  $I_1, \dots, I_{\lceil n/2 \rceil}$ , where each  $I_i$  has size 2 except possibly  $I_{\lceil n/2 \rceil}$  (which has size 1 if  $n$  is odd). Then,  $\phi(G)$  is obtained by “contracting” each  $I_i$  to a single vertex. Specifically, we include  $e \in E(K_{\lceil n/2 \rceil}^{(r)})$  as an edge of  $\phi(G)$  if and only if  $\{i_1, \dots, i_r\} \in E(G)$  for some  $i_1 \in I_{e[1]}, i_2 \in I_{e[2]}, \dots, i_r \in I_{e[r]}$ . Note that this contraction operation cannot create copies of  $Q$ .

Now, fixing  $G' \in \mathcal{G}_{\lceil n/2 \rceil}$ , we are going to estimate the number of  $G \in \mathcal{G}_n$  such that  $\phi(G) = G'$ . For all possible edges  $e \in E(K_n^{(r)})$  of  $G$ , say that  $e$  is *contractible* if  $e[1] < \dots < e[r]$  (i.e., if the vertices of  $e$  lie in different  $I_i$ , in the definition of  $\phi(G)$ ).

First, every edge  $e \in E(G')$  arose by contracting some edge  $f_e \in E(G)$ . There are  $2^r$  possibilities for  $f_e$ , and at least one must have been present in  $G$ , so given that  $\phi(G) = G'$  there are at most  $(2^{2^r} - 1)^{\epsilon(G')} < 2^{2^r \text{ex}_<(n', Q)}$  ways to choose the contractible edges in  $G$ .

On the other hand, if  $e \in E(G)$  is not contractible, then  $e[k] + 1 = e[k + 1]$  for some  $1 \leq k < r$ . By [Lemma 2.2](#) and [Proposition 7.2](#), the number of  $e \in E(K_n^{(r)})$  which have this property is

$$\binom{n}{r} - \binom{(n-r)_+}{r} < 2^r \left( \binom{\lceil n/2 \rceil}{r} - \binom{(\lceil n/2 \rceil - r)_+}{r} \right) \leq 2^r \text{ex}_<(\lceil n/2 \rceil, Q),$$

assuming that  $n$  is sufficiently large in terms of  $r$  (say,  $n \geq n_r$ ). So, the total number of ways to choose the non-contractible edges in  $G$  is at most  $2^{2^r \text{ex}_<(\lceil n/2 \rceil, Q)}$ . An  $r$ -graph  $G$  is specified by its contractible and non-contractible edges, so we have

$$|\mathcal{G}_n| \leq |\mathcal{G}_{\lceil n/2 \rceil}| \cdot 2^{2^r \text{ex}_<(\lceil n/2 \rceil, Q)} \cdot 2^{2^r \text{ex}_<(\lceil n/2 \rceil, Q)} \leq |\mathcal{G}_{\lceil n/2 \rceil}| \cdot \exp\left(2^{r+1} \text{ex}_<(\lceil n/2 \rceil, Q)\right).$$

Iterating this inequality, and using that  $\text{ex}_<(\lceil n/2^k \rceil, Q) \leq 2^{-(k-1)} \text{ex}_<(\lceil n/2 \rceil, Q)$  (by [Proposition 7.2](#)), we have

$$|\mathcal{G}_n| \leq |\mathcal{G}_{n_r}| \prod_{k=1}^{\lceil \log_2(n/n_r) \rceil} \exp\left(2^{r+1} \text{ex}_<(\lceil n/2^k \rceil, Q)\right) \leq |\mathcal{G}_{n_r}| \cdot \exp\left((2^{r+2} - 1) \text{ex}_<(\lceil n/2 \rceil, Q)\right).$$

When  $n$  is sufficiently large in terms of  $r$ , [Proposition 7.2](#) implies  $|\mathcal{G}_{n_r}| \leq 2^{\binom{n_r}{r}} \leq \exp(\text{ex}_<(\lceil n/2 \rceil, Q))$ . Then,  $|\mathcal{G}_n| < \exp(2^{r+2} \text{ex}_<(\lceil n/2 \rceil, Q))$ , as desired.  $\square$

We are now ready to complete the proof of [Theorem 7.1](#).

*Proof of Theorem 7.1.* Let  $\mathcal{M}_n$  be the set of ordered  $r$ -matchings on the vertex set  $\{1, \dots, rn\}$  that are  $Q$ -free, and for any  $n'$  let  $\mathcal{G}_{n'}$  be the set of  $r$ -graphs on the vertex set  $\{1, \dots, n'\}$  that are  $Q$ -free.

As in the proof of [Lemma 7.4](#), we will define a “contraction” operation (though this time we will contract larger intervals, and we will not need to iterate our operation). For some  $B$ , whose value we will specify shortly, let  $n' = \lceil rn/B \rceil$  and partition  $\{1, \dots, rn\}$  into  $n'$  contiguous intervals  $I_1, \dots, I_{n'}$ , where each  $I_i$  has size  $B$  except possibly  $I_{n'}$  (which is smaller if  $n$  is not divisible by  $B$ ). For  $M \in \mathcal{M}_n$  let  $\psi(M) \in \mathcal{G}_{n'}$  be the graph on the vertex set  $\{1, \dots, n'\}$  obtained by contracting each interval  $I_i$  to a single vertex (specifically, we include  $e \in E(K_{n'}^{(r)})$  as an edge of  $\psi(G)$  if and only if  $\{i_1, \dots, i_r\} \in E(M)$  for some  $i_1 \in I_{e[1]}, \dots, i_r \in I_{e[r]}$ ). Now, we specify  $B$  to be the largest integer such that  $B^{r-1} \leq \alpha r^n n^{(r-2)}$ . Note that  $B^{r-1} \geq \alpha r^n n^{r-2}/2$  and  $(n')^{r-1} \leq 2(rn/B)^{r-1} = 2r^{r-1} n^{r-1}/B^{r-1} \leq 4n/(r\alpha)$ ; and  $(n')^{r-1} \geq (rn/B)^{r-1} = r^{r-1} n^{r-1}/B^{r-1} \geq n/(r\alpha)$ . By [Lemma 7.4](#) we have

$$|\mathcal{G}_{n'}| \leq \exp\left(2^{r+2} \text{ex}_<(\lceil n'/2 \rceil, Q)\right) \leq \exp\left(2^{r+3} \alpha (n'/2)^{r-1}\right) \leq \exp\left(16\alpha (n')^{r-1}\right) \leq e^{64n/r} \leq e^{32n},$$

assuming  $n$  is sufficiently large with respect to  $r$ . (In this deduction we assumed that  $\alpha \leq 4n^{3/4}/r$  and  $n$  is sufficiently large with respect to  $n_Q$  so that  $(n/(r\alpha))^{1/(r-1)}/2 \geq n_Q$ , thus  $\lceil n'/2 \rceil \geq (n/(r\alpha))^{1/(r-1)}/2 \geq n_Q$ . Note that if  $\alpha > 4n^{3/4}/r$  then  $(n')^{r-1} \leq 4n/r\alpha < n^{1/4}$ , hence  $|\mathcal{G}_{n'}| \leq 2^{\binom{n'}{r}} \leq 2^{(n')^r} < e^{32n}$ . Now, fixing  $G \in \mathcal{G}_{n'}$ , let us estimate the number of  $M \in \mathcal{M}_n$  with  $\psi(M) = G$ . Viewing each edge  $e \in E(G)$  as an ordered  $r$ -tuple  $(e[1], \dots, e[r])$ , let

$$\mathcal{T} = E(G) \cup \{(t_1, \dots, t_r) \in [n']^r : t_1 \leq t_2 \leq \dots \leq t_r \text{ and there exists } k \text{ such that } t_k = t_{k+1}\}.$$

The idea is that  $\mathcal{T}$  specifies the ‘‘possible places where edges of  $M$  can lie’’: if  $\psi(M) = G$  then for each edge  $e \in M$  there must be some tuple  $(t_1, \dots, t_r) \in \mathcal{T}$  such that  $e[1] \in I_{t_1}, \dots, e[r] \in I_{t_r}$  (say that  $e$  is *consistent* with  $\vec{t}$ ). Then,

$$\begin{aligned} |\mathcal{T}| &= |E(G)| + \binom{n' + r - 1}{r} - \binom{n'}{r} \leq \text{ex}_{<}(n', Q) + 2 \binom{n'}{r} - 2 \binom{(n' - r)_+}{r} \\ &\leq 3 \text{ex}_{<}(n', Q) \leq 3\alpha(n')^{r-1} \leq \frac{12n}{r} \leq 6n. \end{aligned}$$

(In the deduction above we assumed that  $\alpha \leq 4n^{3/4}/r$  and  $n$  is sufficiently large with respect to  $n_Q$  and  $r$  so that the first inequality holds, and  $n' \geq (n/(r\alpha))^{1/(r-1)} \geq n_Q$  so that the third inequality holds. In the second inequality we used [Proposition 7.2](#). Also note that the inequality  $|\mathcal{T}| \leq 6n$  holds trivially if  $\alpha > 4n^{3/4}/r$  as this implies that  $n' < n^{1/(4(r-1))}$  and thus  $|\mathcal{T}| \leq (n')^r < n$ .)

Now, for every  $M \in \mathcal{M}_n$  with  $\psi(M) = G$ , we consider the number  $a_{\vec{t}}$  of edges  $e \in M$  which are consistent with  $\vec{t}$ . Since  $\sum_{\vec{t} \in \mathcal{T}} a_{\vec{t}} = n$ , the number of possibilities for  $(a_{\vec{t}} : \vec{t} \in \mathcal{T})$  is at most

$$\binom{n + |\mathcal{T}| - 1}{n} < \binom{n + 6n}{n} \leq 2^{7n} \leq e^{5n}.$$

Then, given any choice of  $(a_{\vec{t}} : \vec{t} \in \mathcal{T})$ , the number of ways to choose the edges of  $M$  is at most

$$\prod_{\vec{t} \in \mathcal{T}} (B^r)^{a_{\vec{t}}} \leq B^{rn}.$$

Thus,

$$|\mathcal{M}_n| \leq |\mathcal{G}_{n'}| \cdot e^{5n} \cdot B^{rn} \leq e^{37n} (\alpha r^r n^{r-2})^{(r/(r-1))n},$$

as desired.  $\square$

Now, we turn our attention to the lower bound in [Theorem 1.15](#). Actually, we are able to prove the desired bound even when restricting our attention to  $r$ -partite  $r$ -matchings (i.e.,  $r$ -dimensional orders). Let  $N_{P,m}^{\text{part}}(n)$  be the number of  $r$ -partite  $r$ -matchings  $M$  on the vertex set  $\{1, \dots, rn\}$  with  $L_P(M) < m$ .

**Theorem 7.5.** *Let  $r \geq 2$  and  $P$  be an  $r$ -partite  $r$ -pattern. Also consider any  $m \geq 2$  such that  $m - 1$  is divisible by  $r - 1$ , and consider any integer  $b \geq r(m - 1)/(r - 1)$ . Then, defining*

$$n = \binom{b + r - (m - 1)/(r - 1)}{r} - \binom{b + r - r(m - 1)/(r - 1)}{r},$$

we have

$$N_{P,m}^{\text{part}}(n) \geq e^{-rn} (r!)^{-(r/(r-1))n} (m - 1)^{(r/(r-1))n} n^{(r-1-1/(r-1))n}.$$

*Remark 7.6.* If  $P$  is an  $r$ -partite  $r$ -pattern with block representation  $|\mathfrak{B}_1| |\mathfrak{B}_2| \dots |\mathfrak{B}_r|$ , and  $M$  is an  $r$ -partite  $r$ -matching (with edges viewed as  $r$ -tuples in  $\{1, \dots, n\}$ ), then the condition  $L_P(M) < m$  is equivalent to the condition that there is no sequence of edges  $e_1, \dots, e_m$  which is increasing in all coordinates  $i$  with  $\mathfrak{B}_i = \text{AB}$ , and decreasing in all coordinates  $i$  with  $\mathfrak{B}_i = \text{BA}$ . So,  $N_{P,m}^{\text{part}}(n)$  does not actually depend on  $P$ , and is precisely equal to the number of  $(r - 1)$ -tuples of permutations of  $\{1, \dots, n\}$  which contain no common increasing subsequence of length  $m$ .

*Remark 7.7.* The bound in [Theorem 7.5](#) is only for  $n$  of a certain form, but it’s easy to deduce almost as strong of a bound for all  $n$  (if we view  $r$  as a constant).

Indeed, first note that if (say)  $b \geq 2m$ , the resulting value of  $n$  scales like  $mb^{r-1}$ . So, for any desired order of magnitude (greater than  $m^r$ ) we can choose  $b$  such that  $n$  has our desired order of magnitude.

Then, note that  $N_{P,m}^{\text{part}}(n + 1) \geq N_{P,m}^{\text{part}}(n)$  for all  $n$  (i.e.,  $N_{P,m}^{\text{part}}(n)$  is monotone increasing in  $n$ ). This is because for any  $(r - 1)$ -tuple of permutations of  $\{1, \dots, n\}$  which contain no common increasing

subsequence of length  $m$ , we can simply prepend each permutation with “ $n+1$ ” to obtain an  $(r-1)$ -tuple of permutations of  $\{1, \dots, n+1\}$  which still contain no common increasing subsequence of length  $m$ .

*Proof of Theorem 7.5.* As we have just discussed (in Remark 7.6), it suffices to consider the case where  $P$  has representation  $|\text{AB}|\text{AB}|\dots|\text{AB}|$ , in which case the condition  $L_P(M) < m$  is equivalent to the condition that there is no sequence of  $m$  edges that is increasing in every coordinate.

We start by constructing an  $r$ -partite ordered  $r$ -graph  $G$ , which we will “uncontract” to obtain our desired matchings  $M$  (by analogy with the contraction operation in the proof of Theorem 7.1 earlier in this section). Let  $a = (m-1)/(r-1)$ , and let  $U_1, \dots, U_r$  be disjoint sets of size  $b-a+1$  (we write  $U_i = \{u_a^i, \dots, u_b^i\}$ ). Define  $G$  to be the graph on the vertex set  $U_1 \cup \dots \cup U_r$  (ordered such that  $U_1 < U_2 < \dots < U_r$ ), where  $\{u_{j_1}^1, u_{j_2}^2, \dots, u_{j_r}^r\}$  is included as an edge if and only if

$$b+1 \leq j_1 + \dots + j_r \leq b+m-1.$$

By symmetry, for any  $j \in \{a, \dots, b\}$ , the vertices  $u_j^1, \dots, u_j^r$  have the same degree, which we denote as  $d_j$ . Note that each  $d_j \geq 1$ , since for any integer  $k$  in the range  $[a, b] = [(m-1)/(r-1), b]$ , there is always some way to add  $r-1$  more integers in this range to  $k$ , to obtain a number in the range  $[b+1, b+m-1]$ . Also, observe that  $b \geq r(m-1)/(r-1) = ra$ . Writing  $n = d_a + \dots + d_b = e(G)$ , we have

$$\begin{aligned} n &= \#\left((j_1, \dots, j_r) \in \{a, \dots, b\}^r : b+1 \leq j_1 + \dots + j_r \leq b+(m-1)\right) \\ &= \#\left((x_1, \dots, x_r) \in \{0, \dots, b-a\}^r : b+1-ra \leq x_1 + \dots + x_r \leq b+(m-1)-ra = b-a\right) \\ &= \#\left((x_1, \dots, x_r) \in \mathbb{Z}_{\geq 0}^r : x_1 + \dots + x_r \leq b-a\right) \\ &\quad - \#\left((x_1, \dots, x_r) \in \mathbb{Z}_{\geq 0}^r : x_1 + \dots + x_r \leq b-ra\right) \\ &= \binom{b+r-(m-1)/(r-1)}{r} - \binom{b+r-r(m-1)/(r-1)}{r}. \end{aligned} \tag{7.1}$$

Starting from the vertex set  $U_1 \cup \dots \cup U_r$  of  $G$ , “uncontract” each vertex  $u_j^i$  into an interval  $I_j^i$  of length  $d_j$ , to obtain an ordered set  $V$  of  $rn$  vertices such that  $I_a^1 < \dots < I_b^1 < I_a^2 < \dots < I_b^{r-1} < I_a^r < \dots < I_b^r$ . Then, we consider each matching on the vertex set  $V$  which “yields  $G$  after contracting the intervals  $I_j^i$  to single vertices”. Specifically, we consider each matching  $M$  on  $V$  with the property that for  $e \in E(M)$  there is a unique edge  $f \in E(G)$  with  $e[1] \in I_{f[1]}^1, e[2] \in I_{f[2]}^2, \dots, e[r] \in I_{f[r]}^r$  (and every  $f \in E(G)$  corresponds to precisely one  $e \in E(M)$ ).

It is not hard to see there are exactly  $\prod_{v \in V(G)} d_G(v) = \left(\prod_j d_j!\right)^r$  ways to choose ordered  $r$ -matchings  $M$  as above. Note that each such matching  $M$  satisfies  $L_P(M) < m$ . To see this, recall that every edge of  $G$  (and therefore every edge of  $M$ ) can be identified by a unique  $r$ -tuple  $(j_1, \dots, j_r)$  with  $b+1 \leq j_1 + \dots + j_r \leq b+m-1$ . For two edges  $e, e' \in E(M)$  (corresponding to two different  $r$ -tuples  $(j_1, \dots, j_r)$  and  $(j'_1, \dots, j'_r)$ ), we can only have  $e[i] > e'[i]$  when  $j_i \geq j'_i$ . In particular, if  $e[i] > e'[i]$  for all  $i$ , then  $j_1 + \dots + j_r > j'_1 + \dots + j'_r$  (note that  $(j_1, \dots, j_r) \neq (j'_1, \dots, j'_r)$  when  $e \neq e'$ ). For our  $r$ -tuples  $(j_1, \dots, j_r)$  under consideration,  $j_1 + \dots + j_r$  is constrained to an interval of  $m-1$  integers, so there cannot be any sequence of  $m$  edges which is increasing in every coordinate (i.e.,  $L_P(M) < m$ ).

It remains to lower-bound  $\left(\prod_j d_j!\right)^r$  in terms of  $n, m, r$ . Recall that the Gamma function  $\Gamma$  is log-convex and satisfies  $\Gamma(x) \geq (x/e)^{x-1}$ . Using that  $d_a + \dots + d_b = n$ , we have

$$\begin{aligned} \left(\prod_{j=a}^b d_j!\right)^r &= \left(\prod_{j=a}^b \Gamma(d_j+1)\right)^r \geq \left(\Gamma\left(\frac{n}{b-a+1}+1\right)\right)^{(b-a+1)r} \\ &> \left(\left(\frac{n}{\epsilon(b-a+1)}\right)^{n/(b-a+1)}\right)^{(b-a+1)r} \\ &= e^{-rn} n^{rn} \left(b - \frac{m-1}{r-1} + 1\right)^{-rn}. \end{aligned} \tag{7.2}$$

On the other hand, recalling (7.1),

$$\begin{aligned}
n &= \binom{b+r-(m-1)/(r-1)}{r} - \binom{b+r-r(m-1)/(r-1)}{r} \\
&= \frac{1}{r!} \left[ \prod_{i=1}^r \left( b - \frac{m-1}{r-1} + i \right) - \prod_{i=1}^r \left( b - \frac{r(m-1)}{r-1} + i \right) \right] \\
&> \frac{1}{r!} \cdot (m-1) \prod_{i=1}^{r-1} \left( b - \frac{m-1}{r-1} + i \right) \\
&> \frac{m-1}{r!} \left( b - \frac{m-1}{r-1} + 1 \right)^{r-1}.
\end{aligned}$$

In other words,  $b - (m-1)/(r-1) + 1 < (r!n/(m-1))^{1/(r-1)}$ . Plugging this into (7.2), we have

$$\begin{aligned}
N_{P,m}^{\text{part}}(n) &\geq \prod_{i=1}^r \prod_{j=a}^b d_j! \geq e^{-rn} n^{rn} \left( \frac{r!n}{m-1} \right)^{-(rn)/(r-1)} \\
&= e^{-rn} (r!)^{-(r/(r-1))n} (m-1)^{(r/(r-1))n} n^{(r-1-1/(r-1))n},
\end{aligned}$$

as desired.  $\square$

## 8. PARTITIONING ORDERED HYPERGRAPHS

A well-known result of Erdős and Kleitman [24] says that every  $r$ -graph with  $m$  edges has an  $r$ -partite subgraph with at least  $(r!/r^r)m$  edges (here “ $r$ -partite” means that we can divide the vertex set into  $r$  parts such that every edge has exactly one vertex in each part). This theorem is enormously useful in extremal hypergraph theory, as it allows one to prove a result about  $r$ -partite  $r$ -graphs and “transfer” it to general  $r$ -graphs.

It would be very useful if the same would be possible in the setting of ordered hypergraphs (where our notion of  $r$ -partiteness should impose that the  $r$  parts are each *intervals* according to our ordering). Although it is not possible to prove a direct analogue of the Erdős–Kleitman theorem, Füredi, Jiang, Kostochka, Mubayi and Verstraëte [31] managed to prove a result that is almost as good: in a certain sense every ordered  $r$ -graph  $G$  has a subgraph  $G'$  which is “nearly”  $r$ -partite, in such a way that the “relative density” of  $G'$  is not too much less than the relative density of  $G$ .

In this section we prove a refinement of the Füredi–Jiang–Kostochka–Mubayi–Verstraëte theorem, with essentially optimal dependence on  $r$ , which might be of independent interest. This will be used in our proof of Theorem 1.18(2). To state our theorem we need some definitions.

**Definition 8.1.** Let  $1 \leq k \leq r$  and  $\vec{a}$  be a positive integer-valued vector of length  $k$  with sum of entries  $\|\vec{a}\|_1 = r$ . We say an ordered  $r$ -graph  $G$  is  $\vec{a}$ -*equipartite* if the following holds.

- (1)  $G$  must have exactly  $kn$  vertices.
- (2) Partition the vertex set of  $G$  into contiguous intervals  $I_1, \dots, I_k$  of size  $n$ . For each  $i \in \{1, \dots, k\}$  and each edge  $e$  of  $G$ , we must have  $|e \cap I_i| = a_i$ .

If  $G$  is  $\vec{a}$ -equipartite with order  $n$ , then we define its  $\ell$ -*density* to be its number of edges divided by  $n^\ell / \prod_i a_i!$  (this is a twisted notion of density, which can take values between 0 and  $n^{r-\ell}$ ).

Moreover, say that  $G$  is  $k$ -equipartite if it is  $\vec{a}$ -equipartite for some  $\vec{a}$  with  $\|\vec{a}\|_0 = k$ , and say that  $G$  is  $(\geq t)$ -equipartite if it is  $k$ -equipartite for some  $k \geq t$ .

Note in particular that there is only one possible  $\vec{a}$  with  $\|\vec{a}\|_0 = r$ : namely, the all-ones vector of length  $r$ . There are  $r-1$  essentially equivalent  $\vec{a}$  with  $\|\vec{a}\|_0 = r-1$ : namely, the the vectors of length  $r-1$ , whose all entries are “1” except one “2”. So, being  $(\geq r-1)$ -equipartite corresponds to being  $\vec{a}$ -equipartite for one of the above two types of  $\vec{a}$ .

Now, our refinement of the Füredi–Jiang–Kostochka–Mubayi–Verstraëte theorem is as follows.

**Theorem 8.2.** Fix  $r \in \mathbb{N}$  and  $\varepsilon > 0$ , and suppose  $n$  is sufficiently large in terms of  $r$  and  $\varepsilon$ . Let  $p \geq \varepsilon$  and let  $G$  be an ordered  $r$ -graph with of  $n$  vertices and at least  $(p/n) \binom{n}{r}$  edges. Then,  $G$  contains a  $(\geq r-1)$ -equipartite subgraph  $G'$  of order  $n^{\Omega_r(1)}$  with  $(r-1)$ -density at least  $\Omega(p/r^2)$ .

We also show that the multiplicative factor  $O(r^{-2})$  in the “loss of density” in [Theorem 8.2](#) is essentially best possible.

**Theorem 8.3.** *Fix a constant  $r \in \mathbb{N}$ . There exists an ordered  $r$ -graph  $G$  with  $(1 + o(1))r(r-1)n^{r-1}/r!$  edges such that any order- $m$  ( $\geq r-1$ )-equipartite subgraph  $G' \subseteq G$  has  $(r-1)$ -density at most  $2 + o_{m \rightarrow \infty}(1)$ .*

*Remark 8.4.* Instead of making an  $(r-1)$ -density assumption, one could make an  $\ell$ -density assumption for any  $\ell \leq r$ . In general, such an assumption would allow one to prove an analogue of [Theorem 8.2](#) in which we guarantee a  $(\geq \ell)$ -equipartite subgraph with reasonably large  $\ell$ -density. (Such a result was previously proved in [\[31\]](#), and our proof approach can always attain a better dependence on  $r$ ). However, since  $(r-1)$ -density is the relevant parameter in most applications, we prefer to keep our proof as simple as possible, and only treat the case  $\ell = r-1$ .

To prove [Theorem 8.2](#), the following lemma turns out to be crucial. Roughly speaking, given any  $\vec{a}$ -equipartite  $G$  (for some  $\vec{a}$ ), we can either obtain a subgraph of  $G$  with more parts or a subgraph of  $G$  with the same number of parts but higher density.

**Lemma 8.5.** *Suppose  $1 \leq k \leq r$  and  $\vec{a}$  is a positive integer-valued vector of length  $k$  with  $\|\vec{a}\|_1 = r$ . Let  $G$  be an  $\vec{a}$ -equipartite ordered  $r$ -graph with  $(r-1)$ -density  $p$ , and consider any  $R \in \mathbb{N}$ . Suppose that  $n$  is sufficiently large in terms of  $r, R$ . Then, for some  $t \in \mathbb{N}$  satisfying  $\|\vec{a}\|_0 \leq t \leq r$ , we can find a  $t$ -equipartite subgraph of order  $\lfloor n/R \rfloor$  in  $G$ , with  $(r-1)$ -density at least*

$$\frac{p}{4R} \left( \frac{R}{(r-k+1)^2} \right)^{r-t}.$$

We remark that in practice we will apply [Lemma 8.5](#) with  $R$  a large constant.

*Proof.* Write  $I_1, \dots, I_k$  for the  $k$  parts of  $G$ . First, it is convenient to reduce to a subgraph whose order is divisible by  $R$ . So, let  $n' = R \lfloor n/R \rfloor$ . By averaging, there’s a way to delete  $n - n' < R$  vertices from each  $I_i$  to obtain an  $\vec{a}$ -equipartite subgraph  $G'$ , such that at most  $(rR/n)e(G)$  edges are deleted. So, the  $(r-1)$ -density of  $G'$  is at least  $p(1 - rR/n) \geq p/2$  for  $n \geq 2rR$ .

Now, we divide each  $I_i$  into  $R$  contiguous intervals  $I_{i,1}, \dots, I_{i,R}$  of the same length  $n/R$ . Write  $\mathbb{1} \in \mathbb{R}^R$  be the all-ones vector of length  $R$ , and for every  $t \in \{k, k+1, \dots, r\}$ , define  $\mathcal{B}_{\vec{a},t}^R$  to be the set of  $k \times R$  matrices  $b \in \mathbb{Z}_{\geq 0}^{k \times R}$  with  $\|b\|_0 = t$  nonzero entries, satisfying  $b\mathbb{1} = \vec{a}$ . Then, for each  $b \in \mathcal{B}_{\vec{a},t}^R$ , we define a  $t$ -equipartite subgraph  $G_b$  of order  $\lfloor n/R \rfloor$ , as follows. The vertex set of  $G_b$  is obtained by including part  $I_{i,j}$  whenever  $b_{ij} \neq 0$ , and the edge set is obtained by including each  $e \in G'$  if  $|e \cap I_{i,j}| = b_{ij}$  for all  $i, j$ . Note that the subgraphs  $G_b$  (as  $b$  varies in  $\mathcal{B}_{\vec{a},t}^R$ , while  $t$  varies in  $\{k, \dots, r\}$ ) partition the edges of  $G'$ .

Now, if none of the subgraphs  $G_b$  satisfied the conclusion of the lemma, the total number of edges in  $G'$  would be

$$e(G') < \sum_{t=k}^r \sum_{b \in \mathcal{B}_{\vec{a},t}^R} \frac{p}{4R} \left( \frac{R}{(r-k+1)^2} \right)^{r-t} \frac{(n/R)^{r-1}}{\prod_{i,j} b_{ij}!} = \sum_{t=k}^r \frac{p \cdot R^{-t}}{4(r-k+1)^{2(r-t)}} \frac{n^{r-1}}{\prod_i a_i!} \sum_{b \in \mathcal{B}_{\vec{a},t}^R} \frac{\prod_i a_i!}{\prod_{i,j} b_{ij}!}.$$

To show that this leads to contradiction, we need the following combinatorial inequality.

**Claim 8.6.** *For every  $k \leq t \leq r$ , we have  $\sum_{b \in \mathcal{B}_{\vec{a},t}^R} \left( \prod_i a_i! / \prod_{i,j} b_{i,j}! \right) \leq R^t ((r-k+1)^2/2)^{r-t}$ .*

We will prove [Claim 8.6](#) momentarily, but first we see how to use it to conclude the proof of [Lemma 8.5](#): recalling our estimate for  $e(G')$  above, we have

$$e(G') < \sum_{t=k}^r \frac{p \cdot R^{-t}}{4(r-k+1)^{2(r-t)}} \frac{n^{r-1}}{\prod_i a_i!} \cdot R^t \left( \frac{(r-k+1)^2}{2} \right)^{r-t} = \sum_{t=k}^r \frac{p}{4} \cdot 2^{-(r-t)} \cdot \frac{n^{r-1}}{\prod_i a_i!} < \frac{(p/2)n^{r-1}}{\prod_i a_i!},$$

contradicting the fact that  $G'$  has  $(r-1)$ -density at least  $p/2$ .  $\square$

Now we prove [Claim 8.6](#).

*Proof of Claim 8.6.* We prove this inequality combinatorially, relating both sides of the inequality to certain counts of tuples of set partitions.

In this proof we write  $[n] = \{1, \dots, n\}$ , and write  $\|f\|_0$  for the number of elements in the image of a function  $f$ . For  $1 \leq q \leq n$ , let  $\mathcal{P}_{n,q}$  be the collection of (unlabelled) partitions of  $[n]$  into  $q$  nonempty



subsets (so  $|\mathcal{P}_{n,q}| = \left\{ \begin{matrix} n \\ q \end{matrix} \right\}$  is a Stirling number of the second kind). Let  $\mathcal{P}_n = \bigcup_{q=1}^n \mathcal{P}_{n,q}$  be the collection of partitions of  $[n]$  into any number of subsets, and let  $\|P\|_0$  be the number of parts of a partition  $P \in \mathcal{P}_n$ . Then, define

$$\mathcal{P}_{\vec{a},t} = \left\{ (P_1, \dots, P_k) \in \mathcal{P}_{a_1} \times \dots \times \mathcal{P}_{a_k} \mid \sum_{i=1}^k \|P_i\|_0 = t \right\}.$$

Note that

$$\sum_{b \in \mathcal{B}_{\vec{a},t}^R} \frac{\prod_i a_i!}{\prod_{i,j} b_{i,j}!} = \sum_{b \in \mathcal{B}_{\vec{a},t}^R} \prod_i \binom{a_i}{b_{i,1} \dots b_{i,R}} \leq R^t |\mathcal{P}_{\vec{a},t}|. \quad (8.1)$$

(To see this, observe that the right-hand side of (8.1) can be interpreted as the number of ways to choose a tuple of partitions  $(P_1, \dots, P_k) \in \mathcal{P}_{\vec{a},t}$  together with a labelling of the parts in these partitions, in such a way that for each  $1 \leq i \leq k$ , each of the parts in  $P_i$  is assigned a distinct label from  $[R]$ ).

Now, note that every  $(P_1, \dots, P_k) \in \mathcal{P}_{\vec{a},t}$  gives rise to a partition  $P \in \mathcal{P}_{r,t}$ : first,  $P_1$  describes a partition of  $\{1, \dots, a_1\}$ , then  $P_2$  describes a partition of  $\{a_1 + 1, \dots, a_1 + a_2\}$ , and  $P_3$  describes a partition of  $\{a_1 + a_2 + 1, \dots, a_1 + a_2 + a_3\}$ , and so on. We can then “collapse”  $P$  into a partition  $P' \in \mathcal{P}_{r-k+1, t-k+1}$ , by sequentially merging the parts containing  $a_1$  and  $a_1 + 1$ , then merging the parts containing  $a_1 + a_2$  and  $a_1 + a_2 + 1$ , and so on. Since this mapping from  $(P_1, \dots, P_k)$  to  $P'$  is injective, we deduce that

$$|\mathcal{P}_{\vec{a},t}| \leq |\mathcal{P}_{r-k+1, t-k+1}| = \left\{ \begin{matrix} r-k+1 \\ t-k+1 \end{matrix} \right\}.$$

Then, note that for all  $1 \leq q \leq n$  we have

$$\left\{ \begin{matrix} n \\ q \end{matrix} \right\} \leq \prod_{i=0}^{n-q-1} \binom{n-i}{2} \leq \binom{n}{2}^{n-q},$$

because we can generate a partition  $P \in \mathcal{P}_{n,q}$  by starting with the trivial partition  $P_0 \in \mathcal{P}_{n,n}$  into  $n$  parts, and iteratively merging pairs of parts ( $n - q$  times). We deduce that

$$|\mathcal{P}_{\vec{a},t}| \leq \binom{r-k+1}{2}^{r-t} < \left( \frac{(r-k+1)^2}{2} \right)^{r-t}.$$

Combining this with (8.1), the desired result follows.  $\square$

*Proof of Theorem 8.2.* Let  $C$  be a large constant ( $C = 200$  will do). Let  $R(k) = C \cdot (r - k + 1)^2$ .

The plan is to repeatedly apply Lemma 8.5 until we find the desired subgraph. First, let  $G_0 = G$ ,  $p_0 = p/2$ ,  $n_0 = n$  and  $k_0 = 1$ . Observe that  $G_0$  (like any ordered  $r$ -graph) is  $\vec{a}$ -equipartite for  $\vec{a} = (r)$  (i.e., 1-equipartite). It has order  $n$  and  $(r-1)$ -density at least  $(p/n) \binom{n}{r} / (n^{r-1}/r!) > p/2$  (assuming  $n$  is sufficiently large in terms of  $r$ ). Then, for each  $i \geq 1$  (“at step  $i$ ”), we take the  $k_{i-1}$ -equipartite subgraph  $G_{i-1}$  with  $(r-1)$ -density  $p_{i-1}$ , and apply Lemma 8.5 with  $R = R(k_{i-1})$  to obtain a  $k_i$ -equipartite subgraph  $G_i$  with order  $n_i = \lfloor n_{i-1}/R \rfloor$  and with  $(r-1)$ -density at least

$$p_i = \frac{p_{i-1}}{4R} \left( \frac{R}{(r - k_{i-1} + 1)^2} \right)^{r-k_i} = \frac{p_{i-1}}{4(r - k_{i-1} + 1)^2} C^{r-k_i-1}, \quad (8.2)$$

for some  $k_i \geq k_{i-1}$ . We stop when  $k_i \geq r-1$  (i.e., when we have found a  $(\geq r-1)$ -equipartite subgraph), and we also abort the process if at any point  $G_{i-1}$  has too few vertices to apply Lemma 8.5 with the desired  $R$ . Let  $\tau$  be the total number of steps taken.

We first prove that  $p_j \geq p/(8Cr^2)$  for  $j \leq \tau$ . In fact, we prove, by reverse induction, the stronger statement that for each step  $i < j$  we have

$$p_j \geq \frac{p_i}{4C(r - k_i + 1)^2}.$$

The base case (where  $i = j-1$ ) follows from (8.2), so consider some  $i < j-1$  and suppose as our induction hypothesis that the statement is true for  $i+1$ . Since  $i+1 \leq j-1 \leq \tau-1$ , we have  $k_{i+1} \leq r-2$ . Hence,

$$p_j \geq \frac{p_{i+1}}{4C(r - k_{i+1} + 1)^2} = \frac{p_i}{(4C)^2(r - k_{i+1} + 1)^2(r - k_i + 1)^2} C^{r-k_{i+1}} \geq \frac{p_i}{(4C)(r - k_i + 1)^2},$$

using the induction hypothesis and (8.2) (for  $i+1$ ), and the fact that  $C^{r-k}/(4C(r-k+1)^2) \geq 1$  for  $k \leq r-2$  and  $C \geq 36$ .

It now suffices to show that this process cannot continue for too many steps (in particular, we want to make sure that the graphs  $G_i$  never get too small). For each  $1 \leq k \leq r-2$ , let  $i(k)$  be the first step  $i$  such that  $k_i \geq k$ , and let  $N(k)$  be the number of steps  $i > i(k)$  such that  $k_i = k$ . Note that for each step  $i(k) < i \leq i(k) + N(k)$  we have  $p_i \geq (C/40)p_{i-1}$  by (8.2) (because  $C^{r-k}(r-k+1)^{-2} \geq C^2/10$  for  $k \leq r-2$ ) and  $n_i \leq n_{i-1}/R(k)$ . This gives  $p_{i(k)+N(k)} \geq (C/40)^{N(k)}p_{i(k)}$  and  $n_{i(k)+N(k)} \leq n_{i(k)}/R(k)^{N(k)}$ . But then, due to Definition 8.1, the  $(r-1)$ -density of  $G_{i(k)+N(k)}$  is at most  $n_{i(k)+N(k)}$ , so  $p_{i(k)+N(k)} \leq n_{i(k)+N(k)}$ . This, plus the fact that  $p_j \geq p/(8Cr^2) \geq \varepsilon/(8Cr^2)$  for  $j \leq \tau$ , implies

$$(C/40)^{N(k)} \cdot \varepsilon/(8Cr^2) \leq (C/40)^{N(k)}p_{i(k)} \leq p_{i(k)+N(k)} \leq n_{i(k)+N(k)} \leq n_{i(k)}/R(k)^{N(k)}.$$

So we must have  $N(k) \leq \log(8Cr^2n_{i(k)}/\varepsilon)/\log(R(k)C/40) \leq \log(n_{i(k)})/\log(4R(k))$ , where in the latter inequality we assume  $n_{i(k)}$  is sufficiently large in terms of  $r$  and  $\varepsilon$  and  $C$  is a large constant. Consequently,  $n_{i(k)+N(k)} \geq (2R(k))^{-N(k)} \cdot n_{i(k)} \geq n_{i(k)}^{\alpha(k)}$ , where  $\alpha(k) = 1 - \log(2R(k))/\log(4R(k)) > 0$ , and  $n_{i(k+1)} = n_{i(k)+N(k)+1} \geq n_{i(k)+N(k)}/(2R(k)) \geq n_{i(k)}^{\alpha(k)}/(2R(k))$ .

So,  $n_i$  is at least  $n^{\alpha(1)\dots\alpha(r-2)}/(2^{r-2}R(1)\dots R(r-2))$  for each  $i$ . Note that  $\alpha(1)\dots\alpha(r-2) > 0$  and  $R(1)\dots R(r-2)$  only depend on  $r$ , so if  $n$  is sufficiently large in terms of  $r$  then we can guarantee that each  $n_i$  is always large enough (in terms of  $r$  and  $R(k_i)$ ) to apply Lemma 8.5 (i.e., the process never aborts). The process therefore terminates with the  $(\geq r-1)$ -equipartite subgraph  $G_\tau$  with  $(r-1)$ -density  $p_\tau \geq p/(8Cr^2)$ . This concludes the proof.  $\square$

Now we prove Theorem 8.3.

*Proof of Theorem 8.3.* Let  $G$  be the ordered  $r$ -graph on the vertex set  $\{1, \dots, n\}$ , with every possible edge that contains two consecutive vertices. By Lemma 2.2, the number of edges in  $G$  is

$$\binom{n}{r} - \binom{(n-r+1)_+}{r} = (1 + o(1))r(r-1)\frac{n^{r-1}}{r!}.$$

- (1) If  $G'$  is an  $r$ -equipartite subgraph with parts  $I_1 < \dots < I_r$  of size  $m$ , then for every edge  $e$  there must be some pair of parts  $I_i, I_{i+1}$  such that  $e$  contains the last element of  $I_i$  and the first element of  $I_{i+1}$  (and these two elements must be consecutive). So, the number of edges in  $G'$  is at most  $(r-1)m^{r-2}$ , meaning that  $G'$  has  $(r-1)$  density at most  $(r-1)/m = o_{m \rightarrow \infty}(1)$ .
- (2) If  $\vec{a}$  satisfies  $\|\vec{a}\|_0 = r-1$  and  $G'$  is an  $\vec{a}$ -equipartite subgraph with parts  $I_1 < \dots < I_{r-1}$  of size  $m$ , then we can divide the edges of  $G'$  into two types.
  - First, we could have some edges  $e$  containing the last element of some  $I_i$  and the first element of  $I_{i+1}$ , as in (1). The number of such edges is at most  $(r-2)m^{r-2}$ .
  - Second, writing  $j$  for the single index with  $a_j = 2$ , we could have some edges  $e$  containing two consecutive vertices in  $I_j$ . The number of such edges is at most  $(m-1)m^{r-2}$ .
All in all, the number of edges in  $G'$  is at most  $(1 + o_{m \rightarrow \infty}(1))m^{r-1}$ , meaning that  $G'$  has  $(r-1)$ -density at most  $2 + o_{m \rightarrow \infty}(1)$ .  $\square$

## 9. EXTREMAL NUMBERS FOR $P$ -CLIQUES

In this section, we prove Theorems 1.17 and 1.18. First, the following theorem is a generalisation of Theorem 1.18(1). It will also be used for the lower bound in Theorem 1.17.

**Theorem 9.1.** *Let  $r, n \geq 1$ , let  $P$  be an  $r$ -pattern, and let  $m \geq 1$ . If  $P$  is collectable or  $m \in \{1, 2\}$  then*

$$\text{ex}_{<}(n, P^{(m)}) \geq \binom{n}{r} - \binom{(n-r(m-1))_+}{r}.$$

(Note that Theorem 9.1 allows  $P$  to be non-collectable if  $m \in \{1, 2\}$ . We do not need to worry about non-collectable patterns in our proofs of Theorems 1.17 and 1.18 the extra generality is so that Proposition 7.2 (which we stated without proof in Section 7) is a direct corollary.

*Proof.* Note that  $\text{ex}_{<}(n, P^{(1)}) = 0$  and  $\text{ex}_{<}(n, P^{(m)}) = \binom{n}{r}$  if  $rm > n$ . We may assume  $2 \leq m \leq n/r$  for the rest of the proof.

Write  $f_1, f_2$  for the two edges of  $P$ , with  $f_1[1] < f_2[1]$ . For each  $1 \leq i < r$ , let  $s_i = \#\{j \in [r] : f_1[i] < f_2[j] < f_1[i+1]\}$  be the number of vertices of  $f_2$  lying between the  $i$ -th and  $(i+1)$ -th vertices of  $f_1$ .

Let  $s_r = \#\{j \in [r] : f_2[j] > f_1[r]\}$  be the number of vertices of  $f_2$  after all the vertices of  $f_1$ . Clearly,  $s_1 + \dots + s_r = r$ .

Let  $G$  be the ordered  $r$ -graph with vertex set  $\{1, 2, \dots, n\}$ , where we include an edge  $e$  whenever  $e[i+1] - e[i] \leq s_i \cdot (m-1)$  for some  $1 \leq i < r$  or  $e[r] \geq n - s_r \cdot (m-1) + 1$ . We first claim that  $G$  is  $P^{(m)}$ -free. Indeed, suppose for the purpose of contradiction that the edges  $e_1, \dots, e_m$  in  $G$  form  $P^{(m)}$ , with  $e_1[1] < \dots < e_m[1]$ . Fix  $2 \leq j \leq m$ . As  $e_1, e_j$  form the pattern  $P$ , we know that for all  $1 \leq i < r$ , there are  $s_i$  vertices of  $e_j$  lying between  $e_1[i], e_1[i+1]$ , and that there are  $s_r$  incident vertices of  $e_j$  lying after  $e_1[r]$ . By considering all  $j$ , it holds that  $e_1[i+1] - e_1[i] > s_i \cdot (m-1)$  for  $1 \leq i < r$  and  $e_1[r] \leq n - s_r \cdot (m-1)$ . But then,  $e_1$  should not have been put into  $G$ , a contradiction.

Now, it suffices to count the number of edges in  $G$ . It is convenient to instead count *non-edges*: an  $r$ -subset  $e$  of  $\{1, \dots, n\}$  is a non-edge of  $G$  if and only if  $e[i+1] - e[i] > s_i \cdot (m-1)$  for all  $1 \leq i < r$  and  $e[r] \leq n - s_r \cdot (m-1)$ . By [Lemma 2.2](#), the number of such  $e$  is exactly

$$\binom{n - s_r(m-1) - \sum_{i=1}^{r-1} s_i(m-1)}{r} = \binom{n - r(m-1)}{r}.$$

Thus, the number of edges in  $G$  is

$$\binom{n}{r} - \binom{n - r(m-1)}{r},$$

so this is a lower bound on  $\text{ex}_{<}(n, P^{(m)})$ , as desired.  $\square$

*Remark 9.2.* In fact, the above lower bound works in a stronger setting: given any  $r$ -pattern  $P$  and any  $r$ -matching  $Q$  with edges  $e_1, \dots, e_m$  (ordered such that  $e_1[1] < e_2[1] < \dots < e_m[1]$ ), such that  $e_1, e_i$  form the pattern  $P$  for all  $2 \leq i \leq m$ , we have

$$\text{ex}_{<}(n, Q) \geq \binom{n}{r} - \binom{(n - r(m-1))_+}{r}.$$

We next prove [Theorem 1.18\(3\)](#).

*Proof of Theorem 1.18(3).* Without loss of generality, we may assume that  $m \leq n/r$ , as otherwise  $\text{ex}_{<}(n, P^{(m)}) = \binom{n}{r}$ , as desired. Let  $P$  be the ‘‘alternating pattern’’ represented by  $|\text{AB}|\text{BA}|\text{AB}|\text{BA}|\dots|$ , and let  $\tilde{G}$  be the ordered  $r$ -graph in the proof of [Theorem 1.18\(1\)](#) above, i.e.,  $e$  is an edge of  $G$  if  $e[2i] - e[2i-1] \leq 2(m-1)$  for some  $1 \leq i \leq r/2$  or if  $e[r] \geq n - (m-1) + 1$  and  $r$  is odd. The number of edges in  $\tilde{G}$  is

$$\binom{n}{r} - \binom{n - r(m-1)}{r}.$$

Consider any  $n$ -vertex  $P^{(m)}$ -free ordered  $r$  graph  $G$ . We are going to show that  $e(G) \leq e(\tilde{G})$ . Our strategy is to partition the collection of all possible edges into subsets, and to separately show that in each subset the number of edges of  $G$  is at most the number of edges of  $\tilde{G}$ .

Let  $E(K_n^{(r)})$  be the collection of all subsets of  $\{1, \dots, n\}$  (i.e., all edges of the complete  $r$ -graph). We define an equivalence relation  $\sim$  on  $E(K_n^{(r)})$ : for  $e, f \in E(K_n^{(r)})$ , we write  $e \sim f$  if there is  $\ell \in \mathbb{Z}$  such that  $f[i] = e[i] + \ell$  whenever  $i$  is odd, and  $f[i] = e[i] - \ell$  whenever  $i$  is even. Fix an equivalence class  $A$  of  $\sim$ . It is easy to see that any two edges in  $A$  form the pattern  $P$ , and thus all the edges (in  $K_n^{(r)}$ ) in  $A$  form a  $P$ -clique<sup>4</sup>.

Now, it suffices to show that  $|E(G) \cap A| \leq |E(\tilde{G}) \cap A|$  for every equivalence class  $A$  of  $\sim$ . Suppose there exists some  $e \in A \cap (E(G) \setminus E(\tilde{G}))$  (otherwise  $|E(G) \cap A| \leq |E(\tilde{G}) \cap A|$  holds trivially). Since  $e \notin E(\tilde{G})$ , we have  $e[2i] - e[2i-1] > 2(m-1)$  for all  $1 \leq i \leq r/2$ , and  $e[r] \leq n - (m-1)$  if  $r$  is odd. For  $\ell \in \mathbb{N}$ , let

$$f_\ell(e) := \{e[1] + \ell, e[2] - \ell, e[3] + \ell, e[4] - \ell, \dots\}.$$

As we increase  $\ell$ , we are ‘‘pushing each  $e[2i-1]$  and  $e[2i]$  towards each other’’, and ‘‘pushing  $e[r]$  towards  $n$ ’’ (if  $r$  is odd). At some point, there will be a ‘‘collision’’: let  $\ell^*$  be the ‘‘time just before this collision’’, i.e., the largest  $\ell$  for which  $e[r] + (-1)^{r+1}\ell \leq n$  and  $e[2i-1] + \ell < e[2i] - \ell$  for all  $1 \leq i \leq r/2$ . Note that  $\ell^* \geq m-1$  (since  $e \notin E(\tilde{G})$ ) and that  $f_\ell(e) \in E(\tilde{G})$  for all  $\ell \in \{\ell^* - (m-2), \dots, \ell^*\}$  (the condition for  $f$  to be an edge of  $\tilde{G}$  is the condition that  $f$  is ‘‘within  $m-1$  steps of a collision’’). Note that  $f_\ell(e) \sim e$

<sup>4</sup>This is the crucial fact we are using about the alternating pattern (for all other patterns  $P$ , one can define similar equivalence classes, but it is not true that all equivalence classes are  $P$ -cliques).

for each  $\ell \leq \ell^*$ , so  $|E(\tilde{G}) \cap A| \geq m - 1$ . On the other hand,  $E(G) \cap A$  is a  $P$ -clique as every two edges in  $A$  form the pattern  $P$ . Since  $G$  is  $P^{(m)}$ -free, we have  $|E(G) \cap A| \leq m - 1$ . This proves that  $|E(G) \cap A| \leq |E(\tilde{G}) \cap A|$ , as desired.  $\square$

Now, we prove [Theorem 1.17](#) (giving the exact extremal numbers for all  $r$ -partite  $r$ -patterns  $P$ ).

*Proof of Theorem 1.17.* We only need to prove the upper bound, as the lower bound is given by the  $m = 2$  case of [Theorem 1.18\(1\)](#).

We proceed by induction on  $r$  (the base case  $r = 1$  is trivial). Fix an  $r$ -partite  $r$ -pattern  $P$ , and write  $|\mathfrak{B}_1| \cdots |\mathfrak{B}_r|$  for its block representation, where  $\mathfrak{B}_1 = \text{AB}$  and  $\mathfrak{B}_i \in \{\text{AB}, \text{BA}\}$  for  $2 \leq i \leq r$ . Let  $G$  be an  $n$ -vertex  $P$ -free ordered  $r$ -graph. We may assume that  $n \geq r$ ; our goal is to prove that

$$e(G) \leq \binom{n}{r} - \binom{n-r}{r}.$$

Let  $\hat{P}$  be the pattern with block representation  $|\mathfrak{B}_1| |\mathfrak{B}_2| \cdots |\mathfrak{B}_{r-1}|$ , and for an edge  $e = \{e[1], \dots, e[r]\} \in E(K_n^{(r)})$ , let  $\hat{e} = \{e[1], \dots, e[r-1]\} \in E(K_{n-1}^{(r-1)})$ . Let  $G'$  be obtained from  $G$  as follows. For each  $f \in E(K_{n-1}^{(r-1)})$ , consider all edges  $e \in E(G)$  with  $\hat{e} = f$ . If there are any such edges, then delete the one with the largest value of  $e[r]$ . Note that  $e(G') \geq e(G) - \binom{n-1}{r-1}$ .

Then, for  $v \in \{1, \dots, n\}$ , let  $\hat{G}_v$  be the graph with vertex set  $\{1, \dots, v-1\}$  and edge set

$$\{\hat{e} : e \in E(G'), e[r] = v\}.$$

We claim that  $\hat{G}_v$  is  $\hat{P}$ -free. To see this, suppose for the purpose of contradiction that  $\hat{e}_1, \hat{e}_2 \in E(\hat{G}_v)$  form pattern  $\hat{P}'$  with  $e_1[1] < e_2[1]$ . By the definition of  $G'$ , there are  $x, y \in \{v+1, \dots, n\}$  such that  $\{e_1[1], \dots, e_1[r-1], x\}, \{e_2[1], \dots, e_2[r-1], y\} \in E(G)$ . But note that

- if  $\mathfrak{B}_r = \text{AB}$  then  $\{e_1[1], \dots, e_1[r-1], v\}, \{e_2[1], \dots, e_2[r-1], y\}$  form pattern  $P$ ;
- if  $\mathfrak{B}_r = \text{BA}$  then  $\{e_1[1], \dots, e_1[r-1], x\}, \{e_2[1], \dots, e_2[r-1], v\}$  form pattern  $P$ .

In either case, we have found  $P$  in  $G$ , which is a contradiction. So, each  $\hat{G}_v$  is  $\hat{P}$ -free and by induction we have

$$e(\hat{G}_v) \leq \binom{v-1}{r-1} - \binom{(v-1-(r-1))_+}{r-1} = \binom{v-1}{r-1} - \binom{(v-r)_+}{r-1}.$$

Also, note that  $e(\hat{G}_n) = 0$ , because any  $e \in E(G)$  with  $e[r] = n$  must have been deleted in the construction of  $G'$ . Consequently,

$$\begin{aligned} e(G) &\leq \binom{n-1}{r-1} + e(G') = \binom{n-1}{r-1} + \sum_{v=1}^{n-1} e(\hat{G}_v) \\ &\leq \binom{n-1}{r-1} + \sum_{v=1}^{n-1} \left[ \binom{v-1}{r-1} - \binom{(v-r)_+}{r-1} \right] \\ &\leq \binom{n-1}{r-1} + \sum_{k=0}^{n-2} \binom{k}{r-1} - \sum_{k=0}^{n-r-1} \binom{k}{r-1} \\ &= \binom{n-1}{r-1} + \binom{n-1}{r} - \binom{n-r}{r} = \binom{n}{r} - \binom{n-r}{r}, \end{aligned}$$

as desired.  $\square$

The last result we prove in this section is [Theorem 1.18\(2\)](#). Starting from a  $P^{(m)}$ -free  $r$ -graph  $G$ , we will apply [Theorem 8.2](#) to pass to a  $(\geq r-1)$ -equipartite subgraph (whose  $(r-1)$ -density is comparably large, and which is still  $P^{(m)}$ -free). To this end, we need a generalisation of  $\text{ex}_{<}(n, F)$ .

**Definition 9.3.** Let  $r \geq 1$  and  $\vec{a}$  be a positive integer-valued vector with sum of entries  $\|\vec{a}\|_1 = r$ . For an ordered  $r$ -graph  $H$ , we write  $\text{ex}_{<}^{\vec{a}}(n, H)$  for the maximum number of edges in an  $H$ -free  $\vec{a}$ -equipartite  $r$ -graph of order  $n$ .

We separately consider the  $r$ -equipartite ( $\|\vec{a}\|_0 = r$ ) and  $(r-1)$ -equipartite ( $\|\vec{a}\|_0 = r-1$ ) cases. First, we consider the  $r$ -equipartite case.

**Lemma 9.4.** *Let  $n, m, r \geq 1$ . Let  $\vec{a} = (1, 1, \dots, 1)$  be the all-one vector of length  $r$ , and  $P$  be an  $r$ -partite  $r$ -pattern. Then,*

$$\text{ex}_{<}^{\vec{a}}(n, P^{(m)}) = n^r - ((n - m + 1)_+)^r \leq r(m - 1)n^{r-1}.$$

*Proof.* Let  $V_1, \dots, V_r$  be disjoint copies of  $\{1, \dots, n\}$ , and let  $\mathcal{V} = V_1 \times \dots \times V_r$ . Note that an  $\vec{a}$ -equipartite  $r$ -graph  $G$  can be viewed as a set of tuples  $e \in \mathcal{V}$ . From this perspective, a copy of  $P^{(m)}$  (where  $P$  has block representation  $|\mathfrak{B}_1|\mathfrak{B}_2|\dots|\mathfrak{B}_r|$ ) corresponds to a sequence of edges  $e_1, \dots, e_m \in E(G)$  which is increasing in all coordinates  $i$  with  $\mathfrak{B}_i = \text{AB}$ , and decreasing in all coordinates  $i$  with  $\mathfrak{B}_i = \text{BA}$ . By symmetry,  $\text{ex}_{<}^{\vec{a}}(n, P^{(m)})$  depends only on  $n$  and  $m$  (i.e., it does not depend on  $P$ ). Therefore, it suffices to consider the case where  $P$  has representation  $|\text{AB}|\text{AB}|\dots|\text{AB}|$  (i.e.,  $P$  is a “generalised crossing”).

For the lower bound, we simply consider the  $\vec{a}$ -equipartite  $r$ -graph  $\tilde{G}$  consisting of all possible edges  $e \in \mathcal{V}$  such that some  $e[i] < m$ . By the above discussion, any copy of  $P^{(m)}$  must contain an edge with all coordinates at least  $m$ . Thus,  $\tilde{G}$  is  $P^{(m)}$ -free, and it has the desired number of edges.

For the upper bound, consider a  $P^{(m)}$ -free  $n$ -vertex graph  $G$  with parts  $V_1, \dots, V_r$ . Define an equivalence relation  $\sim$  on  $\mathcal{V}$  by taking  $e \sim f$  when there is some  $\ell \in \mathbb{Z}$  such that each  $e[i] = f[i] + \ell$ . We proceed very similarly to the proof of [Theorem 1.18\(3\)](#), showing that  $|A \cap E(G)| \leq |A \cap E(\tilde{G})|$  for each equivalence class  $A$  of  $\sim$ .

Consider some equivalence class  $A$ , and suppose there is  $e \in A \cap (E(G) \setminus E(\tilde{G}))$  (otherwise trivially  $|A \cap E(G)| \leq |A \cap E(\tilde{G})|$ ). Since  $e \notin E(\tilde{G})$ , each  $e[i] \geq m$ . Define  $f_\ell(e) = (e[1] - \ell, \dots, e[r] - \ell)$ , and let  $\ell^* = \min_i e[i] - 1 \geq m - 1$ . Note that  $f_\ell(e) \in E(\tilde{G})$  for all  $\ell \in \{\ell^* - (m - 2), \dots, \ell^*\}$ , and note that  $f_\ell(e) \sim e$  for each  $\ell \leq \ell^*$ . So,  $|A \cap E(\tilde{G})| \geq m - 1$ . On the other hand, note that  $A \cap E(G)$  is a  $P$ -clique so  $|A \cap E(G)| \leq m - 1$ . This proves that  $|A \cap E(G)| \leq |A \cap E(\tilde{G})|$ , as desired.  $\square$

For the  $(r - 1)$ -partite case, we will make a reduction to the setting where  $r = 2$ . As mentioned in the introduction, in this setting the extremal number is known exactly, as follows.

**Theorem 9.5.** *Let  $n, m \geq 1$  and let  $P$  be the crossing 2-pattern or the nesting 2-pattern. Then,*

$$\text{ex}_{<}(n, P^{(m)}) = \binom{n}{2} - \binom{(n - 2(m - 1))_+}{2} \leq 2(m - 1)n.$$

The crossing case ( $|\text{AB}|\text{AB}|$ ) of [Theorem 9.5](#) is due to Capocyleas and Pach [\[16\]](#) while the nesting case ( $|\text{AB}|\text{BA}|$ ) is a special case of [Theorem 1.18\(3\)](#). (The nesting case also follows from results on queue-numbers of graphs due to Pemmaraju [\[52\]](#) and to Dujmović and Wood [\[23\]](#); see [\[23, Lemma 8\]](#)). Now, our result in the  $(r - 1)$ -partite setting is as follows.

**Lemma 9.6.** *Let  $n, m, r \geq 1$ . Let  $\vec{a} = \{1, 2\}^{r-1}$  be a vector with exactly one “2” (so all other entries are “1”). Let  $P$  be an  $r$ -partite  $r$ -pattern. Then,*

$$\text{ex}_{<}^{\vec{a}}(n, P^{(m)}) \leq 4r(m - 1)n^{r-1}.$$

*Proof.* Fix an  $r$ -partite  $r$ -pattern  $P$ , and write  $|\mathfrak{B}_1|\dots|\mathfrak{B}_r|$  for its block representation. Let  $V_1, \dots, V_{r-1}$  be disjoint copies of  $\{1, \dots, n\}$ , and let  $i^*$  be the unique index such that  $a_{i^*} = 2$ . Note that an  $\vec{a}$ -equipartite  $r$ -graph  $G$  can be viewed as a set of tuples

$$e \in V_1 \times \dots \times V_{i^*-1} \times E(K_n) \times V_{i^*+1} \times \dots \times V_{r-1},$$

where  $K_n$  is the complete 2-graph on the vertex set  $V_{i^*}$ . From this perspective, a copy of  $P^{(m)}$  corresponds to a sequence of edges  $e_1, \dots, e_m \in E(G)$  which is increasing in all coordinates  $i \neq i^*$  with  $\mathfrak{B}_i = \text{AB}$ , and decreasing in all coordinates  $i \neq i^*$  with  $\mathfrak{B}_i = \text{BA}$ , and in coordinate  $i^*$ , the corresponding 2-edges have a consistent 2-partite pattern (the possibilities can be represented as ABAB, BABA, ABBA and BAAB, i.e., a crossing-clique going left, a crossing-clique going right, a nesting-clique going inwards, or a nesting-clique going outwards).

By symmetry, we may assume that  $i^* = r - 1$ . In addition, by exchanging the “A”s and the “B”s in each  $\mathfrak{B}_i$ , if necessary, we may assume that  $|\mathfrak{B}_{r-1}|\mathfrak{B}_r| = |\text{AB}|\text{AB}|$  or  $|\text{AB}|\text{BA}|$ , and finally (as in [Lemma 9.4](#)) we may assume that  $\mathfrak{B}_i = \text{AB}$  for  $i \leq r - 2$ . In other words,  $P$  has block representation  $|\text{AB}|\text{AB}|\dots|\text{AB}|\text{AB}|$  or  $|\text{AB}|\text{AB}|\dots|\text{AB}|\text{BA}|$ .

Now, we want to reduce the upper bound of  $\text{ex}_{<}^{\vec{a}}(n, P^{(m)})$  to the graph ( $r = 2$ ) case and apply [Theorem 9.5](#). Fix an  $\vec{a}$ -equipartite  $P^{(m)}$ -free  $r$ -graph  $G$  with parts  $V_1, \dots, V_{r-1}$ . Write  $Q$  for the 2-partite 2-pattern with block representation  $|\mathfrak{B}_{r-1}|\mathfrak{B}_r|$ . Observe that if  $G$  has two edges  $e, f$  such that

$e[i] - e[r-1] = f[i] - f[r-1]$  for all  $i \in \{1, \dots, r-2\}$ , and such that the last two coordinates of  $e$  and the last two coordinates of  $f$  (both are viewed as edges in  $K_n$ ) form pattern  $Q$ , then  $e, f$  form pattern  $P$ .

Now, for each  $\vec{x} = (x_1, \dots, x_{r-2}) \in \{1-n, 2-n, \dots, n-1\}^{r-2}$ , we define  $G_{\vec{x}}$  to be the subgraph of  $G$  consisting of all edges  $e \in E(G)$  such that  $e[i] - e[r-1] = x_i$  for all  $i \in \{1, \dots, r-2\}$  (i.e., the edges of the form  $(x_1 + u, x_2 + u, \dots, x_{r-2} + u, uv)$  with  $u < v$ ). Write  $G_{\vec{x}}^{(2)}$  for the 2-graph consisting of all  $uv$  such that  $uv$  appears as the last coordinate of some edge in  $G_{\vec{x}}$ . Note that there is a natural bijection between edges in  $G_{\vec{x}}$  and edges in  $G_{\vec{x}}^{(2)}$ , and by the above discussion, any  $Q$ -clique in  $G_{\vec{x}}^{(2)}$  corresponds to a  $P$ -clique in  $G_{\vec{x}}$  of the same size. Thus,  $G_{\vec{x}}^{(2)}$  is  $Q^{(m)}$ -free. Also, as  $G = \bigcup_{\vec{x}} G_{\vec{x}}$ , we know  $e(G) \leq \sum_{\vec{x}} e(G_{\vec{x}}) = \sum_{\vec{x}} e(G_{\vec{x}}^{(2)})$ .

Then, [Theorem 9.5](#) implies  $e(G_{\vec{x}}) = e(G_{\vec{x}}^{(2)}) \leq 2(m-1)n$  for all possible  $\vec{x}$ . Naïvely, there are  $(2n-1)^{r-2}$  possible choices for  $\vec{x}$ , but we can do better. Note that for each edge  $uv \in G_{\vec{x}}^{(2)}$ , we must have  $\max_i x_i + u \leq n$  and  $\min_i x_i + u \geq 1$ . So, for  $G_{\vec{x}}$  to be nonempty, we must have  $\max_i x_i - \min_i x_i < n$ . The number of choices of  $\vec{x}$  which satisfy this inequality is at most  $(2n-1)rn^{r-3} < 2rn^{r-2}$  (enumerating over all choices of  $\min_i x_i$ , and all choices of an  $i$  for which  $x_i$  is minimal). So,

$$e(G) \leq 2rn^{r-2} \cdot 2(m-1)n = 4r(m-1)n^{r-1},$$

and the desired result follows.  $\square$

*Remark 9.7.* Let  $i^*$  be the only  $i \in [r-1]$  with  $a_i = 2$  and let  $|\mathfrak{B}_1| |\mathfrak{B}_2| \cdots |\mathfrak{B}_r|$  be the block representation of  $P$ . Similarly to the proof of [Theorem 1.18\(3\)](#), one can show that if  $\mathfrak{B}_{i^*} \neq \mathfrak{B}_{i^*+1}$ , then

$$\text{ex}_{<}^{\vec{a}}(n, P^{(m)}) = n^{r-2} \binom{n}{2} - (n - (m-1))^{r-2} \binom{(n - 2(m-1))_+}{2}.$$

We wonder if a more general statement is true. Let  $n, m, r \geq 1$ , let  $P$  be an  $r$ -partite  $r$ -pattern, and let  $\vec{a}$  be a vector of positive integers with sum of entries  $\|\vec{a}\|_1 = r$ . Then, is it always the case that

$$\text{ex}_{<}^{\vec{a}}(n, P^{(m)}) = \prod_{i=1}^k \binom{n}{a_i} - \prod_{i=1}^k \binom{(n - (m-1)a_i)_+}{a_i}?$$

(Cf. [Conjecture 1.20](#)).

We are finally ready to prove [Theorem 1.18\(2\)](#) (giving a general upper bound on  $\text{ex}_{<}(n, P^{(m)})$ ) using [Theorem 8.2](#) to reduce to the  $(\geq r-1)$ -equipartite case.

*Proof of [Theorem 1.18\(2\)](#).* Let  $G$  be an  $n$ -vertex  $P^{(m)}$ -free ordered  $r$ -graph, with  $(p/n) \binom{n}{r}$  edges. Our goal is to prove that  $p = O(r^3(m-1))$ .

By [Theorem 8.2](#), we can find a  $(\geq r-1)$ -equipartite subgraph  $G' \subseteq G$ , of order  $n' = n^{\Omega_r(1)}$ , which has  $(r-1)$ -density at least  $\Omega(p/r^2)$ . Note that  $G'$  is still  $P^{(m)}$ -free.

- If  $G'$  is  $r$ -partite, then by [Lemma 9.4](#), we have  $e(G') \leq r(m-1)n^{r-1}$ . That is to say,  $G'$  has  $(r-1)$ -density at most  $r(m-1)$ .
- If  $G'$  is  $(r-1)$ -partite, then by [Lemma 9.6](#), we have  $e(G') \leq 4r(m-1)n^{r-1}$ . That is to say,  $G'$  has  $(r-1)$ -density at most  $8r(m-1)$ .

In both cases,  $G'$  has  $(r-1)$ -density  $O(r(m-1))$ ; the desired result follows.  $\square$

## REFERENCES

- [1] D. Aldous and P. Diaconis, *Hammersley's interacting particle process and longest increasing subsequences*, Probab. Theory Related Fields **103** (1995), no. 2, 199–213. [1.3](#), [1.6.2](#)
- [2] N. Alon and E. Friedgut, *On the number of permutations avoiding a given pattern*, J. Combin. Theory Ser. A **89** (2000), no. 1, 133–140. [1.4](#)
- [3] B. Aronov, V. Dujmović, P. Morin, A. Ooms, and L. F. Schultz Xavier da Silveira, *More Turán-type theorems for triangles in convex point sets*, Electron. J. Combin. **26** (2019), no. 1, Paper No. 1.8, 26. [1.5](#)
- [4] J. Baik, P. Deift, and K. Johansson, *On the distribution of the length of the longest increasing subsequence of random permutations*, J. Amer. Math. Soc. **12** (1999), no. 4, 1119–1178. [1.3](#)
- [5] J. Baik and R. Jenkins, *Limiting distribution of maximal crossing and nesting of Poissonized random matchings*, Ann. Probab. **41** (2013), no. 6, 4359–4406. [1](#), [1.3](#)
- [6] J. Baik and E. M. Rains, *The asymptotics of monotone subsequences of involutions*, Duke Math. J. **109** (2001), no. 2, 205–281. [1](#), [1.3](#)
- [7] B. Bollobás and P. Winkler, *The longest chain among random points in Euclidean space*, Proc. Amer. Math. Soc. **103** (1988), no. 2, 347–353. [1.3](#), [5](#), [5](#), [5](#)

- [8] P. Braß, G. Rote, and K. J. Swanepoel, *Triangles of extremal area or perimeter in a finite planar point set*, Discrete Comput. Geom. **26** (2001), no. 1, 51–58. [1.5](#)
- [9] P. Braß, *Turán-type extremal problems for convex geometric hypergraphs*, Towards a theory of geometric graphs, Contemp. Math., vol. 342, Amer. Math. Soc., Providence, RI, 2004, pp. 25–33. [1.5](#)
- [10] P. Braß, G. Károlyi, and P. Valtr, *A Turán-type extremal theory of convex geometric graphs*, Discrete and computational geometry, Algorithms Combin., vol. 25, Springer, Berlin, 2003, pp. 275–300. [1.5](#)
- [11] G. Brightwell, *Random  $k$ -dimensional orders: width and number of linear extensions*, Order **9** (1992), no. 4, 333–342. [1.4](#), [1.4](#)
- [12] G. Brightwell, *Models of random partial orders*, Surveys in combinatorics, 1993 (Keele), London Math. Soc. Lecture Note Ser., vol. 187, Cambridge Univ. Press, Cambridge, 1993, pp. 53–83. [1](#)
- [13] M. Bucić, B. Sudakov, and T. Tran, *Erdős-Szekeres theorem for multidimensional arrays*, J. Eur. Math. Soc. (JEMS) (2022). [1](#)
- [14] B. Bukh and J. Matoušek, *Erdős-Szekeres-type statements: Ramsey function and decidability in dimension 1*, Duke Math. J. **163** (2014), no. 12, 2243 – 2270. [1](#)
- [15] H. Burkill and L. Mirsky, *Monotonicity*, J. Math. Anal. Appl. **41** (1973), 391–410. [1.2](#), [4](#)
- [16] V. Capoleas and J. Pach, *A Turán-type theorem on chords of a convex polygon*, J. Combin. Theory Ser. B **56** (1992), no. 1, 9–15. [1.5](#), [1.5](#), [9](#)
- [17] W. Y. C. Chen, E. Y. P. Deng, R. R. X. Du, R. P. Stanley, and C. H. Yan, *Crossings and nestings of matchings and partitions*, Trans. Amer. Math. Soc. **359** (2007), no. 4, 1555–1575. [1](#)
- [18] J. Cibulka and J. Kynčl, *Better upper bounds on the Füredi-Hajnal limits of permutations*, Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms, SIAM, Philadelphia, PA, 2017, pp. 2280–2293. [1.4](#), [7](#)
- [19] G. de B. Robinson, *On the Representations of the Symmetric Group*, Amer. J. Math. **60** (1938), no. 3, 745–760. [1](#)
- [20] A. de Mier,  *$k$ -noncrossing and  $k$ -nonnesting graphs and fillings of Ferrers diagrams*, Combinatorica **27** (2007), no. 6, 699–720. [1.6.4](#)
- [21] A. Dudek, J. Grytczuk, and A. Ruciński, *Erdős-Szekeres type theorems for ordered uniform matchings*, arXiv preprint arXiv:2301.02936 (2023). [1](#), [1.2](#), [1.2](#), [1.3](#), [1.6.1](#), [3](#), [3.1](#), [4](#)
- [22] A. Dudek, J. Grytczuk, and A. Ruciński, *Ordered unavoidable sub-structures in matchings and random matchings*, arXiv preprint arXiv:2210.14042 (2023). [1](#), [1.1](#), [1.2](#), [1.3](#), [1.6.1](#), [3](#), [4](#), [4](#), [4.2](#), [4](#)
- [23] V. Dujmović and D. R. Wood, *On linear layouts of graphs*, Discrete Math. Theor. Comput. Sci. **6** (2004), no. 2, 339–357. [1.5](#), [9](#)
- [24] P. Erdős and D. J. Kleitman, *On coloring graphs to maximize the proportion of multicolored  $k$ -edges*, J. Combinatorial Theory **5** (1968), 164–169. [8](#)
- [25] P. Erdős and G. Szekeres, *A combinatorial problem in geometry*, Compositio Math. **2** (1935), 463–470. [1](#)
- [26] A. Errera, *Analysis situs: un problème d'énumération*, Mémoires de la Classe des sciences. Académie royale de Belgique. Collection 8 **11** (1931), no. 6, 3–26. [1](#)
- [27] J. Fox, *Stanley-Wilf limits are typically exponential*, arXiv preprint arXiv:1310.8378 (2013). [1.4](#)
- [28] J. Fox, J. Pach, B. Sudakov, and A. Suk, *Erdős-Szekeres-type theorems for monotone paths and convex bodies*, Proc. Lond. Math. Soc. (3) **105** (2012), no. 5, 953–982. [1](#)
- [29] Z. Füredi, T. Jiang, A. Kostochka, D. Mubayi, and J. Verstraëte, *Tight paths in convex geometric hypergraphs*, Adv. Comb. (2020), Paper No. 1, 14. [1.5](#)
- [30] 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. [1.5](#), [1.5](#), [1.5](#)
- [31] Z. Füredi, T. Jiang, A. Kostochka, D. Mubayi, and J. Verstraëte, *Partitioning ordered hypergraphs*, J. Combin. Theory Ser. A **177** (2021), Paper No. 105300, 18. [1.5](#), [8](#), [8.4](#)
- [32] Z. Füredi, D. Mubayi, J. O’Neill, and J. Verstraëte, *Extremal problems for pairs of triangles*, J. Combin. Theory Ser. B **155** (2022), 83–110. [1.5](#)
- [33] A. Girão, G. Kronenberg, and A. Scott, *A multidimensional Ramsey theorem*, arXiv preprint arXiv:2210.09227. [1](#)
- [34] W. T. Gowers and J. Long, *The length of an  $s$ -increasing sequence of  $r$ -tuples*, Combin. Probab. Comput. **30** (2021), no. 5, 686–721. [1.5](#)
- [35] P. Groeneboom, *Ulam’s problem and Hammersley’s process*, Ann. Probab. **29** (2001), no. 2, 683–690. [1.3](#)
- [36] S. Janson, T. Łuczak, and A. Ruciński, *Random graphs*, Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley-Interscience, New York, 2000. [2](#)
- [37] K. Johansson, *The longest increasing subsequence in a random permutation and a unitary random matrix model*, Math. Res. Lett. **5** (1998), no. 1-2, 63–82. [1.3](#)
- [38] J. Jonsson and V. Welker, *A spherical initial ideal for Pfaffians*, Illinois J. Math. **51** (2007), no. 4, 1397–1407. [1.6.4](#)
- [39] J. Justicz, E. R. Scheinerman, and P. Winkler, *Random intervals*, Amer. Math. Monthly **97** (1990), no. 10, 881–889. [1](#), [1.3](#), [5](#)
- [40] K. Kalmanson, *On a theorem of Erdős and Szekeres*, J. Combinatorial Theory Ser. A **15** (1973), 343–346. [1.2](#), [4](#)
- [41] 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, Research Paper 33, 12. [1](#)
- [42] M. Klazar, *The Füredi-Hajnal conjecture implies the Stanley-Wilf conjecture*, Formal power series and algebraic combinatorics (Moscow, 2000), Springer, Berlin, 2000, pp. 250–255. [1.4](#)
- [43] M. Klazar, *On identities concerning the numbers of crossings and nestings of two edges in matchings*, SIAM J. Discrete Math. **20** (2006), no. 4, 960–976. [1](#)

- [44] D. E. Knuth, *Permutations, matrices, and generalized Young tableaux*, Pacific J. Math. **34** (1970), 709–727. [1](#)
- [45] C. Krattenthaler, *Growth diagrams, and increasing and decreasing chains in fillings of Ferrers shapes*, Adv. in Appl. Math. **37** (2006), no. 3, 404–431. [1](#)
- [46] N. Linial and M. Simkin, *Monotone subsequences in high-dimensional permutations*, Combin. Probab. Comput. **27** (2018), no. 1, 69–83. [1](#)
- [47] B. F. Logan and L. A. Shepp, *A variational problem for random Young tableaux*, Advances in Math. **26** (1977), no. 2, 206–222. [1.3](#)
- [48] L. Lovász and M. D. Plummer, *Matching theory*, North-Holland Mathematics Studies, vol. 121, North-Holland Publishing Co., Amsterdam; North-Holland Publishing Co., Amsterdam, 1986, Annals of Discrete Mathematics, 29. [1](#)
- [49] C. McDiarmid, *Concentration*, Probabilistic methods for algorithmic discrete mathematics, Algorithms Combin., vol. 16, Springer, Berlin, 1998, pp. 195–248. [2](#)
- [50] L. Mirsky, *A dual of Dilworth’s decomposition theorem*, Amer. Math. Monthly **78** (1971), 876–877. [2](#)
- [51] J. Pach and R. Pinchasi, *How many unit equilateral triangles can be generated by  $N$  points in convex position?*, Amer. Math. Monthly **110** (2003), no. 5, 400–406. [1.5](#)
- [52] S. V. Pemmaraju, *Exploring the powers of stacks and queues via graph layouts*, ProQuest LLC, Ann Arbor, MI, 1992, Thesis (Ph.D.)–Virginia Polytechnic Institute and State University. [1.5](#), [9](#)
- [53] D. Romik, *The surprising mathematics of longest increasing subsequences*, Institute of Mathematical Statistics Textbooks, vol. 4, Cambridge University Press, New York, 2015. [1](#), [2](#), [5](#)
- [54] C. Schensted, *Longest increasing and decreasing subsequences*, Canadian J. Math. **13** (1961), 179–191. [1](#)
- [55] T. Seppäläinen, *Large deviations for increasing sequences on the plane*, Probab. Theory Related Fields **112** (1998), no. 2, 221–244. [1.3](#)
- [56] R. P. Stanley, *Increasing and decreasing subsequences and their variants*, International Congress of Mathematicians. Vol. I, Eur. Math. Soc., Zürich, 2007, pp. 545–579. [1](#), [1.3](#)
- [57] R. P. Stanley, *Catalan addendum (version of 25 may 2013)*, <http://www-math.mit.edu/%7Erstan/ec/catadd.pdf>., 2013. [1](#)
- [58] J. M. Steele, *Limit properties of random variables associated with a partial ordering of  $R^d$* , Ann. Probability **5** (1977), no. 3, 395–403. [5](#)
- [59] T. Szabó and G. Tardos, *A multidimensional generalization of the Erdős-Szekeres lemma on monotone subsequences*, Combin. Probab. Comput. **10** (2001), no. 6, 557–565. [1](#)
- [60] G. Tardos, *Extremal theory of ordered graphs*, Proceedings of the International Congress of Mathematicians: Rio de Janeiro 2018, World Scientific, 2018, pp. 3235–3243. [1.5](#)
- [61] A. M. Veršik and S. V. Kerov, *Asymptotic behavior of the Plancherel measure of the symmetric group and the limit form of Young tableaux*, Dokl. Akad. Nauk SSSR **233** (1977), no. 6, 1024–1027. [1.3](#)

## APPENDIX A. FURTHER LOWER BOUNDS ON RAMSEY PARAMETERS

In this appendix we prove that  $L_4(n) \geq n^{1/15}/4$ , completing the proof of [Theorem 1.8](#). It is convenient to introduce some notation.

**Definition A.1.** For a pair of  $r$ -patterns  $P, Q$ , we write  $P \circ Q$  for the set of all  $r$ -patterns that can be obtained as a “composition” of  $P$  and  $Q$ . Specifically, we put  $R \in P \circ Q$  if it is possible to have three edges  $e, f, g$  (with  $e[1] < f[1] < g[1]$ ) in an ordered  $r$ -matching such that  $e, f$  form  $P$ ,  $f, g$  form  $Q$  and  $e, g$  form  $R$ .

For example, in the case  $r = 2$ , writing  $\alpha$  for the alignment pattern AABB, we clearly have  $\alpha \circ \alpha = \{\alpha\}$ . As a few more examples, writing  $\nu$  for the nesting pattern ABBA and  $\kappa$  for the crossing pattern ABAB, it is not hard to see that  $\alpha \circ \nu = \{\alpha\}$  and  $\nu \circ \alpha = \{\alpha, \nu, \kappa\}$ .

**Definition A.2.** Say that a collection of  $r$ -patterns  $\mathcal{P}$  is *left-dominated* if  $(P \circ Q) \cap \mathcal{P} \subseteq \{P\}$  for all  $P, Q \in \mathcal{P}$ . Say that  $\mathcal{P}$  is *right-dominated* if  $(P \circ Q) \cap \mathcal{P} \subseteq \{Q\}$  for all  $P, Q \in \mathcal{P}$ .

Roughly speaking, the idea is that, if  $\mathcal{P}$  is left-dominated, then in a  $\mathcal{P}$ -clique, if we compose a pattern  $P$  with any pattern we always get  $P$ . (Similarly, if  $\mathcal{P}$  is right-dominated, then when we compose any pattern with  $P$  we always get  $P$ ). For example, as implicitly observed in the proof of the lower bound on  $L_3(n)$  in [Section 3.4](#), the set  $\{\psi(\alpha\nu), \psi(\nu\alpha)\}$  is left-dominated. The key property of a left- or right-dominated set  $\mathcal{P}$  (which also implicitly appeared in the proof of the lower bound on  $L_3(n)$ ) is that every  $\mathcal{P}$ -clique always contains a very large  $P$ -clique for some  $P \in \mathcal{P}$ , as follows.

**Lemma A.3.** *If  $\mathcal{P}$  is left-dominated or right-dominated, then for every  $\mathcal{P}$ -clique  $M$  of size  $n$  we have  $L(M) \geq n/|\mathcal{P}|$ .*

*Proof.* Suppose  $\mathcal{P}$  is left-dominated (the right-dominated case is similar), and let  $M$  be a  $\mathcal{P}$ -clique of size  $n$ . Write the edges of  $M$  as  $e_1, \dots, e_n$  with  $e_1[1] < \dots < e_n[1]$ . Now, for every index  $i$ , there is a single pattern  $P \in \mathcal{P}$  such that  $e_i$  and  $e_j$  form  $P$  for each  $j > i$ ; say that  $i$  is of “type  $P$ ”. By the



Pattern name	Representation	Pattern name	Representation
$\psi(\alpha\alpha\alpha)$	AAAA BBBB	$\psi(\nu\alpha\kappa)$	AB BBAA BA
$P_{\alpha\alpha\alpha}^*$	AAABABBB	$\psi(\nu\nu\alpha)$	AB BA AABB
$P_{\alpha\alpha\alpha}^{**}$	AAABBABB	$\psi(\nu\nu\nu)$	AB BA AB BA
$P_{\alpha\alpha\alpha}^{***}$	AABAABBB	$\psi(\nu\nu\kappa)$	AB BA AB AB
$P_{\alpha\alpha\alpha}^{****}$	AABABABB	$\psi(\nu\kappa\alpha)$	AB BA BBAA
$\psi(\alpha\alpha\nu)$	AAABBB BA	$\psi(\nu\kappa\nu)$	AB BA BA AB
$P_{\alpha\alpha\nu}^*$	AABABBBBA	$\psi(\nu\kappa\kappa)$	AB BA BA BA
$\psi(\alpha\alpha\kappa)$	AAABBB AB	$\psi(\kappa\alpha\alpha)$	AB AAABBB
$P_{\alpha\alpha\kappa}^*$	AABABBAB	$P_{\kappa\alpha\alpha}^*$	ABAABABB
$\psi(\alpha\nu\alpha)$	AABB BBAA	$\psi(\kappa\alpha\nu)$	AB AABB BA
$\psi(\alpha\nu\nu)$	AABB BA AB	$\psi(\kappa\alpha\kappa)$	AB AABB AB
$\psi(\alpha\nu\kappa)$	AABB BA BA	$\psi(\kappa\nu\alpha)$	AB AB BBAA
$\psi(\alpha\kappa\alpha)$	AABB AABB	$\psi(\kappa\nu\nu)$	AB AB BA AB
$\psi(\alpha\kappa\nu)$	AABB AB BA	$\psi(\kappa\nu\kappa)$	AB AB BA BA
$\psi(\alpha\kappa\kappa)$	AABB AB AB	$\psi(\kappa\kappa\alpha)$	AB AB AABB
$\psi(\nu\alpha\alpha)$	AB BBBAAA	$\psi(\kappa\kappa\nu)$	AB AB AB BA
$P_{\nu\alpha\alpha}^*$	ABBBABAA	$\psi(\kappa\kappa\kappa)$	AB AB AB AB
$\psi(\nu\alpha\nu)$	AB BBAA AB		

**Table 2.** All 35 different 4-patterns. The collectable patterns are named as  $\psi(W)$  for their corresponding weak pattern  $W$  (as in Section 3). The non-collectable patterns are given ad-hoc names with superscript “\*”s.

pigeonhole principle, at least  $n/|\mathcal{P}|$  of the indices have the same type (say,  $P$ ); the corresponding edges form a  $P$ -clique.  $\square$

We also need to introduce a new notion of “one-sided”  $P$ -freeness.

**Definition A.4.** Given an  $r$ -matching  $M$  and an  $r$ -pattern  $P$ , say that an edge  $e \in E(M)$  is *left- $P$ -free* if there is no  $f \in E(M)$  with  $f[1] < e[1]$  such that  $f$  and  $e$  form pattern  $P$ . Similarly, say that  $e$  is *right- $P$ -free* if there is no  $f \in E(M)$  with  $f[1] > e[1]$  such that  $e$  and  $f$  form pattern  $P$ .

Note that if we have a matching  $M$  whose every edge is left- $P$ -free, or whose every edge is right- $P$ -free, then  $M$  itself is  $P$ -free.

Now we are ready to complete the proof of Theorem 1.8.

*Proof of the lower bound on  $L_4(n)$  in Theorem 1.8.* We use the notions of weak patterns and signatures introduced in Section 3. There are 35 different 4-patterns, including eight non-collectable patterns (which are all given names in Table 2).

Define

$$\begin{aligned} \mathcal{P}_1 &= \phi^{-1}(\alpha\alpha\alpha) = \{\psi(\alpha\alpha\alpha), P_{\alpha\alpha\alpha}^*, P_{\alpha\alpha\alpha}^{**}, P_{\alpha\alpha\alpha}^{***}, P_{\alpha\alpha\alpha}^{****}\}, & \mathcal{P}_2 &= \{\psi(\alpha\alpha\kappa), P_{\alpha\alpha\kappa}^*, \psi(\alpha\kappa\alpha), \psi(\kappa\alpha\alpha), P_{\kappa\alpha\alpha}^*\}, \\ \mathcal{P}_3 &= \{\psi(\alpha\alpha\nu), \psi(\alpha\nu\alpha), \psi(\nu\alpha\alpha)\}, & \mathcal{P}_4 &= \{P_{\alpha\alpha\nu}^*, P_{\nu\alpha\alpha}^*\}, & \mathcal{P}_5 &= \{\psi(\alpha\kappa\kappa), \psi(\kappa\alpha\kappa), \psi(\kappa\kappa\alpha)\}, \\ \mathcal{P}_6 &= \{\psi(\alpha\nu\nu), \psi(\nu\alpha\nu), \psi(\nu\nu\alpha)\}, & \mathcal{P}_7 &= \{\psi(\alpha\kappa\nu), \psi(\alpha\nu\kappa), \psi(\nu\alpha\kappa), \psi(\nu\kappa\alpha), \psi(\kappa\alpha\nu), \psi(\kappa\nu\alpha)\}. \end{aligned}$$

For each  $i \in \{1, \dots, 7\}$  we write  $\preceq_i$  to denote the relation  $\preceq_{\mathcal{P}_i}$  (see Definition 3.1 for the definition of  $\preceq_{\mathcal{P}}$ ). Some of these relations give rise to posets, as follows.

- (1) Consider any matching  $M$ . By the proof of Lemma 3.2(A):
  - $\preceq_1$  is always a partial order (with corresponding signature  $(3, \emptyset)$ );
  - $\preceq_2$  is a partial order if  $M$  is  $\mathcal{P}_1$ -free (with corresponding signature  $(2, \emptyset)$ );
  - $\preceq_5$  is a partial order if  $M$  is  $\mathcal{P}_1 \cup \mathcal{P}_2$ -free (with corresponding signature  $(1, \emptyset)$ ).
- (2) In any matching,  $\preceq_3$  is a partial order (this can be observed by direct case-checking, analogous to  $\mathcal{P}_3$  in the proof for  $r = 3$  case of Theorem 1.8; the patterns in  $\mathcal{P}_3$  can be interpreted as “generalised nestings” where one of the two edges of the pattern is fully contained between two consecutive vertices of the other edge).

Also, according to Lemma 3.8,

- (3) for  $i \in \{1, 2, 5\}$  (which each correspond to specific signatures), every  $\mathcal{P}_i$ -clique  $M$  of size  $n$  has  $L(M) \geq n/3$ .

Now, consider any matching  $M = M_1$  of size  $n$ . We show how to find a  $P$ -clique of size at least  $n^{1/15}/4$ , proceeding in stages.

*Stage 1.* By Mirsky's theorem applied to  $\preceq_1$  on  $M_1$  (recalling (1)), there is either a  $\mathcal{P}_1$ -clique of size at least  $n^{1/15}$ , or a  $\mathcal{P}_1$ -free submatching  $M_2$  of size at least  $n^{14/15}$ . In the former case, we are done by (3); in the latter case move on to Stage 2.

*Stage 2.* By Mirsky's theorem applied to  $\preceq_2$  on  $M_2$  (recalling (1)), there is either a  $\mathcal{P}_2$ -clique of size at least  $n^{1/15}$ , or a  $(\mathcal{P}_1 \cup \mathcal{P}_2)$ -free submatching  $M_3$  of size at least  $n^{13/15}$ . In the former case, we are done by (3); in the latter case move on to Stage 3.

*Stage 3.* By Mirsky's theorem applied to  $\preceq_3$  on  $M_3$  (recalling (2)), there is either a  $\mathcal{P}_3$ -clique of size at least  $n^{1/15}$ , or a  $(\mathcal{P}_1 \cup \mathcal{P}_2 \cup \mathcal{P}_3)$ -free submatching  $M_4$  of size at least  $n^{12/15}$ . In the former case, we observe that  $\mathcal{P}_3$  is left-dominated, so we are done by [Lemma A.3](#); in the latter case move on to stage 4.

*Stage 4.* In this stage we wish to eliminate the non-collectable patterns  $P_{\alpha\alpha\nu}^*$  and  $P_{\nu\alpha\alpha}^*$  in  $\mathcal{P}_4$  without seriously decreasing the size of our matching. List all the edges of  $M_4$  as  $e_1, \dots, e_m$  with  $m = |M_4| \geq n^{12/15}$  and  $e_1[1] < \dots < e_m[1]$ . We claim that for any  $1 \leq i < j < k \leq m$ , at most one of the pairs  $(e_i, e_j)$  ( $e_j, e_k$ ) can form  $P_{\alpha\alpha\nu}^*$ . Indeed, if both pairs were to form  $P_{\alpha\alpha\nu}^*$ , then  $(e_i, e_k)$  would also form  $P_{\alpha\alpha\nu}^*$  by [Lemma 3.7](#) (recalling that  $M_4$  is  $\psi(\alpha\alpha\nu)$ -free). But then,  $e_i, e_j, e_k$  would form a  $P_{\alpha\alpha\nu}^*$ -clique, which is impossible as  $P_{\alpha\alpha\nu}^*$  is not collectable.

By the above claim, every edge of  $M_4$  is left- $P_{\alpha\alpha\nu}^*$ -free or right- $P_{\alpha\alpha\nu}^*$ -free. So, by the pigeonhole principle we can find a submatching  $M'_5$  of at least  $n^{12/15}/2$  edges which are either all left- $P_{\alpha\alpha\nu}^*$ -free, or all right- $P_{\alpha\alpha\nu}^*$ -free, meaning that  $M'_5$  itself is  $P_{\alpha\alpha\nu}^*$ -free.

Repeating all this analysis with the pattern  $P_{\nu\alpha\alpha}^*$ , we can find a matching  $M_5$  of size at least  $n^{12/15}/4$  that is  $(\mathcal{P}_1 \cup \dots \cup \mathcal{P}_4)$ -free.

We have now eliminated all non-collectable patterns, so we will refer to each pattern by its corresponding weak pattern (e.g., we simply write  $\alpha\nu\nu$  instead of  $\psi(\alpha\nu\nu)$ ). Move on to stage 5.

*Stage 5.* By Mirsky's theorem applied to  $\preceq_5$  on  $M_5$  (recalling (1)), there is either a  $\mathcal{P}_5$ -clique of size at least  $n^{1/15}$ , or a  $(\mathcal{P}_1 \cup \dots \cup \mathcal{P}_5)$ -free submatching  $M_6$  of size at least  $n^{11/15}/4$ . In the former case, we are done by (3); in the latter case move on to Stage 6.

*Stage 6.* The analysis now starts to get a bit more intricate, and more tedious casework becomes necessary. Recall that  $\mathcal{P}_6 = \{\alpha\nu\nu, \nu\alpha\nu, \nu\nu\alpha\}$ . It is not necessarily the case that  $\preceq_6$  is a poset; in [Table 3](#) we tabulate the possible compositions of patterns in  $\mathcal{P}_6$  (among patterns that can be present in  $M_6$ ). These can all be determined by direct case-checking. Crucially, we have

$$(\alpha\nu\nu \circ \nu\nu\alpha) \setminus (\mathcal{P}_1 \cup \dots \cup \mathcal{P}_5) = \emptyset.$$

That is to say, every edge in  $M_6$  is left- $\alpha\nu\nu$ -free or right- $\nu\nu\alpha$ -free. So, by the pigeonhole principle we can find a submatching  $M'_6$  of size at least  $n^{11/15}/8$  which is  $\alpha\nu\nu$ -free or  $\nu\nu\alpha$ -free.

Suppose that  $M'_6$  is  $\alpha\nu\nu$ -free (the  $\nu\nu\alpha$ -free case is handled basically symmetrically), and let  $\mathcal{P}'_6 = \{\nu\alpha\nu, \nu\nu\alpha\}$ . [Table 3](#) shows that  $\preceq_{\mathcal{P}'_6}$  is a poset on  $M'_6$  (consider the sub-table induced by  $\nu\alpha\nu$  and  $\nu\nu\alpha$ ). So, by Mirsky's theorem we can find a  $\mathcal{P}'_6$ -clique of size  $n^{1/15}$  or a  $(\mathcal{P}_1 \cup \dots \cup \mathcal{P}_6)$ -free submatching  $M_7$  of size at least  $n^{10/15}/8$ . In the former case, [Table 3](#) shows that  $\mathcal{P}'_6$  is right-dominated and we are done by [Lemma A.3](#). In the latter case, move on to stage 7.

*Stage 7.* Recall that  $\mathcal{P}_7 = \{\alpha\kappa\nu, \alpha\nu\kappa, \nu\alpha\kappa, \nu\kappa\alpha, \kappa\alpha\nu, \kappa\nu\alpha\}$ . We tabulate in [Table 4](#) the possible compositions of patterns in  $\mathcal{P}_7$  (among patterns that can be present in  $M_7$ ). Note in particular that  $\preceq_P$  is a partial order for each  $P \in \mathcal{P}_7$ .

Note that there are ten different pairs of patterns  $P, Q \in \mathcal{P}_7$  such that

$$(P \circ Q) \setminus (\mathcal{P}_1 \cup \dots \cup \mathcal{P}_6) = \emptyset.$$

(Call such pairs "special pairs"). For each special pair  $(P, Q)$ , we know that every edge in  $M_7$  is either left- $P$ -free or right- $Q$  free. By the pigeonhole principle, we can find a submatching  $M'_7$  of size at least  $(n^{10/15}/8)/2^{10}$  such that for each special pair  $(P, Q)$ , our submatching  $M'_7$  is either  $P$ -free or  $Q$ -free. Let  $\mathcal{P}'_7$  be the collection of all patterns present in  $M'_7$ . By carefully considering all cases, we can check that

$P \backslash Q$	$\alpha\nu\nu$	$\nu\alpha\nu$	$\nu\nu\alpha$
$\alpha\nu\nu$	$\alpha\nu\nu$	$\alpha\nu\nu$	$\emptyset$
$\nu\alpha\nu$	$\nu\alpha\nu$	$\nu\alpha\nu$	$\nu\nu\alpha$
$\nu\nu\alpha$	?	$\nu\alpha\nu$	$\nu\nu\alpha$

**Table 3.** Possible compositions of the patterns in  $\mathcal{P}_6$ . The cell indexed by  $(P, Q)$  indicates the patterns in  $P \circ Q$ , apart from the patterns in  $\mathcal{P}_1 \cup \dots \cup \mathcal{P}_5$ . For each  $P, Q \in \mathcal{P}_6$ , there is only a single pattern in  $(P \circ Q) \setminus (\mathcal{P}_1 \cup \dots \cup \mathcal{P}_5)$ , unless  $(P, Q) = (\alpha\nu\nu, \nu\nu\alpha)$  or  $(P, Q) = (\nu\nu\alpha, \alpha\nu\nu)$ . (In the former case there are no valid possibilities, and in the latter case there are several (including some not in  $\mathcal{P}_6$ ), which we do not need to carefully enumerate).

$P \backslash Q$	$\alpha\nu\kappa$	$\alpha\kappa\nu$	$\nu\alpha\kappa$	$\nu\kappa\alpha$	$\kappa\alpha\nu$	$\kappa\nu\alpha$
$\alpha\nu\kappa$	$\alpha\nu\kappa$	$\alpha\nu\kappa, \alpha\kappa\nu$	$\alpha\nu\kappa$	$\emptyset$	$\alpha\nu\kappa, \alpha\kappa\nu$	$\emptyset$
$\alpha\kappa\nu$	$\alpha\nu\kappa, \alpha\kappa\nu$	$\alpha\kappa\nu$	$\alpha\nu\kappa$	$\alpha\nu\kappa, \alpha\kappa\nu$	$\emptyset$	$\alpha\nu\kappa, \alpha\kappa\nu$
$\nu\alpha\kappa$	$\nu\alpha\kappa$	?	$\nu\alpha\kappa$	$\emptyset$	?	$\emptyset$
$\nu\kappa\alpha$	$\nu\kappa\alpha, \kappa\nu\alpha$	$\nu\kappa\alpha, \kappa\nu\alpha$	$\emptyset$	$\nu\kappa\alpha$	$\kappa\nu\alpha$	$\nu\kappa\alpha, \kappa\nu\alpha$
$\kappa\alpha\nu$	$\emptyset$	$\emptyset$	?	?	$\kappa\alpha\nu$	$\kappa\alpha\nu$
$\kappa\nu\alpha$	$\emptyset$	$\emptyset$	$\nu\kappa\alpha, \kappa\nu\alpha$	$\nu\kappa\alpha, \kappa\nu\alpha$	$\kappa\nu\alpha$	$\kappa\nu\alpha$

**Table 4.** Possible compositions of the patterns in  $\mathcal{P}_7$ . The cell indexed by  $(P, Q)$  indicates the patterns in  $P \circ Q$ , apart from the patterns in  $\mathcal{P}_1 \cup \dots \cup \mathcal{P}_6$ . There are ten pairs  $(P, Q)$  for which there are no valid compositions, and four cells in which there are many valid compositions (which we do not need to carefully enumerate).

either  $|\mathcal{P}'_7| \leq 2$ , or  $\mathcal{P}'_7 = \{\alpha\nu\kappa, \alpha\kappa\nu, \nu\alpha\kappa\}$ , or  $\mathcal{P}'_7 = \{\nu\kappa\alpha, \kappa\nu\alpha, \kappa\alpha\nu\}$ . The second and third cases are basically symmetric to each other, so we just describe how to handle the first two cases.

- If  $|\mathcal{P}'_7| \leq 2$ , then we can apply Mirsky's theorem once or twice to obtain a  $P$ -clique of size at least  $n^{1/15}/4$  (for some  $P \in \mathcal{P}'_7$ ) or a  $(\mathcal{P}_1 \cup \dots \cup \mathcal{P}_7)$ -free submatching  $M_8$  of size at least  $n^{8/15}/2^9$ . In the former case we are done; in the latter case move on to stage 8.
- If  $\mathcal{P}'_7 = \{\alpha\nu\kappa, \alpha\kappa\nu, \nu\alpha\kappa\}$ , then let  $\mathcal{P}''_7 = \{\alpha\nu\kappa, \nu\alpha\kappa\}$ . Observe from Table 4 that  $\mathcal{P}''_7$  is left-dominating and that  $\preceq_{\mathcal{P}''_7}$  is a partial order. Applying Mirsky's theorem twice,  $M'_7$  has a  $\alpha\kappa\nu$ -clique of size at least  $n^{1/15}/4$  or a  $\mathcal{P}''_7$ -clique of size at least  $n^{1/15}/2$  or a  $(\mathcal{P}_1 \cup \dots \cup \mathcal{P}_7)$ -free submatching  $M_8$  of size at least  $n^{8/15}/2^{10}$ . In the first case we are immediately done, in the second case we are done by Lemma A.3, and in the third case we move on to stage 8.

*Stage 8.* Now, we have a matching  $M_8$  of size at least  $n^{8/15}/2^{10} \geq n^{8/15}/4^8$  which contains only the eight 4-partite 4-patterns. By the proof of Lemma 3.2(A), for each 4-partite 4-pattern  $P$ , the relation  $\preceq_P$  is a partial order (each pattern corresponds to a signature of the form  $(0, S)$  for  $S \subseteq \{1, 2, 3\}$ ). So, by Mirsky's theorem we have  $L(M_8) \geq n^{1/15}/4$ , and we are done.  $\square$

INSTITUTE OF SCIENCE AND TECHNOLOGY AUSTRIA (ISTA).  
Email address: michael.anastos@ist.ac.at

ETH ZÜRICH  
Email address: zhihan.jin@math.ethz.ch

INSTITUTE OF SCIENCE AND TECHNOLOGY AUSTRIA (ISTA).  
Email address: matthew.kwan@ist.ac.at

ETH ZÜRICH  
Email address: benjamin.sudakov@math.ethz.ch