

# On the Turán number of ordered forests

Dániel Korándi <sup>\*</sup>      Gábor Tardos <sup>†</sup>      István Tomon <sup>\*</sup>      Craig Weidert <sup>‡</sup>

## Abstract

An ordered graph  $H$  is a simple graph with a linear order on its vertex set. The corresponding Turán problem, first studied by Pach and Tardos, asks for the maximum number  $\text{ex}_{<}(n, H)$  of edges in an ordered graph on  $n$  vertices that does not contain  $H$  as an ordered subgraph. It is known that  $\text{ex}_{<}(n, H) > n^{1+\varepsilon}$  for some positive  $\varepsilon = \varepsilon(H)$  unless  $H$  is a forest that has a proper 2-coloring with one color class totally preceding the other one. Making progress towards a conjecture of Pach and Tardos, we prove that  $\text{ex}_{<}(n, H) = n^{1+o(1)}$  holds for all such forests that are “degenerate” in a certain sense. This class includes every forest for which an  $n^{1+o(1)}$  upper bound was previously known, as well as new examples. Our proof is based on a density-increment argument.

## 1 Introduction

An *ordered graph*  $G$  is defined as a triple  $(V, E, <)$  where  $(V, E)$  is a simple graph and  $<$  is a linear order on the vertex set  $V$ . We say that an ordered graph  $H = (V', E', <')$  is an *ordered subgraph* of  $G$  if there is an order preserving embedding of  $V'$  into  $V$  that maps edges to edges. If  $G$  does not contain  $H$  as an ordered subgraph then we say that  $G$  is  *$H$ -free*.

The following Turán-type question arises naturally: For a fixed ordered graph  $H$ , what is the maximum number of edges that an  $H$ -free ordered graph on  $n$  vertices can have? This maximum, called the *extremal* or *Turán number* of  $H$ , is denoted by  $\text{ex}_{<}(n, H)$ . The systematic study of extremal numbers was initiated by Pach and Tardos in [7].

For usual (unordered) graphs, the Erdős-Stone-Simonovits theorem says that the extremal number of a graph is controlled by its chromatic number. As it turns out, ordered graphs exhibit a similar phenomenon. The *interval chromatic number*  $\chi_{<}(H)$  of an ordered graph  $H = (V, E, <)$  is the smallest integer  $r$ , such that  $V$  can be split into  $r$  intervals (i.e., sets of consecutive vertices in the ordering) such that no edge of  $H$  has both endpoints in the same interval. It is not hard to show (see [7]) that  $\text{ex}_{<}(n, H) = \left(1 - \frac{1}{\chi_{<}(H)-1}\right) \binom{n}{2} + o(n^2)$ . This asymptotically determines the extremal number of  $H$  when  $\chi_{<}(H) \geq 3$ . However, much like for usual graphs, the problem becomes more difficult when  $\chi_{<}(H) = 2$ .

---

<sup>\*</sup>Institute of Mathematics, EPFL, Lausanne, Switzerland. Research supported in part by SNSF grants 200020-162884 and 200021-175977. Emails: {daniel.korandi,istvan.tomon}@epfl.ch.

<sup>†</sup>Rényi Institute, Budapest, Hungary. Research supported in part by the “Lendület” project of the Hungarian Academy of Sciences and by the National Research, Development and Innovation Office, NKFIH, projects K-116769 and SNN-117879. Part of this research was done while visiting EPFL. Email: tardos@renyi.hu.

<sup>‡</sup>Email: craig.weidert@gmail.com

Our work focuses on the general problem of classifying ordered graphs that have close to linear extremal numbers, i.e., satisfy  $\text{ex}_{<}(n, H) = n^{1+o(1)}$ . Note that if  $H$  contains some cycle of length  $k$  then  $\text{ex}_{<}(n, H) \geq \Omega(n^{1+1/k})$ . Indeed, it is well-known (see, e.g., [1]) that there are  $n$ -vertex (unordered) graphs with  $\Omega(n^{1+1/k})$  edges that do not contain any  $k$ -cycle.<sup>1</sup> So if  $\text{ex}_{<}(n, H) = n^{1+o(1)}$  holds then  $H$  is acyclic with interval chromatic number 2. Pach and Tardos conjectured that the converse holds, as well. More precisely, they made the following stronger conjecture:

**Conjecture 1.1** (Pach–Tardos [7]). *Let  $H$  be an acyclic ordered graph such that  $\chi_{<}(H) = 2$ . Then*

$$\text{ex}_{<}(n, H) = n(\log n)^{O(1)}.$$

In this paper we make some progress towards proving this conjecture by showing that  $\text{ex}_{<}(n, H) = n^{1+o(1)}$  holds for a large class of ordered forests  $H$ .

It will be more convenient to state our result in the language of pattern-avoiding matrices. Let us now describe the analogous problem in this context. A 0-1 matrix  $A$  is a matrix with all entries from  $\{0, 1\}$ . Its *weight*  $w(A)$  is the number of 1-entries in it. We say that  $A$  *contains* another 0-1 matrix  $B$  if either  $B$  is a submatrix of  $A$  or it can be obtained from a submatrix of  $A$  by changing some 1-entries to 0-entries. In other words,  $A$  contains  $B$  if we can get  $B$  by deleting some rows, columns and 1-entries from  $A$ . We denote this relation by  $B \prec A$ . A pattern in our context is just a fixed 0-1 matrix  $P$ , and the corresponding Turán-type problem is then to maximize the weight of an  $n$ -by- $n$  0-1 matrix that does not contain  $P$ . Let  $\text{ex}(n, P)$  denote this maximum weight. Note that this extremal function is not defined for patterns with zero weight.

We can think of a pattern  $P$  as the bipartite adjacency matrix of some ordered graph  $H_P$  of interval chromatic number 2, where the order of the vertices is inherited from the order of the corresponding rows and columns of  $P$ , and row vertices precede column vertices. Then  $\text{ex}(n, P)$  translates to the maximum number of edges in an  $H_P$ -free ordered graph  $G$  on  $2n$  vertices, such that all edges of  $G_A$  connect the first  $n$  vertices to the last  $n$  vertices.<sup>2</sup> If not for this last condition, this would be the exact same quantity as  $\text{ex}(2n, H_P)$ . And indeed, as it was observed in [7], the two functions are closely related:

$$\text{ex}(\lfloor n/2 \rfloor, P) - O(n) \leq \text{ex}_{<}(n, H_P) = O(\text{ex}(n, P) \log n). \quad (1)$$

Let us call a pattern  $P$  *acyclic* if  $H_P$  does not contain any cycle. It is easy to see that  $P$  is acyclic if and only if it contains no submatrix  $P'$  such that every row and column of  $P'$  contains at least two 1-entries. Equation (1) shows that Conjecture 1.1 can be stated in the following, equivalent form:

**Conjecture 1.2** (Pach–Tardos [7]). *Let  $P$  be an acyclic pattern. Then  $\text{ex}(n, P) = n(\log n)^{O(1)}$ .*

A stronger variant of this conjecture had earlier been proposed by Füredi and Hajnal [3], who thought  $\text{ex}(n, P) = O(n \log n)$  holds for every acyclic pattern  $P$ . However, this conjecture was refuted by Pettie [8], who constructed an acyclic pattern  $P_0$  such that  $\text{ex}(n, P_0) = \Omega(n \log n \log \log n)$ .

<sup>1</sup>For odd  $k$  we even have  $\text{ex}_{<}(n, H) = \Omega(n^2)$ .

<sup>2</sup>For equality to hold, we actually need the mild extra assumption that there is a 1-entry in the last row of  $P$  and also in the first column of  $P$ . Otherwise  $H_P = H_Q$  for some pattern  $Q \neq P$ , so avoiding  $H_P$  in the ordered graph means avoiding both  $P$  and  $Q$  in the 0-1 matrix.

There are several patterns that are known to satisfy Conjecture 1.2. For example, it is shown in [7] that if  $P$  is obtained from a pattern  $P'$  by appending a new last column with a single 1-entry, then

$$\text{ex}(n, P) = O(\text{ex}(n, P') \log n) \tag{2}$$

holds, except for the trivial cases when  $w(P) = 0$  (and thus  $\text{ex}(n, P)$  is not defined) or  $P = (1)$  is the 1-by-1 matrix of weight 1 (when  $\text{ex}(n, P) = 0$ ). Using this and similar operations, the conjecture has been verified for a large family of matrices [3, 7, 9]. These include all patterns of weight up to 6, with essentially two exceptions (omitting 0-entries for clarity):

$$Q_1 = \left( \begin{array}{c} 1 \text{---} 1 \\ 1 \text{---} \text{---} 1 \\ \quad 1 \text{---} 1 \end{array} \right) \quad \text{and} \quad Q_2 = \left( \begin{array}{c} 1 \text{---} 1 \\ 1 \text{---} 1 \\ 1 \text{---} 1 \end{array} \right)$$

In certain special cases, e.g., when  $P$  is a permutation matrix [6] or a double permutation matrix (obtained from a permutation matrix by doubling each column) [4], even the stronger bound  $\text{ex}(n, P) = O(n)$  is known to hold.

Let us now define a broad class of patterns that includes, up to transposing, all matrices that are known to satisfy Conjecture 1.2. We say that an  $l$ -by- $k$  0-1 matrix  $P$  is *vertically separable* if it can be cut along a horizontal line without separating the 1-entries in more than one column. In other words, if there exists  $1 \leq a < l$  such that for all but at most one column  $1 \leq y \leq k$ , we have  $P(x, y) = 0$  either for every  $x \leq a$  or for every  $x > a$ . In this case, we say that  $P$  is *separated* into the upper part induced by the first  $a$  rows and the lower part induced by the last  $l - a$  rows.

We call a matrix  $P$  *vertically degenerate* if it can be partitioned into its rows using vertical separations. Equivalently,  $P$  is vertically degenerate if every submatrix  $P'$  of  $P$  either has a single row or is a vertically separable matrix. Note that vertically degenerate matrices are always acyclic.

The reader might find it helpful to visualize a pattern  $P$  as a graph whose vertices are the 1-entries of  $P$ , and two 1-entries are connected if they are in the same row or column with only 0-entries between them (see  $Q_1$  and  $Q_2$  above). A pattern is acyclic if and only if this graph is acyclic. Vertically separable means that the matrix can be split into two parts along a horizontal line that cuts through at most one (vertical) edge. We consider this edge to be destroyed by the cut and this may make the way for subsequent cuts. A pattern is vertically degenerate if it can be split into its rows by applying a series of such cuts.

Our main result is the following theorem that says that every vertically degenerate pattern has close to linear extremal number. Some preliminary results from this work establishing  $\text{ex}(n, Q_1) \leq n^{1+o(1)}$  have previously appeared in the Master's thesis of the fourth author [10].

**Theorem 1.3.** *Let  $P$  be a vertically degenerate matrix. Then  $\text{ex}(n, P) \leq n^{1+o(1)}$ .*

This theorem can be thought of as a common generalization of all previously known results about acyclic patterns, albeit with a somewhat weaker upper bound. It also applies to many new matrices, including all  $3 \times k$  acyclic patterns and among them  $Q_1$  and  $Q_2$ . For  $Q_2$ , no bound better than

$O(n^{5/3})$  was previously known. (Note that the bound  $\text{ex}(n, P) = O(n^{5/3})$  follows from the Kővári-Sós-Turán theorem [5] for all 3-by- $k$  patterns  $P$ , even for non-acyclic ones.) As discussed above, this also implies that every pattern  $P$  of weight up to 6 satisfies  $\text{ex}(n, P) \leq n^{1+o(1)}$ .

We will make the  $o(1)$  term in Theorem 1.3 explicit by showing  $\text{ex}(n, P) = n2^{O(\log^c n)}$  for a constant  $c \leq 1 - 1/l$  (see Theorem 3.1). The proof is essentially a density-increment argument that starts with an  $n \times n$  matrix  $A$  with large weight and shows that either  $A$  contains  $P$  or it contains a significantly denser submatrix. Section 2 contains the heart of the inductive proof. Here we restrict our attention to certain well-structured matrices. Our general result is then reduced to this special case in Section 3 by finding the necessary well-structured submatrix inside any 0-1 matrix with large enough weight. We finish the paper with some remarks in Section 4.

**Notation.** Throughout this paper,  $\log$  stands for the binary logarithm. As usual,  $[n]$  denotes the set  $\{1, \dots, n\}$  and  $[m, n] = \{m, m+1, \dots, n\}$ . We write  $\{0, 1\}^{m \times n}$  for the set of 0-1 matrices with  $m$  rows and  $n$  columns. For  $A \in \{0, 1\}^{m \times n}$  and  $I \subseteq [m]$ ,  $J \subseteq [n]$  we write  $A(I \times J)$  to denote the submatrix of  $A$  induced by the rows in  $I$  and columns in  $J$ .

## 2 The special case of $(k, u)$ -complete 0-1 matrices

Let  $k, m, n, u \in \mathbb{N}$  and let  $A$  be an  $m$ -by- $kn$  0-1 matrix. We consider  $A$  to be the union of  $k$  vertical blocks  $A([m] \times [(j-1)n+1, jn])$  for  $j \in [k]$ . We say that  $A$  is  $(k, u)$ -complete, if among the  $n$  entries in the intersection of any row and any vertical block, one always finds at least  $u$  1-entries.

Let  $A$  be an  $m$ -by- $kn$   $(k, u)$ -complete matrix and let  $Q \in \{0, 1\}^{l \times k}$  be an  $l$ -by- $k$  pattern. If  $Q \prec A$  in such a way that each column of  $Q$  ‘‘comes from’’ a different vertical block of  $A$ , we say  $Q$  has a *block-respecting embedding* in  $A$ . More precisely, a block-respecting embedding of  $Q$  is a pair of functions  $(f, g)$  such that  $f : [l] \rightarrow [m]$  is strictly increasing,  $g : [k] \rightarrow [kn]$  satisfies  $(j-1)n < g(j) \leq jn$  for all  $j \in [k]$  and (in order to make this an embedding) all the 1-entries of  $Q$  map to a 1-entry in  $A$ , that is  $A(f(i), g(j)) = 1$  whenever  $Q(i, j) = 1$ .

The key property of this notion is that block-respecting embeddings of the upper and lower parts of a vertically separated pattern can be easily combined into such an embedding of the whole pattern, as shown by the following lemma.

**Lemma 2.1.** *Let  $P \in \{0, 1\}^{l \times k}$  be a pattern and let  $a \in [l-1]$ . Suppose  $P' = P([a] \times [k])$  has a block-respecting embedding  $(f', g')$  into a  $(k, u)$ -complete matrix  $A \in \{0, 1\}^{m \times kn}$ , and  $P'' = P([a+1, l] \times [k])$  has a block-respecting embedding  $(f'', g'')$  in  $A$ . If  $f'(a) < f''(1)$ , and  $g'(b) = g''(b)$  whenever both  $P'$  and  $P''$  have a 1-entry in some column  $b \in [k]$ , then  $P$  also has a block-respecting embedding in  $A$ .*

*Proof.* The two embeddings can be combined into a single block-respecting embedding  $(f, g)$  of  $P$  in  $A$  as follows:

$$f(i) = \begin{cases} f'(i) & \text{if } i \leq a \\ f''(i-a) & \text{otherwise} \end{cases} \quad g(j) = \begin{cases} g'(j) & \text{if } P(i, j) = 1 \text{ for some } i \in [a] \\ g''(j) & \text{otherwise.} \end{cases}$$

□

We will find block-respecting embeddings of vertically degenerate patterns  $P$  in  $(k, u)$ -complete matrices inductively, starting with single rows and gradually combining them using the above lemma. For this, we will need that  $A$  does not have a submatrix that is too dense. Let us start with classifying vertically degenerate patterns.

We call a 0-1 matrix  $P$  a *class-0* matrix if it has a single row. For  $s > 0$  we call a 0-1 matrix *class- $s$*  if it has a single row or it can be partitioned into two class- $(s - 1)$  patterns by a vertical separation. Clearly, all class- $s$  patterns are vertically degenerate and all  $l$ -by- $k$  vertically degenerate patterns are class- $(l - 1)$ .

Let  $h : \mathbb{N} \rightarrow \mathbb{R}$  be a function. We call a 0-1 matrix  $A$   *$h$ -sparse* if for every positive integer  $n$ , all  $n$ -by- $n$  submatrices of  $A$  have weight at most  $n \cdot h(n)$ . We will use this definition for the function  $h(n) = h_{b,c,d}(n) = d2^{b \log^c n}$  for some positive constants  $b, d$  and  $0 < c < 1$ .

The following inequality is easy to verify for every  $0 < m < n$  using elementary calculus:

$$\frac{h_{b,c,d}(n)}{h_{b,c,d}(m)} > \left(\frac{n}{m}\right)^{bc \log^{c-1} n} \quad (3)$$

**Lemma 2.2.** *Let  $h = h_{b,c,d}$  for some fixed positive constants  $b, d$  and  $0 < c < 1$ . Let  $k, m, n, s$  and  $u$  be positive integers satisfying  $m \leq n$  and  $u \leq h(n)$  such that  $x = bc \log^{c-1} n \leq 1/10$ , and suppose  $A \in \{0, 1\}^{m \times kn}$  is a  $(k, u)$ -complete  $h$ -sparse matrix. If*

$$m \geq 40n^{1-x^s} \left(\frac{h(n)}{u}\right)^2$$

*then every class- $s$  pattern with  $k$  columns has a block-respecting embedding in  $A$ .*

*Remark.* The bound on  $m$  in the lemma can be reformulated as

$$\frac{n}{m} \left(\frac{h(n)}{u}\right)^2 \leq \frac{n^{x^s}}{40}.$$

Both the ratios  $n/m$  and  $h(n)/u$  are assumed to be at least 1. This inequality bounds them from above. If the right-hand side dips below 1, then the condition is not satisfiable. Therefore, in the proof below, we assume  $n^{x^s} \geq 40$ .

*Proof.* We proceed by induction on  $s$ . A pattern  $P$  consisting of a single row clearly has a block-respecting embedding in every row of  $A$ . We may therefore assume that  $s > 0$ , and fix a class- $s$  pattern  $P \in \{0, 1\}^{l \times k}$  that is vertically separable at  $a \in [l - 1]$  into two class- $(s - 1)$  patterns: the upper part  $P' = P([a] \times [k])$  and the lower part  $P'' = P([a + 1, l] \times [k])$ . Let us choose a  $b \in [k]$  so that no column  $b'$  with  $b \neq b' \in [k]$  has a 1-entry in both parts.

Set  $m^* = \left\lceil 3^3 \cdot 40n^{1-x^{s-1}}(h(n)/u)^2 \right\rceil$  and  $\beta = \lfloor m/m^* \rfloor$ . Here  $m \geq m^*$  and hence  $\beta \geq 1$ , because  $n^{x^{s-1}-x^s} = n^{(1/x-1)x^s} \geq n^{9x^s} \geq 40^9 \geq 3^3$  (using  $x < 1/10$  and  $n^{x^s} \geq 40$ ). Recall that  $A$  is partitioned into  $k$  vertical blocks. We also partition most of  $A$  into  $\beta$  horizontal blocks  $A_i = A([(i - 1)m^* + 1, im^*] \times [kn])$  for  $i \in [\beta]$ . Note that each horizontal block is a  $(k, u)$ -complete  $h$ -sparse matrix with enough rows for the induction hypothesis to ensure by that both  $P'$  and  $P''$  have block-respecting embeddings in it.

Let  $S'_i = \{g(b) : (f, g) \text{ is a block-respecting embedding of } P' \text{ in } A_i\}$  be the set of column indices from the  $b$ 'th vertical block that appear in a block-respecting embedding of  $P'$  in  $A_i$ . We define the matrix  $A'_i$  contained in  $A_i$  as follows. As a first step, we turn every 1-entry of the columns in  $S'_i$  into 0. In the second step, we delete all rows that have fewer than  $u/3$  1-entries in the  $b$ 'th vertical block. The first step ensures that  $P'$  has no block-respecting embedding in  $A'_i$ , while the second step makes  $A'_i$  a  $(k, \lceil u/3 \rceil)$ -complete matrix. As  $A'_i$  is contained in  $A$ , it is also  $h$ -sparse. The inductive hypothesis then implies that  $A'_i$  has fewer than  $m^*/3$  rows, so more than  $2m^*/3$  rows have been removed.

We define  $S''_i$  and  $A''_i$  analogously to  $S'_i$  and  $A'_i$ , but using the block-respecting embeddings of  $P''$ . We conclude that the number of rows in  $A''_i$  is also less than  $m^*/3$ .

Now let  $S_i = S'_i \cap S''_i$ . If the sets  $S_i$  are *not* pairwise disjoint, then we have a block-respecting embedding of  $P$  in  $A$ , as required. Indeed, suppose that for some  $i' < i''$  and  $j$  we have  $j \in S_{i'} \cap S_{i''}$ . Then  $j \in S'_{i'}$ , and hence, by definition, we have a block-respecting embedding  $(f', g')$  of  $P'$  in  $A_{i'}$  with  $g'(b) = j$ . Similarly,  $j \in S''_{i''}$  means that we have a block-respecting embedding  $(f'', g'')$  of  $P''$  in  $A_{i''}$  with  $g''(b) = j$ . These embeddings (more precisely, the corresponding embeddings in  $A$ ) satisfy the conditions of Lemma 2.1, and therefore can be combined into a block-respecting embedding of  $P$  in  $A$ .

Thus, we may assume that the sets  $S_i$  are pairwise disjoint. Note that these sets are all contained in the interval of length  $n$  corresponding to the  $b$ 'th vertical block, so their average size is at most  $n/\beta$ . Fix an  $i$  with  $|S_i| \leq n/\beta$  and let  $B = A_i(\lceil m^* \rceil \times S_i)$ . If a row was removed from  $A_i$  in the second step of the construction of both  $A'_i$  and  $A''_i$ , then this row contains more than  $u/3$  1-entries in  $B$ . Indeed, any row of the  $(k, u)$ -complete matrix  $A_i$  contains at least  $u$  1-entries in the  $b$ 'th vertical block, but if a row is removed during the construction of  $A'_i$ , then fewer than  $u/3$  of these 1-entries are outside  $S'_i$ , and similarly, fewer than  $u/3$  of them are outside  $S''_i$  if the row was also removed during the construction of  $A''_i$ .

As we have seen above, at least  $2m^*/3$  rows of  $A_i$  were removed in the second step of constructing each of  $A'_i$  and  $A''_i$ , so more than  $m^*/3$  rows were removed for both of them. Hence  $B$  contains more than  $m^*/3$  rows with more than  $u/3$  1-entries, implying  $w(B) > um^*/9$ . If the number of columns  $|S_i|$  of  $B$  is less than  $m^*$ , then let  $B^*$  be an  $m^*$ -by- $m^*$  submatrix of  $A$  that contains  $B$ . Clearly,  $w(B^*) \geq w(B)$ . Otherwise, let  $B^*$  be the  $m^*$ -by- $m^*$  submatrix of  $B$  of maximum weight. By averaging, we have  $w(B^*) \geq (m^*/|S_i|)w(B) \geq (\beta m^*/n)w(B)$  in this case. Either way, the weight of our  $B^*$  is at least  $(\beta m^*/n)w(B) \geq um^2\beta/(9n)$ . We will show that this value is more than  $m^*h(m^*)$ . This will then contradict our assumption that the matrix  $A$  is  $h$ -sparse (as  $B^*$  is a submatrix of  $A$ ) and finish the proof of the lemma.

The inequality  $m^*h(m^*) < um^2\beta/(9n)$ , or equivalently,  $h(m^*)/u < m^*\beta/(9n)$  is now easy to check. In fact, all the bounds in the statement of the lemma were chosen to facilitate this calculation. First of all, (3) implies  $h(m^*) < h(n)(m^*/n)^x$  and hence

$$\frac{h(m^*)}{u} < \frac{h(n)}{u} \left( \frac{m^*}{n} \right)^x.$$

Then using the definition of  $m^*$ , the assumption  $x \leq 1/10$ , the lower bound on  $m$ , and finally

$m^*\beta \geq m/2$ , we get

$$\begin{aligned} \frac{h(m^*)}{u} &\leq \frac{h(n)}{u} \left( 1200n^{-x^{s-1}} \left( \frac{h(n)}{u} \right)^2 \right)^x = 1200^x n^{-x^s} \left( \frac{h(n)}{u} \right)^{2x+1} \\ &< 2.1n^{-x^s} \left( \frac{h(n)}{u} \right)^2 < \frac{m}{18n} \leq \frac{m^*\beta}{9n} \end{aligned}$$

as needed.  $\square$

### 3 Reduction to the special case

In this section we prove an explicit version of Theorem 1.3:

**Theorem 3.1.** *For any class- $s$  pattern  $P$ , we have*

$$\text{ex}(n, P) \leq n2^{O(\log^{\frac{s}{s+1}} n)}.$$

*Proof.* For class-0 (single-row) patterns  $P$ , the theorem claims  $\text{ex}(n, P) = O(n)$ . This clearly holds for all such patterns, because any  $n$ -by- $n$  matrix of weight at least  $kn$  contains a row with  $k$  1-entries. Let  $s \geq 1$  and fix a class- $s$  pattern  $P \in \{0, 1\}^{l \times k}$  with  $l, k \geq 2$ . We will prove the theorem by showing that  $\text{ex}(n, P) \leq nh(n)$  holds for all  $n$  with  $h(n) = h_{b,c,d}(n) = d2^{b \log^c n}$ , where  $c = s/(s+1)$ , and  $b = b(c, P)$  and  $d = d(b, c, P)$  are constants to be chosen later as follows.

Our proof proceeds by analyzing a (hypothetical) minimal counterexample, and finds a contradiction if  $b$  is large enough in terms of  $c$  and  $P$ . However, this argument only works if this minimal counterexample has size at least  $n_0 = n_0(b, c, P)$  that only depends on  $b, c$  and  $P$ . By choosing  $d$  large enough, we can make sure that  $h(n) \geq n$  and hence  $\text{ex}(n, P) \leq nh(n)$  holds for every  $n \leq n_0$ , guaranteeing that a minimal counterexample has size at least  $n_0$ .

So suppose for contradiction that  $\text{ex}(n, P) \leq nh(n)$  does not always hold, and let  $n$  be the smallest integer violating this inequality. Then there is a matrix  $A \in \{0, 1\}^{n \times n}$  such that  $P$  is not contained in  $A$ , the weight of  $A$  is  $w(A) > nh(n)$ , but any proper submatrix of  $A$  is  $h$ -sparse.

Let  $x = bc \log^{c-1} n$  be the exponent in the estimate (3) and set  $n^* = \lfloor n/(6k)^{1/x} \rfloor$ . Then (3) implies  $h(n^*) < \frac{h(n)}{6k}$ , and we also have  $x \leq 1$  and hence  $1 \leq n^* < n/k$  if  $n$  is large enough. Set  $\alpha = \lfloor n/n^* \rfloor \geq k$  and let  $B$  be the  $n$ -by- $\alpha n^*$  submatrix of  $A$  formed by the  $\alpha n^*$  columns of largest weight in  $A$ . Then  $w(B) \geq \frac{\alpha n^*}{n} w(A) > \frac{nh(n)}{2}$ . We partition the matrix  $B$  into  $\alpha$  vertical blocks, each consisting of  $n^*$  consecutive columns. We call the intersection of a vertical block and a row a *box*, so the matrix  $B$  consists of  $\alpha n$  boxes. The *weight* of a box is the number of 1-entries in it. We distinguish three classes of boxes according to their weight: We say that a box is *light* if its weight is below  $h(n^*)/\alpha$ , *heavy* if its weight is over  $h(n)/(6k)$ , and *regular* if it is neither heavy, nor light.

The total weight of all light boxes is at most  $\alpha n(h(n^*)/\alpha) = nh(n^*) < nh(n)/(6k)$ .

If some vertical block contained  $n^*$  heavy boxes, then we could form an  $n^*$ -by- $n^*$  submatrix of weight over  $n^*h(n)/(6k) > n^*h(n^*)$ , contradicting the assumption that each proper submatrix of  $A$  is  $h$ -sparse. Therefore, we can form  $n^*$ -by- $n^*$  submatrices, one in each vertical block, that together

contain all the heavy boxes. By the same sparsity condition, each of these matrices have weight at most  $n^*h(n^*)$ . This makes the total weight of the heavy boxes at most  $\alpha n^*h(n^*) < nh(n)/(6k)$ .

The total weight of regular boxes is  $w(B)$  minus the total combined weight of heavy and light boxes. By the above calculations, this is at least  $nh(n)/6$ , therefore the number of regular boxes is at least  $\frac{nh(n)}{6} / \frac{h(n)}{6k} = kn$ . Let  $r_i$  be the number of regular boxes in row  $i$ , so we have  $\sum_{i=1}^n r_i \geq kn$ . Let us say that a row is good for a  $k$ -set  $X$  of vertical blocks if it has a regular box in each block in  $X$ . Note that a row  $i$  is good for  $\binom{r_i}{k} \geq r_i - (k-1)$   $k$ -sets, so on average

$$\frac{\sum_{i=1}^n \binom{r_i}{k}}{\binom{\alpha}{k}} \geq \frac{\sum_{i=1}^n r_i - (k-1)n}{\binom{\alpha}{k}} \geq \frac{n}{\alpha^k}.$$

rows are good for a  $k$ -set of vertical blocks. We fix a  $k$ -set  $X$  of blocks such that at least  $m^* = \lceil n/\alpha^k \rceil$  rows are good for it. Let  $T$  be a set of  $m^*$  such rows and let  $S$  be the set of columns in the vertical blocks of  $X$ . Then  $C = B(T, S)$  is an  $m^*$ -by- $kn^*$   $(k, u^*)$ -complete submatrix for  $u^* = \lceil h(n^*)/\alpha \rceil$ .

We finish the proof of the theorem by showing that the matrix  $C$  satisfies the conditions of Lemma 2.2, yet it violates its statement.

We have already observed that  $C$  is  $(k, u^*)$ -complete and (as a proper submatrix of  $A$ ) it is also  $h$ -sparse. On the other hand,  $P$  does not have a block-respecting (or any) embedding in  $C$ , as otherwise  $A$  would also contain  $P$ , which we assumed not to be the case. Note that we also have  $u^* = \lceil h(n^*)/\alpha \rceil \leq h(n^*)$  and  $m^* = \lceil n/\alpha^k \rceil \leq \lceil n/(2\alpha) \rceil \leq n^*$ . So to get a contradiction from Lemma 2.2, it is enough to show that for an appropriately chosen  $b$ ,

$$m^* \geq 40n^{*1-x^*s} \left( \frac{h(n^*)}{u^*} \right)^2,$$

with  $x^* = bc \log^{c-1} n^* < 1/10$ .

Using  $n^* < n$ ,  $x^* > x$ ,  $u^* \geq h(n^*)/\alpha$  and  $m^* \geq n/\alpha^k$ , we obtain

$$\frac{40n^{*1-x^*s} \left( \frac{h(n^*)}{u^*} \right)^2}{m^*} \leq \frac{40n^{1-x^*s} \left( \frac{h(n^*)}{h(n^*)/\alpha} \right)^2}{n/\alpha^k} = 40\alpha^{k+2}n^{-x^*s}.$$

To prove that the right-hand side does not exceed 1, we will show that its logarithm is negative. In the calculation we use  $\alpha \leq n/n^* < 2(10k)^{1/x}$  and  $x = bc \log^{c-1} n = bc/\log^{1/(s+1)} n$ :

$$\begin{aligned} \log(40\alpha^{k+2}n^{-x^*s}) &< 6 + (k+2) \log \alpha - x^s \log n \\ &< 6 + (k+2) \log(20k)/x - x^s \log n \\ &= 6 + \left( \frac{(k+2) \log(20k)}{bc} - b^s c^s \right) \log^{\frac{1}{s+1}} n. \end{aligned}$$

Then indeed, choosing  $b = (k+2) \log(20k)/c$  makes this last expression negative for any parameters  $k, s \geq 1$  and for all  $n > 1$ .

Finally, the condition  $x^* < 1/10$  is equivalent to  $n^* > 2^{(10bc)^{s+1}}$ , which clearly holds if  $n$  (and hence  $n^*$ ) is large enough.

This brings us to a contradiction whenever  $n \geq n_0$ , where the threshold  $n_0$  only depends on  $b, c$  and  $P$ . But as we mentioned above, we can choose the parameter  $d$  so that a minimal counterexample would definitely satisfy  $n \geq n_0$ .  $\square$



## 4 Concluding remarks

Note that both of the acyclic weight-6 patterns  $Q_1$  and  $Q_2$  are class-2 patterns but not class-1 patterns. Theorem 3.1 gives an  $n2^{O(\log^{2/3} n)}$  bound for their extremal functions. As mentioned in the introduction, no non-trivial bound was known for  $\text{ex}(n, Q_2)$ , but as a preliminary result of this present work, the bound  $\text{ex}(n, Q_1) = n2^{O(\sqrt{\log n \log \log n})}$  was proved in the Master's thesis of the fourth author [10]. The slightly better bound of

$$\text{ex}(n, Q_1) = n2^{O(\sqrt{\log n})} \tag{4}$$

can also be proved using the techniques of Section 3 if instead of Lemma 2.2, we use the following simple and elementary fact: *If  $m > 2n/u$ , then  $Q_1$  has a block-respecting embedding in every  $(4, u)$ -complete matrix  $A \in \{0, 1\}^{m \times 4n}$ .*

In fact, we believe that our methods can be used to prove a slightly stronger variant of Theorem 3.1 where the same bounds are claimed but the pattern classes are defined in the following, more relaxed, way. Once again, class-0 are those with a single row, but for  $s \geq 1$ , we call all patterns class- $s$  that are partitioned into class- $(s - 1)$  patterns when *all* the possible vertical separations (i.e., the horizontal cut that each destroy at most one vertical edge) are applied simultaneously. With this definition,  $Q_1$  is class-1 because it can be partitioned into its three rows in one step, so (4) is implied by this stronger variant of our theorem. However  $Q_2$  is still not class-1. To keep this paper simple, we decided not to work out the details of this argument.

It is clear that transposing a matrix does not change its extremal number:  $\text{ex}(n, P) = \text{ex}(n, P^T)$ . In particular, Theorem 1.3 holds for the analogously defined *horizontally degenerate* patterns, as well. The smallest acyclic pattern that is neither horizontally, nor vertically degenerate (and hence not covered by our theorem) is the following 4-by-4 “pretzel”-like matrix (again omitting 0-entries for clarity):

$$R = \left( \begin{array}{cccc} & 1 & & 1 \\ & | & & | \\ & 1 & 1 & \\ & & | & \\ & & 1 & 1 \end{array} \right)$$

It would be very interesting to obtain nontrivial estimates on  $\text{ex}(n, R)$ .

$R$  is neither horizontally nor vertically separable. However, our approach already breaks down for matrices that are degenerate in a weaker sense, namely that every submatrix is separable, but whether it is horizontally or vertically separable depends on the submatrix. This class of patterns, including the following spiral-like pattern  $S$ , might be easier to handle than  $R$ , and estimating their extremal numbers would certainly be a natural next step towards Conjecture 1.2.

$$S = \left( \begin{array}{cccc} & 1 & 1 & 1 \\ & | & 1 & \\ & 1 & & \\ & & | & \\ & & 1 & 1 \end{array} \right)$$

In terms of ordered graphs of interval chromatic number 2, vertically separable means that the vertices of the first interval  $A$  can be split into two subintervals  $A_1$  and  $A_2$  such that there is at most one vertex in the second interval  $B$  that has neighbors in both  $A_1$  and  $A_2$ . An ordered graph  $G$  is then vertically degenerate if every ordered subgraph of  $G$  is vertically separable.

Pach and Tardos [7] conjectured that every unordered graph  $G_0$  has an ordering such that  $\text{ex}_{<}(n, G) \leq O(\text{ex}(n, G_0) \log n)$ , where  $\text{ex}(n, G_0)$  stands for the usual, unordered Turán number of  $G_0$ . In similar spirit, note that in any acyclic ordered graph of interval chromatic number 2, we can rearrange the vertices of the first interval to get a vertically degenerate graph. Or in terms of matrices, we can make any acyclic pattern vertically degenerate by permuting its rows. This implies the following:

**Corollary 4.1.** *Let  $P$  be an acyclic pattern. We can rearrange the rows of  $P$  such that the resulting pattern  $P'$  satisfies  $\text{ex}(n, P') \leq n^{1+o(1)}$ .*

Note that if we were allowed to permute the columns, as well, then we could easily get  $\text{ex}(n, P') \leq n(\log n)^{O(1)}$ . Indeed, we would then be able to make the last column (or row) end up with a single 1-entry, and then apply induction using (2).

## References

- [1] B. Bollobás, *Extremal Graph Theory*, Academic Press, London, 1978.
- [2] P. Erdős, A. H. Stone, *On the structure of linear graphs*, Bull. Amer. Math. Soc. 52 (1946), 1087-1091.
- [3] Z. Füredi, P. Hajnal, *Davenport-Schinzel theory of matrices*, Discrete Mathematics 103 (1992), 233-251.
- [4] J. T. Geneson, *Extremal functions of forbidden double permutation matrices*, Journal of Combinatorial Theory, Series A 116 (7) (2009), 1235-1244.
- [5] T. Kővári, V. Sós, P. Turán, *On a problem of K. Zarankiewicz*, Colloquium Math. (1954), 50–57.
- [6] A. Marcus, G. Tardos, *Excluded permutation matrices and the Stanley-Wilf conjecture*, Journal of Combinatorial Theory, Ser. A 107 (2004), 153-160.
- [7] J. Pach, G. Tardos, *Forbidden paths and cycles in ordered graphs and matrices*, Israel Journal of Mathematics 155 (2006), 359-380.
- [8] S. Pettie, *Degrees of nonlinearity in forbidden 0-1 matrix problems*, Discrete Mathematics 311 (2011), 2396-2410.
- [9] G. Tardos, *On 0-1 matrices and small excluded submatrices*, Journal of Combinatorial Theory, Ser. A 111 (2005), 266-288.
- [10] C. Weidert, *Extremal problems in ordered graphs*, Master's thesis, Simon Fraser University, 2009.