On 0-1 matrices and small excluded submatrices

Gábor Tardos*
Rényi Institute
Pf. 127, H-1354 Budapest, Hungary
tardos@renyi.hu

Abstract

We say that a 0-1 matrix A avoids another 0-1 matrix (pattern) P if no matrix P' obtained from P by increasing some of the entries is a submatrix of A. Following the lead of [7, 2, 8] and other papers we investigate n by n 0-1 matrices avoiding a pattern P and the maximal number $ex(n,P)$ of 1 entries they can have. Finishing the work of [8] we find the order of magnitude of $ex(n,P)$ for all patterns P with four 1 entries. We also investigate certain collections of excluded patterns. These sets often yield interesting extremal functions different from the functions obtained from any one of the patterns considered.

1 Introduction

In this paper we consider 0-1 matrices and their submatrices. We start with introducing our terminology.

A *submatrix* of a matrix A is obtained from A by deleting rows and columns but without permuting the remaining rows and columns. The *weight* $w(A)$ of a 0-1 matrix A is the number of 1 entries in A. A *pattern* is a 0-1 matrix of weight at least 1. *Deleting* a 1 entry in a matrix means replacing it with 0. The 0-1 matrix A *represents* the same size pattern P if $A = P$ or P can be obtained by deleting a few 1 entries in A. We say that the 0-1 matrix A *contains* the pattern P if a submatrix of A represents P, otherwise we say that A *avoids* P.

For a collection of patterns $\mathcal{P}$ and a positive integer n let $ex(n,\mathcal{P})$ stand for the maximum weight of an n by n 0-1 matrix avoiding all patterns $P \in \mathcal{P}$. We call $ex(\cdot,\mathcal{P})$ the *extremal function* of the collection $\mathcal{P}$. For the extremal function of a single pattern P we write $ex(n,P) = ex(n,\{P\})$.

This problem corresponds to the standard question of Turán type extremal graph theory for bipartite *ordered* graphs. Indeed, consider the n by n matrix to be the adjacency matrix of a bipartite graph between n red and n blue vertices. Here the red and blue vertices are ordered among themselves (no order relation

*Partially supported by the Hungarian National Science Fund grants OTKA T037846 and T046234

is given between a red and a blue vertex). A submatrix corresponds to a full subgraph inheriting its vertex coloring and vertex orders from the original graph. Avoiding a pattern corresponds to avoiding a certain colored, ordered subgraph. In this context $ex(n, \mathcal{P})$ can be interpreted as the maximum number of edges of a simple ordered bipartite graph on n linearly ordered red and a n linearly ordered blue vertices avoiding the colored, ordered subgraphs that correspond to the patterns in $\mathcal{P}$.

Z. Füredi and P. Hajnal [8] study this extremal problem extensively. Earlier works on the subject include [7] and [2]. M. Klazar [12] studies a very similar problem, avoiding (not necessarily bipartite) graphs with given linear order on the vertices. P. Brass, Gy. Károlyi and P. Valtr [3] study another variant, where a *cyclic order* is given on the vertices. An early reference to an extremal problem of graphs on an ordered set of vertices avoiding certain ordered subgraphs is the paper of Czipszer, Erdős, and Hajnal [4] considering infinite graphs. M. Klazar [9] showed that the Stanley-Wilf conjecture (a widely circulated enumerative conjecture on permutations) follows from the Füredi-Hajnal conjecture [8], which states that $ex(n, P) = O(n)$ for any permutation matrix P. Recently A. Marcus and G. Tardos [14] settled the Stanley-Wilf conjecture by proving the Füredi-Hajnal conjecture.

Füredi [7] and Bienstock and Győri [2] study $ex(n, P)$ for some specific patterns P of weight 4, namely Q_1, Q_2, and Q_3 (see Table 1 defining these patterns). Füredi and Hajnal [8] systematically consider all patterns P with weight at most 4. For all but a few of these patterns P they find the extremal function $ex(n, P)$ up to a constant factor. In Section 2 we find $ex(n, P)$ up to a constant factor for all remaining patterns P of weight 4. We also find precise asymptotics (the constant factor) for some of the patterns (namely Q_1 and Q_2). In Section 3 we consider pairs of excluded patterns of weight 4. These pairs often yield extremal functions far from the extremal function of any one of the patterns, or for any pattern of weight 4. Note that similar phenomenon is not known to exist in standard (unordered) Turán type extremal graph theory (although recent results of Faudree and Simonovits [6] point in this direction), but it does exist in the extremal theory of 3-uniform hypergraphs as follows from [16]. In Section 4 we consider all collections of excluded patterns of weight at most 4 and list the eleven "unresolved" cases, where the order of the magnitude of the extremal function is not known. Finally Section 5 contains more open problems and concluding remarks.

We admit however, that the results in this paper are still sporadic. It would be desirable to establish a general Turán type extremal theory of graphs with a linear order on the vertices. M. Klazar in [12] considers a few extremal problems of this type. The results in [3] (or rather their extensions to a linearly ordered vertex set) gives the asymptotics in cases the excluded ordered graph is not bipartite with one color class preceding the other in the ordering. The remaining case of excluded bipartite graphs is closely related to the matrix problem considered in this paper. See more on ordered graphs and this relation in [15].

Research on excluded submatrices is largely motivated by problems in discrete geometry, where order relation between points arise naturally. The motivation of the papers [2, 7] is to give a bound on the number of unit length diagonals of convex n-gons. In the recent paper [15] our bound on $ex(n, L_1)$ is applied to another geometric problem. See more on this in Section 5.

All of our investigations (except a few remarks on pattern R) are about patterns corresponding to cycle free graphs. Thus, all these extremal functions are on the low end of the spectrum. In fact all the extremal functions considered here (except those of R and the empty set) are bounded by $O(n \log n)$. Even with this limited scope, the results here serve as further evidence of the large complexity the extremal theory of ordered (bipartite) graphs.

2 Single excluded patterns of weight 4

Füredi and Hajnal [8] found the order of magnitude of $ex(n, P)$ for all but a few patterns P of weight at most 4. In this section we finish their work by doing the same for the remaining few patterns. We have to deal with two specific patterns L_1 and Q_3. Refer to Table 1 defining the patterns considered. In the table we use dots for the 1 entries and blank spaces for the 0 entries.

$Q_1 = \begin{pmatrix} & \bullet & \bullet \\ \bullet & & \bullet \end{pmatrix}$	$L_1 = \begin{pmatrix} \bullet & & \bullet \\ \bullet & & \\ & \bullet & \end{pmatrix}$
$Q_2 = \begin{pmatrix} & \bullet & \bullet \\ & & \bullet \\ \bullet & & \end{pmatrix}$	$L_2 = \begin{pmatrix} & \bullet & & \bullet \\ \bullet & & & \\ & & \bullet & \end{pmatrix}$
$Q_3 = \begin{pmatrix} \bullet & & \bullet \\ & \bullet & \\ & & \bullet \end{pmatrix}$	$L_3 = \begin{pmatrix} & \bullet & & \\ & & & \bullet \\ \bullet & & & \\ & & \bullet & \end{pmatrix}$
$R = \begin{pmatrix} \bullet & \bullet \\ \bullet & \bullet \end{pmatrix}$	$L_4 = \begin{pmatrix} \bullet & & \bullet & \\ & & & \bullet \\ \bullet & & & \\ & \bullet & & \end{pmatrix}$
$S_1 = \begin{pmatrix} \bullet & & \bullet & \\ & \bullet & & \bullet \end{pmatrix}$	$L_5 = \begin{pmatrix} & \bullet & \bullet & \\ \bullet & & & \bullet \\ & \bullet & & \end{pmatrix}$
$S_2 = \begin{pmatrix} \bullet & & & \\ & & \bullet & \\ & \bullet & & \bullet \end{pmatrix}$	$L_6 = \begin{pmatrix} \bullet & & & \bullet \\ \bullet & & & \\ & \bullet & \bullet & \end{pmatrix}$
$S_3 = \begin{pmatrix} \bullet & & & & \\ & & & \bullet & \\ & \bullet & \bullet & & \bullet \end{pmatrix}$	$T = \begin{pmatrix} \bullet & \bullet & \\ & & \bullet \\ \bullet & & \bullet \end{pmatrix}$

Table 1.
The patterns considered

Theorem 2.1.
$$ex(n, L_1) \le 5n$$

Proof: Let $A = (a_{ij})$ be an n by n 0-1 matrix avoiding the pattern L_1. For a column j of A let $l(j)$ denote the row index of the last 1 in column j, i.e., $a_{l(j)j} = 1$ but $a_{ij} = 0$ for $i > l(j)$. If column j does not contain a 1 entry we do not define $l(j)$. Let j' be a column of A containing at least a single 1 and let j be the largest index with $j < j'$ and $l(j) \ge l(j')$. We say that column j' *finds* the entry a_{ij}, where i is the largest index with $i < l(j')$ and $a_{ij} = 1$. If there is no index j or i satisfying the conditions, then column j' does not find any entry. Clearly, any column j' finds at most one entry of A.

We claim that each entry $a_{ij} = 1$ of A falls into one of the categories below.

(i) a_{ij} is the last or the second last 1 in row i.

(ii) a_{ij} is the last 1 in column j.

(iii) $a_{ij'} = 1$ is the last 1 in some column $j' > j$ and there is no 1 in row i between a_{ij} and $a_{ij'}$.

(iv) a column $j' > j$ finds the entry a_{ij}.

At most $2n$ entries fall into category (i) and at most n entries fall in each of the categories (ii), (iii), and (iv). The claim above implies the theorem.

To prove the claim fix an entry $a_{ij} = 1$ of A and assume it does not fall into any of the categories (i), (ii), and (iii). We need to show that some column j' finds a_{ij}. As a_{ij} is not in category (i) there exist entries $a_{ij_1} = a_{ij_2} = 1$ with $j < j_1 < j_2$. We can choose a_{ij_1} to be the first 1 entry in row i after a_{ij}. With this choice we have $l(j_1) > i$ as otherwise a_{ij} is in category (iii). As a_{ij} is not in category (ii) we can find the smallest index $i_1 > i$ with $a_{i_1 j} = 1$.

Let j' be smallest index with $j' > j$ and $l(j') > i$. As j_1 is such an index j' exists and we have $j' \leq j_1$. We must have $l(j') \leq i_1$. Indeed, otherwise the rows $i < i_1 < l(j')$ and the columns $j < j' < j_2$ would determine a submatrix representing L_1. We have $j < j'$ and $l(j) \geq i_1 \geq l(j')$ but no index $j < j'' < j'$ satisfies $l(j'') \geq l(j')$ by the choice of j'. So column j' finds the last 1 entry on column j before row $l(j')$. As $i < l(j') \leq i_1$ and a_{ij} and $a_{i_1 j}$ are consecutive 1 entries of column j, column j' finds the entry a_{ij}. This proves the claim and the theorem. □

Theorem 2.2.

$$ex(n, Q_3) = \Omega(n \log n)$$

A matching upper bound follows from the same bound on $ex(n, Q_1)$ by Lemma 2.3 below. The bound $ex(n, Q_1) = \Theta(n \log n)$ was proved in both of the papers [2, 7]. Bienstock and Győri [2] observed the consequence $ex(n, Q_3) = O(n \log n)$ and gave a construction establishing $ex(n, Q_3) = \Omega(n \log n / \log \log n)$. Later [8] gives a simplified construction for the same bound. Here we give an improved construction.

Let i and j be strings of equal length over an ordered set of letters. We use $<$ to denote the *lexicographic ordering*, i.e., we have $i < j$ if and only if the letter of i is smaller than that of j at the *first* position where i and j differ. We use $<^*$ to denote the *anti-lexicographic ordering*, i.e., we have $i <^* j$ if and only if the letter of i is smaller than that of j at the *last* position where i and j differ. We use the relations $>$, $\leq$, $\geq$, $>^*$, $\leq^*$, and $\geq^*$ between strings with their obvious meaning.

Proof: We construct an n by n 0-1 matrix C_n of weight $\Theta(n \log n)$ avoiding the pattern Q_3 for $n = 2^m$ ($m \geq 1$). For other values of n simply pad our construction for the largest power of 2 below n by adding zero rows and columns.

We index the rows and the columns of the matrix $C_n = (c_{ij})$ with the 0-1 strings of length m. We define C_n by letting $c_{ij} = 1$ if and only if the strings i and j differ in a single coordinate u with $i_u = 0$ and $j_u = 1$.

The weight of the matrix constructed is $w(C_n) = nm/2 = \Theta(n \log n)$ as needed.

So far we have not defined the order of the rows and the columns. Recall that row and column indices are 0-1 strings. If we use the standard lexicographic order for both the rows and the columns, we obtain a matrix avoiding Q_2. This is, indeed, one of the standard constructions of a matrix avoiding that pattern. But the matrix constructed with the lexicographic order *does not avoid* Q_3. We need a different order.

We order the rows of the matrix C_n lexicographically according to their index. For the columns we use the anti-lexicographic order of their indices.

In the rest of the proof we prove that the matrix C_n constructed avoids the pattern Q_3. This means establishing that for row indices $i < i' < i''$ and column indices $j >^* j' >^* j''$ we cannot have $c_{ij} = c_{ij''} = c_{i''j} = c_{i'j'} = 1$. We prove that the four equalities cannot hold simultaneously even if we only have $i < i' \leq i''$ and $j >^* j' \geq^* j''$. Thus, we establish that the matrix constructed avoids the patterns $\overline{Q}_1$, $\overleftarrow{Q}_1$ and R besides Q_3. Refer to Table 2 for $\overline{Q}_1$ and $\overleftarrow{Q}_1$. (Note however, that any matrix avoiding the pattern Q_3 can be turned into one avoiding all these patterns by simply deleting the first 1 in every row and the last 1 in every column. See Lemma 2.3.)

Assume that the row indices $i < i' \leq i''$ and the column indices $j >^* j' \geq^* j''$ satisfy $c_{ij} = c_{ij''} = c_{i''j} = c_{i'j'} = 1$. Our goal is to find a contradiction proving our assumption wrong. Let $1 \leq u \leq m$ be the only position where the sequences i and j differ. We have $i_u = 0$, $j_u = 1$, and $i_z = j_z$ for $z \neq u$. For the sole position v where i'' and j differ $i < i''$ implies $u < v$. Thus, $i_z = i''_z$ for $z < u$ and from $i < i' \leq i''$ we also have $i_z = i'_z = i''_z$ for $z < u$. Similarly, for the sole position w where j'' and i differ, $j >^* j''$ implies $w < u$. Thus, $j_z = j''_z$ for $z > u$ and from $j >^* j' \geq^* j''$ we have $j_z = j'_z = j''_z$ for $z > u$. From $a_{i'j'} = 1$ we have $i'_z \leq j'_z$ for all z. As $i' > i$ there must exist a position z with $i'_z > i_z$. Since we have $i'_z = i_z$ for $z < u$ and $i'_z \leq j'_z = j_z = i_z$ for $z > u$ we must have $i'_u > i_u$. Similarly, from $j' <^* j$ there must exist a position z with $j'_z < j_z$. Since $j'_z = j_z$ for $z > u$ and $j'_z \geq i'_z = i_z = j_z$ for $z < u$ we must have $j'_u < j_u$. But $i_u < i'_u \leq j'_u < j_u$ is a contradiction as all these values are binary. The contradiction proves the correctness of the construction. □

The bipartite graph corresponding to the construction above is disconnected, it has 2 isolated vertices and m nontrivial connected components.

We call the 1 by 1 pattern of a single 1 entry *trivial*. The rotation or reflection of a matrix is called a *geometric transformation*. We use $\overline{A}$, $A^{|}$, $A^{/}$, and $A^{\backslash}$ for the matrices obtained by reflecting the matrix through a horizontal, vertical, or diagonal line as indicated. We denote by $\overleftarrow{A}$, $\overrightarrow{A}$, and $\dot{A}$ the matrixes obtained from A via rotation in the positive, negative direction, and via central reflection, respectively. For the less geometric-minded reader geometric transformations are the reversing of the order of the rows and/or columns and/or taking the transpose. If a pattern P' is obtained from P by a geometric transformation, then P and P' are called *equivalent*. A pattern has at most eight equivalents counting itself. See Table 2 for the equivalents of Q_1, Q_2 and Q_3. If the collection $\mathcal{P}'$ is obtained by applying the same geometric transformation to all patterns in $\mathcal{P}$, then $\mathcal{P}$ and $\mathcal{P}'$ are called *equivalent*. Removing all blank rows and columns of a pattern is called *reducing* the pattern. By *reducing* a collection $\mathcal{P}$ of patterns we mean reducing each pattern in $\mathcal{P}$ and taking the collection of the resulting reduced patterns.

The next lemma collects a few simple observations connecting the extremal functions $ex(n, \mathcal{P})$ for different sets $\mathcal{P}$.

Lemma 2.3. [8] *Let P and P' be patterns and let $\mathcal{P}$ and $\mathcal{P}'$ be sets of patterns.*

a) *If $\mathcal{P} \subseteq \mathcal{P}'$, then $ex(n, \mathcal{P}) \geq ex(n, \mathcal{P}')$.*

b) *If P contains P', then $ex(n, \mathcal{P} \cup \{P'\}) \leq ex(n, \mathcal{P} \cup \{P\})$.*

c) *If $\mathcal{P}$ and $\mathcal{P}'$ are equivalent, then $ex(n, \mathcal{P}) = ex(n, \mathcal{P}')$.*

d) *If P' is obtained from P by adding a first column to P with a single 1 entry next to a 1 entry of P, then $ex(n, \mathcal{P} \cup \{P\}) \leq ex(n, \mathcal{P} \cup \{P'\}) \leq ex(n, \mathcal{P} \cup \{P\}) + n$.*

e) *If P is non-trivial, then $ex(n, P) \geq n$.*

f) *If $\mathcal{P}'$ is finite and reduces to $\mathcal{P}$, then $ex(n, \mathcal{P}) \leq ex(n, \mathcal{P}') = O(ex(n, \mathcal{P}) + n)$.*

g) *If P' is obtained from the pattern P by adding an extra column containing a single 1 entry between two columns of P and the newly introduced 1 entry has 1 next to it on both sides, then $ex(n, \mathcal{P} \cup \{P\}) \leq ex(n, \mathcal{P} \cup \{P'\}) \leq 2ex(n, \mathcal{P} \cup \{P\})$.*

$Q_1 = \begin{pmatrix} & \bullet & \bullet \\ \bullet & & \bullet \end{pmatrix}$	$Q_2 = Q_2^{/} = \begin{pmatrix} & \bullet & \bullet \\ & & \bullet \\ \bullet & & \end{pmatrix}$
$\overline{Q}_1 = \begin{pmatrix} \bullet & & \bullet \\ & \bullet & \bullet \end{pmatrix}$	$\overline{Q}_2 = \overrightarrow{Q}_2 = \begin{pmatrix} \bullet & & \\ & & \bullet \\ & \bullet & \bullet \end{pmatrix}$
$Q_1^{\vert} = \begin{pmatrix} \bullet & \bullet & \\ \bullet & & \bullet \end{pmatrix}$	$Q_2^{\vert} = \overleftarrow{Q}_2 = \begin{pmatrix} \bullet & \bullet & \\ \bullet & & \\ & & \bullet \end{pmatrix}$
$\dot{Q}_1 = \begin{pmatrix} \bullet & & \bullet \\ \bullet & \bullet & \end{pmatrix}$	$\dot{Q}_2 = Q_2^{\backslash} = \begin{pmatrix} & & \bullet \\ \bullet & & \\ \bullet & \bullet & \end{pmatrix}$
$Q_1^{/} = \begin{pmatrix} \bullet & \bullet \\ & \bullet \\ \bullet & \end{pmatrix}$	$Q_3 = Q_3^{/} = \begin{pmatrix} \bullet & & \bullet \\ & \bullet & \\ & & \bullet \end{pmatrix}$
$\overrightarrow{Q}_1 = \begin{pmatrix} \bullet & \\ & \bullet \\ \bullet & \bullet \end{pmatrix}$	$\overline{Q}_3 = \overrightarrow{Q}_3 = \begin{pmatrix} & & \bullet \\ & \bullet & \\ \bullet & & \bullet \end{pmatrix}$
$\overleftarrow{Q}_1 = \begin{pmatrix} \bullet & \bullet \\ \bullet & \\ & \bullet \end{pmatrix}$	$Q_3^{\vert} = \overleftarrow{Q}_3 = \begin{pmatrix} \bullet & & \bullet \\ & \bullet & \\ \bullet & & \end{pmatrix}$
$Q_1^{\backslash} = \begin{pmatrix} & \bullet \\ \bullet & \\ \bullet & \bullet \end{pmatrix}$	$\dot{Q}_3 = Q_3^{\backslash} = \begin{pmatrix} \bullet & & \\ & \bullet & \\ \bullet & & \bullet \end{pmatrix}$

Table 2.
Equivalents of the patterns Q_1, Q_2, and Q_3

Parts a), b), c), d), and e) of the above lemma comes from [8]. Combining part c) (symmetry) with part d), one can use part d) for situations when a last column, a first row, or a last row is introduced. This extension makes part d) of the lemma applicable very often. Typical applications of parts b), c), and d) include

$$ex(n, Q_2) \leq ex(n, Q_1) + n,$$
$$ex(n, Q_3) \leq ex(n, Q_1) + n.$$

Part e) also follows from the more general statement of Theorem 4.1. Part f) is introduced to handle patterns with blank rows or columns. Part g) is similar to part d). It is not used later in this paper.

Proof: Parts a), b), and c) are trivial.

The first inequalities in parts d), f), and g) are special cases of part b).

For the second inequality in part d) consider a matrix A avoiding P' and the patterns in $\mathcal{P}$ and delete the first 1 entry in every non-blank row. Clearly, $w(B) \geq w(A) - n$ for the resulting matrix B and it still avoids the patterns in $\mathcal{P}$. If a submatrix of B represents P, then the same submatrix of A can be extended to a larger submatrix representing P'. The contradiction proves that B avoids P.

Part e) follows for patterns P containing a 1 entry outside the first column by noticing that any matrix not having a 1 outside the first column avoids P. For other patterns use symmetry.

For the second inequality in part f) let k be large enough so that no pattern in $\mathcal{P}'$ has k consecutive blank rows or columns. Take a maximal weight n by n 0-1 matrix A avoiding the patterns in $\mathcal{P}'$. Delete the 1 entries in the first and last k rows and columns of A. We obtain the matrix B and lose at most $4kn$ in the weight. For $0 \leq a, b < k$ let B_{ab} be the matrix obtained from B by deleting all 1 entries except those with row and column indices i and j satisfying that $i \bmod k = a$ and $j \bmod k = b$. The weights $w(B_{ab})$ sum to $w(B)$. It is easy to see that every matrix B_{ab} avoids the patterns in $\mathcal{P}$. We have $ex(n, \mathcal{P}') = w(A) \leq w(B) + 4kn \leq k^2 ex(n, \mathcal{P}) + 4kn = O(ex(n, P) + n)$ as claimed.

Finally for the second inequality of part g) consider a matrix A avoiding P' and the patterns in $\mathcal{P}$. Let B be obtained from A by deleting every other 1 entry in every row. Clearly, $w(B) \geq w(A)/2$ and B avoids P and the patterns in $\mathcal{P}$. □

The extremely slow growing inverse Ackermann function is denoted by $\alpha(n)$.

Corollary 2.4. *If P is a pattern with $w(P) \leq 4$ we have*

$$ex(n, P) = \begin{cases} 0 & \text{or} \\ \Theta(n) & \text{or} \\ \Theta(n\alpha(n)) & \text{or} \\ \Theta(n \log n) & \text{or} \\ \Theta(n^{\frac{3}{2}}). \end{cases}$$

We have $ex(n, P) = 0$ for the trivial pattern only.

We have $ex(n, P) = \Theta(n\alpha(n))$ for the patterns P that reduce to an equivalent of S_1 or S_2.

We have $ex(n, P) = \Theta(n \log n)$ for the patterns P that reduce to an equivalent of Q_1, Q_2, or Q_3.

We have $ex(n, P) = \Theta(n^{3/2})$ for the patterns P that reduce to R.

For all the rest of the patterns P with $w(P) \leq 4$ (including all non-trivial patterns of weight at most 3) we have $ex(n, P) = \Theta(n)$.

Proof: Most of this characterization was proved by Füredi and Hajnal [8]. They established the order of magnitude of the extremal function of the patterns P with $w(P) \leq 4$ not containing blank rows and columns except for the patterns equivalent with L_1, L_2, L_3, and Q_3. The missing upper bound for L_1 is stated in Theorem 2.1. The missing lower bound for Q_3 is stated in Theorem 2.2. Theorem 2.1 and Lemma 2.3 imply $ex(n, L_2) \leq 6n$ and $ex(n, L_3) \leq 7n$ providing the missing upper bounds for these patterns. Finally Lemma 2.3/f extends this classification to patterns with blank rows or columns. For non-trivial patterns of weight 1 we need to use Lemma 2.3/e too. □

As we shall see in Section 4 the little variety of extremal functions is a result of excluding only a *single* pattern. If we exclude several patterns, still with weight at most 4, then several new extremal functions show up.

We finish this section by proving asymptotically tight bounds for the extremal function of the single excluded patterns Q_1 and Q_2. It seems that the pattern Q_1 plays an important role both in the applications and in implications through Lemma 2.3. It was also the very first excluded pattern considered in this line of research, see [7]. Although the order of magnitude of the extremal function for Q_1 has already been established, the upper bound and the lower bound were a constant factor apart. We close this gap by a pair of an upper bound and a construction.

Theorem 2.5. *With the binary* log *function we have*

$$ex(n, Q_1) = n \log n + O(n),$$

$$ex(n, Q_2) = n \log n + O(n).$$

Proof: We start with the construction. Let $D_n = (d_{ij})$ be the n by n 0-1 matrix given by $d_{ij} = 1$ if and only if $j - i = 2^k$ for some integer k. Here $0 \leq k < \log n$ and a given k contributes $n - 2^k$ to the weight. We have $w(D_n) = \sum_{k=0}^{\lfloor \log n \rfloor} (n - 2^k) \geq n \log n - n$.

We prove that for $i < i' \leq i''$ and $j > j' \geq j''$ we do not have $a_{ij} = a_{ij'} = a_{i'j} = a_{i''j''} = 1$. This establishes that A avoids the patterns Q_1, Q_2, $Q_1^{/}$, and R. Assume that the indices $i < i'$ and $j > j'$ satisfy $a_{ij} = a_{ij'} = a_{i'j} = 1$. Our goal is to prove that for $i'' \geq i'$ and $j'' \leq j'$ we have $a_{i''j''} = 0$. We have $j - i = 2^k$ for some integer k. The values $j - i'$ and $j' - i$ are both less than $j - i$ and also powers of 2, so we have $j - i' \leq 2^{k-1}$ and $j' - i \leq 2^{k-1}$. We have $j'' - i'' \leq j' - i' = (j' - i) + (j - i') - (j - i) \leq 2^{k-1} + 2^{k-1} - 2^k = 0$. As $j'' \leq i''$ we have $d_{i''j''} = 0$ as needed.

We turn to the upper bound. Let $A = (a_{ij})$ be an n by n 0-1 matrix avoiding the pattern Q_1. We need to bound $w(A)$.

Let $f(i)$ be the index of the first 1 in row i of A, i.e., $a_{if(i)} = 1$, but $a_{ij} = 0$ for $j < f(i)$. Let $p(i, j)$ be the index of the last 1 in row i of A preceding column j, i.e., $p(i, j) < j$ and $a_{ip(i,j)} = 1$ but $a_{ij'} = 0$ for $p(i, j) < j' < j$. If row i does not contain a 1 or does not contain a 1 before column j we do not define these values. For each 1 entry in A except for the first two in each row we define two weight functions as follows. Let S be the set of the pairs (i, j) satisfying $a_{ij} = 1$ and $p(i, j) > f(i)$. Clearly, $|S| \geq w(A) - 2n$. We define the weight functions for $(i, j) \in S$ by letting

$$w_1(i,j) = \log\left(\frac{j - f(i)}{p(i,j) - f(i)}\right);$$

$$w_2(i,j) = \log\left(\frac{j - f(i)}{j - p(i,j)}\right).$$

Take a row i of A containing at least two 1 entries. Summing for j with $(i, j) \in S$ we have $\sum_j w_1(i, j) = \log((j_0 - f(i))/(j_1 - f(i))) \leq \log n$, where j_0 and j_1 are the column indices of the last and second 1 entries, respectively, in row i. For the total weight we have

$$\sum_{(i,j)\in S} w_1(i,j) \leq n \log n.$$

To bound the total weight distributed by w_2 we need to use that A avoids Q_1. Consider a column j of A containing some positions in S. For indices $i < i'$ with $(i, j) \in S$ and $(i', j) \in S$ we have $p(i, j) \leq f(i')$, as otherwise the rows $i < i'$ and the columns $f(i') < p(i, j) < j$ determine a submatrix representing the pattern Q_1. For a fixed column j we have $\sum_i w_2(i, j) \leq \log(j - f(i_0)) \leq \log n$, where the summation extends for indices i with $(i, j) \in S$ and i_0 is the smallest such index. For the total weight we have

$$\sum_{(i,j)\in S} w_2(i,j) \leq n \log n.$$

For $(i, j) \in S$ we have $w_1(i, j) + w_2(i, j) = -\log(t - t^2) \geq 2$, where $t = (j - p(i, j))/(j - f(i))$. Thus, bounding the total weights also bounds the size of S:

$$|S| \leq \frac{1}{2}\left(\sum_{(i,j)\in S} w_1(i,j) + \sum_{(i,j)\in S} w_2(i,j)\right) \leq n \log n.$$

Adding the first two 1 entries in each row we obtain $w(A) \leq n \log n + 2n$, proving the first statement of theorem. Using Lemma 2.3 one has $ex(n, Q_2) \leq ex(n, Q_1) + n$ proving the second statement. □

We can bring the upper and lower estimates for $ex(n, Q_1)$ somewhat closer to each other. For a better upper bound notice that for fixed j we have $\sum_i w_2(i, j) \leq \log(j - f(i_0)) \leq \log(j - 1)$, so $\sum_{(i,j)\in S} w_2(i, j) \leq \sum_{j=1}^{n-1} \log j \leq n \log n - n \log e$, where e is the base of the natural logarithm. The construction can also be improved by introducing 1 entries in the main diagonal. The n new 1 entries introduce the pattern R, but the matrix still avoids the patterns Q_1, Q_2, and $Q_1^{/}$. This gives

$$n \log n \leq ex(n, Q_1) \leq n \log n + (2 - \frac{\log e}{2})n.$$

One can further introduce 1 entries on the diagonal just below the main diagonal: this results in a matrix not avoiding Q_1 or $Q_1^{/}$ but still avoiding Q_2. We have

$$n \log n + n - 1 \le ex(n, Q_2) \le n \log n + (3 - \frac{\log e}{2})n.$$

As we mostly deal with orders of magnitude in this paper the base of the logarithm rarely matters. In the few places where it does (like in the estimates above) we use the binary logarithm.

Further improvement in the construction is possible by "shifting" the matrix, defining $d_{ij} = 1$ if and only if $j - i + \lfloor \sqrt{n} \rfloor$ is a power of 2 or zero (or -1 in case of avoiding Q_2). But the improvement here is only $\Theta(\sqrt{n})$.

Note that Q_3 (the third non-equivalent pattern with extremal function $\Theta(n \log n)$) seems to be harder to handle. The construction in Theorem 2.2 gives $ex(n, Q_3) \ge \frac{n \log n}{2} - O(n)$ for $n = 2^m$. The same lower bound can easily be extended to arbitrary values of n. The upper bound comes from the bound on Q_1 and gives only $ex(n, Q_3) \le n \log n + O(n)$. The upper and lower bounds are a factor of 2 apart.

The upper and lower bounds on the extremal functions $ex(n, S_1)$ and $ex(n, S_2)$ are also a constant factor apart. Standard extremal graph theory gives $ex(n, R) = n\lceil \sqrt{n} \rceil$ if n is the number of points in a finite plane (and $ex(n, R) = n^{3/2} + O(n^{1.2625})$ in general from the prime gap bound in [1]).

For certain patterns P with a linear extremal function $ex(n, P)$ can be found asymptotically, or even exactly. For the pattern L_1 considered in Theorem 2.1 we have $ex(n, L_1) \ge 4n - 4$: the matrix with 1 entries in the last two rows and columns and 0 entries elsewhere avoids L_1. (It also avoids the pattern obtained by deleting all but the top left 1 entry from L_1). There is still a constant factor gap between this and the upper bound of $5n$ in Theorem 2.1.

3 Pairs of excluded patterns

In Turán type extremal graph theory $ex(n, \mathcal{G})$ stands for the maximum number of edges a simple graph on n vertices can have without containing a subgraph isomorphic to a graph in the collection $\mathcal{G}$. For many graphs G_1 and G_2 we have

$$ex(n, \{G_1, G_2\}) = \Theta(\min(ex(n, \{G_1\}), ex(n, \{G_2\})).$$

It is an open problem in classical extremal graph theory if the above holds for all pairs of graphs. Recent results of Faudree and Simonovits [6] suggests certain graphs where the above relation may fail. The famous result of I. Ruzsa and E. Szemerédi [16] shows that the analogous relation does not hold for some 3-uniform hypergraphs. As we shall see, the similar relation for the 0-1 matrix problem considered in this paper fails very often. Considering more than a single excluded pattern leads to new interesting problems. Even if we restrict our attention to patterns with weight at most 4 several new extremal functions are obtained by excluding several patterns simultaneously.

Our most interesting results are about forbidding pairs of patterns equivalent to Q_1, Q_2, or Q_3, these are the patterns P in Corollary 2.4 with $ex(n,P) = \Theta(n \log n)$. See discussion in Section 4 on excluding different patterns or more than 2 patterns.

As the first example showing that excluding two patterns can lead to a significant decrease of the extremal function, consider Q_1 and $\overline{Q}_1$. By Theorem 2.5 and symmetry we have

$$ex(n,Q_1) = ex(n,\overline{Q}_1) = n \log n + O(n).$$

Excluding both of these patterns yields a linear bound:

Theorem 3.1.

$$ex(n,\{Q_1,\overline{Q}_1\}) = 3n-2$$

Proof: For the upper bound consider an n by n matrix A avoiding Q_1 and $\overline{Q}_1$ and delete the first two 1 entries in each row. It is easy to see that two 1 entries cannot remain in the same column, as together with two deleted entries in their rows they would represent either Q_1 or $\overline{Q}_1$. So we have at most a single 1 entry remaining in each column and clearly, no 1 entries in the first two columns. So after deleting at most $2n$ 1 entries from A at most $n-2$ such entries remain, so the weight of A is at most $3n-2$ as stated.

Consider an n by n matrix with the first two columns consisting of all 1 entries and each of the remaining columns containing a single 1 entry. This matrix avoids Q_1 and $\overline{Q}_1$ and it has weight $3n-2$. It is also straight forward to check that these are the only extremal matrices.

Note that neither the upper bound nor the construction makes sense for $n=1$ but the statement of the theorem is true in this case too. □

Theorem 3.1 gives an example where excluding a pair of patterns yields an extremal function much smaller than the extremal function of either one of these patterns. But the linear function in Theorem 3.1 is hardly a new and exciting extremal function, it coincides with the extremal function of any two by three pattern of weight one. We proceed by investigating another pairs of excluded patterns equivalent to Q_1. Although the excluded set of patterns in Theorems 3.2 and 3.5 are very similar to each other, the extremal function they determine differs widely. Neither of these functions appear in Corollary 2.5 characterizing single excluded patterns of weight at most 4.

The following definitions will come handy in the proofs of both Theorems 3.2 and 3.5. Assume the 0-1 matrix $A = (a_{ij})$ is fixed. The row and column indices are positive integers. If row i of A is not blank we use $f(i)$ and $l(i)$ to denote the column indices of the first and last 1 entries in that row, respectively. The entries $a_{il(i)} = 1$ of A are called *last entries*, these are the last 1 entries in their row. There is a single last entry in each non-blank row. If the $a_{ij} = 1$ entry is not last, let $n(i,j)$ denote the column index of the next 1 entry in the row i of A, i.e., the smallest value $j' > j$ with $a_{ij'} = 1$. We call an $a_{ij} = 1$ entry a *left entry of* A if $j - f(i) \le l(i) - j$, i.e., if j is in the left half of the interval $[f(i), l(i)]$. We call A left-leaning if at least half of the 1 entries in A are left entries.

Theorem 3.2.

$$ex(n, \{Q_1, Q_1^{|}\}) = \Theta(n \log n / \log \log n)$$

Proof: We start with the construction of the n by n matrix E_n proving the lower bound. First assume $n = k^k$ for some $k \geq 1$. We use sequences from $\{1, \ldots, k\}^k$ as column indices and sequences from $\{0, 1 \ldots, k\}^k$ containing exactly one 0 as row indices. Notice that the number of row indices, as well as the number of column indices, are exactly $k^k = n$. We order both the rows and the columns lexicographically according to their index. We define $E_n = (e_{ij})$ by setting $e_{ij} = 1$ if and only if the sequences i and j differ at a single position. Clearly, i has 0 at this position, so we must have $i < j$ if $e_{ij} = 1$.

We have $w(E_n) = kn = \Theta(n \log n / \log \log n)$.

Assume that $e_{ij} = e_{i'j} = 1$ for indices $i < i'$ and j. Let u be the position where i has 0 and let v be the position where i' has 0. Clearly, i and i' agree with j and with each other except for these two positions. Since $i < i'$ we must have $u < v$.

If $e_{ij'} = 1$ for some $j' < j$, then j and j' differ at position u, while i' and j agree agree up to position v, so we have $j' < i'$. For a row index $i'' \geq i'$ and a column index $j'' \leq j'$ we have $i'' \geq i' > j' \geq j''$ and therefore $e_{i''j''} = 0$. This shows that E_n avoids the patterns Q_2, Q_1, $Q_1^{/}$, and R. Similarly, if $e_{ij'} = e_{i'j''} = 1$ for some indices $j' > j$ and j'', then j and j' differ at position u, while j and j'' agree up to position v, so we have $j' > j''$. This shows that E_n avoids the pattern $Q_1^{|}$.

To extend our construction to values of n not of the form k^k it is not enough to pad the matrices constructed above by blank rows and columns. Let $n > 0$ be arbitrary and let k be the largest integer with $n' = k^k \leq n$. Let $t = \lfloor n/n' \rfloor$. Let E_n contain t diagonally arranged blocks, each containing $E_{n'}$, and 0 entries outside these blocks. We have $w(E_n) = \Theta(n \log n / \log \log n)$ and E_n avoids the patterns Q_2, Q_1, $Q_1^{/}$, R, and $Q_1^{|}$.

For the upper bound let $A = (a_{ij})$ be an n by n 0-1 matrix avoiding both patterns Q_1 and $Q_1^{|}$. The row and column indices i and j are integers $1 \leq i, j \leq n$. We need to bound $w(A)$.

Consider the matrix $A^{|}$ obtained by reflecting A through a vertical line. Clearly, $w(A^{|}) = w(A)$. As this geometric transformation maps Q_1 and $Q_1^{|}$ to each other, $A^{|}$ also avoids these patterns. Clearly, any 1 entry of A is either a left entry of A or its image is a left entry of $A^{|}$. Therefore, one of A and $A^{|}$ is left-leaning. We assume without loss of generality that A is left-leaning. In the rest of the proof we use only that A avoids $Q_1^{|}$, the pattern Q_1 plays no further role.

We assume $n \geq 5$ and let $t = \log n / \log \log n > 2$.

We call an entry $a_{ij} = 1$ of A *long* if it is not last and $n(i,j) - j \geq (l(i) - f(i))/t$. A 1 entry of A is long if the gap to the next 1 in the same row is at least a $1/t$ fraction of the distance between the first and last 1 entries in the row. Clearly, there are no more than t long entries in any row of A, no more than nt long entries in total.

We call a left entry $a_{ij} = 1$ of A *short* if it is neither last, nor long. For a short entry $a_{ij} = 1$ we have $l(i) - j \geq (l(i) - f(i))/2$, $n(i,j) - j < (l(i) - f(i))/t$, and therefore $\frac{l(i)-j}{n(i,j)-j} > \frac{t}{2}$.

Let us fix j and consider the entries $a_{ij} = a_{i'j} = 1$ in column j with $i < i'$. If we have $l(i') > n(i,j)$, then the rows $i < i'$ and the columns $j < n(i,j) < l(i')$ determine a submatrix representing $Q_1^{|}$. As A avoids $Q_1^{|}$ we have $l(i') \leq n(i,j)$ if $n(i,j)$ exists. Let j be still fixed and consider the set S of row indices i of the non-last entries $a_{ij} = 1$ of A. Let i_0 and i_1 be the minimal and maximal elements of S. Using the above observation one has

$$\prod_{i \in S} \frac{l(i) - j}{n(i,j) - j} \leq \frac{l(i_0) - j}{n(i_1) - j} < n.$$

Each factor of this product is at least 1. The factors corresponding to short entries are larger than $t/2$. Thus, for the number k of short entries in column j we have $(t/2)^k < n$ and therefore $k < \log n / \log(t/2)$. The total number of short entries in A is less than $n \log n / \log(t/2)$.

Each left entry in A is either short or long or the only 1 entry in its row. Thus A has at most $nt + n \log n / \log(t/2) + n = O(n \log n / \log \log n)$ left entries. As A is left-leaning we also have $w(A) = O(n \log n / \log \log n)$. This finishes the proof of the theorem. □

Corollary 3.3.

$$ex(n, \{Q_2, Q_1^{|}\}) = \Theta(n \log n / \log \log n)$$

$$ex(n, \{Q_3, \dot{Q}_1\}) = \Theta(n \log n / \log \log n)$$

Proof: The upper bound follows from Theorem 3.2. Indeed, by Lemma 2.3 we have $ex(n, \{Q_2, Q_1^{|}\}) \leq ex(n, \{Q_1, Q_1^{|}\}) + n$ and $ex(n, \{Q_3, \dot{Q}_1\}) \leq ex(n, \{Q_1, Q_1^{|}\}) + n$.

For the lower bound in the first statement notice that the n by n matrix E_n constructed in the proof of Theorem 3.2 one has $w(E_n) = \Theta(n \log n / \log \log n)$ and E_n avoids Q_2 and $Q_1^{|}$.

For the lower bound on the extremal function of $\{Q_3, \dot{Q}_1\}$ we use the equivalent collection $\{Q_3^{|}, \overline{Q}_1\}$ instead. We construct matrices avoiding both $Q_3^{|}$ and $\overline{Q}_1$ by modifying the matrices E_n constructed in the proof of Theorem 3.2. First assume that n is of the form $n = k^k$. We permute the rows of E_n by ordering them anti-lexicographically according to their indices. Recall that row and column indices are sequences. We keep the lexicographic order on the columns. We still have e_{ij} for the entry in row i and column j of the resulting matrix E'_n. Clearly, $w(E'_n) = w(E_n) = \Theta(n \log n / \log \log n)$.

Assume that $e_{ij} = e_{i'j} = 1$ for indices $i <^* i'$ and j. Let u be the position where i has 0 and let v be the position where i' has 0. Clearly, i and i' agree with j and with each other except for these two positions. Since $i <^* i'$ we must have $u > v$.

To show that E'_n avoids $Q_3^|$ (and also $Q_1^/$, $\dot{Q}_1$ and R) we need to show that for the row index $i <^* i'' \leq^* i'$ and the column indices $j < j'' \leq j'$ we cannot have $e_{ij'} = e_{i''j''} = 1$. Assume the contrary. The sequence j' differs from j only at position u. As $j < j'' \leq j'$ the sequence j'' must agree with both j and j' up to position u. We have $i <^* i'' \leq^* i'$, so i'' agrees with both of i and i' on positions larger than u. These positions do not contain 0, and j'' differs from i'' only where the latter has 0. Thus, j'' agrees with i'', and so also with i and j, on all positions larger than u. Now j and j'' can only differ at position u. From $j < j''$ we have $j_u < j''_u$. We get i'' by replacing a digit of j'' with 0. If this digit is before position u, then we have $i'' >^* i'$. If the 0 digit of i'' is at position u, then we get $i'' = i$. If the 0 digit of i'' is after position u, then we get $i'' <^* i$. All of these possibilities contradict to our assumptions. Thus, the matrix E'_n avoids the patterns $Q_3^|$, $Q_1^/$, $\dot{Q}_1$, and R.

Still assume $a_{ij} = a_{i'j} = 1$ for indices $i <^* i'$ and j and let u be the position of 0 in i, v the position of 0 in i'. We have $u > v$. Further assume that for column indices j' and $j'' < j$ we have $e_{ij'} = e_{i'j''} = 1$. As $j'' < j$ and j and j'' differ only at position v while j' agrees with i and j up to position u we have $j'' < j$. This shows E'_n avoids $\overline{Q}_1$.

For values of n not of the form k^k we construct E'_n just as E_n was constructed in the proof of Theorem 3.2. Let k be the largest integer with $n' = k^k \leq n$. Let $t = \lfloor n/n' \rfloor$. The n by n matrix E'_n contains t diagonally arranged submatrices equal to $E'_{n'}$ and 0 entries outside these submatrices. We have $w(E'_n) = w(E_n) = \Theta(n \log n / \log\log n)$ and E'_n avoids the patterns $Q_3^|$, $Q_1^/$, $\dot{Q}_1$, R, and $\overline{Q}_1$. □

For the lower bound construction in Theorem 3.5 we use a recent result of [17]. We state the result here. Note that the result is proved in [17] using an involved randomized procedure called *lexicographic thinning*.

On a *bipartite graph* we mean a triple $G = (A, B, E)$, where A and B are disjoint sets of vertices and $E \subseteq A \times B$ is the set of edges. A *proper edge m-coloring* is a function $\chi : E \to \{1, \ldots, m\}$ with $\chi(e) \neq \chi(e')$ for adjacent edges e and e'. A walk of length 4 in G starting in B and going through the edges e_1, e_2, e_3, and e_4 is a *slow walk* with respect to χ if $\chi(e_2) < \chi(e_3) < \chi(e_4)$ and $\chi(e_2) < \chi(e_1) \leq \chi(e_4)$.

Theorem 3.4. [17] *Let $G = (A, B, E)$ be a bipartite graph. If χ is a proper edge m-coloring, then there exists a subgraph $G' = (A, B, E')$ of G with $|E'| \geq \frac{\log m}{480m}|E|$ such that G' does not contain a slow walk with respect to χ.*

Theorem 3.5.

$$ex(n, \{Q_1, \dot{Q}_1\}) = \Theta(n \log\log n)$$

Proof: We start with the construction. Let $D_n = (d_{ij})$ be the n by n matrix from the proof of Theorem 2.5. Recall that $d_{ij} = 1$ if and only if $j - i = 2^k$ for some integer k. As we saw it in the proof of Theorem 2.5 the matrix D_n avoids the patterns Q_2, Q_1, $Q_1^/$, and R, and has weight $w(D_n) \geq n \log n - n$. Consider the bipartite graph $G = (A, B, E)$ with adjacency matrix D_n. Here

$A = \{r_i | 1 \leq i \leq n\}$ is the set of rows of D_n, $B = \{c_i | 1 \leq i \leq n\}$ is the set of columns of D_n. We have $(r_i, c_j) \in E$ if and only if $d_{ij} = 1$. We set $m = \lceil \log n \rceil$ and define an edge m-coloring $\chi : E \to \{1, \ldots, m\}$ on G by setting $\chi(r_i, c_j) = k+1$ if $j - i = 2^k$. This is a proper edge coloring. By Theorem 3.4 there is a subgraph $G' = (A, B, E')$ of G that has no slow path with respect to this coloring and has $|E'| \geq \frac{\log m}{480m}|E|$. We let F_n be the adjacency matrix of G', i.e., $f_n = (f_{ij})$ is an n by n 0-1 matrix with $f_{ij} = 1$ if and only if $(r_i, c_j) \in E'$.

We have $|E| = w(M) \geq n \log n - n$ and $w(F_n) = |E'| \geq \frac{\log m}{480m}|E| = \Omega(n \log\log n)$.

As F_n is obtained from D_n by deleting 1 entries, it avoids all the patterns avoided by D_n, namely Q_2, Q_1, $Q_1^{/}$, and R.

Consider any 2 by 3 submatrix of F_n consisting of the rows $i < i'$ and the columns $j < j' < j''$. This submatrix cannot represent $\dot{Q}_1$, as otherwise the walk $(c_{j'}, r_{i'}, c_j, r_i, c_{j''})$ is a slow walk in G', a contradiction. This shows that F_n avoids $\dot{Q}_1$. The lower bound of the theorem is established.

For the upper bound let $A = (a_{ij})$ be an n by n 0-1 matrix avoiding Q_1 and $\dot{Q}_1$. The row and column indices i and j are integers $1 \leq i, j \leq n$. We need to bound $w(A)$.

Consider the matrix $\dot{A}$ obtained by reflecting A centrally. Clearly, $w(\dot{A}) = w(A)$. As this geometric transformation maps Q_1 and $\dot{Q}_1$ to each other $\dot{A}$ also avoids these patterns. Clearly, any 1 entry of A is either a left entry of A or its image is a left entry of $\dot{A}$. Therefore, one of A and $\dot{A}$ is left-leaning. We assume without loss of generality that A is left-leaning.

Recall that the functions $f(i)$, $l(i)$ and $n(i,j)$ (as well as the notions of a left entry and a left-leaning matrix) were defined before Theorem 3.2. For any non-last $a_{ij} = 1$ entry of A we associate the point $P_{ij} = (x_{ij}, y_{ij})$ in the real plane with

$$x_{ij} = \frac{j - f(i)}{l(i) - f(i)} \quad \text{and} \quad y_{ij} = \frac{l(i) - n(i,j)}{l(i) - f(i)}.$$

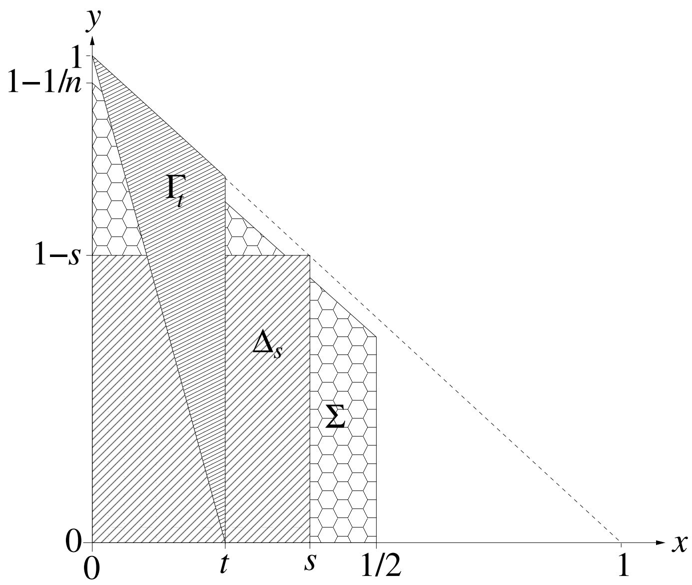

Figure 1.
The trapezoid Σ, the rectangle Δ_s and the triangle Γ_t

We have $x_{ij} \geq 0$, $y_{ij} \geq 0$ and $x_{ij} + y_{ij} \leq 1 - 1/n$. These inequalities define a triangle containing P_{ij}. For left entries a_{ij} of A we further have $x_{ij} \leq 1/2$ limiting the region for P_{ij} to a trapezoid Σ. See Figure 1.

For an arbitrary $0 < s < 1$ consider the rectangle Δ_s given by $0 \leq x \leq s$ and $0 \leq y < 1 - s$. If i is a row of A containing at least two 1 entries, then this row contains exactly one entry $a_{ij} = 1$ with $P_{ij} \in \Delta_s$. Indeed, $P_{ij} \in \Delta_s$ means that $j \leq f(i) + s(l(i) - f(i)) < n(i,j)$. This happens if and only if a_{ij} is the 1 entry with the largest column index not exceeding $f(i) + s(l(i) - f(i))$. So any rectangle Δ_s contains at most n points P_{ij}. (Here several of the points P_{ij} may coincide, we count them with multiplicities, i.e., we count the corresponding 1 entries of A.)

For $0 < t < 1$ consider the triangle Γ_t given by $x \leq t < x/(1-y)$ and $x + y \leq 1$. For any column index j there is at most a single row index i with $P_{ij} \in \Gamma_t$. Assume the contrary, that for $i < i'$ both P_{ij} and $P_{i'j}$ are in Γ_t. This implies $x_{ij} > 0$ and thus $f(i) < j$. Similarly, we have $f(i') < j$. Furthermore, $f(i) \leq f(i')$, as otherwise the submatrix of A consisting of rows $i < i'$ and columns $f(i') < f(i) < j$ would represent Q_1. Similarly, we have $l(i) \leq n(i',j)$, as otherwise the submatrix of A consisting of rows $i < i'$ and columns $j < n(i',j) < l(i)$ would represent $\dot{Q}_1$. We have

$$x_{ij} \leq t < \frac{x_{i'j}}{1 - y_{i'j}}$$

since the points P_{ij} and $P_{i'j}$ are in Γ_t. But we also have

$$\frac{x_{i'j}}{1-y_{i'j}} = \frac{j-f(i')}{n(i',j)-f(i')} = \frac{1}{1+\frac{n(i',j)-j}{j-f(i')}} \leq \frac{1}{1+\frac{l(i)-j}{j-f(i)}} = \frac{j-f(i)}{l(i)-f(i)} = x_{ij}.$$

The contradiction in the last two inequalities proves that for fixed t and j the triangle Γ_t contains P_{ij} for at most a single index i. So Γ_t contains at most n of the points P_{ij} (again, counting with multiplicities).

Let $z = \lfloor \log\log n \rfloor + 1$. The final part of the proof of the upper bound is geometric in nature: we show that $z+1$ of the rectangles Δ_t and z of the triangles Γ_t together cover Σ. Consider the pair Δ_t and Γ_t for some $0 < t < 1$. Together they cover the part of Σ bounded by $t^2 < x \leq t$. We have $x \leq 1/2$ for all of Σ. Thus, the collection $\{\Delta_t, \Gamma_t | t = 2^{-2^k}, k = 0, \ldots, z-1\}$ collectively cover all points of Σ with $x > 2^{-2^z}$. Adding Δ_t to the collection with $t = 2^{-2^z}$ they cover the entire trapezoid Σ. Indeed, the uncovered points would have $y \geq 1 - 2^{-2^z}$. But $2^{-2^z} < 1/n$ by the choice of z and $x \geq 0$, $x + y \leq 1 - 1/n$ implies that $y \leq 1 - 1/n$ for the points of Σ. As a consequence Σ contains at most $(2z+1)n$ of the points P_{ij}. For all the left entries of A (except for the 1 entries in rows containing a single 1 entry) we defined a point P_{ij}, so we have at most $(2z+2)n$ left entries in A. As A is left-leaning, we have $w(A) \leq (4z+4)n = O(n \log\log n)$. This finishes the proof of the theorem. □

Corollary 3.6.

$$ex(n, \{Q_2, \dot{Q}_1\}) = \Theta(n \log\log n)$$

$$ex(n, \{Q_2, \dot{Q}_1, \overrightarrow{Q}_1\}) = \Theta(n \log\log n)$$

$$ex(n, \{Q_3, Q_1^{|}\}) = \Theta(n \log\log n)$$

Proof: The upper bound follows from Theorem 3.5. Indeed, by Lemma 2.3 we have $ex(n, \{Q_2, \dot{Q}_1, \overrightarrow{Q}_1\}) \leq ex(n, \{Q_2, \dot{Q}_1\}) \leq ex(n, \{Q_1, \dot{Q}_1\}) + n$ and $ex(n, \{Q_3, Q_1^{|}\}) \leq ex(n, \{\overline{Q}_1, Q_1^{|}\}) + n = ex(n, \{Q_1, \dot{Q}_1\}) + n$.

For the lower bound in the first statement notice that the n by n matrix F_n constructed in the proof of Theorem 3.5 one has $w(F_n) = \Theta(n \log\log n)$ and F_n avoids Q_2 and $\dot{Q}_1$.

For the second statement we delete the last 1 entry in every row of F_n. We show that the resulting matrix avoids $\overrightarrow{Q}_1$. Assume the contrary, let the submatrix consisting of rows $i'' < i' < i$ and columns $j < j''$ represent $\overrightarrow{Q}_1$. The $f_{i''j} = 1$ entry of F_n is not the last 1 entry in its row, so we have $f_{i''j'} = 1$ for some $j' > j$. As F_n avoids $\dot{Q}_1$ we have $j' \leq j''$. As F_n also avoids R we have $j' < j''$. Thus the submatrix of F_n consisting of rows $i'' < i' < i$ and columns $j < j' < j''$ represents T (see Table 1 defining T). We constructed F_n by deleting entries of the matrix D_n. The contradiction comes from the fact that even D_n avoids T.

Assume for a contradiction that the submatrix of D_n consisting of the rows $i'' < i' < i$ and columns $j < j' < j''$ represents T. We have $d_{i'j''} = d_{ij''} = 1$, so both of $j'' - i'$ and $j'' - i$ are powers of 2 and since $j'' - i' > j'' - i$ we must have $j'' - i' \geq 2(j'' - i)$. Using this and $j > i$ from $d_{ij} = 1$ we have $i - i'' > i - i' \geq j'' - i > j'' - j$. Similarly, $j' - i'' \geq 2(j - i'')$ implies $j'' - j > j' - j \geq j - i'' > i - i''$. The contradiction shows that D_n avoids T as claimed. It also shows that the matrix obtained by deleting the last 1 entry in every row of F_n avoids $\overrightarrow{Q}_1$. As it also avoids Q_2 and $\dot{Q}_1$ and its weight is $\Theta(n \log \log n)$ the second statement of the theorem follows.

For the lower bound in the third statement we need a modified construction. In the proof of Theorem 3.5 we applied the thinning procedure (Theorem 3.4) to the matrix D_n from the proof of Theorem 2.5. Here we apply the same procedure to the matrix C_n from the proof of Theorem 2.2 instead. Let us assume that $n = 2^m$ for some integer $m \geq 1$. For other values of n simply pad our construction for the largest power of 2 below n by adding blank rows and columns.

Recall that the matrix $C_n = (c_{ij})$ has binary strings of length m as row and column indices. Rows are ordered lexicographically according to their indices, while columns are ordered anti-lexicographically. We have $c_{ij} = 1$ if and only if the binary strings i and j differ at a single position u with $i_u = 0$ and $j_u = 1$. The matrix C_n constructed avoids the patterns Q_3, $\overline{Q}_1$, $\overleftarrow{Q}_1$, and R and we have $w(C_n) = nm/2 = n \log n/2$.

Let $G = (A, B, E)$ be the bipartite graph with adjacency matrix C_n, i.e., let $A = \{r_i | i \in \{0,1\}^m\}$, $B = \{c_j | j \in \{0,1\}^m\}$, $E = \{(r_i, c_j) | c_{ij} = 1\}$. We define the edge m-coloring $\chi : E \to \{1, \ldots, m\}$ by setting $\chi(r_i, c_j) = u$, where u is the only position where the sequences i and j differ. This is a proper edge m-coloring. We apply Theorem 3.4 for G and obtain a subgraph $G' = (A, B, E')$ without a slow path and with $|E'| \geq \frac{\log m}{480m}|E|$. We let F'_n be the adjacency matrix of G', i.e., $F'_n = (f'_{ij})$ is an n by n 0-1 matrix with $f'_{ij} = 1$ if and only if $(r_i, c_j) \in E'$.

We have $|E| = w(C_n) = n \log n/2$ and $w(F'_n) = |E'| \geq \frac{\log m}{480m}|E| = \Omega(n \log \log n)$.

As F'_n is obtained from C_n by deleting 1 entries, it avoids all the patterns avoided by C_n, namely Q_3, $\overline{Q}_1$, $\overleftarrow{Q}_1$, and R.

Consider any 2 by 3 submatrix consisting of the rows $i < i'$ and the columns $j <^* j' <^* j''$ of F'_n. This submatrix cannot represent $Q_1^|$, as otherwise the walk $(c_{j'}, r_i, c_j, r_{i'}, c_{j''})$ must be a slow walk in G', a contradiction. This shows that F'_n avoids $Q_1^|$. This finishes the proof of the corollary. □

Our last result in this section does not establish the order of magnitude for a pair of excluded patterns but comes very close to doing so. Refer to Table 1 for the patterns considered. We have $ex(n, S_3) \geq ex(n, S_2) = \Theta(n\alpha(n))$ as S_2 is a submatrix of S_3. The upper bound $ex(n, S_3) \leq n2^{(\alpha(n))^{O(1)}}$ follows from a result of Klazar [10] on Davenport-Schinzel sequences. In fact, this bound holds for all patterns P satisfying that P has only a single 1 entry in every column. The stronger bound $ex(n, S_3) = O(n\alpha(n))$ follows from a related conjecture of Klazar [11, Problem 6]. If proved, this settles the order of magnitude of $ex(n, \{Q_2, Q_3^|\})$.

Theorem 3.7.

$$ex(n, S_2) - 2n \leq ex(n, \{Q_2, Q_3^{|}\}) \leq ex(n, S_3) + n$$

Proof: Instead of the pair $\{Q_2, Q_3^{|}\}$, we use the equivalent pair $\{\overline{Q}_2, \dot{Q}_3\}$ (refer to Table 2 for these patterns). Notice that S_2 can be obtained from $\dot{Q}_3$ by first adding a new first column with a single 1 entry in the first row, and then deleting another 1 entry. Similarly, S_2 can be obtained from $\overline{Q}_2$ by adding a new last column with a single 1 entry in the last row, and then deleting another 1 entry. Thus, by Lemma 2.3 we have $ex(n, \{Q_2, Q_3^{|}\}) = ex(n, \{\overline{Q}_2, \dot{Q}_3\}) \geq ex(n, \{\overline{Q}_2, S_3\}) - n \geq w(B) \geq ex(n, S_2) - 2n$. This proves the first inequality of the theorem.

Now let $A = (a_{ij})$ be a maximum weight n by n matrix avoiding $\dot{Q}_3$ and $\overline{Q}_2$. Let us obtain B by deleting the first 1 entry in every non-blank column of A. Clearly, $w(B) \geq w(A) - n$. Let us assume that the submatrix consisting of the rows $i_1 < i_2 < i_3$ and the columns $j_1 < j_2 < j_3 < j_4 < j_5$ represent S_3. We have $a_{i_3 j_3} = 1$ and this is not the first 1 entry in column j_3 of A. So we have $a_{ij_3} = 1$ for some $i < i_3$. In case $i > i_1$ the submatrix of A consisting of the rows $i_1 < i < i_3$ and $j_1 < j_2 < j_3$ represent $\overline{Q}_2$, a contradiction. If $i < i_2$ we get a contradiction by considering the submatrix of A consisting of the rows $i < i_2 < i_3$ and columns $j_3 < j_4 < j_5$, this submatrix represents $\dot{Q}_3$. We must either have $i > i_1$ or $i < i_2$, so the contradiction is unavoidable. This proves that B avoids S_3.

We have $ex(n, S_3) \geq w(B) \geq w(A) - n = ex(n, \{\overline{Q}_2, \dot{Q}_3\}) - n = ex(n\{Q_2, Q_3^{|}\}) - n$. This proves the second inequality in the theorem. □

4 All collections of small excluded patterns

First we characterize the sets of excluded patterns with bounded extremal functions. Let I_n stand for identity matrix of size n, and let J_n stand for the n by n matrix with 1 entries in the first row and 0 entries elsewhere.

Theorem 4.1. *Let $\mathcal{P}$ be a set of patterns. If there exists a size n_0 such that none of the four equivalents of J_{n_0} avoids all patterns in $\mathcal{P}$ and neither of the two equivalents of I_{n_0} avoids all patterns in $\mathcal{P}$, then $ex(n, \mathcal{P}) = O(1)$. Otherwise $ex(n, P) \geq n$.*

Proof: We prove the bound $ex(n, \mathcal{P}) \leq (n_0 - 1)^3$ in case the assumption of the theorem holds. Let $A = (a_{ij})$ be an n by n matrix of weight $w(A) > (n_0 - 1)^3$. We need to show that A does not avoid the patterns in $\mathcal{P}$. We show this by finding an equivalent of I_{n_0} or J_{n_0} contained in A.

We define two partial orders on the 1 entries of A. Among the entries $a_{ij} = 1$ and $a_{i'j'} = 1$ the former is larger in the first partial order if $i > i'$ and $j > j'$, while the first is larger in the second partial order if $i < i'$ and $j > j'$. A chain of n_0 entries in the first partial order determine a submatrix representing I_{n_0}. A chain of n_0 entries in the second partial order determine a submatrix representing $\overline{I}_{n_0}$, another matrix equivalent to I_{n_0}. Repeated applications of Dilworth's theorem show that if neither partial order contains a chain of length n_0, then there is a set of n_0 entries pairwise not comparable in either partial order. This set must come from either the same row or the same column. If $n \geq 2n_0 - 2$ this set can be extended to a submatrix representing one of the patterns equivalent with J_{n_0}. If $n < 2n_0 - 2$ the bound holds trivially.

Note that the above analysis is tight since for $n \geq 2(n_0 - 1)^2$ there exist n by n matrices with weight $(n_0 - 1)^3$ that avoid all equivalents of I_{n_0} and J_{n_0}.

The second statement of the theorem holds as $w(I_n) = w(J_n) = n$. □

Füredi and Hajnal tried to find $ex(n, P)$ for all patterns of weight not exceeding 4. We extend their research and try to find the order of magnitude for $ex(n, \mathcal{P})$ for all collection $\mathcal{P}$ of patterns of weight not exceeding 4. With the few sporadic cases studied in Section 3 we came surprisingly close to this goal.

Theorem 4.1 settles the extreme low end of the spectrum of extremal functions. It tells us when the extremal function is bounded and claims that it is at least linear otherwise. Thus we know the order of magnitude of $ex(n, \mathcal{P})$ for all collections $\mathcal{P}$ that contains a single pattern P with $ex(n, P) = O(n)$. The same holds if $\mathcal{P}$ contains a subcollection like $\{Q_1, \overline{Q}_1\}$ with a linear extremal function.

By Lemma 2.3/f we can disregard patterns with blank rows and columns. If we restrict attention to patterns of weight not exceeding 4 we have to consider only equivalents of Q_1, Q_2, Q_3, R, S_1 and S_2.

As $\alpha(n)$ is extremely slow growing, it is not unreasonable to consider a collection settled if we have an *almost linear* upper bound: one of the form $nf(\alpha(n))$ with a function $f(\alpha) \leq 2^{\alpha^{O(1)}}$. As we have mentioned a result of Klazar [10] implies an almost linear upper bound for $ex(n, P)$ if the pattern P contains a single 1 entry in every column.

We have almost linear bounds for collections containing equivalents of S_1 or S_2, or even the collection $\{Q_2, Q_3^{|}\}$. We are left with equivalents of Q_1, Q_2, Q_3, and R only. By exhaustive search of the remaining finite (but huge) number of excluded collections one finds that the results in this paper establish the order of magnitude for most of them. More precisely, we have the following:

Corollary 4.2. *For each collection $\mathcal{P}$ of patterns of weight not exceeding* 4 *our results imply one of the following:*

(i) *Either $ex(n, \mathcal{P})$ is (constant, linear or) almost linear or*

(ii) $ex(n, \mathcal{P}) = \Theta(n \log \log n)$ *or*

(iii) $ex(n, \mathcal{P}) = \Theta(n \log n / log \log n)$ *or*

(iv) $ex(n, \mathcal{P}) = \Theta(n \log n)$ *or*

(v) $ex(n, \mathcal{P}) = n^{3/2} + o(n^{3/2})$ *or*

(vi) $ex(n, \mathcal{P}) = n^2$ *or*

(vii) *$ex(n, \mathcal{P}) = \Theta(ex(n, \mathcal{Q}_i))$ for one of the eleven exceptional collections $\mathcal{Q}_i$ in Table 3.*

Note that $ex(n, \mathcal{P}) = n^2$ if and only if $\mathcal{P} = \emptyset$, while $ex(n, \mathcal{P}) = n^{3/2} + o(n^{3/2})$ if and only if $\mathcal{P}$ reduces to $\{R\}$.

The proof of this corollary is an exhaustive search and careful reduction of the cases based on Lemma 2.3. For example we have $ex(n, \mathcal{P}) = \Theta(ex(n, \mathcal{P} \cup \{R\}))$ for any non-empty collection $\mathcal{P}$ of patterns of weight at most 4. Thus R can be disregarded except for the singleton case in (v). We leave the details to the interested readers. We would have twelve exceptional families, but one of them is taken care of in [8]:

Theorem 4.3. Füredi-Hajnal [8]

$$ex(n, \{Q_2, \dot{Q}_2\}) = O(n)$$

In Table 3 we list the eleven exceptional collections with the best bounds known for them.

$$\begin{array}{rcccl}
n & \leq & ex(n, \{Q_2, \dot{Q}_3\}) & = & O(n \log n / \log\log n) \\
n & \leq & ex(n, \{Q_3, \dot{Q}_3\}) & = & O(n \log\log n) \\
n & \leq & ex(n, \{Q_2, \dot{Q}_1, Q_1^{\backslash}\}) & = & O(n \log\log n) \\
n \log\log n & = & O(ex(n, \{Q_2, Q_1^{|}, \overrightarrow{Q}_1\})) & = & O(n \log n / \log\log n) \\
n & \leq & ex(n, \{Q_3, Q_1^{|}, \overrightarrow{Q}_1\}) & = & O(n \log\log n) \\
n & \leq & ex(n, \{Q_3, Q_1^{|}, Q_1^{\backslash}\}) & = & O(n \log\log n) \\
n & \leq & ex(n, \{Q_3, \dot{Q}_1, Q_1^{\backslash}\}) & = & O(n \log n / \log\log n) \\
n \log n / \log\log n & = & O(ex(n, \{Q_1, \overrightarrow{Q}_1\})) & = & O(n \log n) \\
n & \leq & ex(n, \{Q_1, \dot{Q}_1, Q_1^{/}, Q_1^{\backslash}\}) & = & O(n \log\log n) \\
n & \leq & ex(n, \{Q_1, \dot{Q}_1, \overrightarrow{Q}_1, \overleftarrow{Q}_1\}) & = & O(n \log\log n) \\
n \log\log n & = & O(ex(n, \{Q_1, Q_1^{|}, Q_1^{/}, \overrightarrow{Q}_1\})) & = & O(n \log n / \log\log n)
\end{array}$$

Table 3.
List of the 11 exceptional collections of patterns of weight at most 4

All bounds stated in Table 3 follow from our results via reductions using Lemma 2.3 except for the lower bounds in the fourth and last lines. The latter of these lower bounds follows from the former. The former one follows from the pair of the result stated in Theorem 3.4, also from [17]. We state the relevant result from [17] and the give the sketch of the reduction below. Unfortunately, we do not have a good upper bound in this case.

Consider a bipartite graph $G = (A, B, E)$, with proper edge coloring given by $\chi : E \to \{1, \ldots, m\}$. We call a walk of length 4 in G through the edges e_1, e_2, e_3, and e_4 a *fast walk* with respect to χ if it satisfies $\chi(e_2) < \chi(e_3) < \chi(e_4) \leq \chi(e_1)$.

Theorem 4.4. [17] *Let $G = (A, B, E)$ be a bipartite graph. If χ is a proper edge m-coloring, then there exists a subgraph of $G' = (A, B, E')$ of G with $|E'| \geq \frac{\log m}{480m}|E|$ such that G' contains no fast walk with respect to χ.*

Theorem 4.5.

$$ex(n, \{Q_2, Q_1^{|}, \vec{Q}_1\}) = \Omega(n \log \log n)$$

Proof: This proof is the same as the construction in the proof of Theorem 3.5, except we use Theorem 4.4 in place of Theorem 3.4 for thinning. Let $D_n = (d_{ij})$ be the n by n matrix from the proof of Theorem 2.5. Recall that $d_{ij} = 1$ if and only if $j - i = 2^k$ for some integer k. Consider the bipartite graph $G = (A, B, E)$ with adjacency matrix D_n. We set $m = \lceil \log n \rceil$ and define an edge m-coloring $\chi : E \to \{1, \ldots, m\}$ on G by setting $\chi(e) = \log(j - i) + 1$ if $d_{ij} = 1$ is the entry in D_n corresponding to the edge e. We applied Theorem 3.4 to this edge colored graph in the construction for Theorem 3.5. Now we apply Theorem 4.4 and find a subgraph $G'' = (A, B, E'')$ of G without any fast walk, but with $|E''| = \Omega(n \log \log n)$ edges. The adjacency matrix of G'' has weight $|E''|$. It avoids Q_2 since it is contained in D_n. It avoids $Q_1^{|}$ and $\vec{Q}_1$ because G'' avoids fast walks. □

5 Concluding remarks

Bienstock and Győri [2] and Füredi [7] were first to consider excluded submatrix problems. Their motivation to estimate $ex(n, Q_1)$ came from discrete geometry: they proved an $O(ex(n, Q_1)) = O(n \log n)$ bound on the number of unit length diagonals of a convex n-gon. Connections to problems in discrete geometry remains a main source of motivation for this type of problems. See [15] for a recent paper on these connections. Here we sketch a single application only.

Efrat and Sharir [5] consider *critical placements* of a convex n-gon Δ in a hippodrome H. The hippodrome H is the region of the plane consisting of points within radius r of an interval I. A critical placement of Δ is a placement inside the hippodrome H with three of the vertices on the boundary. We assume the radius r is generic, i.e., no vertex of Δ is at distance $2r$ from the line of an edge of Δ and no three vertices of Δ are on a common circle of radius r. Efrat and Sharir prove that the number of critical placements is $O(ex(n, Q_1))$ and thus, by the results of [2, 7], it is $O(n \log n)$. Pach and Tardos [15] prove that the number of critical placements is $O(ex(n, L_1))$. Using our Theorem 2.1 they give a linear bound on the number of critical placements.

Most of the small excluded patterns considered here have small extremal functions and many of them have a linear extremal function. A problem raised by Füredi and Hajnal in [8] is to characterize all patterns with $ex(n, P) = O(n)$. A recent related result of A Marcus and G. Tardos [14] is a step in this direction. Proving a conjecture of Füredi and Hajnal they establish that the extremal function of *permutation matrices* is linear.

Another class of patterns with linear extremal function are bitonic patterns. Let $f : \{1, \ldots, n\} \to \{1, \ldots, m\}$ be function consisting of an increasing interval followed by a decreasing interval. We do not assume strict monotonicity. We call the m by n pattern $P = (p_{ij})$ *bitonic* if $p_{f(j)j} = 1$ and all other entries are 0. By the results of M. Klazar and P. Valtr [13] on generalized Davenport-Schinzel sequences the extremal functions of bitonic patterns are linear.

As noted in [14] all pattern P for which $ex(n, P)$ is known to be super-linear contain such a pattern of weight 4. It would be desirable to find other *minimal* (with respect to containment) non-linear patterns, or (even more interesting) to prove they do not exist.

If one looks for minimal non-linear patterns among patterns of weight 5 there are only a few choices. The following lemma rules out one them. Refer to Table 1 defining L_4.

Lemma 5.1.

$$ex(n, L_4) = O(n)$$

This is a generalization of Theorem 2.1, and can be proved very similarly. We omit the proof here. The patterns L_5 and L_6 are prime candidates for minimal non-linear patterns of weight 5. It would be desirable to find their extremal function.

M. Klazar [12] also asked for the characterization of the patterns with almost linear extremal functions. His definition of almost linear is somewhat more relaxed than the one we used in Section 4.

An even higher threshold is between functions like $n \log n$ and $n^{1.1}$. From standard extremal graph theory we know that if the pattern P is the adjacency matrix of a bipartite graph containing a cycle, then $ex(n, P) = \Omega(n^c)$ for some $c > 1$. Füredi and Hajnal [8] conjecture a strong converse: if P is the adjacency matrix of a cycle free bipartite graph, then we have $ex(n, P) = O(n \log n)$. Even proving $ex(n, P) = O(n^{1+\epsilon})$ for every cycle free P and every $\epsilon > 0$ would be a breakthrough. See related results in [15].

References

[1] R. Baker, G. Harman, J. Pintz, The difference between consecutive primes II, *Proc. London Math. Soc.* **83** (3) (2001), 532–562.

[2] D. Bienstock, E. Győri, An extremal problem on sparse 0-1 matrices, *SIAM J. Discrete Math.* **4** (1991), 17–27.

[3] P. Brass, Gy. Károlyi, P. Valtr, A Turán-type extremal theory of convex geometric graphs, in: *Discrete and Computational Geometry—The Goodman-Pollack Festscrift* (B. Aronov et al. eds.), Springer 2003, 277–302.

[4] J. Czipszer, P. Erdős, A. Hajnal, Some extremal problems on infinite graphs, *Magyar Tud. Akad. Mat. Kutató Int. Közl.* **7** (1962), 441–457.

[5] A. Efrat and M. Sharir: A near-linear algorithm for the planar segment center problem, *Discrete Comput. Geom.* **16** (1996), 239–257.

[6] R. Faudree, M. Simonovits, On a class of degenerate extremal graph problems II, in preparation.

[7] Z. Füredi, The maximum number of unit distances in a convex n-gon, *J. Combinatorial Theory Ser. A* **55** (1990), 316–320.

[8] Z. Füredi, P. Hajnal, Davenport-Schinzel theory of matrices, *Discrete Mathematics* **103** (1992), 233–251.

[9] M. Klazar, The Füredi-Hajnal conjecture implies the Stanley-Wilf conjecture, in: *Formal Power Series and Algebraic Combinatorics* (D. Krob, A. A. Mikhalev, and A. V. Mikhalev eds.), Springer, Berlin, 2000, pp. 250–255.

[10] M. Klazar, A general upper bound in extremal theory of sequences, *Comment. Math. Univ. Carolinae* **33** (1992), 737–747.

[11] M. Klazar, Generalized Davenport-Schinzel sequences: results, problems, and applications, *Integers* **2** (2002), A11, 39 pages.

[12] M. Klazar, Extremal problems for ordered (hyper)graphs: applications of Davenport-Schinzel sequences, *European J. of Combinatorics* **25** (2004), 125–140.

[13] M. Klazar and P. Valtr, Generalized Davenport-Schinzel sequences, *Combinatorica* **14** (1994), 463–476.

[14] A. Marcus, G. Tardos, Excluded permutation matrices and the Stanley-Wilf conjecture, *J. Combinatorial Theory Ser. A* **107** (2004), 153–160.

[15] J. Pach, G. Tardos, Forbidden patterns and unit distances, SoCG 2005, submitted, available from `http://www.renyi.hu/~tardos/papers.html`.

[16] I. Ruzsa, E. Szemerédi, Triple systems with no six points carrying three triangles, *Combinatorics (Proc. Fifth Hungarian Colloq., Keszthely, 1976)*, Vol. II, Colloq. Math. Soc. János Bolyai **18**, North-Holland, Amsterdam, New York, 1978, pp. 939–945.

[17] G. Tardos, Construction of locally plane graphs, submitted to *SIAM J. on Discrete Mathematics*, 2004, available from `http://www.renyi.hu/~tardos/papers.html`.