Forbidden submatrices in 0-1 matrices

THESIS

WRITTEN BY: Balázs Keszegh

SUPERVISOR: Gábor Tardos

Eötvös Loránd University
Faculty of Sciences
2005

Contents

1 Introduction

1.1 Definitions

A 0-1 matrix is a matrix with just 1's and 0's (blanks) at its entries. *Deleting* a 1 entry means replacing it with a 0. We usually call the 'small' matrix we forbid a *pattern*. 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. If a submatrix of A represents P, then we say A *contains* P, otherwise it *avoids* P. The *weight* $w(P)$ of a pattern P is the number of 1 entries in it.

For a pattern P with positive weight and a positive integer n, $ex(n, P)$ is the maximum weight of an n by n 0-1 matrix avoiding P. This can be considered as a function on positive integers, and is called the *extremal function* of P. Sometimes we refer to a pattern with linear extremal function as a *linear* pattern, and similarly we can say that a pattern is *non-linear*.

1.2 Relations with other problems

Regard an n by n 0-1 matrix M as an adjacency matrix of a bipartite graph $G(n)$ between n red and n blue vertices, and there is a linear ordering on the red vertices and on the blue vertices as well (note that there is no order relation between a red and a blue vertex). Here, a submatrix corresponds to a full colored subgraph of $G(n)$. Avoiding a pattern means avoiding a certain colored, ordered subgraph. Similarly $ex(n, P)$ is the maximum number of edges in an ordered bipartite graph on $n + n$ vertices avoiding the colored, ordered subgraph corresponding to the pattern P. This problem is an ordered variant of the classical Turán extremal graph theory for bipartite graphs. P. Brass, Gy. Károlyi and P. Valtr [3] study a very similar problem, where a cyclic order is given on the vertices.

For the case of permutation matrix patterns P, the Füredi-Hajnal conjecture [6], which states that $ex(n, P) = O(n)$, was proved by A. Marcus and G. Tardos [13]. Without going into details, we mention that the Stanley-Wilf conjecture, as M. Klazar [10] showed, follows from this conjecture. We will discuss this topic more deeply in section 2.2.

Generalized Davenport-Schinzel sequences (DS sequences) are finite sequences over an alphabet with n symbols avoiding a certain subsequence and with no close repetition (for details see section 2.3). In the classic version the subsequence is an alternating sequence with 2 symbols. There are many results in this topic (for a survey see [9]), and we will see that sometimes they have implications for 0-1 matrices too. We will discuss this topic mainly in section 3.2.

1.3 Results

This thesis can be regarded as a survey on 0-1 matrices with results mainly from [6],[8],[13] and [15]. We present some new results as well, which are detailed below.

Our main aim is to determine the order of magnitude of $ex(n, P)$ for certain patterns. The patterns with weight 4 are systematically considered by Füredi and Hajnal [6]. They found the extremal function up to a constant factor for almost all such patterns. The remaining patterns with weight 4 are examined by Tardos [15].

There are two type of proofs for upper bounds. For some patterns we have straightforward proofs and for the rest of the patterns we have reductions. If we can obtain a pattern with a certain transformation from another, then we can compute bounds for the new pattern too. In section 3.1 and 3.2 we present two new such transformations.

For lower bound proofs, there are several constructions. One of these comes from the construction for DS-sequences and the others are mostly the variations of a diagonal construction.

Patterns with linear extremal function are particularly interesting. It is clear that if a pattern P contains a pattern Q, then $ex(n, P) \geq ex(n, Q)$, because if a matrix avoids Q, then it avoids P too. So if a pattern P has linear extremal function, then every Q, which is contained by P, is linear as well (with the exception of the trivial pattern all pattern has at least linear extremal function). So we can define the set of minimal (for containment) non-linear patterns. So far no minimal non-linear pattern with weight bigger then 4 was found. In chapter 4 we present a weight 5 pattern with that property. Moreover, we present infinite number of candidates for being minimal non-linear patterns too. We can't prove that these non-linear patterns are minimal, but in section 4.1 we show some results suggesting that statement.

1.4 Acknowledgements

I would like to thank my supervisor Gábor Tardos for his innumerable advices and remarks about the Thesis.

2 Examples

2.1 Simple bounds

The 1 by 1 pattern with a single 1 entry is called *trivial.* If a matrix contains a 1 entry, then it contains the trivial pattern, so $ex(n, P) = 0$ for the trivial pattern P. Otherwise, if the pattern is *non-trivial*, then $ex(n, P) \geq n$. Indeed, if the pattern contains a 1 entry outside its first row, then an n by n matrix with 1 entries only in its first row avoids it. For other cases use symmetry. Let us begin with a trivial proposition already mentioned in the introduction:

Proposition 2.1. *If a pattern P contains a pattern Q, then $ex(n, Q) \leq ex(n, P)$.*

Proof. If a matrix avoids Q, then avoids P too and so the proposition follows. $\square$

Rotating or *reflecting* a pattern P into pattern P' does not affect its extremal function, as if a matrix M avoids P, then the matrix M' obtained by the same transformation from M, avoids P', so $ex(n, P) \leq ex(n, P')$ and considering also the reverse of this transformation we get the equality. These patterns are called the *equivalents* of P.

Removing all blank rows and columns is called *reducing* the pattern.

Theorem 2.2. *If P reduces to P', then $ex(n, P) \leq ex(n, P') = O(ex(n, P) + n)$.*

Proof. The first inequality is true as P contains P'. For the second inequality let k be large enough so that P has no k consecutive blank rows or columns. Take a maximal weight n by n 0-1 matrix M' avoiding P'. Delete the 1 entries in the first and last k rows and columns of M'. We obtain the matrix M and lose at most $4kn$ in the weight. For $0 \leq a, b < k$ let M_{ab} be the matrix obtained from M 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 $w(M_{ab})$ sum to $w(M)$. It is easy to see that here every matrix M_{ab} avoids P. We have $ex(n, P') = w(M') \leq w(M) + 4kn \leq k^2 ex(n, P) + 4kn = O(ex(n, P) + n)$, as claimed. $\square$

As one could notice, these transformations can not increase the weight of the pattern. In the following we present some transformations that increase the weight but does not increase the extremal function too much. The first appeared in [6]:

Theorem 2.3. *If P' can be obtained from P by attaching an extra column or row to the boundary of P and placing a single 1 entry in the new column or row next to an existing one in P, then $ex(n, P) \leq ex(n, P') \leq ex(n, P) + n$.*

Proof. The first inequality is trivial. For the second we can assume by symmetry that the extra row or column is a column at the left side of the matrix. Let M' be a matrix avoiding P' with maximal weight. Deleting the first 1 entry in every row we obtain the matrix M. Here $w(M) \geq w(M') - n$. The matrix M avoids P, as a representing submatrix in M together with the column of the appropriate deleted 1 entry would represent P' in M'. So $ex(n, P) \geq w(M) \geq w(M') - n = ex(n, P') - n$. $\square$

The natural way to apply this theorem is that an upper bound on P gives an upper bound on P'. In many cases we even delete the 1 entry next to the new one, which does not increase the extremal function. For example Theorem 2.7 stating that $ex(n, L_1) = O(n)$, implies that $ex(n, L_2) = O(n)$.

With a slight generalization of the result of Tardos in [15] we get a similar theorem for putting 1 entries in between already exiting ones:

Theorem 2.4. *If P' is obtained from the pattern P by adding k extra columns between two columns of P, each containing a single 1 entry and the newly introduced 1 entries has 1 next to them on both sides, then $ex(n, P) \leq ex(n, P') \leq (k+1)ex(n, P)$.*

Proof. The first inequality is trivial. For the second let M' be a maximum weight matrix avoiding P'. In each row number the 1 entries from left to right and delete the 1's whose number is not equal to 1 modulo $(k+1)$. Clearly, the obtained M matrix has weight at least $w(M')/(k+1)$. Moreover M avoids P, because a representation of P with the columns of the appropriate deleted entries would be a representation of P' in M'. We got $ex(n, P') = w(M') \leq (k+1)w(M) \leq (k+1)ex(n, P)$. □

The following theorem was proved in [14].

Theorem 2.5. *For any pattern P super-additivity holds for any positive integers n and m, that is*

$$ex(n+m, P) \geq ex(n, P) + ex(m, P).$$

If the pattern has no blank entries, then the bipartite graph analogon is no further order-sensitive, so we can transform the results of (non-ordered) extremal graph theory to 0-1 matrices in this case. For example the weight 4 2 by 2 matrix R corresponds to the four-cycle, and so by classical results in extremal graph theory, we get:

Theorem 2.6. $f(n, R) = \Theta(n^{3/2})$.

2.2 Linear bounds

In this section we present some direct proofs for upper bounds. The first we mention was proved by Tardos in [15]:

Theorem 2.7. $ex(n, L_1) \leq 5n$

Proof. Let $A = (a_{ij})$ be an n by n 0-1 matrix avoiding the pattern L_1. For a column j containing at least one 1 entry, let $l(j)$ be the index of the row containing the last 1 entry in column j. For such a column j' let be j the largest column index with $j < j'$ and $l(j) \geq l(j')$. We say that column j' *finds* the entry a_{ij}, where i is the largest row index with $i < l(j')$ and $a_{ij} = 1$. If there is no such index j or i then column j' does not find any entry, so each column finds at most one 1 entry. Now delete the last and the second last 1 entries in each row, the last 1 entry in each column and the 1 entries whose right neighbour (i.e. the next 1 entry in the same row with larger column index) is the last 1 entry in its column. Clearly, we deleted at most $4n$ 1 entries. We claim that the remaining 1 entries are found by some column and so the theorem follows.

To prove the claim fix a remaining entry $a_{ij} = 1$. There exists entries $a_{ij_1} = a_{ij_2} = 1$ with $j < j_1 < j_2$ (the deleted last two 1 entries in row i). Further, we can choose a_{ij_1} to be the first 1 entry in row i after a_{ij}. If a_{ij_1} would be the last 1 entry in its column then a_{ij} would be deleted earlier, so $l(j_1) > i$. Besides, there is a 1 entry $a_{i_1 j}$ in column j with the smallest row index $i_1 > i$ otherwise the entry a_{ij} would be the last in its column.

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

Now we turn to the case of permutation matrices and present the proof of the Füredi-Hajnal conjecture, which appeared in [13] (when proving the theorem we used also the version in [16], which uses a different terminology):

Theorem 2.8. *For all permutation matrices P we have $ex(n, P) = O(n)$.*

Proof. Let P be a k by k permutation pattern. Let A be a maximum weight n by n matrix avoiding P. Let a be a positive integer constant, its exact value determined later. First assume that n divides a. Partition the rows and the columns into groups of size a (each group contains a consecutive rows or columns). Let B_{ij} be an n/a by n/a 0-1 matrix with $b_{ij} = 1$ if and only if the *block* S_{ij} of A defined by the ith group of rows and the jth group of columns has at least one 1 entry.

We claim that B avoids P too. Indeed, assume not and take the k 1 entries of a submatrix of B representing P. For each 1 entry in B there is at least one 1 entry in A in the corresponding block S_{ij}. Take one for each of the k 1 entries, these represent P in A, a contradiction.

We say that a block is *wide* if it contains 1 entries in at least k different columns. Similarly, a block is *tall* if it contains 1 entries in at least k different rows. We claim that for a any i there are at most $(k-1)\binom{a}{k}$ tall blocks in the ith group of rows. Indeed, otherwise there would be k different blocks with 1 entries in the same set of k rows by the pigeonhole principle (there are altogether $\binom{a}{k}$ possibilities). Let these rows be $c_1 < c_2 < \ldots < c_k$ and let these blocks be $S_{id_1}, \ldots, S_{id_k}$ with $1 \leq d_1 < d_2 < \ldots < d_k \leq n/a$. For each 1 entry p_{rs} in P choosing a 1 entry in row c_r of S_{id_s} we obtain a submatrix of A representing P which is a contradiction. Similarly, there are at most $(k-1)\binom{a}{k}$ wide blocks in each group of columns. Note that the weight of a block is at most $(k-1)^2$ unless it is tall or wide.

Counting the 1 entries separately in the tall and the wide blocks and for the remaining blocks using that B avoids P we get:

$$ex(n, P) = w(A) \leq 2 \cdot n/a \cdot a^2 \cdot (k-1)\binom{a}{k} + ex(n/a, P) \cdot (k-1)^2$$

and so

$$ex(n, P) \leq (k-1)^2 \cdot ex(n/a, P) + 2ak\binom{a}{k} \cdot n.$$

Now with the above recursion we are able to prove that

$$ex(n, P) \le \frac{4a^2k\binom{a}{k}}{a-(k-1)^2} \cdot n$$

for any a greater then $(k-1)^2$ and for any n (note that a is not necessarily divisible by n).

We proceed by induction on n. The cases $n \le a$ are trivial as the left side is at most n^2 and the right is at least an. Now assume the hypothesis to be true for all $n < n_0$ and consider the case $n = n_0$. We let n' be the largest integer less than or equal to n which is divisible by a. Then by the previous claim we have:

$$ex(n, P) \le ex(n', P) + 2an \le (k-1)^2 ex(n'/a, P) + 2ak\tbinom{a}{k}n' + 2an \le$$
$$\tfrac{(k-1)^2\left[4a^2k\binom{a}{k}\frac{n'}{a}\right]}{a-(k-1)^2} + 2ak\tbinom{a}{k}n + 2an \le \tfrac{(a-(a-(k-1)^2))\left[4ak\binom{a}{k}n\right]}{a-(k-1)^2} + 2ak\tbinom{a}{k}n + 2an \le$$
$$\tfrac{4a^2k\binom{a}{k}}{a-(k-1)^2}n - 4ak\tbinom{a}{k}n + 2ak\tbinom{a}{k}n + 2an \le \tfrac{4a^2k\binom{a}{k}}{a-(k-1)^2}n.$$

Now take any a greater then $(k-1)^2$ and the theorem follows. In [13] Marcus and Tardos took $a = k^2$. $\square$

2.3 Quasi-linear bounds

In this section we present the results of [6] about some patterns with extremal function $\Theta(n\alpha(n))$, where $\alpha(n)$ denotes the extremely slowly growing inverse Ackermann function.

Following the notation of [8], $S(u)$ is the set of symbols which appear in sequence u. We denote the *length* of a sequence u by $|u|$ and the size of $S(u)$ by $\|u\|$. Two sequences $u = a_1a_2\dots a_n$ and $v = b_1b_2\dots b_n$ of the same length are *equivalent* if there exists a bijection $f : S(u) \to S(v)$ such that $f(a_i) = b_i$ for all $i = 1, 2, \dots, n$. We say that u *contains* v, if v is equivalent to a subsequence of u, otherwise it *avoids* u. We call the occurrences of a symbol $a \in S(u)$ in the sequence u *a-occurrences.*

Similarly to 0-1 matrices, $ex(n, u)$ is the maximum length of a string on n symbols with no *close* repetition (any at most $\|u\|$ consecutive elements of the string must be different from each other) and avoiding the string u. Note that if $\|u\| = 2$ then immediate repetitions are forbidden.

We will need the following classic result about DS-sequences of Hart and Sharir in [7] that resolved a long standing open problem:

Theorem 2.9. *For the string $u = ababa$ we have $ex(n, u) = \Theta(n\alpha(n))$.*

From this theorem we can derive the following upper bound:

Theorem 2.10. *For the extremal function of the pattern S_1 we have $ex(n, S_1) = O(n\alpha(n))$.*

Proof. Let M be a maximum weight n by n matrix avoiding S_1. Delete the first and last 1 entries in each row. Number the rows from top to bottom with the numbers $1, \dots, n$. For every i change every 1 entry in the ith row to i. Change 0 entries to blanks. Let u denote the string that we obtain if we read the columns of M from left to right and from top to bottom in each column. If there are any two neighbouring positions with the same letter, then we delete the earlier letter from the pair. We repeat this step until there are no repetitions. Each repetition corresponds to n entry in M (the earlier (or left) entry of the pair and $n-1$ consequent blank entries after that), and in each step we deleted at most one letter. Therefore we deleted at most $n^2/n = n$ letters. In this way there is no immediate repetition in the string u and $w(M) - 3n \leq |u| \leq w(M)$ and $\|u\| \leq n$.

Now we can apply Theorem 2.9 and so it is enough to prove that u avoids $ababa$. Suppose u contains $ababa$. Then it must contain a subsequence $abab$ with $a > b$ as well. Examine the submatrix in M corresponding to this subsequence. Clearly, this must be a representation of one of the patterns R, S_1, Q_1 or Q_1'. If it is the pattern S_1, then the theorem follows. In case it is R regard the deleted first 1 entry in the bth row and the deleted last entry in the ath row. These together with the appropriate 1 entries from the representation of R represent S_1 in M. Cases Q_1 and Q_1' are similar. □

Following the idea of Tardos we can apply again Theorem 2.9 quite simply to obtain an upper bound for $ex(n, S_1)$.

Theorem 2.11. *For the extremal function of the pattern S_1 we have $ex(n, S_1) = \Omega(n\alpha(n))$.*

Proof. For any positive integer n by Theorem 2.9 there is a string u, which has no immediate repetition, avoids $ababa$. Further, $\|u\| \leq n$ and $|u| = \Omega(n\alpha(n))$. We can assume that $\|u\| = n$ holds too. Change the letters of u to numbers from 1 to n in order of their first appearance (the leftmost letter of u will be 1, the second leftmost which is not 1 will be 2 etc.). Partition u into $3n$ blocks of size $\left\lfloor \frac{|u|}{3n} \right\rfloor$ (the last block is possibly bigger). For each block if it contains the first or last appearance of any letter, then we delete this block. We deleted at most $2n$ blocks this way. If in a remaining block there are two positions with the same letter b, then there is one letter a between them (as there is no immediate repetition in u). Together with the first and last appearance of a (which are in other blocks) we obtain $ababa$, contradiction. We take n remaining blocks and build the matrix M. In the jth column the ith entry is 1 if and only if the jth remaining block of u contains the letter i. This matrix avoids the pattern S_1 as otherwise a representation of S_1 (with entries in the ath and bth row, where $a < b$) would represent $ababa$ in u together with the first a entry of u ($a < b$ and the construction of the matrix implies that this entry precedes the two b entries). As there are no two positions with the same letter in any block, we can deduce for the weight of the matrix M that $w(M) \geq n \cdot \left\lfloor \frac{|u|}{3n} \right\rfloor \geq n \cdot \frac{|u|-3n}{3n} = \frac{|u|}{3} - n = \Omega(n\alpha(n))$. □

By Theorem 2.3 for the pattern S_2 the same upper bound follows from Theorem 2.10: $ex(n, S_2) = O(n\alpha(n))$. Now we prove a lower bound for this pattern too, but in this case we cannot prove the lower bound so easily. In [6] there is a construction of n by n matrices with $\Theta(n\alpha(n))$ 1 entries, avoiding this pattern. This construction is derived from the construction of sequences proving the lower bound in Theorem 2.9. We present a modified version of this construction. The construction for matrices uses double recursion.

The matrices we are constructing have two parameters, s and t. We refer to them as $M(s,t)$. The recursion which gives the construction is for s and inside that is another recursion for t. Note that these are not square matrices.

Let us begin with some properties of the matrix $M(s,t)$:

(a) Its size is $tC(s,t) \times (t+2)C(s,t) - 2$, where the following recursion defines $C(s,t)$:

$$\begin{aligned} C(1,t) &= 1 && (t \geq 1) \\ C(s,1) &= 2C(s-1,2) && (s > 1) \\ C(s,t) &= C(s,t-1)C(s-1,C(s,t-1)) && (s,t > 1). \end{aligned}$$

(b) There are exactly s 1 entries in each row of $M(s,t)$.

(c) The rows are divided into $C(s,t)$ horizontal blocks, where each block contains t rows. Let H_i be the ith block.

(d) Inside H_i the first 1 entry in each row is in the same column. Let c_i be the index of this column and call these first 1 entries the *leading* 1's.

(e) For the columns containing the leading 1's the following is true: $1 = c_1 < c_2 < \ldots < c_{C(s,t)}$. These columns divide the matrix into vertical blocks. Let V_i be the ith block containing the columns from the c_ith to the $(c_{i+1}-1)$th ($V_{C(s,t)}$ contains the columns from $c_{C(s,t)}$ to the last column).

Finally we can define the matrices. $M(1,t)$ for $t \geq 1$ is an t by t matrix with t 1 entries in its first column and no more 1 entries. For $s > 1$ $M(s,1)$ is obtained from $M(s-1,2)$ by repeating the following procedure for every horizontal block H_i of $M(s-1,2)$. Put two new columns before the leading column c_i. Note that the ith horizontal block has two rows. Put a 2 by 2 identity matrix into the intersection of these two rows and the two new columns. In this way we got $C(s,1) = 2C(s-1,2)$ horizontal blocks, each containing one row with s 1 entries in each row.

In the general case we build $M(s,t)$ using the matrices $S = M(s,t-1)$ and $B = M(s-1,C(s,t-1))$. First regard B, which is the 'big' matrix. It has $C(s-1,C(s,t-1))$ vertical blocks, let v_i be the number of columns in the ith one. S, the 'small' one has $C(s,t-1)$ horizontal blocks (one for each row in a horizontal block of B).

In order to define $M(s,t)$ for $s,t > 1$ take $C(s-1,C(s,t-1))$ copies of S (one for each horizontal block of B). We build M in $C(s-1,C(s,t-1))$ steps starting with the empty matrix. In the ith step we add $(t-1)C(s,t-1) + C(s,t-1)$ new rows and $(t+1)C(s,t-1) - 2 + v_i$ new columns to the bottom right of the already built part of the matrix. The way we do this is the following:

(1) We put $(t-1)C(s,t-1)$ new rows and $(t+1)C(s,t-1)-2$ columns after the already existing ones and in the intersection of these new rows and columns we put a copy of S.

(2) We insert another new row after each horizontal block of this new copy of S and place one 1 entry into these new rows under the leading column of 1's in S.

(3) We add v_i new columns after the existing ones. In the intersection of the new rows added in all foregoing steps type (2) and of these new columns we put the ith vertical block of B. Note that we don't have enough rows for this, but the remaining rows of the ith vertical block of B are empty.

Simple calculations give that the construction maintains property (a). One can check that this construction maintains properties (b)-(e) too. Let $M'(s)$ be the n by n square matrix obtained from $M(s,1)$ by adding $2C(s,1)$ empty rows and 2 empty columns after the existing ones. Here $n = 3C(s,1)$. This matrix has the same weight as $M(s,1)$, which has s 1 entries in every row, therefore $w(M'(s)) = s \cdot C(s,1) = \frac{1}{3}s \cdot n$. Note that it is clear now that this weight is higher than linear as s can be appropriate big. Moreover, for these matrices $w(M'(s)) = \Theta(\beta(n) \cdot n)$, where $\beta(n)$ is the inverse function of $C(s,1)$. $A(s,t)$ denotes the Ackermann function. Now we prove that $C(s,1) < A(s+1,s+1)$ for $s \geq 1$, which implies $\beta(n) > \alpha(n) - 1$ as $\alpha(n)$ is the inverse of $A(s,s)$. For that we need the definition of $A(s,t)$, which uses again a double recursion:

$$\begin{aligned} A(0,t) &= t+1 \qquad (t \geq 0) \\ A(s,0) &= A(s-1,1) \qquad (s > 0) \\ A(s,t) &= A(s-1, A(s,t-1)) \qquad (s,t > 1). \end{aligned}$$

From the definition of $C(s,t)$ we can deduce that $C(3,t) = 2^{t+1}$ for $t \geq 1$.Moreover $A(4,t) = 2^{2^{\cdot^{\cdot^{2}}}} - 3$ (we take $t+3$ 2's) for $t \geq 0$ follows easily too and it is clear that $A(4,t) > t^2 C(3,t)$ for $t \geq 1$. It is also clear that these functions are monotone in s and in t. Now we prove by induction that $t^2 C(s,t) < A(s+1,t)$ for $s \geq 3$ and $t \geq 1$. For $s = 3$ the claim is true. For $t = 1$ we have $1 \cdot C(s,1) = 2C(s-1,2) < A(s,2)/2$ by induction. Moreover $A(s+1,1) = A(s, A(s+1,0)) = A(s, A(s,1)) > A(s,2) > C(s,1)$ where the first inequality follows from $A(s,1) > A(0,1) = 2$ and the monotonicity for t. For the general step $s \geq 4$ and $t \geq 2$ we have $t^2 C(s,t) = t^2 C(s,t-1) C(s-1, C(s,t-1)) < t^2 A(s-1, C(s,t-1))/(C(s,t-1))$ by induction. The right side is smaller than $A(s-1, A(s,t-1)) = A(s,t)$ as $C(s,t-1) > t^2$ trivially for any $s \geq 4$, $C(s,t-1) < A(s,t-1)/(t-1)^2 \leq A(s,t-1)$ and A is monotone in t.

Finally, for any $s \geq 4$ we have $C(s,1) < A(s+1,1) < A(s+1,s+1)$ as needed.

Now we need some notations first. *Ordinary rows* and *ordinary columns* are the rows and columns introduced in step (1) of the construction. The rows introduced in step (2) are the *extra rows*. The columns introduced in step (3) are *extra columns*. The 1 entries introduced in step (1) are the *ordinary* 1 entries, in step (2) we introduced *extra* 1's and in step (3) the *new* 1's. Clearly, for example each extra 1 is in an ordinary column and in an extra row. We need some more simple observations, which can be proved easily by induction:

Lemma 2.12. *The matrix $M(s,t)$ has the following properties:*

1. *The c_ith column contains 1's inside H_i and no other 1's.*
2. *If l is a new 1 and k has at most as big row index and at least as big column index as l, then k is new too.*
3. *If l is an ordinary 1 and k is in the same column or row, then k is ordinary too and they are in the same copy of S with the exception when l is a leading 1 and k is an extra 1 in its column.*

4. *If l is extra or ordinary and k is ordinary with at most as big row index and at least as big column index as l, then l and k are in the same copy of S.*

Theorem 2.13. *The extremal function of the pattern S_2 is $\Theta(n\alpha(n))$.*

Proof. We've seen that the upper bound follows easily from Theorems 2.3 and 2.10. Now we prove that the above constructed matrices $M(s,t)$ avoid S_2. This implies that the matrices $M'(s)$ avoid these patterns too. As these matrices have weight $\Theta(n\beta(n)) \geq \Theta(n\alpha(n))$ the theorem follows.

We need to prove a stronger statement. We prove that M avoids S_1, Q_1' and Q_3 also. We prove this by induction on the parameters s and t.

If $s = 1$ the matrix $M(s,t)$ avoids trivially all the patterns mentioned. In other cases we assume that $M(s,t)$ contains the appropriate pattern and we get that S or B contains this pattern too which contradicts to the induction hypothesis. Take a representation of the pattern, this means four 1 entries in each case. The first row of these entries contains two 1 entries in each case. Call the left a and the right b. Call the leftmost bottom entry x and the remaining entry c.

Now suppose the hypothesis is true for all matrices $M(s',t')$ with $s' < s$ or $s = s'$ and $t' < t$. This implies that the hypothesis is true for the matrices S and B too.

If $t = 1$ then the matrix $M(s,1)$ is constructed from $M(s-1,2) = B'$. Call the 1 entries in B' ordinary, the 1 entries added to B' new. We can assume that x is new, because it is clear that otherwise all 1 entries of the pattern are ordinary and so contained in B', which is a contradiction. The entries a, b and c must be ordinary by construction. Take x' the leftmost ordinary 1 entry in the row of x. This is not right from a again by construction. So these four 1 entries together form a forbidden pattern in B'.

We can assume that x is an extra entry, because if it is not then take the leftmost 1 entry in its row. It can't be new, if it is extra, then call that x, together with a, b and c they represent one of the forbidden patterns. If it is ordinary then it is a leading 1, now take the extra 1 entry below that and call this x. The four 1 entries represent again a forbidden pattern.

There are several cases depending on the type of a.

The entry a can't be extra because x is extra too and its row index is not smaller whereas its column index is not bigger than that of a.

Assume now that a is new. Then b is new too and they were added in a step not earlier then the step when x appeared. This holds for c too as its column is between the column of a and b and its row is not below the row of x. Therefore c is new too. Now take x', which is the leftmost new entry in the row of x. As x' is a leading 1 in B, it can't be right from a. Therefore a, b, c and x' are four 1 entries which represent a forbidden pattern in B, a contradiction.

Assume now that a is ordinary. Then b is ordinary as well and is in the same copy of S. Moreover, c is ordinary too and it is in the same copy of S too as its column is between the column of a and b and it can't be extra because of the position of x. Now take a, b, c and x', which is the leading 1 in the column of x exactly above it (so it is in the same copy of S as the other 1 entries). As x' is the last ordinary leading 1 in its horizontal block, its row can't be above the row of c. Therefore these four 1 entries represent a forbidden pattern in S, a contradiction. $\square$

Another result of Klazar [11] gives:

Theorem 2.14. *For the extremal function of any general DS-sequence u the upper bound $ex(n,u) \leq n \cdot 2^{O(\alpha(n)^{|u|-4})}$ holds.*

This result implies the following theorem for matrices (mentioned in [12]). Note that the function $2^{\alpha(n)^{O(1)}}$ is extremely slowly growing.

Theorem 2.15. *For any pattern P with exactly one 1 entry in each column $ex(n,P) \leq n2^{\alpha(n)^{O(1)}}$ holds.*

Proof. By Theorem 2.2 it is enough to prove the theorem for patterns with no empty rows. We define the string u corresponding to the pattern P. Suppose P has k rows and l columns. Number the rows of P from top to bottom with the numbers $1,\dots,k$. Now $u_{12\dots k} = a_1^2 a_2^2 \dots a_l^2$, where a_i is the row index of the single 1 entry in the ith column of P. Now repeat this step with all possible numbering of the rows with the numbers $1,\dots,k$. To get u concatenate the obtained $k!$ strings. Therefore, $|u| = 2l \cdot k!$, which is a constant depending on P.

Take an n by n matrix M with maximum weight avoiding P. Number its rows from top to bottom with the numbers $1,\dots,n$. For every i change every 1 entry in the ith row to i. The 0 entries are regarded as blanks. Let v denote the string that we obtain if we read the columns of M from left to right and from top to bottom in each column. In this way $|v| = w(M)$ and $\|v\| \leq n$. Regard the temporary matrix M' and the string v' corresponding to it (in the beginning they are equal to M and v). If there is any letter with close (length of $\|u\| \leq k$) repetition, then we delete all the letters from v' between the pair of repeating letters and the earlier letter from the pair. Moreover, we delete from M' the n positions corresponding to the ones starting at the first repeating letter (i.e. a column). In this way we obtain a new M' with one less column. We repeat this step until there are no close repetitions. As M had n columns at the beginning, there were at most n steps and in each step we deleted at most k letters from v'. We obtained the string v' which has no close repetitions and we deleted at most kn letters.

Now we prove that the string v corresponding to the 1 entries in M avoids u. This implies that v' avoids u as well and so $ex(n,P) = w(M) \leq kn + n2^{O(\alpha(n)^{|u|-4})} =\leq n2^{\alpha(n)^{O(1)}}$ by Theorem 2.14, as needed. Indeed, assume that v does not avoid u and consider the entries in M corresponding to the letters representing u in v. Every second is in a new column as two entries with the same index can't be in the same column as they are in the same row. Regard the words obtained from P with the permutations of the numbers $1,\dots,k$. Clearly, the even 1 entries corresponding to the part of u that holds one of these represent the pattern P.

$\square$

2.4 Patterns with higher extremal function

In this section we present patterns with extremal function $O(n \log n)$. These bounds were proved mainly by Füredi. Later Tardos in [15] found the constant factor also. The proof we present appeared in [5] parallel with the proof of Bienstock and Győri in [2].

Theorem 2.16.

$$ex(n, Q_1) = \Theta(n \log n)$$
$$ex(n, Q_2) = \Theta(n \log n)$$

Proof. We start with the construction proving the lower bounds. Let $A_n = (a_{ij})$ be the n by n matrix, where $a_{ij} = 1$ if and only if $j - i = 2^k$ for some integer k. The weight of A_n is $w(A_n) = \sum_{k=0}^{\lfloor \log_2 n \rfloor}(n - 2^k) \geq n \log_2 n - n$. Note that the 1 entries in A_n are arranged in diagonals (one for every k).

We prove that for $i < i' \leq i''$ and $j > j' \geq j''$ we don't have $a_{ij} = a_{ij'} = a_{i'j} = a_{i''j''} = 1$ in D_n. This guarantees that the matrix avoids among others the required patterns (for example R is also avoided). Assume that $a_{ij} = a_{ij'} = a_{i'j} = 1$ and prove that $a_{i''j''} = 1$ can't hold in this case. We have $j - i = 2^k$ for some integer k. The values $j - i'$ and $j' - i$ are both less then $j - i$ and also powers of 2, so they are at most 2^{k-1}. Finally, $j'' - i'' \leq j' - i' = (j' - i) + (j - i') - (j - i) \leq 2^{k-1} + 2^{k-1} - 2^k = 0$ and so $a_{i''j''} = 0$ as claimed.

Now we prove the upper bound for Q_1. By Theorem 2.3 it gives the desired upper bound for Q_2 too. Let $A = (a_{ij})$ be a maximum weight n by n matrix avoiding the pattern Q_1. It is enough to bound the weight of A.

For any 1 entry we call the 1 entry in the same row left to it with the biggest column index its left neighbour. In every row i let $f(i)$ be the index of the column containing the leftmost 1 entry in that row. For each 1 entry a_{ij} except the first two in each row we can define two positive integers. If the column index of the left neighbour of a_{ij} is j' (this exists as the 1 entry is not the first two in its row), then the two integer we need is $p = j' - f(i)$ and $q = j - j'$. We refer to the 1 entry a_{ij} as *far* if and only if $p \leq q$, otherwise we refer to it as *close*.

We claim that in each row there are at most $\log_2 n$ far entries. Indeed, take the far 1 entries in row i with column indices $f(i) < j_i < \ldots < j_k$. For any $1 < l \leq k$ let be j'_l the column index of the left neighbour in A of the 1 entry with column index j_l. Then by definition $j_{l-1} - f(i) \leq j'_l - f(i) \leq j_l - j'_l \leq j_l - j_{l-1}$. This implies $2(j_l - f(i)) \leq (j_{l+1} - f(i))$, which guarantees that $k \leq \log n$ as $j_k - f(i) \leq n$.

Further we claim that in each column there are at most $\log_2 n$ close entries. Indeed, take the close 1 entries in column j with row indices $i_1 < i_2 < \ldots < i_k$. As before let j'_l be the column index of the left neighbour in A of the 1 entry with in position (i_l, j). We get that $j'_l \leq f(i_{l+1})$ for any $1 \leq l < k$ as otherwise the 1 entry in column j with row index i_l and its left neighbour together with the 1 entries in position (i_{l+1}, j) and $(i_{l+1}, f(i_{l+1}))$ would represent Q_1. This implies that $j - f(l+1) \leq j - j'_l \leq j'_l - f(l)$ and so $n \geq j - f(l) = (j - j'_l) + (j'_l - f(l) \geq 2(j - f(l+1))$, which guarantees that $k \leq \log_2 n$.

Counting the first two 1 entries in each row and the close and far 1 entries separately we have at most $2n + n \log_2 n + n \log_2 n$. $\square$

The lower bound for the following theorem was proved by Tardos in [15], though in [2] the weaker lower bound $\Theta(n \log n / \log \log n)$ was proved and conjectured to be tight.

Theorem 2.17.

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

Proof. The upper bound follows from the upper bound for Q_1 using Theorem 2.3. For the lower bound we present a construction.

Let i and j be strings of equal length over an ordered set of letters. Let $<$ denote the *lexicographic ordering*, i. e., $i < j$ if and only if in the *first* position where the two strings differ, i has a smaller letter. Let $<^*$ denote the *anti-lexicographic ordering*, i. e., $i <^* j$ if and only if in the *last* position where the two strings differ, i has a smaller letter. We use the relations $>, \geq, \leq, >^*, \geq^*, \leq^*$ with their obvious meaning.

We construct the matrix C_n of weight $\Theta(n \log n)$ avoiding Q_3 only for values $n = 2^m$ $(m \geq 1)$ and for other values simply take the construction for the largest power of 2 below n and add zero columns and rows.

We index the the rows and the columns of the matrix $C_n = (c_{ij})$ with the 0-1 strings of length m. Order the rows lexicographically and the columns reverse anti-lexicographically according to their index. Let $c_{ij} = 1$ if and only if the strings i and j differ in a single position u and here $i_u = 0$ and $j_u = 1$.

Now we need to prove that the matrix C_n avoids Q_3. Equally, we need prove that cannot occur for any indices $i < i' < i''$ and $j >^* j' >^* j''$ that $c_{ij} = c_{ij''} = c_{i''j} = c_{i'j'} = 1$. We prove a stronger claim, namely that this cannot occur for indices $i < i' \leq i''$ and $j >^* j' \geq^* j''$ either. This means that the construction avoids also (among others) the pattern Q_1'.

Assume on the contrary that for the row indices $i < i' \leq i''$ and $j >^* j' \geq^* j''$ we have $c_{ij} = c_{ij''} = c_{i''j} = c_{i'j'} = 1$. Let $i \leq u \leq m$ be the 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 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, if the position where i and j'' differ is w then $j >^* j''$ 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 $z > u$. Further, $a_{i'j'} = 1$ implies $i'_z \leq j'_z$ for all z. The inequality $i' > i$ guarantees the existence of 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 $i'_u > i_u$. From $j' <^* j$ we get similarly that there must exist a position z with $j'_z < j_z$. Finally, from $j'_z = j_z$ for $z > u$ and from $j'_z \geq i'_z = i_z = j_z$ for $z < u$ we get that $j'_u < j_u$. Putting these together we get that $i_u < i'_u \leq j'_u < j_u$ which is a contradiction as all these values are 0 or 1. □

We determined the order of magnitude of the extremal function for all patterns with weight at most 4 except for the pattern P_2 which will be determined in the next section. There were four types of extremal functions, namely $O(n)$, $O(n\alpha(n))$, $O(n \log n)$ and $O(n^{\frac{3}{2}})$.

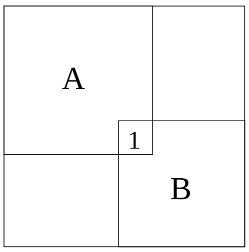
3 New types of reductions

3.1 Two matrices in diagonal arrangement

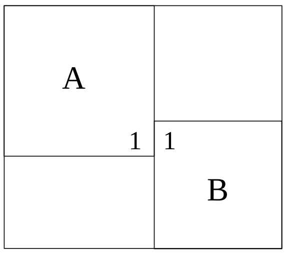
In this section we generate from two patterns one bigger pattern and give a bound on its extremal function.

Theorem 3.1. *Let A and B be two patterns. Assume that pattern A has got a 1 at its lower right and B at its upper left entry. Let C be a pattern consisting of A at its upper left part and B at its lower right part with exactly one common entry, which is the 1 entry mentioned. The other entries are blank. Then* $\max(ex(n, A), ex(n, B)) \leq ex(n, C) \leq ex(n, A) + ex(n, B)$.

Proof. The first inequality is trivial, for the second it is enough to prove that if a matrix M has weight more than $ex(n, A) + ex(n, B)$, then it cannot avoid C. Let M be a matrix with weight $w(M) \geq ex(n, A) + ex(n, B) + 1$, then $w(M) \geq ex(n, A) + 1$ also and so it contains the pattern A. Take one representation and delete the 1 entry corresponding to the lower right entry of A. We can repeat this step $ex(n, B) + 1$ times, until the weight of the matrix goes below $ex(n, A) + 1$. Now consider only the matrix M' which contains only the deleted 1 entries. Clearly, $w(M') \geq ex(n, B) + 1$ and there is a representation of B in M' (and so in M too). By construction the upper left entry of this representation is a lower right entry of a representation of A in M. Putting these representations together we get a representation of the pattern C in M. $\square$



Theorem 3.1



Corollary 3.2

Table 1.
Illustrations for Theorem 3.1 and Corollary 3.2

Corollary 3.2. *Assume that C is a pattern, which can be decomposed diagonally into two patterns A and B with exactly one common corner-entry as in Theorem 3.1. Let the pattern C' be the one consisting of A at its upper left part and B at its lower right part with no common entry and with their corner-entries being next to each other in a common row. In this case* $ex(n, C) \leq ex(n, C') \leq 2ex(n, C) + n$.

Proof. The first inequality is trivial. For the second it is obvious, that $ex(n, A) \leq ex(n, C)$ and $ex(n, B) \leq ex(n, C)$. Let A' be the pattern obtained from A by adding an extra column at its right side with a single 1 entry in its last row. By Theorem 2.3 we have that $ex(n, A') \leq ex(n, A) + n$ and putting together A' and B we obtain C' and Theorem 3.1 implies that $ex(n, C') \leq ex(n, A') + ex(n, B) \leq ex(n, A) + n + ex(n, B) \leq 2ex(n, C) + n$. $\square$

We can observe that this corollary and Theorem 2.3 and 2.4 mean that doubling a 1 entry in a certain position does not affect the order of magnitude of the extremal function.

Some results in [6] can be composed to have a similar look as Theorem 3.1. We start with a lemma and then we present this theorem. A *rectangle* R in the matrix M is the intersection of $v(R)$ consecutive rows and $h(R)$ consecutive columns. Clearly, M is a rectangle itself.

Lemma 3.3. *For any M 0-1 matrix there is a system of rectangles $\{R_i\}$ such that*

(1) all 1 entry is covered by an R_i,

(2) $\sum_i h(R_i) \leq 4h(M)$ and $\sum_i v(R_i) \leq 4v(M)$,

(3) each R_i has a 1 entry in its upper left or bottom right corner.

Proof. Define a partial order between the 1 entries in $M = (m_{ij})$. For two positions with 1 entries we say $(i, j) \leq (k, l)$, if $i \leq k$ and $j \leq l$. If $i < k$ or $j < l$ then $(i, j) < (k, l)$. For two incomparable positions with 1 entries we say that $(i, j) \lhd (k, l)$ if $j < l$ (and so $i > k$). Let $m_1 \lhd m_2 \lhd \ldots \lhd m_k$ be the set of minimal positions of 1 entries for the partial order $<$. Let $M_1 \lhd M_2 \lhd \ldots \lhd M_l$ be the set of maximal positions of 1 entries for the partial order $<$. We can assume that m_1 is in the first column, m_k is in the first row, M_1 is in the last row and M_l is in the last column. Let $m_{i+\frac{1}{2}}$ (for $i = 1, \ldots, k-1$) be the position in the intersection of the row of m_i and the column of m_{i+1}. Let $m_{\frac{1}{2}}$ be the lower left corner of M and $m_{k+\frac{1}{2}}$ be the upper right corner of M. Similarly let $M_{j+\frac{1}{2}}$ (for $j = 1, \ldots, l-1$) be the position in the intersection of the column of M_j and the row of M_{j+1}. Let $M_{\frac{1}{2}} = m_{\frac{1}{2}}$ and $M_{l+\frac{1}{2}} = m_{k+\frac{1}{2}}$. Let $h_i = [m_i, m_{i+\frac{1}{2}}]$ be the horizontal interval of positions in the row of m_i with endpoints m_i and $m_{i+\frac{1}{2}}$. Let v_i be the vertical interval $[m_{i-\frac{1}{2}}, m_i]$. Similarly, $V_i = [M_i, M_{i+\frac{1}{2}}]$ and $H_i = [M_{i-\frac{1}{2}}, M_i]$. The stair shaped curve defined by $v_1, h_1, \ldots, v_k, h_k$ is denote by s and the curve defined by $V_1, H_1, \ldots, V_k, H_k$ is denoted by S. By definition there are no 1 entries above s and below S.

Now we can begin to construct the covering system of rectangles. We will define two types of rectangles, Q_i's with a 1 entry in their bottom right position and P_i's with a 1 entry in their upper left position.

First, let Q_1 be a rectangle with lower left corner at $m_{\frac{1}{2}}$, lower right corner at M_1. Its upper right corner is the lowest position in s above M_1. Let h_i be the first horizontal interval in s not covered by Q_1. Let P_1 be a rectangle with its upper left corner at m_i. Its lower left corner is the highest position in the column of m_i which is in S too. Its upper right corner is the leftmost position in the row of m_i which is in S too.

Assume that we already defined $Q_1, P_1, \ldots, Q_i, P_i$. Let V_j be the first vertical interval in S not covered by $Q_1 \cup \ldots \cup P_i$. Let M_j be the bottom right corner of Q_{i+1}. Its bottom left corner is in the same row the rightmost position in s. Its upper right corner is the lowest position in s in the same column as its lower right corner.

Let h_j be the first horizontal interval in s which is not covered by $Q_1 \cup \ldots \cup P_i \cup Q_{i+1}$. Let m_j be the upper left corner of P_{i+1}. Its bottom left corner is in the same column the highest position in S and its upper right corner is the leftmost position in S in the same row as its upper left corner.

If there is no more suitable h_i (or V_j) then we end the construction.

Now it is enough to prove that these rectangles satisfy the conditions. (3) follows trivially. Call e_i be the rightmost column of a Q_i and f_j the highest row of a P_j. By definition, e_i's are moving to the right and f_j's are moving up. For (1) it is easy to see that the rectangles cover all positions below s and above S and all 1 entries are in such positions. Indeed, the lower right corner of Q_{i+1} is on f_i or below f_i and so $Q_1 \cup \ldots \cup P_i \cup Q_{i+1}$ covers everything left from e_{i+1} between s and S. Similarly $Q_1 \cup \ldots \cup Q_{i+1} \cup P_{i+1}$ covers everything below f_{i+1} and between s and S.

Finally, for (2) by definition the highest row of Q_i is not above f_i. The lower right corner of Q_{i+1} is not above f_i too, but it is the last maximal 1 entry with that property, so Q_{i+2}'s lower right corner is above f_i. This implies that the rows of Q_i and Q_{i+2} are disjoint. Similarly, this is true for the columns and for the P_i's as well. The required bounds follow. □

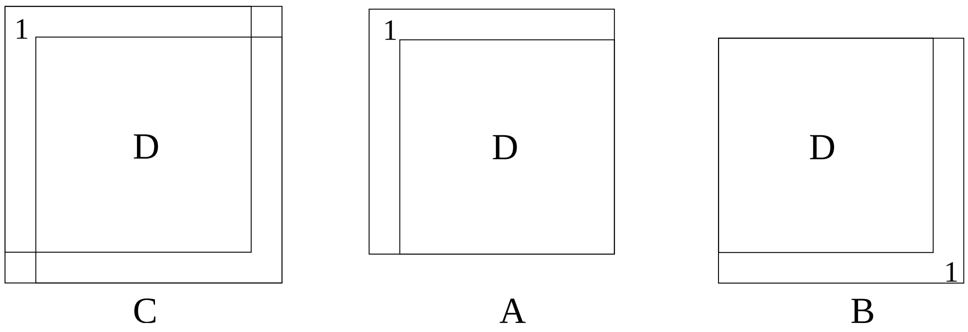


Table 2.
Illustrations for Theorem 3.4

Theorem 3.4. *Let C be a pattern containing exactly one 1 entry in its leftmost column, in its rightmost column, in its first row and in its last row as well. These 1 entries are in the upper left and in the lower right corner positions of C. Let A be the pattern obtained from C by deleting its last row and column. Respectively, B is obtained from C by deleting its first row and column. In this case* $\max(ex(n,A), ex(n,B)) \leq ex(n,C) \leq 16(n + ex(n,A) + ex(n,B))$.

Proof. The first inequality is trivial. For the second let M be a matrix with maximum weight avoiding the pattern C. Construct the set of covering rectangles guaranteed by Lemma 3.3. Let $\{Q_i\}$ be the set of rectangles with 1 entry in the lower right corner. Respectively, $\{P_j\}$ is the set of rectangles with 1 entry in the upper left corner. For any Q_i let Q_i' be the matrix obtained from it by deleting its last row and column. If Q_i' contains the pattern A, then together with the corner entry of Q_i we get a representing matrix of C in M, which is a contradiction. Let be Q_i'' the smallest square matrix obtained from

Q_i' by adding some empty rows or columns before the existing ones. Clearly, Q_i'' avoids A as well. Now we can deduce that $w(Q_i) - 2q_i \leq w(Q_i) - (2q_i' + 1) \leq w(Q_i') = w(Q_i'') \leq ex(q_i', A) \leq ex(q_i, A)$, where $q_i' = max(h(Q_i'), v(Q_i')) = max(h(Q_i), v(Q_i)) - 1$ and $q_i = h(Q_i) + v(Q_i) \geq q_i' + 1$. A similar statement holds for any P_j: $w(P_j) - 2p_j \leq ex(p_j, B)$, where $p_j = h(P_j) + v(P_j)$. Putting these together for the number of 1 entries in M holds: $w(M) \leq \sum_{i=1}^{k} w(Q_i) + \sum_{j=1}^{l} w(P_j) \leq \sum_{i=1}^{k}(ex(q_i, A) + 2q_i) + \sum_{j=1}^{l}(ex(p_j, B) + 2p_j) \leq ex(\sum_{i=1}^{k} q_i, A) + w(\sum_{j=1}^{l} p_j, B)$. The last inequality follows from the repeated application of Theorem 2.5. Therefore by Lemma 3.3, $ex(n, C) = w(M) \leq 16n + ex(\sum_{i=1}^{k} q_i, A) + w(\sum_{j=1}^{l} p_j, B) \leq ex(4n, A) + ex(4n, B) \leq 16(n + ex(n, A) + ex(n, B))$. The last inequality is trivial as by dividing a $4n$ by $4n$ matrix into 16 n by n matrices we get the bound. $\square$

Using Theorems 3.1 and 3.4 in case both of the matrices A and B are linear, we get that C is linear as well.

3.2 Adding two 1's between and below other two

In this section we prove a new reduction using a lemma about DS-sequences, which appeared in [8]. In the original paper (using a rather complicated technical lemma) it is applied to prove a theorem, which can also be applied to 0-1 matrices. First we present this theorem and its consequence on 0-1 matrices. Later we present and prove the lemma's equivalent for 0-1 matrices, which gives upper bounds for a wider range of patterns. As a corollary, it gives also a simpler proof for the theorem mentioned above.

For definitions about DS-sequences check section 2.3.

Theorem 3.5. *Suppose a and b are two symbols and $u = u_1 a^2 u_2 a$ is a sequence such that $b \notin S(u)$. Then $ex(n, u_1 ab^i a u_2 ab^i) = \Theta(ex(n, u))$ for any $i \geq 1$.*

Building up a string with this operation from $u = a_1^6$, which is trivially linear, we get quite simply the following consequence:

Corollary 3.6. *The string $u = a_1^2 a_2^2 \dots a_{k-1}^2 a_k^4 a_{k-1}^2 \dots a_2^2 a_1^4 a_2^2 \dots a_{k-1}^2 a_k^2$, where the symbols $a_1, a_2, \dots, a_n$ are mutually distinct, has linear extremal function for all $k \geq 1$.*

In the proof of the following theorem we will need the following lemma, known as the Erdős-Szekeres Lemma (appeared in [4]):

Lemma 3.7. *Any sequence of numbers of length $(k-1)^2 + 1$ contains a monotone subsequence of length k.*

Theorem 3.8. *Let P_k be a k by $2k$ pattern with exactly one 1 entry in every column such that for $1 \leq i \leq k$ the 1 entry in the ith and $(2k+1-i)$th columns are in the ith row. For every $k \geq 1$ this pattern has extremal function $ex(n, P) = O(n)$.*

Proof. Take an n by n matrix M with maximum weight avoiding P. Number its rows from top to bottom with the numbers $1, \dots, n$. For every i change every 1 entry in the ith row to i. The 0 entries are regarded as blanks. Let u denote the string that we obtain if we read the columns of M from left to right and from top to bottom in each column. In this way $|u| = w(M)$ and $\|u\| \le n$. Regard the temporary matrix M' and the string u' corresponding to it (in the beginning they are equal to M and u). If there is any letter with close (length of l, where l is a constant defined later) repetition, then we delete all the letters from u' between the pair of repeating letters and the earlier letter from the pair. Moreover, we delete from M' the n positions corresponding to the ones starting at the first repeating letter (i.e. a column). In this way we obtain a new M' with one less column. We repeat this step until there are no close repetitions. As M' had n columns at the beginning, there were at most n steps and in each step we deleted at most l letters from u'. We obtained the string u' which has no close repetitions and we deleted at most ln letters.

If we can prove that the obtained string u avoids the string $v = a_1^2 a_2^2 \dots a_{l-1}^2 a_l^4 a_{l-1}^2 \dots a_2^2 a_1^4$ $a_2^2 \dots a_{l-1}^2 a_l^2$ for a constant l, then u' avoids also and so the theorem follows from Corollary 3.6. A suitable l is $(k-1)^2+1$. Indeed, assume u contains v and let s be a subsequence of u equivalent to v. Use Lemma 3.7 on the sequence defined by the first l different letter of s. Assume that we got an increasing sequence of length k. This implies that the subsequence s contains a shorter subsequence $s' = i_1^2 i_2^2 \dots i_{k-1}^2 i_k^4 i_{k-1}^2 \dots i_2^2 i_1^4 i_2^2 \dots i_{k-1}^2 i_k^2$, where the indices $i_1, i_2, \dots, i_k$ are in increasing order. Take in s' the subsequence $s'' = i_1^2 i_2^2 \dots i_{k-1}^2 i_k^4 i_{k-1}^2 \dots i_2^2 i_1^2$. Consider now the 1 entries in M corresponding to these letters. Every second is in new column because two symbols with the same index can't be in the same column as they are in the same row. Moreover the order of the indexes guarantees that taking every other symbol we obtain a representation of P in the matrix M, which is a contradiction. In the second case of the lemma we obtain a decreasing sequence of length k, and so s contains an s' as above but now the indices $i_1, i_2, \dots, i_k$ are in decreasing order. Now consider the 1 entries in M corresponding to the subsequence $s' = i_k^2 i_{k-1}^2 \dots i_2^2 i_1^4 i_2^2 \dots i_{k-1}^2 i_k^2$ and the same argument holds. □

Following again the notations of [8], suppose $u = a_1 a_2 \dots a_m$ is a sequence. An *interval* $I = \langle a_j, a_{j+k} \rangle$ in u is any contiguous subsequence $a_j a_{j+1} \dots a_{j+k}$, $k \ge 1$ of length at least 2. In the case a_j and a_{j+k} are both a-occurrences we call I an *a-interval.* An *ordered sequence* $(u, <)$ is a sequence enriched by a linear order on the alphabet $(S(u), <)$. If u is ordered and a_i is an a-occurrence in u then we say a_i is *covered* (in u), if there is an interval I in u, such that

1. $a_i \in I$

2. there are at most 16 a-occurrences in I

3. there are 2 occurrences of a symbol $x \in S(u)$, $x < a$ in I.

The following lemma is a simplified version of a Lemma in [8], where the definition of being covered consisted two parameters, and so the lemma was also more general.

Lemma 3.9. *Suppose u is an ordered sequence without immediate repetitions. Either $|u| \le 1440\,\|u\|$ or there exists at least $\frac{1}{10}|u|$ occurrences in u, which are covered.*

Proof. We can suppose that $S(u) = \{1, 2, \dots, n\}$ and that the order on them is the standard order of integers. We will define by induction sets $U_0, U_1, \dots, U_n$ of disjoint intervals in u. For each j U_j will contain some k-intervals for $k = 1, 2, \dots, j$. The set U_0 is empty. Suppose that U_{j-1} is defined. Split the j-occurrences in u to 16-tuples $T_1, T_2, \ldots, T_m$ and to T, so that T_1 consists of the 16 leftmost j-occurrences, T_2 the next 16 j-occurrences etc. and T contains the possibly remaining maximum 15 rightmost j-occurrences. Define

$$\mathcal{S}_j = \{ \text{ i} \mid \text{at least one } j\text{-occurrences of } T_i \text{ is not covered } \}.$$

Group all elements of T_i for all $i \in \mathcal{S}_j$ into 8 pairs (x, x') of consecutive elements generating j-intervals $\langle x, x' \rangle$. Put all these intervals into U_j and the intervals of U_{j-1} not intersecting them. By definition, some j-occurrences of $\mathcal{S}_j$ are not covered and so for any $i \in \mathcal{S}_j$ the 8 intervals corresponding to T_i intersect at most 2 intervals of U_{j-1}. Indeed, more intersecting intervals would give one pair of symbols in the interval defined by the leftmost and rightmost element of T_i, while in this interval there are just 16 j-occurrences, so every element of T_i would be covered, contradicting the definition. We got that

$$|U_j| \geq 8\,|\mathcal{S}_j| + |U_{j-1}| - 2\,|\mathcal{S}_j| = 6\,|\mathcal{S}_j| + |U_{j-1}|$$

and so

$$|U_n| \geq 6 \textstyle\sum_{j=1}^{n} |\mathcal{S}_j|.$$

There are no immediate repetitions, so the disjoint k-intervals in U_n have length at least 3, and so $|u| \geq 3\,|U_n| \geq 18 \sum_{j=1}^{n} |\mathcal{S}_j|$, that is

$$\frac{|u|}{18} \geq \sum_{j=1}^{n} |\mathcal{S}_j|\,.$$

Assume that $|u| \geq 1440n$, that is $\frac{|u|}{1440} \geq n$. Counting the occurrences in u, which are not covered, we have at most 16 times the number of T_i with $i \in \mathcal{S}_j$ and some of the occurrences in the residual Ts for every j. So the number of occurrences, which are covered, is at least

$$|u| - 16 \textstyle\sum_{j=1}^{n} |\mathcal{S}_j| - 15n \geq |u| - \frac{16}{18}\,|u| - \frac{1}{90}\,|u| = \frac{1}{10}\,|u|.$$

□

Now we can turn to the main result of this section.

Theorem 3.10. *Let A be a pattern which has two 1 entries in its first row in column i and $i+1$ for a given i. Let A' be the pattern obtained from A by adding two new columns between the ith and the $(i+1)$th column and a new row before the first row with exactly two 1 entries in the intersection of the new row and columns. Then $ex(n, A') = O(ex(n, A))$.*

Proof. Suppose M is a maximum weight n by n matrix avoiding A'. As in 3.8 we generate a sequence with the letters $1, 2, \ldots, n$ from the 1 entries of M. Number the rows of M from top to bottom with the numbers $1, \ldots, n$. For every i change every 1 entry in the ith row to i. The 0 entries are considered blanks. Let u' denote the string that we obtain if we read the columns of P from left to right and from top to bottom in each column. Again, $|u'| = w(M)$ and $\|u'\| \leq n$. If there is any immediate repetition, then we delete one letter from it and we repeat this step until there are no repetitions. Again, we deleted at most n letters. Call the string obtained u.

Use Lemma 3.9 on this string. If the length of u is at most $1440n$, then $w(M) \leq n + 1440n = O(n) = O(ex(n, A))$. Otherwise there are at least $\frac{1}{10}|u|$ letters in u, which are covered. Consider the matrix M' with the 1 entries of M corresponding to the covered letters. We claim that M' avoids the pattern A'', which is obtained from A by putting 33 new columns between the ith and the $(i+1)$th column with exactly one 1 entry in the first row. This would prove the theorem as by Theorem 2.4 $ex(n, A') = w(M) \leq n + 10w(M') \leq n + 10ex(n, A'') \leq n + 10 \cdot 34ex(n, A) = O(ex(n, A))$. Suppose M' does not avoid A''. Consider one representation of A'' in M' and consider the middle entry of the 33 new 1 entries in this copy of A''. The corresponding letter in u is some positive number a, and is covered, so there exist an interval I with maximum 16 a-occurrences and with two b-occurrences for some b, where $b < a$ by the definition of covering. This interval must not contain the two letters on the left and the two letters on the right end of the 35 letters being next to each other in A'', because then the interval would contain more than 16 a-occurrences. So between the 2nd and the 34th a-occurrence there are two b-occurrences. The corresponding 1 entries of these b-s are not in common columns with the 1 entries corresponding to the 1st and the 35th a-occurrences, because there is one more a-occurrence between them (the 2nd and the 34th). Moreover the 1 entries corresponding to the b-s are in a row above the row corresponding to a. Now we can deduce that the representation of A'' without the middle 33 entries in row a and together with the two 1 entries in row b represent A', which is a contradiction. □

It is clear that this theorem can be similarly applied to the last row or to the first or last column by rotating back and forth the matrix. From the 1 by 2 pattern with two 1 entries (for which the extremal function is trivially linear) we can build up using this theorem on P_k, proving Theorem 3.8. For example the pattern P_2 has linear extremal function, which was proved with a direct method too in [6]. Another example where we can make use of this theorem is to prove that $ex(n, L_3) = O(n)$. It follows from the fact that L_3 can be obtained from L_1 using this theorem on the first column and after that deleting the appropriate 1 entry. Using Theorem 2.3 on L_3 we can find some more linear matrices with weight 5.

4 On minimum non-linear matrices

4.1 Patterns avoided by the diagonal construction

First we define the patterns $H_k = (h_{ij})$ for $k \geq 0$. The pattern H_k for $k \geq 0$ has m rows and m columns for $m = (3k+4)$ and it is symmetrical to the line going from its top right corner to the bottom left corner. All entries are blank except the following ones:

$$h_{41} = h_{12} = h_{13} = h_{(m-1)m} = h_{(m-2)m} = 1,$$
$$h_{(3l+4)(3l+1)} = h_{(3l-1)(3l+3)} = h_{(3l)(3l+2)} = 1 \qquad (1 \leq l \leq k).$$

See the Appendix for H_0 and H_1.

First we prove that the weight 5 pattern H_0 has extremal function $\Theta(n \log n)$. It is easy to see that by deleting any 1 entry from it we obtain a weight 4 pattern with linear extremal function. We call these type of patterns *minimal non-linear* patterns. So far, this is the only pattern with weight more than 4, known to be the member of this class of patterns. We conjecture that the patterns H_k for $k \geq 1$ are minimal non-linear patterns as well.

Theorem 4.1. *For the pattern H_0 we have $ex(n, H_0) = \Theta(n \log n)$.*

Proof. The upper bound follows from applying Theorem 3.10 for the last column of pattern Q_2 and after that deleting the 4th row.

For the lower bound we use the construction in the proof of Theorem 2.16. Let $A_n = (a_{ij})$ be the n by n matrix, where $a_{ij} = 1$ if and only if $j - i = 2^k$ for some integer k. The weight of A_n is $w(A_n) = \sum_{k=0}^{\lfloor \log_2 n \rfloor}(n - 2^k) \geq n \log_2 n - n$. Note that the 1 entries in A are arranged in diagonals (one for every k).

We need that the matrix A_n avoids H_0 for any n. We prove that for $i_1 < i_2 < i_3 < i_4$ and $j_1 < j_2 < j_3 < j_4$ we don't have $a_{i_4 j_1} = a_{i_1 j_2} = a_{i_1 j_3} = a_{i_2 j_4} = a_{i_3 j_4} = 1$ in D_n. This guarantees that the matrix avoids H_0.

Assume on contrary that $a_{i_4 j_1} = a_{i_1 j_2} = a_{i_1 j_3} = a_{i_2 j_4} = a_{i_3 j_4} = 1$. Therefore we have $j_1 - i_4 > 0$ and $j_3 - i_1 = 2^{k_3}$, $j_2 - i_1 = 2^{k_2}$ for some integers $k_3 > k_2$ (as $j_3 > j_2$). Thus $i_3 - i_2 < i_4 - i_1 < j_2 - i_1 + i_4 - j_1 < j_2 - i_1 = 2^{k_2} \leq (j_3 - i_1 - (j_2 - i_1)) = j_3 - j_2$. Symmetrical argument shows that $j_3 - j_2 < i_3 - i_2$, which is a contradiction. □

The above defined A_n contains H_2 and so for the general case we need a little modified version of the construction in Theorem 4.1. From now on let $A_n = (a_{ij})$ be the n by n matrix, where $a_{ij} = 1$ if and only if $j - i = 3^k$ for some integer k. The weight of A_n is $w(A_n) = \sum_{k=0}^{\lfloor \log_3 n \rfloor}(n - 3^k) \geq n \log_3 n - n$. Note that the 1 entries in A are arranged in diagonals (one for every k).

Lemma 4.2. *If for the row indices $i_1 \leq i_2 < i_3 < i_4 < i_5$ and column indices $j_1 < j_2 < j_3 < j_4 \leq j_5$ in A_n we have $a_{i_1 j_3} = a_{i_2 j_2} = a_{i_3 j_5} = a_{i_4 j_4} = a_{i_5 j_1} = 1$ and $j_3 - j_2 > i_2 - i_1$, then $j_5 - j_4 > i_4 - i_3$.*

Proof. Assume that $a_{i_1 j_3} = a_{i_2 j_2} = a_{i_3 j_5} = a_{i_4 j_4} = a_{i_5 j_1} = 1$. Therefore by the definition of A_n we have $j_3 - i_1 = 3^{k_1}$ and $j_2 - i_2 = 3^{k_2}$ for some positive integers $k_1 > k_2$ (as

$j_3 - i_1 > j_2 - i_2$). Similarly $j_5 - i_3 = 3^{l_1}$ and $j_4 - i_4 = 3^{l_2}$ for some positive integers $l_1 > l_2$ (as $j_5 - i_3 > j_4 - i_4$). Finally, $j_1 - i_5 = 3^{k_3} \geq 1$ for some positive integer k_3 which is smaller then k_2 and l_2 (as $j_2 - i_2 > j_1 - i_5$ and $j_4 - i_4 > j_1 - i_5$).

As $(j_5 - j_4) + (i_4 - i_3) = (j_5 - i_3) - (j_4 - i_4) = 3^{l_1} - 3^{l_2} \geq 2 \cdot 3^{l_2}$ and $j_3 - j_2 < (j_4 - j_1) + (i_5 - i_4) = (j_4 - i_4) - (j_1 - i_5) = 3^{l_2} - 3^{k_3} < 3^{l_2}$ we have $(j_5 - j_4) + (i_4 - i_3) > 2(j_3 - j_2) > (j_3 - j_2) + (i_2 - i_1)$. Similarly $(j_3 - j_2) + (i_2 - i_1) = (j_3 - i_1) - (j_2 - i_2) = 3^{k_1} - 3^{k_2} \geq 2 \cdot 3^{k_2}$ and $i_4 - i_3 < (i_5 - i_2) + (j_2 - j_1) = (j_2 - i_2) - (j_1 - i_5) = 3^{k_2} - 3^{k_3} < 3^{k_2}$ implies that $(j_3 - j_2) + (i_2 - i_1) > 2(i_4 - i_3)$. Putting these together we have $(j_5 - j_4) + (i_4 - i_3) > (j_3 - j_2) + (i_2 - i_1) > 2(i_4 - i_3)$ and so $j_5 - j_4 > i_4 - i_3$ as claimed. □

Theorem 4.3. *For any $k \geq 0$ for the pattern H_k we have $ex(H_k, n) = \Omega(n \log n)$.*

Proof. It is enough to prove that A_n avoids H_k for $k \geq 0$. Suppose on contrary that A_n contains H_k. Take a submatrix of A_n representing H_k. Its row and column indices are $i_1 < i_2 < \ldots < i_m$ and $j_1 < j_2 < \ldots < j_m$. Using Lemma 4.2 for the 1 entries in the first 4 rows we got that $j_3 - j_2 > i_1 - i_1 = 0$ implies $j_6 - j_5 > i_3 - i_2$. Now we repeat this lemma for all of the diagonal pairs of ones in H_k. In each step we take the two 1 entries in the positions defined by the last inequality we got and take three 1 entries defined in the next step of H_k. We use the lemma for these five 1 entries. For example the second step uses the lemma for the 1 entries in positions $(i_2, j_6), (i_3, j_5), (i_7, j_4), (i_5, j_9), (i_6, j_8)$ and implies that $j_9 - j_8 > i_6 - i_5$. In this way we got that $j_{3l+3} - j_{3l+2} > i_{3l} - i_{3l-1}$ for $1 \leq l \leq k$. In the last step we use the lemma for the 1 entries in the last 5 columns. Therefore $j_{3k+3} - j_{3k+2} > i_{3k} - i_{3k-1}$ implies that $0 = j_{3k+4} - j_{3k+4} > i_{3k+3} - i_{3k+2} > 0$, which is a contradiction. □

As mentioned earlier, we conjecture that these patterns are minimal non-linear patterns. The above theorems guarantee that these patterns are non-linear. Thus, to prove the conjecture we need that by deleting any 1 entry from H_k for any k we get a linear pattern. We call a pattern *quasilinear* if its extremal function is $O(n2^{\alpha(n)^{O(1)}})$. Clearly, H_k is not quasilinear for any k by the above theorem. A pattern with no empty rows and columns is *minimal non-quasilinear* if it is not quasilinear and by deleting any 1 entry from it we obtain a quasilinear pattern. A weaker conjecture claims that the patterns H_k are minimal non-quasilinear patterns. To prove this conjecture we need that by deleting any 1 entry from H_k for any k we get a quasilinear pattern. We cannot prove this either, though we are able to prove this for some of the 1 entries in H_k:

Lemma 4.4. *There are at least $k + 5$ 1 entries in H_k from which deleting any one gives a quasilinear pattern.*

Proof. The matrix obtained from H_k by deleting the first two column has at most 1 entry in every row, so by Theorem 2.15 and Theorem 2.2 it is quasilinear. If we delete the 1 entry in the first column, we obtain the matrix H_k^0. By Theorem 2.3 and from the quasilinearity of the above matrix its quasilinearity follows. Deleting the 1 entry in the second or in the third column gives a quasilinear matrix by Theorem 2.15. Symmetrical argument gives the quasilinearity of the matrices obtained by deleting one of the 1 entries in the last three rows. Note that for $k = 0$ two of these six 1 entries coincide, giving the quasilinearity for every matrix obtained by deleting 1 entry. As earlier mentioned, in this case linearity holds as well.

Now we continue the proof with the 1 entries in position $((3l+4),(3l+1))$ for $1 \leq l \leq k-1$ (if $k \geq 2$). Denote by H_k^l the matrix obtained from H_k by deleting the entry in position $((3l+4),(3l+1))$. It can be decomposed into two smaller matrices A' and B' in diagonal arrangement with no 1 entries out of these matrices. Indeed, A' is the matrix which is in the intersection of the first $3l+1$ rows and $3l+3$ columns of H_k^l and B' is the matrix in the intersection of the remaining rows and columns, i.e. the last $3(k-l)+3$ rows and $3(k-l)+1$ columns. Let us obtain A from A' by adding one row after the last row and one column after the last column of A' with exactly one 1 entry in their intersection. In each column of A there is at most one 1 entry and so by Theorem 2.15 for its extremal function we have $ex(n,A) = O(n2^{\alpha(n)^{O(1)}})$. Similarly, B is obtained from B' by adding one row before the first row and one column before the first column of B' with exactly one 1 entry in their intersection. Again, each row of B contains at most one 1 entry and so $ex(n,B) = O(n2^{\alpha(n)^{O(1)}})$.

Applying Theorem 3.1 on the matrices A and B we get that for the pattern C defined in the theorem we have $ex(n,C) \leq ex(n,A) + ex(n,B) = O(n2^{\alpha(n)^{O(1)}})$. Clearly, H_k^l can be obtained from C by deleting the row and the column containing the common 1 entry of A and B. This implies that $ex(n,H_k^l) \leq ex(n,C) = O(n2^{\alpha(n)^{O(1)}})$, that is H_k^l is quasilinear.

We showed for $k+5$ 1 entries that by deleting any of them we obtain a quasilinear pattern. □

Theorem 4.5. *There exist infinitely many H_k' pairwise different minimal non-quasilinear patterns.*

Proof. By deleting 1 entries (and empty rows and columns) we can obtain a minimal non-quasilinear pattern H_k' from H_k. By Lemma 4.4 this algorithm gives patterns for which $3k+5 = w(H_k) \geq w(H_k') \geq k+5$. These bounds on the weight guarantee that there are infinite many different matrices among H_k'. □

4.2 Conjectures

If we want to use the proof in Theorem 4.5 to obtain infinite many minimal non-linear patterns, then we need that the patterns A and B in Lemma 4.4 have linear extremal function. Note that the shape of A and B is symmetrical, thus it is enough to prove this for A. To make it more precise, let G_k be the matrix obtained from H_k by deleting the column containing the 1 entry in the last row, the last column and the last three rows. Clearly, any A which can appear in the above proof is contained in a G_k for some k. Thus, if $ex(n,G_k) = O(n)$ would be true for every k then the proof would give that the patterns H_k reduce to infinite many pairwise different minimal non-linear patterns.

At the end of section 2.4 we mentioned that the patterns with weight at most 4 are classified. Though, there are some patterns with weight 5 whose extremal function is not determined yet. At the end of section 3.2 we proved that L_3 is linear. In the previous section we proved that H_0 has extremal function $\Theta(n \log n)$. For the weight 5 pattern G_1 the extremal function is not determined yet.

Conjecture 4.6.

1. *For the pattern G_1 we have $ex(n, G_1) = O(n)$.*

2. *For the pattern G_k obtained from H_k by deleting the last three rows and the last column we have $ex(n, G_k) = O(n)$ $(k \geq 1)$.*

As already mentioned in section 4.1, the patterns H_k are not only prime candidates for containing infinite many non-linear patterns, but the patterns H_k can be minimal non-linear patterns themselves:

Conjecture 4.7.

1. *There are infinite many minimal non-linear patterns.*

2. *The patterns H_k are minimal non-linear patterns.*

Note that Conjecture 4.6 would prove the first statement of this conjecture.

Notice that the patterns H_k can be obtained from a permutation pattern by doubling the column containing the 1 entry in its first row. Permutation patterns have linear extremal function by Theorem 2.8. It may be true that by doubling one of its columns the extremal function remains linear. A weaker claim would be enough, namely that by doubling the column containing the 1 entry in its first row the extremal function remains linear. Note that these are not true for arbitrary patterns, as H_0 can be obtained from a linear pattern by doubling the column containing the 1 entry in its first row, yet its extremal function is $\Theta(n \log n)$. Besides, it is also necessary to put the new column next to the one which was doubled. Indeed, S_2 can be obtained from a permutation pattern by adding the copy of the column containing the 1 entry in the first row after the existing columns, though $ex(S_2, n) = \Theta(n\alpha(n))$. For permutation patterns even the stronger claim, that we can double all columns without increasing the order of magnitude, may be true.

Conjecture 4.8.

1. *For any permutation pattern by doubling the column containing the 1 entry in its first row we obtain a pattern with linear extremal function.*

2. *By doubling one column of a permutation pattern we obtain a pattern with linear extremal function.*

3. *By doubling every column of a permutation pattern we obtain a pattern with linear extremal function.*

5 Appendix

5.1 The patterns considered

In the table we use dots for the 1 entries and blank spaces for the 0 entries.

$$R = \begin{pmatrix} \bullet & \bullet \\ \bullet & \bullet \end{pmatrix} \qquad P_2 = \begin{pmatrix} \bullet & & & \bullet \\ & \bullet & \bullet & \end{pmatrix}$$

$$L_1 = \begin{pmatrix} \bullet & & \bullet \\ \bullet & & \\ & \bullet & \end{pmatrix} \qquad L_2 = \begin{pmatrix} & \bullet & & \bullet \\ \bullet & & & \\ & & \bullet & \end{pmatrix}$$

$$L_3 = \begin{pmatrix} & \bullet & & \bullet \\ \bullet & & & \\ \bullet & & & \\ & & \bullet & \end{pmatrix} \qquad Q_1 = \begin{pmatrix} & \bullet & \bullet \\ \bullet & & \bullet \end{pmatrix}$$

$$Q_2 = \begin{pmatrix} & \bullet & \bullet \\ & & \bullet \\ \bullet & & \end{pmatrix} \qquad Q_3 = \begin{pmatrix} \bullet & & \bullet \\ & \bullet & \\ \bullet & & \end{pmatrix}$$

$$Q_1' = \begin{pmatrix} \bullet & & \bullet \\ \bullet & \bullet & \end{pmatrix} \qquad H_0 = \begin{pmatrix} & \bullet & \bullet & \\ & & & \bullet \\ & & & \bullet \\ \bullet & & & \end{pmatrix}$$

$$S_1 = \begin{pmatrix} & \bullet & & \bullet \\ \bullet & & \bullet & \end{pmatrix} \qquad S_2 = \begin{pmatrix} & \bullet & & \bullet \\ & & \bullet & \\ \bullet & & & \end{pmatrix}$$

$$H_1 = \begin{pmatrix} & \bullet & \bullet & & & & \\ & & & & & \bullet & \\ & & & & \bullet & & \\ \bullet & & & & & & \\ & & & & & & \bullet \\ & & & & & & \bullet \\ & & & \bullet & & & \end{pmatrix} \qquad G_1 = \begin{pmatrix} & \bullet & \bullet & & \\ & & & & \bullet \\ & & & \bullet & \\ \bullet & & & & \end{pmatrix}$$

References

[1] Noga Alon, E. Friedgut, On the number of permutations avoiding a given pattern, *Journal of Combinatorial Theory Ser. A* **89** (2000), 133–140.

[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, *Discrete and Computational Geometry—The Goodman-Pollack Festscrift* (B. Aronov et al. eds.), Springer, Berlin (2003), 275–300.

[4] P. Erdős and G. Szekeres, A combinatorial problem in geometry, *Compocito Math.* **2** (1935), 464–470.

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

[6] Zoltán Füredi, Péter Hajnal, Davenport-Schinzel theory of matrices, *Discrete Mathematics* **103** (1992), 233–251.

[7] S. Hart and M. Sharir, Nonlinearity of Davenport-Schinzel sequences and of generalized path compression schemes, *Combinatorica* **6** (1986), 151–177.

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

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

[10] 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), 250–255.

[11] M. Klazar, A general upper bound in Extremal theory of sequences, *Comment. Math. Un. Carolinae* **33** (1992), 737–746.

[12] M. Klazar, Enumerative and extremal combinatorics of a containment relation of partitions and hypergraphs, *Habilitation Thesis* (2001)

[13] Adam Marcus, Gábor Tardos, Excluded permutation matrices and the Stanley-Wilf conjecture, *Journal of Combinatorial Theory Ser. A* **107** (2004), 153–160.

[14] J. Pach, G. Tardos, Forbidden paths and cycles in ordered graphs and matrices, *Israel Journal of Mathematics*, to appear

[15] Gábor Tardos, On 0-1 matrices and small excluded submatrices, *Journal of Combinatorial Theory Ser. A*, to appear

[16] Doron Zeilberger, A loving rendition of the Marcus-Tardos amazing proof of the Füredi-Hajnal conjecture,

http://www.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/martar.html