Forbidden Patterns in 0–1 Matrices
Semester Thesis

Samuel Zürcher*

March 2, 2006

Abstract

We consider a setting that was first introduced by Füredi in [6] and by Bienstock and Győri in [1] and that was further investigated in [7], [20], and [9]. For 0–1 matrices P and A_n, we are interested in the maximum number of 1 entries contained in a matrix A_n that avoids P and denote this number by $ex(n, P)$.

First, we introduce $ex(n, P)$ together with results from previous papers and compare $ex(n, P)$ with Turán type problems in extremal graph theory. Secondly, we look at the concept of minimally non-linear patterns as mentioned in [14] and investigate a pattern that comes up in the context of this concept. In the end, we consider a class of patterns and conjecture an upper bound on $ex(n, P)$ for patterns in this class.

1 Introduction

Let us start with the definition of a submatrix:

Definition 1.1. *A and B are 0–1 matrices of arbitrary dimension. B is a submatrix of A if and only if B can be obtained from A by deleting rows and columns but* not *permuting rows or columns*

Based on the definition of a submatrix, Füredi and Hajnal present two relations on matrices:

*sam@orelias.ch

Definition 1.2. *A, B, and P are 0–1 matrices of arbitrary dimension:*

1. *A contains P if and only if there is a submatrix B of A so that* $\forall i, j :$ $P_{i,j} = 1 \Rightarrow B_{i,j} = 1$.
2. *A avoids P if and only if A does not contain P.*

Let us first have a look at an example for this definition. We consider the pattern L_1[1]

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

and the square matrix A_5

$$A_5 = \begin{pmatrix} \bullet & \bullet & & & \\ & \bullet & & \bullet & \bullet \\ \bullet & \bullet & & \bullet & \\ & \bullet & \bullet & & \bullet \\ \bullet & & \bullet & \bullet & \end{pmatrix}$$

A_5 *contains* P because we can find a submatrix B that clearly *contains* the pattern P.

$$A_5 = \begin{pmatrix} \circ & \bullet & & & \\ & \bullet & & \circ & \bullet \\ \circ & \circ & & \circ & \\ & \bullet & \bullet & & \bullet \\ \circ & & \circ & \circ & \end{pmatrix} \qquad B = \begin{pmatrix} \bullet & & \\ \bullet & & \bullet \\ \bullet & \bullet & \bullet \end{pmatrix}$$

The matrix A'_5 *avoids* the pattern P because we cannot find such a submatrix. It is easy to give a certificate for the fact that A_5 *contains* P; it is not possible to give such a short certificate for the fact that A'_5 *avoids* P. Instead, one possibly has to test all submatrices of A'_5.

$$A'_5 = \begin{pmatrix} \bullet & \bullet & & & \\ & \bullet & & \bullet & \\ \bullet & \bullet & & & \\ & \bullet & & & \bullet \\ \bullet & & & & \bullet \end{pmatrix}$$

[1] To present 0–1 matrices, we use bullets for 1 entries and leave blanks for 0 entries.

1.1 Different Point of View

Let us interpret 0–1 matrices as adjacency matrices of a bipartite graphs:

Definition 1.3. *$A_{m,n}$ is a 0–1 matrix of dimension $m \times n$. Let $G(A_{m,n}) = (V, E)$ be a bipartite graph with $V = \{r_i \mid 1 \leq i \leq m\} \cup \{c_i \mid 1 \leq i \leq n\}$ and $E = \{\{r_i, c_j\} \mid A_{i,j} = 1\}$.*

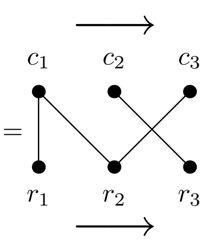
Based on the bipartite graph $G(A_{m,n})$, we define a bipartite graph $\overrightarrow{G}(A_{m,n})$:

Definition 1.4. *Let $\overrightarrow{G}(A_{m,n})$ be the graph $G(A_{m,n})$ together with the ordering relations $r_1 < r_2 < ... < r_m$ and $c_1 < c_2 < ... < c_n$ on V.*

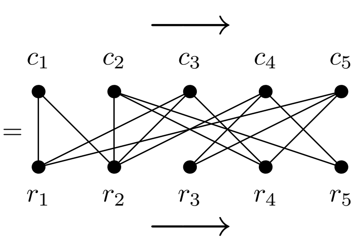
We introduce the graph $\overrightarrow{G}(A_{m,n})$ because it matches Definition 1.1 for submatrices. The vertices $r_i \in V(\overrightarrow{G}(A_{m,n}))$ represent the rows of $A_{m,n}$ and the vertices $c_i \in V(\overrightarrow{G}(A_{m,n}))$ the columns of $A_{m,n}$. Definition 1.1 states that to obtain a submatrix B from $A_{m,n}$, one is not allowed to permute rows or columns. This is equivalent to say that there is an ordering relation on the rows and columns of $A_{m,n}$. Notice that the definition of *contains* and *avoids* use Definition 1.1 — we can conclude that *contains* and *avoids* operate on the graph $\overrightarrow{G}(A_{m,n})$.

Let us show the bipartite graphs of the so far presented 0–1 matrices. Notice that for a graph $\overrightarrow{G}(A)$, we hint at the ordering relation with an arrow — other than that, the graph $\overrightarrow{G}(A)$ is exactly the same as $G(A)$.

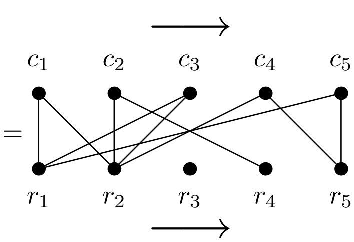
$\overrightarrow{G}(L_1) =$



$\overrightarrow{G}(A_5) =$



$\overrightarrow{G}(A'_5) =$



1.2 Extremal Function

Let us first define the weight of a 0–1 matrix:

Definition 1.5. *Assume that A is a 0–1 matrix. We define the weight $w(A)$ of A as the number of 1 entries that are contained in A.*

We define now the function $ex(n, P)$ that is based on the presented *contains* and *avoids* relations and that will be the basis for the investigations in the rest of this paper:

Definition 1.6. *P is a 0–1 pattern and A_n is an $n \times n$ 0–1 matrix. We define $ex(n, P)$ to be the maximum weight of the matrix A_n avoiding the pattern P.*

Let us state a few properties for $ex(n, P)$ that can be easily seen and that will be used in the rest of the paper:

1. **Trivial Lower Bound:** It is possible that an $n \times n$ matrix A_n contains zero 1 entries. The pattern Z attains this bound — it is easy to see that as soon as A_n contains a 1 entry, the pattern Z is contained in A_n.

   $$Z = \begin{pmatrix} \bullet \end{pmatrix}$$

   For all patterns that are not equivalent to Z, we have a trivial lower bound that is linear. A pattern that is not equivalent to Z has to contain at least two 1 entries. Assume first that the two 1 entries are in different rows. A matrix A_n with 1 entries in only one row clearly avoids the pattern. If all 1 entries of the pattern are in one row, then at least two different columns of the pattern have to contain 1 entries. In this case, we just fill one column of A_n with 1 entries. Combining the two cases, we can conclude that a linear lower bound for all patterns not equivalent to Z has been established.

2. **Trivial Upper Bound:** The extremal function $ex(n, P)$ is equivalent to the weight of an $n \times n$ matrix A_n that avoids P. A_n has maximum weight n^2 and this is the trivial upper bound for $ex(n, P)$. It is important to notice that there is no pattern that attains this trivial upper bound.

3. **Reflected Patterns:** For a pattern P of dimension $m \times n$, let us define the vertically reflected pattern $P^{|}$ as

$$\forall i,j : P^{|}_{i,j} = P_{i,n-j+1}$$

and the horizontally reflected pattern $\overline{P}$ as

$$\forall i,j : \overline{P}_{i,j} = P_{m-i+1,j}$$

There are two ways of reflecting a pattern P on the diagonal:

$$\forall i,j : P^{/}_{i,j} = P_{j,m-i+1}$$

and

$$\forall i,j : P^{\backslash}_{i,j} = P_{j,i}$$

The last reflection presented here is the central reflection:

$$\forall i,j : \dot{P}_{i,j} = P_{m-i+1,n-j+1}$$

All patterns P' that can be obtained from P by one of the above reflections are called *equivalent* and all those patterns have the same extremal function as P. This holds because the same transformation that is applied to P can as well be applied to the matrix A_n that avoids P with maximum weight. We get a transformed matrix A_n that avoids the transformed P.

4. **Containment:** If a pattern P contains a pattern Q — that means

$$\forall i,j : Q_{i,j} = 1 \Rightarrow P_{i,j} = 1$$

— then we have for the extremal function

$$ex(n,Q) \leq ex(n,P)$$

This holds because a matrix A_n avoiding Q avoids as well P.

5. **Removing Empty Rows and Columns:** If a pattern P can be obtained from a pattern Q by removing empty rows and columns, we have

$$ex(n,P) \leq ex(n,Q) = O(ex(n,P) + n)$$

We do not give a proof here but refer to [9], Theorem 2.2.

The extremal function $ex(n, P)$ has been investigated for some patterns. Füredi and Hajnal started in [7] a systematic investigation of $ex(n, P)$ for patterns with weight at most four and determined the asymptotical value of the $ex(n, P)$ for most of them. Tardos finished this work in [20] by providing the missing bounds. Together with Marcus, Tardos investigated in [14] $ex(n, P)$ of permutation matrices. Similar work has been done in [19] by Brass, Károlyi, and Valtr.

In [7], Füredi and Hajnal introduced reductions between patterns. By reducing a pattern P to a pattern Q, we get information about $ex(n, P)$ based on information about $ex(n, Q)$. This helps us to deduce information about new patterns based on patterns for which we already have an asymptotical value for $ex(n, Q)$. Keszegh introduced in [9] additional reductions. We list the most important reductions from the two mentioned papers in Appendix A.

It is important to notice that reductions are a first step towards kind of a theory. Together with a base set of patterns for which we know $ex(n, P)$, reductions can be used to explore the space of patterns. Although we know $ex(n, P)$ for all patterns with weight at most 4 (see Section 1.5) and therefore have a base set, we are not able to explore the whole space of patterns. One has to find more reductions and determine $ex(n, P)$ for whole classes of patterns. For a start on this, have a look at Section 1.4.

1.3 Turán Type Problems

Let us refer to a well studied problem — Turán type problems in extremal graph theory. For a graph G, one asks for the maximum number of edges that a graph with n vertices can have at most under the condition that it does not contain G as a subgraph. Let us denote this number by $ex(n, G)$. Note that the second parameter G is a graph opposed to Definition 1.6, where P is a pattern.

Theorem 1.7. *[4, 3] For a graph G, we have*

$$ex(n, G) = \left(1 - \frac{1}{\chi(G) - 1}\right) \binom{n}{2} + o(n^2)$$

with $\chi(G)$ the chromatic number of G.

This theorem is known as the Erdős-Stone-Simonovits Theorem and is important because it determines the asymptotical value of $ex(n, G)$ for all graphs that are not bipartite. It follows that the problem of determining the asymptotical value of $ex(n, G)$ is an open question only for bipartite graphs. In Section 1.1, we have seen that all patterns P correspond to a bipartite graph G.

It is important to notice that the Turán type problem is based on a graph G that has no ordering relation imposed on its vertices. This is in fact the difference to $ex(n, P)$ — which is based on a graph $\overrightarrow{G}$ with an ordering relation.

1.3.1 Bipartite Graphs

The graphs we obtain from patterns have not only an ordering relation on the vertices but they are as well bipartite. For a bipartite graph G without an ordering relation on the vertices, the only information we get from Theorem 1.7 is $ex(n, G) = o(n^2)$. A lot of work has been done to obtain a better result for $ex(n, G)$ but still many problems remain unsolved. Let us mention one result by Alon, Krivelevich, and Sudakov:

Theorem 1.8. *[17] For a bipartite graph G with two partite vertex sets of size m and n, $ex(n, G) \leq O(n^{2-\frac{1}{\min(m,n)}})$ holds.*

1.4 Permutation Matrix Patterns

Marcus and Tardos investigate in [14] permutation matrices P.

Definition 1.9. *A permutation matrix P has dimension $n \times n$, contains 0 and 1 entries, and has exactly one 1 entry per row and exactly one 1 entry per column.*

Marcus and Tardos show that for a pattern P that is a permutation matrix, $ex(n, P) = \Theta(n)$ holds. The key idea in their proof is to reduce the problem $ex(n, P)$ to $ex(\frac{n}{k^2}, P)$ — for some constant k. Assume that A_n is the matrix that avoids P with maximum weight. A_n is partitioned into blocks of size $k^2 \times k^2$ and every block is mapped to one entry — giving us a new matrix $A_{\frac{n}{k^2}}$. Marcus and Tardos show that if A_n avoids the pattern P, then so does $A_{\frac{n}{k^2}}$. This reduction gives the recursion

$$ex(n, P) \leq (k-1)^2 ex\left(\frac{n}{k^2}, P\right) + 2k^3 \binom{k^2}{k} n$$

that has the solution

$$ex(n,P) \le 2k^4 \binom{k^2}{k} n = O(n)$$

for a constant k.

The proof as presented by Marcus and Tardos is very short and simple. This is remarkable considering that $ex(n,P)$ for permutation matrices has been unknown for years and that the problem was considered to be important because both the Stanley–Wilf conjecture and the Alon–Friedgut conjecture [16] are closely related. Notice as well that this proof not only determines the extremal function for one pattern but rather for a class of patterns.

1.5 Patterns with at Most Four Entries

Füredi and Hajnal start the investigation of patterns with weight at most four in [7] and Tardos finishes this investigation in [20]. We characterize the patterns according to $ex(n,P)$ and notice that the patterns belong to one of five categories.

For all P with $w(P) \le 4$:

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

As we have noticed during the investigation of the trivial lower bound for the extremal function, there is exactly one pattern with $ex(n,P) = 0$ — it is the pattern Z that has exactly one 1 entry. The value of the extremal function for the two patterns S_1 and S_2 was proved by Füredi and Hajnal in [7] and we have $ex(n,S_i) = \Theta(n\alpha(n))$, where $\alpha(n)$ is the inverse Ackermann function. The proof uses a reduction to Davenport–Schinzel sequences and a result by Hart and Sharir from [8].

The fourth category contains the three patterns Q_1, Q_2, and Q_3 and all three of them have $ex(n,Q_i) = \Theta(n\log(n))$. For Q_1 and Q_2, this bound was proved by Füredi and Hajnal in [7]. For Q_3, the upper bound has been established in the before mentioned paper. The matching lower bound was proved by Tardos in [20].

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

The pattern R is the only one that has $ex(n, R) = \Theta(n^{\frac{3}{2}})$. We refer to Section 2.1 where we will show how the value for $ex(n, P)$ is obtained.

$$Q_1 = \begin{pmatrix} & \bullet & \bullet \\ \bullet & & \bullet \end{pmatrix} \qquad Q_2 = \begin{pmatrix} & \bullet & \bullet \\ & & \bullet \\ \bullet & & \end{pmatrix}$$

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

All patterns up to reflection that have not been listed so far have a linear extremal function. Notice that this includes as well patterns of weight two and three. We can conclude that most of the patterns with $w(P) \leq 4$ have a linear extremal function.

The presented results were not obtained with a theorem such as Theorem 1.7 but rather with a case analysis. Reductions as presented in Appendix A are a first step towards kind of a theory — but this theory is still very incomplete.

1.6 Patterns with Exactly Five Entries

There is no exhaustive analysis of patterns with weight five available but for some patterns, the extremal function is known and we present some of them here. It is important to notice that we do not present patterns that can be obtained from patterns P with $w(P) \leq 4$ using reductions as described in Appendix A.

Tardos investigated in [20] the pattern L_4 and concludes that it has $ex(n, L_4) = \Theta(n)$. In [9], Keszegh investigates the patterns L_3 and H_0 and shows that $ex(n, L_3) = \Theta(n)$ and $ex(n, H_0) = \Theta(n \log(n))$.

$$L_4 = \begin{pmatrix} \bullet & & \bullet & \\ & & & \bullet \\ \bullet & & & \\ & \bullet & & \end{pmatrix} \qquad H_0 = \begin{pmatrix} & \bullet & \bullet & \\ & & & \bullet \\ & & & \bullet \\ \bullet & & & \end{pmatrix}$$

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

Patterns with weight five play a role in two parts of this paper. In Section 3, we show that removing an arbitrary 1 entry from H_0 gives us a pattern with linear extremal function. In Section 4, we investigate a pattern with weight five for which the value of the extremal function is not known.

1.7 Exact Bounds

In this section, we list some of the patterns for which we know not only the asymptotic lower and upper bounds but exact bounds. Notice that the bounds presented here are not completely tight.

1.7.1 Patterns with at Most Three 1 Entries

For a pattern P with only two entries, we can show a lower bound of n and an upper bound of $n+1$. We argue with a case analysis:

1. **P has 1 entries in different rows and in different columns:** A matrix A_n with 1 entries only in the first row avoids the pattern. If the entry in the second row of P is to the left of the entry in the first row, then we can add an entry in the last column of A_n and the pattern is still avoided. If the entry in the second row of P is to the right of the entry in the first row, then we add a 1 entry in the first column of A_n and again, the pattern is still avoided.

2. P **has 1 entries in different rows:** A matrix A_n with 1 entries in only one row avoids the pattern but as soon as we add an additional entry to A_n, the pattern P is contained.

3. P **has** 1 **entries in different columns:** Same argumentation as for the first case but now we place the 1 entries into one column.

Füredi and Hajnal show in [7] that a pattern P with at most three 1 entries has

$$n \leq ex(n, P) \leq 4n.$$

1.7.2 Pattern L_1

For the pattern L_1, we can find a lower bound of $4n - 4$: a matrix A_n with 1 entries in only the last two rows and in the last two columns clearly avoids the pattern L_1. Assume that A_n avoids L_1 with maximum weight. In [20], Tardos shows an upper bound of $5n$. He does so by associating every entry of A_n with one of four categories. Three of those categories contain at most n entries and one contains at most $2n$ entries.

$$4n - 4 \leq ex(n, L_1) \leq 5n$$

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

1.7.3 Patterns Q_1, Q_2, and Q_3

In [20], Tardos shows for the pattern Q_1 the bounds

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

For Q_2, he proofs a bound of

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

For the pattern Q_3, again Tardos shows in [20] that for $n = 2^m$, we have

$$\frac{n\log(n)}{2} - O(n) \leq ex(n, Q_3) \leq n\log(n) + O(n).$$

1.8 Collection of Patterns

Tardos considers in [20] the extremal function of a collection of patterns:

Definition 1.10. *For a collection $\mathcal{P} = \{P_1, P_2, ...\}$ of patterns, we define $ex(n, \mathcal{P})$ to be the maximum weight that a matrix A_n can possibly have under the condition that it has to avoid* ***all*** *patterns in $\mathcal{P}$.*

Tardos shows in [20] that

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

With the fact that

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

and

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

it follows that

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

does **not** have to hold. Further investigations of collections of patterns can be found in Section 3 of [20]. In Section 4.2.3 of this paper, we investigate a collection of patterns.

2 Similar Settings

In this section, we consider two settings that are similar to $ex(n, P)$. By putting them into relation with the extremal function, we hope to get further insights. A proof in one of the considered settings might give us a proof for the setting that we are interested in.

2.1 Turán Setting

As we have mentioned in the introduction, Turán type problems in extremal graph theory are closely related to $ex(n, P)$:

Definition 2.1. *Assume that G is a graph. By $ex(n, G)$, let us denote the maximum number of edges that a a graph with n vertices can have under the condition that it does not contain G as a subgraph.*

Notice that for Turán type problems, the second parameter of $ex(n, G)$ is a graph whereas for $ex(n, P)$, the second parameter is a pattern.

In Definition 2.1, we consider arbitrary graphs with no ordering relation on the vertices and in Section 1.1, we have seen that such graphs can be represented as 0–1 matrices. For the rest of this section, let us assume that we are considering graphs $G(P)$ as defined in Section 1.1. To put $ex(n, G(P))$ into relation with $ex(n, P)$, we notice that for $ex(n, P)$, we are considering a graph $\overrightarrow{G}(P)$ with an ordering relation on the two partite vertex sets (see Section 1.1). This ordering relation does not exist for $ex(n, G(P))$ and we can conclude that in the Turán setting, a graph $G(A_n)$ with maximum number of edges has to avoid all possible *permutations* of a pattern P.

Definition 2.2. *Assume that P is a pattern with dimensions l and k. Further assume that $\pi_1 : \{1, ..., l\} \to \{1, ..., l\}$ and $\pi_2 : \{1, ..., k\} \to \{1, ..., k\}$ are permutations. A pattern P' with the same dimensions as P is called a permutation of P if only if we can find permutations π_1 and π_2 so that $\forall i, j : P_{i,j} = P'_{\pi_1(i),\pi_2(j)}$. For a pattern P, denote all permutations of P with $\Pi(P)$.*

Because the graph $G(A_n)$ has to avoid all possible permutations of $G(P)$, we can make the following proposition:

Proposition 2.3.

$$ex(n, G(P)) \leq \min_{U \in \Pi(P)} ex(n, U)$$

From this proposition, it follows that we obtain a lower bound for $ex(n, P)$ using $ex(n, G(P))$.

For the pattern R as presented in Section 1.5, we notice that all permutations of R are equivalent to R. From this observation, we can conclude that $ex(n, R) = ex(n, G(R))$. The pattern R equals to the complete bipartite graph $K_{2,2}$ with 2 vertices in each partite vertex set. It is well known from literature [2, 18] that $ex(n, G(R)) = ex(n, K_{2,2}) = \Theta(n^{\frac{3}{2}})$. Notice that the upper bound can as well be obtained from Theorem 1.8: $ex(n, K_{2,2}) = O(n^{2-\frac{1}{\min(2,2)}}) = O(n^{\frac{3}{2}})$.

2.2 Exact Match Setting

For the *Exact Match setting*, we replace the *contains* and *avoids* relations with *induced* and *not-induced* relations:

Definition 2.4. *A, B and P are 0–1 matrices of arbitrary dimension:*

1. *Matrix A induces P if and only if there is a submatrix B of A so that* $\forall i, j : P_{i,j} = B_{i,j}$.

2. *Matrix A is not-induced by P if and only if A does not induce P.*

Building on the induced relation, we define the extremal function for the Exact Match setting:

Definition 2.5. *P is a 0–1 pattern and A_n is an $n \times n$ 0–1 matrix. We define $\overline{ex}(n, P)$ to be the maximum weight of A_n not inducing P.*

We are interested in the relation between $ex(n, P)$ and $\overline{ex}(n, P)$. To analyze this relation, let us consider two types of patterns:

1. **The pattern P contains no 0 entries:** For such a pattern, the definitions for $ex(n, P)$ and $\overline{ex}(n, P)$ are exactly the same and $ex(n, P) = \overline{ex}(n, P)$ follows.

2. **The pattern P contains at least one 0 entry:** According to Definition 2.4, a matrix A_n containing n^2 entries does not induce the pattern P — $\overline{ex}(n, P) = n^2$ holds for all patterns P with at least one 1 entry.

From the two investigated cases, it follows that $ex(n, P) \leq \overline{ex}(n, P)$ for all patterns P. For most of the patterns P, the value $\overline{ex}(n, P)$ is the trivial upper bound of n^2 and therefore this setting does not help us to learn more about the extremal function $ex(n, P)$.

To construct a more interesting setting, let us switch to the point of view introduced in Section 1.1: One could restrict the number of edges that are allowed per vertex and require that the resulting graph has to be connected. It is important that we ask for a graph that is connected because it can be shown that otherwise, we can subdivide the problem and get again a trivial upper bound.

In [15], Chung, Jiang, and West consider a very similar setting. In their setting, the number of edges per vertex is restricted and they ask for connected graphs. But as a difference to the setting considered here, they do not have an ordering relation imposed on the vertices.

3 Minimally Non-Linear Patterns

In [20], Tardos asks for minimally non-linear patterns with more than four 1 entries. We start with the definition of minimally non-linear patterns:

Definition 3.1. *Let P be a pattern with $ex(n, P) = \omega(n)$. P is a minimally non-linear pattern if and only if replacing an arbitrary 1 entry with a 0 entry produces a pattern P' with $ex(n, P') = \Theta(n)$.*

As Marcus and Tardos note in [14], any pattern with at most four 1 entries that has a non-linear extremal function is minimally non-linear. This holds because all patterns with three 1 entries have linear extremal function.

Keszegh gives in [9] a construction that leads to a minimally non-linear pattern with more than four 1 entries — the pattern H_0. The construction is general enough so that it could even lead to infinitely many minimally non-linear patterns. We present Keszegh's construction together with the known facts and the open questions.

In [9], it is shown that for all $k \geq 0$, there is a pattern $H_k = (h_{i,j})$ with $ex(n, H_k) = \Theta(n \log(n))$. The pattern H_k has $m = 3k + 4$ rows and columns and 1 entries occur exactly at the following positions:

$$\{h_{4,1}, h_{1,2}, h_{1,3}, h_{m-1,m}, h_{m-2,m}\} \cup \{h_{3l+4,3l+1}, h_{3l-1,3l+3}, h_{3l,3l+2} | \forall l, 1 \leq l \leq k\}$$

All positions that have not been mentioned above have 0 entries.

$$H_0 = \begin{pmatrix} & \bullet & \bullet & \\ & & & \bullet \\ & & & \bullet \\ \bullet & & & \end{pmatrix} \qquad H_1 = \begin{pmatrix} & \bullet & \bullet & & & & \\ & & & & & \bullet & \\ & & & & \bullet & & \\ \bullet & & & & & & \\ & & & & & & \bullet \\ & & & & & & \bullet \\ & & & \bullet & & & \end{pmatrix}$$

3.1 H_0 is a Minimally Non-Linear Pattern

Let us consider the pattern H_0 and show that all patterns obtained from H_0 by removing one 1 entry have a linear extremal function. In the subsequent paragraphs, we will refer to reductions from Appendix A.

Let us start with the pattern R_0 that was obtained from H_0 by removing the bottom left 1 entry. The pattern R_{0a} can be obtained from Z using reductions as described in Theorem A.2 and has $ex(n, R_{0a}) = \Theta(n)$. Because R_0 is contained in R_{0a} (see Theorem A.1) and with the fact that R_0 has more than one 1 entry, we can conclude that $ex(n, R_0) = \Theta(n)$.

$$R_0 = \begin{pmatrix} \bullet & \bullet & \\ & & \bullet \\ & & \bullet \end{pmatrix} \qquad R_{0a} = \begin{pmatrix} \bullet & \bullet & \bullet \\ & & \bullet \\ & & \bullet \end{pmatrix}$$

The pattern R_1 is obtained from H_0 by deleting the first 1 entry in the first row. Notice that if we delete the second 1 entry in the first row, we get the exact same pattern. In [9], Keszegh shows that a pattern of the form R_{1a} has a linear extremal function. R_1 is contained in R_{1a} (see Theorem A.1) and we can therefore conclude that $ex(n, R_1) = \Theta(n)$.

$$R_1 = \begin{pmatrix} & \bullet & \\ & & \bullet \\ & & \bullet \\ \bullet & & \end{pmatrix} \qquad R_{1a} = \begin{pmatrix} \bullet & & \\ & \bullet & \\ & & \bullet \\ & & \bullet \\ & \bullet & \\ \bullet & & \end{pmatrix}$$

For the last two entries that can be removed, we notice that by removing either one of them, we get the same pattern R_2. The argumentation is the same as for R_1 — there is a pattern R_{2a} with linear extremal function that contains R_2 and $ex(n, R_2) = \Theta(n)$ follows.

$$R_2 = \begin{pmatrix} & \bullet & \bullet & \\ & & & \bullet \\ \bullet & & & \end{pmatrix} \qquad R_{2a} = \begin{pmatrix} & & \bullet & \bullet & & \\ & \bullet & & & \bullet & \\ \bullet & & & & & \bullet \end{pmatrix}$$

With the investigation of the patterns R_0, R_1, and R_2 we can conclude that H_0 is a minimally non-linear pattern.

3.2 Minimally Non-Linearity for Patterns H'_k

Let us show how we can use the pattern H_k to construct a pattern H'_k that is a candidate for a minimally non-linear pattern. In a second step, we will list what needs to be proved so that we can indeed conclude that H'_k is a minimally non-linear pattern.

3.2.1 Obtaining H'_k from H_k

Let us construct the matrix H'_k as follows:

1. Set $T_k = H_k$.
2. As long as T_k contains a 1 entry with the property that removing it produces a matrix T_k with $ex(n, T_k) = \omega(n2^{\alpha(n)^{O(1)}})$, then remove such an entry.
3. Set $H'_k = T_k$.

By the way how we constructed H'_k, we know that $ex(n, H'_k) = \omega(n2^{\alpha(n)^{O(1)}})$. In [9], Keszegh shows that there are $k + 5$ 1 entries in H_k with the property that removing any one of them gives us a pattern with a quasi-linear upper bound ($= O(n2^{\alpha(n)^{O(1)}})$) for the extremal function. We know additionally that the matrix H_k contains $3k + 5$ 1 entries. With those observations, we can conclude that $3k+5 = w(H_k) \geq w(H'_k) \geq k+5$. It is important to notice that we do not actually have to construct the H'_k — the above mentioned bounds are fully sufficient.

From the above bounds, we can conclude that there are infinitely many pairwise different patterns H'_k for different $k \in \{1, 2, ...\}$. This holds because we know lower and upper bounds on the weight of the pattern H'_k. So for example, it is not possible that H'_1 and H'_4 are the same pattern. This holds because

$$w(H'_1) \leq 8 < 9 \leq w(H'_4)$$

Notice that H'_k is a minimally non-quasi-linear pattern; i.e. removing an arbitrary entry from H'_k gives us a pattern with a quasi-linear upper bound.

3.2.2 Showing Minimally Non-Linearity for H'_k

For all patterns H'_k with $k \geq 1$, Keszegh shows that proving the property of minimally non-linearity comes down to show that a pattern G_k has linear extremal function. G_k is obtained from H_k by deleting the last three rows, the last column, and the column that contains the entry in the last row. Let us sketch why we can reduce the minimally non-linearity of H'_k to showing that G_k has a linear extremal function:

He basically shows that by removing the i'th 1 entry from H'_k, we get a matrix that is contained in a composition of two matrices A^i_k and B^i_k. All possible A^i_k's and B^i_k's are contained in G_l for some l. It follows that if we can find a linear extremal function for all G_k, then we can show, using a result by Keszegh from [9], that the composition of the two matrices A^i_k and B^i_k is as well linear. This composition contains the matrix that was obtained by removing the i'th 1 entry from H'_k — and it follows that proving minimally non-linearity for H'_k can be reduced to show linearity for G_k. For more details, we refer to [9].

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

Keszegh states in [9] the following conjectures:

Conjecture 3.2. *[9]*

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 $ex(n, G_k) = O(n)$ $(k \geq 1)$.*

3. *There are infinite [sic!] many minimally non-linear patterns.*

4. *The patterns H_k are minimally non-linear patterns.*

Notice that G_1 has almost the form of a permutation matrix; only the column containing the 1 entry in the first row is doubled. We call such a modified permutation matrix an *enriched permutation matrix*. For $k \geq 1$, all

G_k are enriched permutation matrices and therefore proving minimally non-linearity for H_k reduces to showing the linearity of enriched permutation matrices. Showing this linearity for the general enriched permutation matrix gives us infinitely many minimally non-linear patterns.

4 The Pattern G_1

To learn more about enriched permutation matrices, it might make sense to first consider a specific instance. We do that with the smallest instance that plays a role in Keszegh's proof and that is unknown — the pattern G_1.

In this section, we list all the knowledge that we got during the investigation of the pattern G_1. Our investigations did not lead to a tight bound but we would rather like to prevent you from trying the same paths. We start out with the known facts about the pattern G_1. Secondly, we list facts that are not closely related to G_1 but that might play an important role. Thirdly, we point out some proof strategies and try to highlight why, according to our intuition, they did not work out.

4.1 Known Facts

This section gives an overview on the known facts for the pattern G_1. First, we state the known lower and upper bounds — although they are not tight. In a second step, we observe what happens by removing an arbitrary 1 entry from G_1.

4.1.1 Lower Bound

Theorem 4.1. $ex(n, G_1) = \Omega(n)$

Proof. We use the proof idea mentioned for the trivial lower bound of any pattern: G_1 has entries in four rows and five columns. We can fill four columns of a matrix A_n entirely with 1 entries. This results in a matrix with $4n$ entries avoiding G_1 — $ex(n, G_1) = \Omega(n)$ is established. □

Notice that we have only established the trivial lower bound — nothing fancy. The main conclusion from this theorem is that we are still able to achieve a tight bound of $\Theta(n)$. We cannot only establish this lower bound for G_1 but for all enriched permutation matrices.

4.1.2 Upper Bound

In [10], Klazar gives a proof in the context of Davenport-Schinzel sequences that, as he points out in [12], can be used in the context of our setting. He uses a reductions from 0–1 matrices to DS sequences, argues on the maximum length of the sequence obtained from the 0–1 matrix and concludes about the maximum weight of the matrix A_n that avoids the pattern.

Theorem 4.2. *[12] For a pattern P with exactly one 1 entry per column, $ex(n, P) = O(n2^{\alpha(n)^{O(1)}})$ holds.*

The pattern G_1 contains exactly one 1 entry per column and Theorem 4.2 can therefore be applied. It is important to notice that $2^{\alpha(n)^{O(1)}}$ is an extremely slow growing function. We say that $O(n2^{\alpha(n)^{O(1)}})$ is a quasi-linear upper bound.

Again, this upper bound holds not only for G_1 but the very same argument gives us as well a quasi-linear upper bound for an enriched permutation matrix.

4.1.3 Removing a 1 Entry from G_1

Removing an arbitrary 1 entry from G_1 results in a pattern with linear extremal function. To show this, we use reductions from Appendix A and proceed with a case analysis:

The pattern RG_0 that can be obtained from G_1 by removing the left bottom entry has a linear extremal function. This holds because the pattern RG_{0a} is a permutation matrix and we can get RG_0 from RG_{0a} with the reduction from Theorem A.2.

$$RG_0 = \begin{pmatrix} \bullet & \bullet & & \\ & & & \bullet \\ & & \bullet & \end{pmatrix} \qquad RG_{0a} = \begin{pmatrix} \bullet & & \\ & & \bullet \\ & \bullet & \end{pmatrix}$$

We notice that removing either one of the two entries in the first row of G_1 gives us the same pattern RG_1. RG_1 is a permutation matrix and it follows that $ex(n, RG_1) = \Theta(n)$.

$$RG_0 = \begin{pmatrix} & \bullet & & \\ & & & \bullet \\ & & \bullet & \\ \bullet & & & \end{pmatrix}$$

Removing the 1 entry in the second row and removing the 1 entry in the third row produces both the pattern RG_2. RG_2 is contained in RG_{2a} (see Theorem A.1). Keszegh shows in [9] that the pattern RG_{2a} has a linear extremal function. We can therefore conclude that $ex(n, RG_2) = \Theta(n)$.

$$RG_2 = \begin{pmatrix} & \bullet & \bullet & \\ & & & \bullet \\ \bullet & & & \end{pmatrix} \qquad RG_{2a} = \begin{pmatrix} & & \bullet & \bullet & & \\ & \bullet & & & \bullet & \\ \bullet & & & & & \bullet \end{pmatrix}$$

This result can turn out to be interesting if a tight bound of complexity larger than linear is established for G_1. In this case, G_1 is a minimally non-linear pattern. Notice that the results from this section cannot directly be applied to enriched permutation matrices. One could ask the question whether it is possible to establish a similar result for arbitrary enriched permutation matrices.

4.2 Some Not Closely Related Facts about G_1

In this part, some facts are listed that we run into during our investigation of G_1. We do not claim that they have to be related to find a tight bound for G_1.

4.2.1 Surrounding of G_1

We investigate patterns in the surrounding of G_1 and define that the surrounding of a pattern P contains patterns that are obtained by moving one 1 entry in P by one position. To consider the surrounding of a pattern is motivated by the intuition that one could use a reduction, either one presented in Appendix A or one that is newly found, to determine $ex(n, G_1)$.

Permutation Matrices There are two permutation matrices in the surrounding of G_1 — both of them have $ex(n, P) = \Theta(n)$. The bound for the two matrices could be used to prove a linear bound for G_1. Unfortunately, there are no reductions known that can be used to deduce anything about G_1 based on the linear bound for SG_1 and SG_2.

$$SG_1 = \begin{pmatrix} & \bullet & & & \\ & & \bullet & & \\ & & & & \bullet \\ & & & \bullet & \\ \bullet & & & & \end{pmatrix} \qquad SG_2 = \begin{pmatrix} & & \bullet & & \\ & \bullet & & & \\ & & & & \bullet \\ & & & \bullet & \\ \bullet & & & & \end{pmatrix}$$

Pattern with Linear Extremal Function Besides the permutation matrices that we have just seen, the pattern SG_3 has as well a linear extremal function. In [9], Keszegh shows for the pattern SG_{3b} a linear extremal function. A linear extremal function for SG_{3a} can be obtained from SG_{3b} using the reduction mentioned in Theorem A.1. From SG_{3a}, we can obtain a linear bound for SG_3 using the reduction from Theorem A.2. As for the previous two patterns, there is no reduction known from SG_3 to G_1.

$$SG_3 = \begin{pmatrix} & \bullet & \bullet & & \\ & & & \bullet & \bullet \\ \bullet & & & & \end{pmatrix} \qquad SG_{3a} = \begin{pmatrix} & \bullet & \bullet & \\ & & & \bullet \\ \bullet & & & \end{pmatrix}$$

$$SG_{3b} = \begin{pmatrix} & & \bullet & \bullet & & \\ & \bullet & & & \bullet & \\ \bullet & & & & & \bullet \end{pmatrix}$$

Patterns with Quasi-Linear Upper Bound The four patterns SG_4, SG_5, SG_6, and SG_7 have an upper bound of $O(n2^{\alpha(n)^{O(1)}})$ (see Theorem 4.2). SG_5, SG_6, and SG_7 contain the pattern S_2 and we have therefore with Theorem A.1 a lower bound of $\Omega(n\alpha(n))$ established for them. As for the permutation matrices, there are no reductions from any of the three patterns to G_1.

For the pattern SG_4, we have only a linear lower bound and an upper bound of $O(n2^{\alpha(n)^{O(1)}})$.

Theorem 4.3. $ex(n, G_1) \leq ex(n, SG_4)$.

Proof. The reductions from Theorem A.1 and A.2 give us this relation. $\square$

It is important to notice that $ex(n, SG_4) = \Theta(n)$ gives us $ex(n, G_1) = \Theta(n)$.

$$SG_4 = \begin{pmatrix} & \bullet & \bullet & & \\ & & & & \bullet \\ \bullet & & & \bullet & \end{pmatrix} \qquad SG_5 = \begin{pmatrix} & & \bullet & & \\ & \bullet & & & \bullet \\ & & & \bullet & \\ \bullet & & & & \end{pmatrix}$$

$$SG_6 = \begin{pmatrix} & \bullet & & & \\ & & \bullet & & \bullet \\ & & & \bullet & \\ \bullet & & & & \end{pmatrix} \qquad SG_7 = \begin{pmatrix} & \bullet & \bullet & & \bullet \\ & & & \bullet & \\ \bullet & & & & \end{pmatrix}$$

Super-Linear Patterns In the surrounding of G_1, two patterns have a bound of $\Theta(n \log(n))$. Having already an upper bound of $O(n2^{\alpha(n)^{O(1)}})$ for G_1, we cannot expect to deduce anything about G_1 from the two patterns.

$$SG_8 = H_0 = \begin{pmatrix} & \bullet & \bullet & \\ & & & \bullet \\ & & & \bullet \\ \bullet & & & \end{pmatrix} \qquad SG_9 = \begin{pmatrix} & \bullet & \bullet & \\ & & & \bullet \\ & & \bullet & \\ \bullet & & & \end{pmatrix}$$

SG_8 is the same pattern as H_0 and therefore we have a tight bound of $\Theta(n \log(n))$. For SG_9, there is an easy argument that proves a tight bound of $\Theta(n \log(n))$ — let us show the lower and upper bound separately:

- **Lower Bound:** The pattern SG_9 contains the pattern Q_2 and we have therefore with Theorem A.1 a lower bound of $\Omega(n \log(n))$.

- **Upper Bound:** SG_{9a} can be constructed from Q_2 using the reductions described in Theorem A.2 and A.4. It follows that $ex(n, SG_{9a}) = \Theta(n \log(n))$ holds. To get SG_9 from SG_{9a}, we can employ the reduction described in Theorem A.1.

$$Q_2 = \begin{pmatrix} & \bullet & \bullet \\ & & \bullet \\ \bullet & & \end{pmatrix} \qquad SG_{9a} = \begin{pmatrix} & \bullet & \bullet & \\ & & \bullet & \bullet \\ & & \bullet & \\ \bullet & & & \end{pmatrix}$$

Patterns with Completely Unknown Extremal Function We do not have a tight bound for the pattern SG_{10}. The trivial linear lower bound of $\Omega(n)$ can be established — but no upper bound that is close to this lower bound is known. It is important to notice that a linear bound for SG_{10} gives us as well a linear bound for G_1:

Theorem 4.4. *$ex(n, G_1) \le ex(n, G_{10})$.*

Proof. The reductions from Theorem A.1 and A.2 give us this relation. □

$$SG_{10} = \begin{pmatrix} \bullet & \bullet & & \\ & & & \bullet \\ & & \bullet & \\ \bullet & & & \end{pmatrix}$$

4.2.2 Extremal Function for G_1 in the Turán Setting

As mentioned before, the relation

$$ex(n, G(P)) \le \min_{U \in \Pi(P)} ex(n, U)$$

holds between the main setting and the Turán setting. To learn more about the extremal function of G_1 in the Turán setting, we have to consider permutations of G_1. Finding a permutation with a linear extremal function gives us as well a linear extremal function for the Turán setting. This holds because we can argue for the Turán setting in the same way as we have argued for the main setting — we can construct a matrix with linear weight that avoids all permutations of the pattern G_1.

$$PG_1 = \begin{pmatrix} \bullet & \bullet & & & \\ & & & & \bullet \\ & & & \bullet & \\ & & \bullet & & \end{pmatrix}$$

It is easy to see that $PG_1 \in \Pi(G_1)$ can be obtained from a permutation matrix with the reduction described in Theorem A.2. We can conclude that $ex(n, PG_1) = \Theta(n)$ and that therefore $ex(n, G(G_1)) = \Theta(n)$.

4.2.3 Extremal Function of G_1 and $G_1^{|}$

Assume that $G_1^{|}$ is the reflection of G_1 at the vertical axis. In the rest of this section, we investigate the collection $\mathcal{P} = \{G_1, G_1^{|}\}$.

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

Theorem 4.5. *For the two patterns G_1 and $G_1^{|}$, $ex(n, \{G_1, G_1^{|}\}) = \Theta(n)$ holds.*

To prove this theorem, we use ideas mentioned in [9] and a reduction from the main setting to DS sequences[2] as mentioned by Füredi and Hajnal in [7].

Definition 4.6. *Let u be a sequence and $S(u)$ denote the set of symbols that appear in u. The length of u is written as $|u|$ and the number of symbols in u as $|S(u)|$. We say that two sequences $u = u_1u_2...u_n$ and $v = v_1v_2...v_n$ are equivalent if and only if there is a bijection $f : S(u) \rightarrow S(v)$ so that $\forall i \in \{1, ..., n\} : f(u_i) = v_i$ holds. The sequence u contains v if and only if v is equivalent to a subsequence of u. Otherwise, we say that u avoids v.*

As for the extremal function considered in this paper, we define an extremal function for DS sequences:

[2] Davenport-Schinzel sequences

Definition 4.7. *Define a close repetition in a DS sequence as a block of at least* $|S(u)|$ *symbols that are equal. Let* $ex(n, u)$ *be the maximum length of a sequence with* n *symbols that avoids the sequence* u *and that has no close repetition.*

So far, a setup for DS sequences that is similar to the main setting has been presented. We show now a mapping, as presented in [7], from 0–1 matrices to DS sequences and back. Let us first show how to transform a 0–1 matrix into a DS sequence:

Procedure 4.8. *Given a 0–1 matrix* A_n*. Number the rows from 1 to* n *and replace every 1 entry in* A_n *with the number of the row that it is located at — e.g. all 1 entries in the third row are replaced with 3. To construct the DS sequence* v'*, read the columns top-down starting with the first column. Ignore all 0 entries and add all other entries to* v'*.*

Notice that $|v'|$ *is equal to the weight of* A_n *(*$|v'| = w(A_n)$*), that* $|S(v')| \leq n$*, and that* v' *may contain close repetitions. If we have a close repetition, we remove up to one all entries of the block — e.g. abccccd becomes abcd. Using the knowledge of the construction of* v'*, we know that every removed entry can be mapped to a column and that every column has at most one removed entry. Because there are* n *columns, it follows that for the* v *that is constructed from* v' *by removing all close repetitions, we have* $|v| \geq w(A_n) - n$ *and* $|S(v)| \leq n$*.*

The transformation of a Davenport–Schinzel sequences to 0–1 matrices is rather simple:

Procedure 4.9. *To transform a DS sequence* v *to a 0–1 matrix* A_n*, read* v *and start populating the first column with 1 entries according to the row numbers mentioned in* v *— e.g.* $v = 1, 3, 5, ...$ *gives us a first column with 1 entries in the first, third, and fifth row. For the first column, do this as long as an increasing row number can be read from* v*. As soon as the row number is decreasing, start the same process over again with the second column and continue this way.*

The presented mapping helps us to use a theorem presented in [13] in the context of the main setting:

Theorem 4.10. *Suppose that* 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* $ex(n, u_1 a b^i a u_2 a b^i) = \Theta(ex(n, u))$ *for any* $i \geq 1$*.*

It is easy to see that $ex(n, a_1^9) = \Theta(n)$ holds[3]. Starting with a_1^9, we can build the string $u_l = a_1^3 a_2^3 ... a_{l-1}^3 a_l^6 a_{l-1}^3 ... a_2^3 a_1^6 a_2^3 ... a_{l-1}^3 a_l^3$ with the help of Theorem 4.10 and know that $ex(n, u_l) = \Theta(n)$ for all $l \geq 1$.

As a last building block, we need a lemma from [5] that is known as the Erdős-Szekeres Lemma:

Lemma 4.11. *Any sequence of numbers of length $(k-1)^2 + 1$ contains a monotone subsequence of length k.*

After this lengthy setup, we are ready to start with the main theorem of this section. First, we make the following observation:

Observation 4.12. *Assume that A_n is a 0–1 matrix and v has been obtained from A_n as described in Procedure 4.8. The pattern G_1 is contained in A_n if we find the string $g_1 = daaacb$ for $a < b < c < d$ in v. We obtain g_1 by mapping the pattern G_1 to DS sequences and by introducing the third a entry, we make sure that c is located in a column that is to the right of the column where the second a is located at.*

Similar reasoning lets us conclude that $G_1^{|}$ is contained in A_n if $g_1^{|} = bbcaaad$ for $a < b < c < d$ is contained in v.

$$\mathcal{G}_1 = \begin{pmatrix} & a & a & a & \\ & & & & b \\ & & & c & \\ d & & & & \end{pmatrix} \qquad \mathcal{G}_1^{|} = \begin{pmatrix} & & a & a & a \\ b & b & & & \\ & c & & & \\ & & & & d \end{pmatrix}$$

From now on we assume that $ex(n, \{G_1, G_1^{|}\}) = \omega(n)$ and that A_n is a matrix that avoids $\{G_1, G_1^{|}\}$ with maximum weight. In the following paragraphs, we will show a contradiction and can therefore conclude that $ex(n, \{G_1, G_1^{|}\}) = O(n)$. With the observation that there is a trivial lower bound of $ex(n, \{G_1, G_1^{|}\}) = \Omega(n)$, we will be able to conclude that for the collection $\{G_1, G_1^{|}\}$, we have $ex(n, \{G_1, G_1^{|}\}) = \Theta(n)$.

Let us transform the matrix A_n to v as seen in Procedure 4.8; it is known that $|v| \geq w(A_n) - n = \omega(n)$ and $|S(v)| \leq n$. Construct a string

$$u_l = a_1^3 a_2^3 ... a_{l-1}^3 a_l^6 a_{l-1}^3 ... a_2^3 a_1^6 a_2^3 ... a_{l-1}^3 a_l^3$$

[3] a_1^9 means that the symbol a_1 is repeated nine times

for $l = (4-1)^2 + 1 = 10$. We know that $ex(n, u_{10}) = \Theta(n)$ and therefore u_{10} is contained in v.

From Lemma 4.11, it is known that the string $w = a_1^3 a_2^3 ... a_9^3 a_{10}^3$ contains a monotone subsequence of length 4 — but we do not know whether this monotone subsequence is ascending or descending. We use a case analysis:

- **Subsequence is monotone ascending:** Assume that $a_{i(1)}$, $a_{i(2)}$, $a_{i(3)}$ and $a_{i(4)}$ are the four monotone ascending entries — $i(1) < i(2) < i(3) < i(4)$ and $a_{i(1)} < a_{i(2)} < a_{i(3)} < a_{i(4)}$ holds. With the fact $|v| = \omega(n)$, we know that the subsequence

$$u_{asc} = a_{i(1)}^3, a_{i(2)}^3, a_{i(3)}^3, a_{i(4)}^6, a_{i(3)}^3, a_{i(2)}^3, a_{i(1)}^6, a_{i(2)}^3, a_{i(3)}^3, a_{i(4)}^3$$

  is contained in v. Setting

$$\begin{aligned} a_{i(1)} &= a \\ a_{i(2)} &= b \\ a_{i(3)} &= c \\ a_{i(4)} &= d \end{aligned}$$

  we can conclude that the subsequence

$$g_{1,asc} = a^3, b^3, c^3, d^6, c^3, b^3, a^6, b^3, c^3, d^3$$

  is contained in v. But with the observation that the pattern $G_1^{|}$ can be represented as $g_1^{|} = bbcaaad$ for $a < b < c < d$ and from the fact that

$$g_{1,asc} = a^3, \mathbf{b}, \mathbf{b}, b, \mathbf{c}, c^2, d^6, c^3, b^3, \mathbf{a}, \mathbf{a}, \mathbf{a}, a^3, b^3, c^3, \mathbf{d}, d^2$$

  it follows that the pattern $G_1^{|}$ is contained in A_n — a contradiction to the assumption that A_n avoids both G_1 and $G_1^{|}$.

- **Subsequence is monotone descending:** The same argumentation as above is used but this time we make use of G_1. Assume that $a_{i(1)}$, $a_{i(2)}$, $a_{i(3)}$, and $a_{i(4)}$ is a monotone descending sequence — $i(1) < i(2) <$

$i(3) < i(4)$ and $a_{i(1)} > a_{i(2)} > a_{i(3)} > a_{i(4)}$ holds. Again with $|v| = \omega(n)$, we know that the subsequence

$$u_{desc} = a^3_{i(1)}, a^3_{i(2)}, a^3_{i(3)}, a^6_{i(4)}, a^3_{i(3)}, a^3_{i(2)}, a^6_{i(1)}, a^3_{i(2)}, a^3_{i(3)}, a^3_{i(4)}$$

is contained in v. Setting

$$\begin{aligned} a_{i(1)} &= d \\ a_{i(2)} &= c \\ a_{i(3)} &= b \\ a_{i(4)} &= a \end{aligned}$$

we conclude that the subsequence

$$g_{1,cont} = d^3, c^3, b^3, a^6, b^3, c^3, d^6, c^3, b^3, a^3$$

is contained in v. We know that the pattern G_1 can be represented as $g_1 = daaacb$ for $a < b < c < d$ and from the fact that

$$g_{1,cont} = \mathbf{d}, d^2, c^3, b^3, \mathbf{a}, \mathbf{a}, \mathbf{a}, a^3, b^3, \mathbf{c}, c^2, d^6, c^3, \mathbf{b}, b^2, a^3$$

it follows that the pattern G_1 is contained in A_n — a contradiction to the assumption that A_n avoids both G_1 and $G_1^{|}$.

From the two contradictions we can conclude that our initial assumption $ex(n, \{G_1, G_1^{|}\}) = \omega(n)$ is wrong and together with the trivial linear lower bound we can conclude that $ex(n, \{G_1, G_1^{|}\}) = \Theta(n)$ holds.

We can apply the very same proof strategy to other collections of patterns, e.g. we can show that:

Theorem 4.13. *For the two patterns G_1 and $\overline{G}_1$, $ex(n, \{G_1, \overline{G}_1\}) = \Theta(n)$ holds.*

For this collection of patterns, we do not give a proof here but mention that the DS equivalent of the pattern $\overline{G}_1$ is contained in the ascending subsequence.

$$\overline{G}_1 = \begin{pmatrix} \bullet & & & & \\ & & & \bullet & \\ & & & & \bullet \\ & \bullet & \bullet & & \end{pmatrix}$$

4.3 Proof Ideas

We consider in this section possible proof ideas for the pattern G_1. None of them gave us a linear upper bound but we would like to point out the intuition why they failed.

4.3.1 Adaption of the Permutation Matrix Proof

The pattern G_1 is almost a permutation matrix. It is possible that the proof for permutation matrices by Marcus and Tardos [14] can be tweaked in a way so that we can as well prove a linear bound for G_1.

Let us first present the most general ideas of this proof by Marcus and Tardos. In a second step, we will show why it is not obvious how to tweak this proof so that it can be used for G_1.

Original Setup The idea is to reduce the original problem to a problem of smaller size. This produces a recursive definition that we can solve by induction — and that turns out to be linear. For a permutation matrix of size $k \times k$, we reduce the problem of size n to a problem of size $\frac{n}{k^2}$.

Assume that we have a matrix A_n that avoids the pattern with maximum weight. We divide A_n into $(\frac{n}{k^2})^2$ blocks of size $k^2 \times k^2$. The blocks are defined as $S_{i,j} = \{a_{i',j'} : i' \in [k^2 i + 1, k^2(i+1)], j' \in [k^2 j + 1, k^2(j+1)]\}$. A block is called wide (respectively tall) if it contains 1 entries in at least k columns (respectively rows).

Based on A_n and its subdivision into blocks, we define a new matrix B. B has dimension $\frac{n}{k^2} \times \frac{n}{k^2}$ and has a 1 entry at $B_{i,j}$ if and only if there is at least one 1 entry in block $S_{i,j}$.

We present now the three properties that are the basis for the recursion:

1. The matrix B avoids the pattern P.

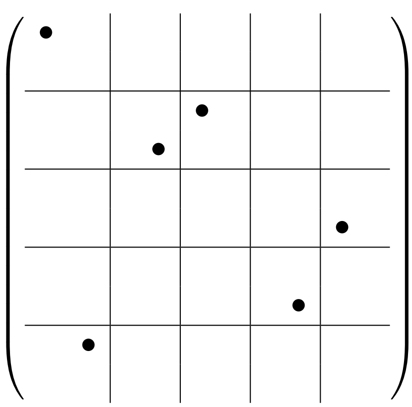
2. The number of blocks in $C_j = \{S_{i,j} : i = 1 ... \frac{n}{k^2}\}$ that are wide is less than $k\binom{k^2}{k}$.

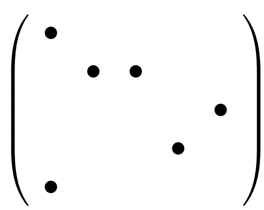
3. The number of block in $R_i = \{S_{i,j} : j = 1 ... \frac{n}{k^2}\}$ that are tall is less than $k\binom{k^2}{k}$.

Adaption for G_1 If we want to adapt the proof by Marcus and Tardos for our case, we have to analyze the three properties that the proof builds on:

1. **The matrix B avoids P:** Marcus and Tardos give the following proof for this property in the context of permutation matrices: Assume that B contains P. Consider the 1 entries that represent P. A 1 entry in B corresponds to a block in A_n that is non-empty. For every 1 entry in B that represents P, choose an arbitrary 1 entry in the corresponding block of A_n. A_n contains P as well — a contradiction.

   With A_{10} and B, we have a simple counter example showing us that this property does not hold for G_1.

   $A_{10} =$ 

   $B =$ 

   Notice that the failure to adapt this first property takes us the possibility to define a recursion. There might be an alternative way of defining a matrix B but intuitively, it is difficult to define an appropriate B because

   (a) we have to make sure that if A_n avoids P, then B avoids P as well

   (b) we have to show that most of the blocks containing 1 entries do not contain a lot of them

We could not find a definition of blocks so that the above conditions are fulfilled.

2. **The number of blocks in $C_j = \{S_{i,j} : i = 1...\frac{n}{k^2}\}$ that are wide is less than $k\binom{k^2}{k}$:** Marcus and Tardos argue with the pigeonhole principle that if there are $k\binom{k^2}{k}$ or more wide blocks in C_j, then we can find k blocks that have a 1 entry in the columns $c_1 < c_2 < ... < c_k$. From those k blocks, we can choose the 1 entries according to the permutation matrix — that means the permutation matrix is contained in A_n and we have a contradiction.

   It seems to be very difficult to adapt this proof for the pattern G_1. Let us argue with the $k = 5$ columns that contain 1 entries in 5 blocks. It is not entirely trivial to establish the condition that in the first of the 5 blocks, the 1 entries in columns c_2 and c_3 have to be in the same row.

3. **The number of blocks in $R_i = \{S_{i,j} : j = 1...\frac{n}{k^2}\}$ that are tall is less than $k\binom{k^2}{k}$:** Marcus and Tardos argue here the same way as they did for property 2. If there are more than $k\binom{k^2}{k}$ tall blocks in R_i, then we have rows $r_1 < r_2 < ... < r_k$ so that there are k blocks with a 1 entry in every of those rows. We can therefore find a representation for the pattern P and have a contradiction.

   In G_1, we have two 1 entries that are in the same row. The above idea guarantees that we can represent G_1 and we can therefore conclude that the third property holds for G_1.

Conclusion Only one of the three necessary properties can be adapted direclty for G_1. With the first property, the most important property cannot be adapted in the same way. The first property is the most important one because it allows us to define a recursion. Because this first property fails, the authors assume that it might be difficult to adapt the proof of permutation matrices for G_1.

4.3.2 Reduction to DS Sequences

The reduction from 0–1 matrices to DS sequences is an often used proof strategy, e.g. Füredi and Hajnal use it in [7] and Keszegh in [9]. This reduction to DS sequences allows us to reuse the theory of DS sequences.

We presented in Section 4.2.3 a proof for a linear bound for the collection $\{G_1, G_1^{|}\}$. This proof uses a reduction from 0–1 matrices to DS sequences and is historically the result of a failed proof for G_1. In the subsequent paragraphs, we will analyze why the proof works for $\{G_1, G_1^{|}\}$ but not for G_1.

In Section 4.2.3, we refer to the Erdős-Szekeres Lemma from which follows that in a sequence of $(k-1)^2+1$ different numbers, we have a monotone sequence of length at least k. We do not know whether this sequence is ascending or descending. With the counter examples $z_1 = 1, 2, 3, ..., (k-1)^2+1$ and $z_2 = (k-1)^2+1, (k-1)^2, ..., 3, 2, 1$ we can easily see that there are sequences where there is no descending respectively no ascending subsequence.

In the proof of a linear bound for the collection $\{G_1, G_1^{|}\}$, we use two building blocks from DS sequence theory:

1. The Erdős-Szekeres Lemma. It say that 'Any sequence of numbers of length $(k-1)^2+1$ contains a monotone subsequence of length k'.

2. The sequence $u_l = a_1^3 a_2^3 ... a_{l-1}^3 a_l^6 a_{l-1}^3 ... a_2^3 a_1^6 a_2^3 ... a_{l-1}^3 a_l^3$ for which we know that $ex(n, u_l) = \Theta(n)$ holds.

We argue with the Erdős-Szekeres Lemma that if we choose $l = (k-1)^2+1$, then we have a monotone subsequence of length k in $u_{build,l} = a_1^3 a_2^3 ... a_{l-1}^3 a_l^3$. The Erdős-Szekeres Lemma does not tell us whether the sequence is ascending or descending.

Assume that A_n is a matrix avoiding G_1 with maximum weight. A_n is mapped to a DS sequence v with $\Theta(w(A_n)) = \Theta(|v|)$. We proceed by case analysis:

1. Let us assume that the monotone subsequence in $u_{build,l} = a_1^3 a_2^3 ... a_{l-1}^3 a_l^3$ is **descending**. As we have shown in Section 4.2.3, the DS representation of G_1 is contained in u_l. Because we know that $ex(n, u_l) = \Theta(n)$ and by the fact that $\Theta(w(A_n)) = \Theta(|v|)$, we can conclude that $ex(n, G_1) = \Theta(n)$ holds for this case.

2. Assume now that the subsequence in $u_{build,l} = a_1^3 a_2^3 ... a_{l-1}^3 a_l^3$ is **ascending**. If we assume that $ex(n, G_1) = \omega(n)$, then we know that

$$g_{1,asc} = a^3, b^3, c^3, d^6, c^3, b^3, a^6, b^3, c^3, d^3$$

for $a < b < c < d$ is contained in v. As mentioned in Section 4.2.3, the DS sequence representation of G_1 is $g_1 = daaacb$ with $a < b < c < d$. It is easy to check that we cannot find g_1 in $g_{1,asc}$. We can conclude that the proof fails here.

There are two possible ways to fix this proof:

1. We can show that we have a descending monotone sequence in $u_{build,l} = a_1^3 a_2^3 ... a_{l-1}^3 a_l^3$. This might be possible using knowledge from our special setting but the intuition tells us that this is not very likely.

2. We can show that $u_{proof,l} = a_1^3 a_2^3 ... a_{l-1}^3 a_l^6 a_{l-1}^3 ... a_2^3 a_1^6 a_2^3 ... a_{l-1}^3 a_l^6 a_{l-1}^3 ... a_2^3 a_1^3$ has $ex(n, u_{proof,l}) = \Theta(n)$. Having this bound established for $u_{proof,l}$, we could argue for the case that we have an ascending subsequence in $u_{build,l}$ in exact the same way as for the descending subsequence. We just use the sequence $u'_l = a_l^3 a_{l-1}^3 ... a_2^3 a_1^6 a_2^3 ... a_{l-1}^3 a_l^6 a_{l-1}^3 ... a_2^3 a_1^3$ instead.

   As far as we know, no such result is known. It is important to notice that there is a result by Klazar from [11] that shows that $7n - 9 \leq ex(n, ab^2a^2b) \leq 8n - 7$. This result bounds a sequence linearly that is similar to the class $u_{proof,l}$.

We can conclude that this specific proof idea failed — although it might be possible to fix this proof. It is important to notice that there might be other proof strategies using reductions to DS sequences. We outlined here only one possibility.

4.3.3 Categorize Entries of A_n

In [20], Tardos proves a linear upper bound for L_1 by categorizing the entries of the matrix A_n that avoids L_1 with maximum weight. He shows that every 1 entry in A_n belongs to one of four categories. Every category contains only a linear number of entries and it follows that there is only a linear number of entries in A_n.

A similar proof idea is used by Füredi in [6]. He shows that every entry belongs to one of two categories and that there are at most $O(n \log(n))$ entries per category. With this observation, he proves an upper bound of $O(n \log(n))$.

We tried both to directly adapt the proof for L_1 and to use the general proof idea. All our trials failed. It is rather difficult to argue why they failed.

The intuition is that we were not able to find the correct categories and that the left lower entry of G_1 produces some difficulties.

4.3.4 Computationally Explore the Space of All Matrices

To computationally explore the space of all matrices with small dimension is not exactly a proof idea. But we mention it here nevertheless.

Assume that A_n avoids the pattern G_1 with maximum weight. The intuition is that if we know how the matrix A_n looks like for small n, we can maybe deduce something about a possible proof strategy. We started an investigation of small cases. For $n = 6$ we have $ex(6, G_1) = 30$. The matrix A_6 is an example for a matrix that attains this bound.

$$A_6 = \begin{pmatrix} \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & & \\ \bullet & \bullet & \bullet & \bullet & & \\ \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ & & \bullet & \bullet & \bullet & \bullet \end{pmatrix}$$

For $n = 7$, we were able to establish $ex(n, G_1) \geq 37$. The matrix A_7' attains this bound.

$$A_7' = \begin{pmatrix} \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \\ \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \\ & & \bullet & \bullet & \bullet & \bullet & \\ & & \bullet & \bullet & \bullet & \bullet & \bullet \\ & & \bullet & \bullet & \bullet & \bullet & \bullet \\ & & & \bullet & \bullet & \bullet & \bullet \end{pmatrix}$$

We were not able to establish an upper bound for $n = 7$. To do so, one had to test $\binom{49}{38} = 29135916264 \approx 29 * 10^9$ matrices. If we could test 10^6 matrices per second, this would take about 8 hours. The reason why we were not able to establish the upper bound is that our implementation currently takes 8 seconds to test 10^6 matrices. But one can conclude that with enough resources, it is possible to establish this upper bound.

For $n = 8$, we have $ex(8, G_1) \geq 44$. The matrix A_8' attains this bound but to establish an upper bound of 44, we had to test $\binom{64}{45} = 8.71987813 * 10^{15}$ matrices. Assuming again that we can test 10^6 matrices per second, this takes about 276 years. It follows that even for very small values of n, it is not possible to computationally explore $ex(n, G_1)$.

$$A'_8 = \begin{pmatrix} \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & & \\ \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & & \\ & & \bullet & \bullet & \bullet & \bullet & & \\ & & \bullet & \bullet & \bullet & \bullet & \bullet & \\ & & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ & & & \bullet & \bullet & \bullet & \bullet & \bullet \\ & & & & \bullet & \bullet & \bullet & \bullet \end{pmatrix}$$

5 Patterns Composed from a Linear Pattern and an Extremal Entry

In this section, we consider patterns that are composed from a pattern L with linear extremal function and from a single 1 entry. The pattern L is expanded with an empty first column and an empty last row. The 1 entry is added in the intersection of the newly added row and column. We say that all those patterns are in class $\mathcal{P}$.

$$\mathcal{P} \ni P = \begin{pmatrix} & \boxed{\;L\;} \\ \bullet & \end{pmatrix}$$

For all $P \in \mathcal{P}$, we try to establish an upper bound for the extremal function. Notice that the patterns S_2, H_0, and Q_2 are in class $\mathcal{P}$. With $ex(n, H_0) = \Theta(n \log(n))$ and $H_0 \in \mathcal{P}$, it follows that we can only hope to establish $ex(n, P) = O(n \log(n))$.

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

$$S_2 = \begin{pmatrix} & \bullet & & \bullet \\ & & \bullet & \\ \bullet & & & \end{pmatrix} \qquad H_0 = \begin{pmatrix} & \bullet & \bullet & \\ & & & \bullet \\ & & & \bullet \\ \bullet & & & \end{pmatrix}$$

Conjecture 5.1. *All patterns $P \in \mathcal{P}$ have $ex(n, P) = O(n \log(n))$.*

To establish an upper bound for the extremal function of patterns from $\mathcal{P}$, we have to consider all possible $n \times n$ matrices and argue that they cannot contain more than $O(n \log(n))$ entries while avoiding an arbitrary pattern $P \in \mathcal{P}$. Let us characterize the $n \times n$ matrices by the last entry per column. For column j, we denote the row index of the last entry in this column by $l(j)$:

$$l(j) = \begin{cases} 0 & : \quad \forall i : 1 \le i \le n \wedge P_{i,j} = 0 \\ i & : \quad P_{i,j} = 1 \wedge \forall k > i : P_{k,j} = 0 \end{cases}$$

In the next two sections, we will investigate two classes of matrices that are characterized with the help of $l(j)$.

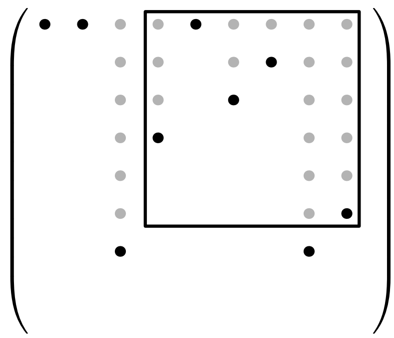
5.1 Extreme Case

For a matrix A_n that avoids a $P \in \mathcal{P}$, we say that $A_n \in \mathcal{E}$ if and only if there is a k with $1 \le k \le n$ so that

$$\forall i < k : l(i) \le 1 \wedge \forall j > k : l(k) \ge l(j)$$

By definition we know that in the first k columns, there are at most k 1 entries. Assume that $(A_n)_{l(k),k}$ represents the left lower 1 entry of P. Then we know that in the intersection of the last $n-k-1$ columns and the first $l(k)-1$ rows, there are at most $O(\max(n-k-1, l(k)-1))$ 1 entries — otherwise, the pattern P is contained in A_n. Only the last 1 entries per column and the 1 entries in the k'th column can be outside of the covered area and it follows that for all matrices $A_n \in \mathcal{E}$ avoiding $P \in \mathcal{P}$, they contain at most $n + n + O(\max(n-k-1, l(k)-1)) = O(n)$ 1 entries.

$\mathcal{E} \ni A_n =$



5.2 $l(j)$ on Diagonal

Let us consider the class of matrices $\mathcal{M}$ that have the last entries $l(j)$ on the diagonal from the left upper to the right lower corner — or said differently:

$$\forall j : l(j) = j.$$

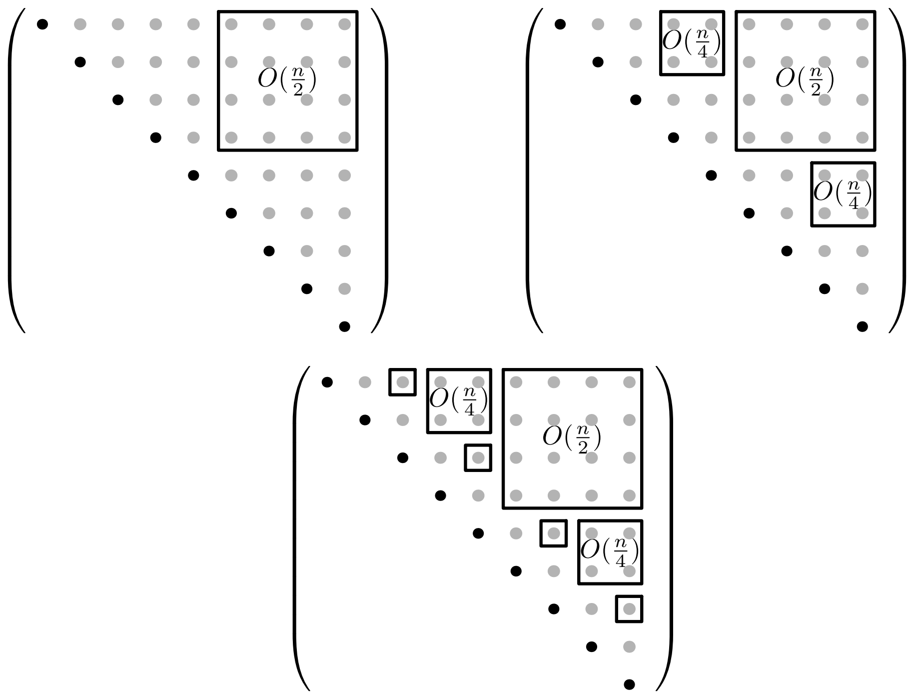
We show that a matrix $M \in \mathcal{M}$ avoiding a pattern $P \in \mathcal{P}$ has at most $O(n \log(n))$ entries.

For an $n \times n$ matrix $A_n \in \mathcal{M}$ avoiding $P \in \mathcal{P}$ with maximum weight, we say that s_i is the square that covers the area of the intersection of the first $i-1$ rows and the last $i-1$ columns. We say that s_i is rooted at the entry $(A_n)_{i,i}$.

We state the key observation that lets us establish an upper bound of $O(n \log(n))$ 1 entries for a matrix A_n avoiding $P \in \mathcal{P}$:

Assume that $A_n \in \mathcal{M}$ and we can conclude that $(A_n)_{l(\frac{n}{2}),\frac{n}{2}} = 1$. Further assume that the left lower entry of P is represented by $(A_n)_{l(\frac{n}{2}),\frac{n}{2}}$. Notice that $s_{\frac{n}{2}}$ lies entirely to the right and above $(A_n)_{l(\frac{n}{2}),\frac{n}{2}}$ — that means that if $s_{\frac{n}{2}}$ contains L, then P is as well contained in A_n and we have a contradiction. L is a pattern with linear extremal function and it follows that there are at most $O(\frac{n}{2})$ 1 entries in an $s_{\frac{n}{2}}$ that avoids L. Notice that for a constant n, there might be a constant number of 1 entries in $s_{\frac{n}{2}}$.

Observe now that the square $s_{\frac{n}{2}}$ does not only give us an upper bound on the number of 1 entries that can be contained in $s_{\frac{n}{2}}$ but that it divides as well the problem into two subproblems — have a look at the figures. In the figures, we denote the 1 entries on the diagonal with black points. Grey points symbolize possible locations of points. The figures show the squares s_i after the first, second, and third level of the recursion.



Let $E(n)$ be the number of 1 entries that are covered with squares s_i. We formulate the above mentioned idea with a recursive formula:

$$E(n) \leq \begin{cases} C & : \quad n \leq n_0 \\ k\frac{n}{2} + 2E(\frac{n}{2}) & : \quad otherwise \end{cases}$$

For constant n's, we have only a constant number of 1 entries in the matrix A_n. Let us say that all $n \leq n_0$ are of constant size. It follows that we can choose a constant C so that $\forall n \leq n_0 : E(n) \leq C$.

For $n > n_0$ and a constant k that is dependent on the pattern L, we have at most $k\frac{n}{2}$ 1 entries in the square $s_{\frac{n}{2}}$. This holds because we know that L is a pattern with $ex(n, L) = \Theta(n)$. To obtain the two subproblems, we basically divide the matrix A_n into four quarters. With the square s_i, we have already considered the right upper quarter. We know that there are no 1 entries in left lower quarter. Remain the left upper and right lower quarter. Those two quarters have the exact same form as the original matrix — we have our two subproblems of size $\approx \frac{n}{2}$. To solve the recursion, we assume that $E(n) \leq An\log(n) + Cn$. For $n \leq n_0$,

$$E(n) \leq C \leq An\log(n) + Cn$$

holds for all $A \geq 0$ and we have therefore establish the base case. It can be shown with induction that there is an $A \geq 0$ so that for arbitrary n,

$$E(n) \leq ... \leq An\log(n) + Cn = O(n\log(n))$$

holds.

So far, we have counted all 1 entries that are covered by the s_i's in the recursive process. In total, we have three categories of entries:

1. The entries that are covered with squares s_i. As we have shown, there are at most $O(n\log(n))$ such 1 entries.

2. The entries $A_{l(j),j}$. There is exactly one such entry per column and therefore we have n 1 entries in this category.

3. The entries $A_{l(j)-1,j}$ are not covered by the s_i. Again, we have at most one such 1 entry per column and it follows that there are at most n 1 entries in this category.

We can conclude that a matrix $M \in \mathcal{M}$ avoiding the pattern $P \in \mathcal{P}$ has at most $O(n\log(n))$ 1 entries. To prove $ex(n, P) = O(n\log(n))$ for all $P \in \mathcal{P}$, we had to show that the matrix A_n avoiding P with maximum weight is contained in the class $\mathcal{M}$.

5.3 Adaption for Arbitrary Matrices

We see two possible ways to adapt the result from the last section for arbitrary matrices $A_n \notin \mathcal{M}$:

1. Show that the class $\mathcal{M}$ contains the matrix avoiding $P \in \mathcal{P}$ with maximum weight.

2. Solve the recursion for the general case and show that there are at most $O(n\log(n))$ entries contained in a matrix A_n avoiding P.

Let us consider the two cases and argue why it might be difficult to get a solution using the considered approach:

1. To explore the space of matrices around the class $\mathcal{M}$ seems like a good idea. If one could show that for every $N \notin \mathcal{M}$, there is a series of steps from an $M \in \mathcal{M}$ to N so that we leave the weight unchanged or decrease it in every step, then we knew that the class $\mathcal{M}$ contains the matrix A_n avoiding P with maximum weight. Unfortunately, one soon realizes that there are matrices $N \notin \mathcal{M}$ where we have more 1 entries than for the $M \in \mathcal{M}$ with maximum weight avoiding P. It follows that there cannot be a stepwise argument as mentioned above. The hope remains that there is no matrix that has an asymptotically higher weight than $O(n \log(n))$.

2. Solving the recursion as given in the previous section for the general case does not give us an upper bound of $O(n \log(n))$. The intuition is that using only the squares s_i to estimate the number of 1 entries to the right and above of a 1 entry $A_{l(j),j}$ is not sufficient. One rather has to include the whole region to the right and above the entry $A_{l(j),j}$. This turns out to be difficult because by including the whole region to the left and above, we might count 1 entries more than once. We could not find a good way to employ the inclusion / exclusion principle — but this might be a good starting point.

We were not able to prove Conjecture 5.1 with either of the two mentioned strategies. It might be that there is different strategy that can be used to prove an upper bound for $P \in \mathcal{P}$.

6 Conclusion

We considered an extremal function as it was presented in [7] and gave an overview of the results. The connection to Turán type problems and to an extremal function based on the *induced* relation has been mentioned. We then gave an overview on minimally non-linear patterns and presented results by Keszegh. Motivated by an open problem brought up by Keszegh, we investigated the pattern G_1. We listed several properties of G_1 and sketched a few proof ideas. In the last section, we conjectured an upper bound for a class of patterns and showed bounds on the weight of special matrices.

7 Acknowledgement

The author thanks Robert Berke for the uncountable discussions and remarks concerning this thesis. Thanks a lot!

A Reductions

In this section, we present the known reductions for the setting defined in Definition 1.6. We do not state proofs here but rather point to the papers where the proofs can be found. Assume for the following that P and P' are patterns.

Theorem A.1. *[7] If P contains P', then $ex(n, P') \leq ex(n, P)$.*

Theorem A.2. *[7] 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 $ex(n, P) \leq ex(n, P') \leq ex(n, P) + n$.*

Theorem A.3. *[7] If we get P' by removing all blank rows and columns from P, then $ex(n, P') \leq ex(n, P) \leq O(ex(n, P') + n)$.*

Theorem A.4. *[7] 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 $ex(n, P) \leq ex(n, P') \leq 2ex(n, P)$.*

Theorem A.5. *[9] 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 $\max(ex(n, A), ex(n, B)) \leq ex(n, C) \leq ex(n, A) + ex(n, B)$.*

Theorem A.6. *[9] 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 $\max(ex(n, A), ex(n, B)) \leq ex(n, C) \leq 16(n + ex(n, A) + ex(n, B))$.*

Theorem A.7. *[9] 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 $ex(n, A') = O(ex(n, A))$.*

B Task Formulation

A submatrix of a matrix A is obtained from A by deleting rows and columns. 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. The quantity of interest here is the maximal number of 1 entries in a n by n matrix avoiding a pattern P, let us denote it by $ex(n, P)$. In [14] they show that if P is a permutation matrix then $ex(n, P) = O(n)$. Moreover the order of magnitude of $ex(n, P)$ for all patterns with four 1 entries have been investigated in [7] and in [20].

Goal: The goal of this thesis is to investigate $ex(n, P)$ for other small patterns. Füredi and Hajnal [7] ask for the characterization of all patterns P with $ex(n, P) = O(n)$. Here we would like to find minimal non-linear patterns P with five 1 entries (minimality w.r.t number of 1-entries), $ex(n, P) = \omega(n)$ or show that they don't exist.

Contact Person: Robert Berke, berker@inf.ethz.ch, CAB G37.2, Tel: (044) 632-7182

References

[1] Dan Bienstock and Ervin Györi. An extremal problem on sparse 0-1 matrices. *SIAM J. Discrete Math*, 4:17–27, 1991.

[2] W. G. Brown. On graphs that do not contain a thomsen graph. *Canad. Math. Bull.*, 9:281–285, 1966.

[3] Paul Erdős and Miklós Simonovits. A limit theorem in graph theory. *Stud. Sci. Math. Hungar. 1*, pages 51–57, 1966.

[4] Paul Erdős and Arthur Harold Stone. On the structure of linea graphs. *Bull. Amer. Math. Soc.*, 52:1087–1091, 1946.

[5] Paul Erdős and George Szekeres. A combinatorial problem in geometry. *Compocito Math.*, 2:464–470, 1935.

[6] Zoltán Füredi. The maximum number of unit distances in a convex n-gon. *J. Combin. Theory Ser. A*, 55:316–320, 1990.

[7] Zoltán Füredi and Péter Hajnal. Davenport schinzel theory of matrices. *Discrete Mathematics*, 103:233–251, 1992.

[8] S. Hart and M. Sharir. Nonlinearity of davenport-schinzel sequences and of generalized path compression schemes. *Combinatorica*, 6:151–177, 1986.

[9] Balázs Keszegh. Forbidden submatrices in 0-1 matrices. Master's thesis, Eötvös Loránd University, 2005.

[10] Martin Klazar. A general upper bound in extremal theory of sequences. *Comment. Math. Un. Carolinae*, 33, 1992.

[11] Martin Klazar. Extremal functions for sequences. *Discrete Math.*, 150:195–203, 1996.

[12] Martin Klazar. Enumerative and extremal combinatorics of a containment relation of partitions and hypergraphs. 2003.

[13] Martin Klazar and Pavel Valtr. Generalized davenport-schinzel sequences. *Combinatorica*, 14(4):463–476, 1994.

[14] Adam Marcus and Gábor Tardos. Excluded permutation matrices and the Stanley–Wilf conjecture. *J. Combin. Theory Ser. A*, 107(1):153–160, 2004.

[15] Tao Jiang Myung S. Chung and Doublas B. West. Induced turán problems: Largest p_m-free graphs with bounded degree. 2005.

[16] Ehud Friedgut Noga Alon. On the number of permutations avoiding a given pattern. *J. Combin. Theory, Ser A*, 89:133–140, 2000.

[17] Michael Krivelevich Noga Alon and Benny Sudakov. Turn numbers of bipartite graphs and related ramsey-type questions. *Combinatorics, Probability and Computing*, 12:477–494, 2003.

[18] Alfréd Rényi Paul Erdős and V.T. Sós. On a problem of graph theory. *Stud. Sci. Math. Hungar*, 1:215–235, 1966.

[19] George Károlyi Peter Brass and Pavel Valtr. A turán-type extremal theory of convex geometric graphs. *Discrete and Computational Geometry — The Goodman-Pollack Festschrift*, pages 275–300, 2003.

[20] Gábor Tardos. On 0-1 matrices and small excluded submatrices, manuscript. 2004.