Small Forbidden Configurations II

Richard Anstee*
Ron Ferguson†
Mathematics Department
The University of British Columbia
Vancouver, B.C. Canada V6T 1Z2
anstee@math.ubc.ca

Attila Sali‡
Department of Computer Science
Indiana University-Purdue University Fort Wayne
Fort Wayne, Indiana, 46805-1499

AMS Subject Classification: 05D05 (primary), 05C20, 05C90 (secondary)

Submitted: July 20, 2000; Accepted: October 3, 2000

Abstract

The present paper continues the work begun by Anstee, Griggs and Sali on small forbidden configurations. In the notation of (0,1)-matrices, we consider a (0,1)-matrix F (the forbidden configuration), an $m \times n$ (0,1)-matrix A with no repeated columns which has no submatrix which is a row and column permutation of F, and seek bounds on n in terms of m and F. We give new exact bounds for some $2 \times l$ forbidden configurations and some asymptotically exact bounds for some other $2 \times l$ forbidden configurations. We frequently employ graph theory and in one case develop a new vertex ordering for directed graphs that generalizes Rédei's Theorem for Tournaments. One can now imagine that exact bounds could be available for all $2 \times l$ forbidden configurations. Some progress is reported for $3 \times l$ forbidden configurations. These bounds are improvements of the general bounds obtained by Sauer, Perles and Shelah, Vapnik and Chervonenkis.

*Research supported in part by NSERC

†Research supported in part by NSERC

‡This research was done while the third author visited the University of British Columbia supported by the first author's NSERC grant. Permanent address: Alfréd Rényi Institute of Mathematics, The Hungarian Academy of Sciences, Budapest, P.O.B. 127 H-1364, Hungary, sali@renyi.hu.

1 Introduction

This paper continues investigations of Anstee, Griggs, Sali [4] into extremal set problems arising from forbidding a single configuration. The reader might consider the analogy with the celebrated results of Erdős and Stone [5] who determined asymptotic bounds on the number of edges (in terms of the number of vertices) in a graph avoiding a single given subgraph (based on the chromatic number of the forbidden subgraph). Our results provide bounds that are remarkably accurate for small forbidden configurations but we still have some small forbidden configurations for which we do not know the asymptotic bounds. The results are examples of a general pattern as yet not fully understood to the same extent as the Erdős-Stone results.

A natural notation for these problems is (0,1)-matrices. Forbidden configurations have been studied by various authors for a long time, because a great number of combinatorial objects can be encoded as $(0,1)$-matrices with forbidden substructures. We will use the term *configuration* (the combinatorial equivalent of submatrix) as follows. For a matrix F, we say a matrix A has no configuration F if A has no submatrix which is a row and column permutation of F. Let F be a $k \times l$ $(0,1)$-matrix and let A be an $m \times n$ $(0,1)$-matrix with no repeated columns (such matrix is called *simple*). The maximum number of *columns* of a simple matrix A of m rows with no configuration F will be denoted by $\mathrm{forb}(m, F)$. Obviously $\mathrm{forb}(m, F) \leq 2^m$ but more can be done.

This paper is noteworthy in establishing a number of best possible bounds and some asymptotically best possible bounds for various $2 \times l$ forbidden configurations F. Careful counting arguments are used and then reused in searching for examples close to the bounds. We are getting close to providing exact bounds for any $2 \times l$ configuration F, which before seemed hopeless. Section 2 focuses on $2 \times l$ F yielding linear bounds. Section 3 focuses on $2 \times l$ F yielding quadratic bounds. The final section provides a new result for general F, (multiple copies of the identity matrix) which in particular shows that for

$$F = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

that $\mathrm{forb}(m, F) = O(m^{7/3})$. We would tend to believe that $\mathrm{forb}(m, F) = \Theta(m^2)$, and must leave this as an open problem.

We extensively use graph theory for $2 \times l$ forbidden configurations F. Lemma 2.1, which generalizes Rédei's result for hamiltonian paths in tournaments, provides an interesting new vertex ordering for directed graphs. The applications suggest that maybe there is a hypergraph generalization to aid in obtaining bounds for general forbidden configurations.

The following result of Füredi is a general asymptotic bound.

Theorem 1.1 ([9]) *Let F be a $k \times l$ $(0,1)$-matrix. Then*

$$forb(m, F) = O(m^k).$$

■

This is best possible for F being the $k\times 2$ matrix of ones. To obtain a proof one can use the following fundamental result of Sauer, Perles and Shelah, Vapnik and Chervonenkis. Let K_k denote the $k\times 2^k$ submatrix of all possible columns of size k.

Theorem 1.2 ([10, 12, 13])

$$\mathit{forb}(m, K_k) = \binom{m}{k-1} + \binom{m}{k-2} + \dots \binom{m}{0}.$$

∎

It is easy to see that the bound of Theorem 1.2 is sharp: take A to be the matrix of all columns with $k-1$ or fewer 1's. To obtain the asymptotic estimate for arbitrary forbidden configurations F one can do the following. Given a matrix A with at least

$$(2^k)(l-1)\binom{m}{k} + \binom{m}{k-1} + \binom{m}{k-2} + \dots \binom{m}{0} + 1$$

columns, we take any $\binom{m}{k-1} + \binom{m}{k-2} + \dots \binom{m}{0} + 1 = \mathrm{forb}(m, K_k) + 1$ columns which then contain the configuration K_k in 2^k columns. Permute the columns so that these 2^k columns are the first ones. Repeat on the remaining columns until one has $(l-1)\binom{m}{k} + 1$ such K_k's and hence by the Pidgeonhole Principle, at least l in the same k subset of rows and hence a copy of F.

One can obtain best possible or at least more accurate bounds (see [3]). For example, we have for F being the $k\times 2$ matrix of ones

$$\mathrm{forb}(m, F) = \binom{m}{k} + \binom{m}{k-1} + \dots \binom{m}{0} = \Theta(m^k).$$

Also if F has no repeated columns, then

$$\mathrm{forb}(m, F) \le \binom{m}{k-1} + \binom{m}{k-2} + \dots \binom{m}{0} = \Theta(m^{k-1}),$$

but the bound need not be best-possible. For other results or generalizations see [1, 2, 6, 7].

2 Linear Bounds

Careful structural analysis has resulted in bounds that are new and surprisingly exact. The following graph theory argument is the key to the exact bound in the Theorem 2.2 that follows. One can envision this as a generalization of Rédei's result [11] that a tournament has a hamiltonian path. While this result is for directed graphs, it is interesting to contemplate whether there are k-uniform hypergraph analogues useful for $k\times l$ F and in particular to use in Theorem 4.1.

Lemma 2.1 *Let $D = (N, A)$ be a directed graph. There is an ordering of the vertices N as $1, 2, \ldots m$ where $m = |N|$ and a subset $T \subseteq A$ consisting of a collection of vertex disjoint indirected trees T with the following property. Let D_i denote the subgraph of D induced by the vertices $\{i, i+1, \ldots m\}$. For each pair i, j, $1 \le i < j \le m$ either there is a directed path in D_i from i to j or there is a k with $i \le k \le m$ so that there is a directed path from i to k in D_i and there is no edge in D from k to j.*

Proof: We proceed by constructing a forest of indirected trees T from D and a vertex ordering in the following way. As vertices are deleted, they enter the vertex ordering; namely the kth vertex deleted is labelled k. We start with $T = \emptyset$. At the jth stage we form a directed path P_j from P_{j-1} by extending it. Let $P_j = i_{j,1} \to i_{j,2} \to \ldots \to i_{j,l_j}$. It will have what we call the *two-maximal* property that for a vertex $k \notin P_j$ (and not yet deleted) there is no edge $k \to i_{j,1}$, there is no pair of edges $i_{j,t} \to k$, $k \to i_{j,t+1}$ for all t with $1 \le t \le l_j - 1$ and there is no edge $i_{j,l_j} \to k$. We initially start with $P_0 = \emptyset$.

At the $n+1$st step we look at P_n. In the case $P_n = \emptyset$, we simply choose P_{n+1} as a two-maximal path in the remaining graph (P_{n+1} could consist of a single vertex). For $P_n = i_{n,1} \to i_{n,2} \to \ldots \to i_{n,l_n}$, we find the smallest index k_n such that there is an edge $j \to i_{n,k_n}$ for a vertex j (not in P_n and not already deleted). If there is no vertex k_n, then set $P_{n+1} = \emptyset$ while deleting the vertices $i_{n,1}, i_{n,2}, \ldots i_{n,l_n}$ in turn. Otherwise add the edges $i_{n,1} \to i_{n,2}$, $i_{n,2} \to i_{n,3}, \ldots$ $i_{n,k_n-1} \to i_{n,k_n}$ to T and delete the vertices $i_{n,1}, i_{n,2}, \ldots i_{n,k_n}$ in turn (note that by the two maximal property for P_n that $k_n > 1$).

Extend P_n to a two-maximal path $i_{n+1,1} \to \ldots \to i_{n+1,e_{n+1}} \to i_{n,k_n} \to \ldots \to i_{n,l_n}$. This we relabel as our new path P_{n+1} $i_{n+1,1} \to i_{n+1,2} \to \ldots \to i_{n+1,l_{n+1}}$.

This process continues until there are no vertices left. Does the resulting ordering and set T have the desired property? Note that each path P_n has all its edges eventually enter T and the vertex ordering respects the vertex ordering in the path. Choose any s satisfying $1 \le s \le k_n - 1$. Using the two-maximal property, we verify that for every vertex j not in P_n and yet still undeleted we note that either there is a vertex $i_{n,t}$ in P_n (with $s \le t \le l_n$) where there is no edge (of D) joining $j, i_{n,t}$ or we get a contradiction. This follows since by the choice of k_n, we have $i_{n,s} \to j$ and using two-maximality we get that $i_{n,s+1} \to j, i_{n,s+2} \to j, \ldots, i_{n,l_n} \to j$ and the final arc contradicts the two-maximality of P_n. ∎

We obtain a new exact result.

Theorem 2.2 *For* $F = \left[\begin{array}{c} 0\,\overbrace{0\,0\ldots0}^{p}\,\overbrace{1\,1\ldots1}^{p} \\ 0\,1\,1\ldots1\,0\,0\ldots0 \end{array}\right]$ *and* $F_1 = \left[\begin{array}{c} \overbrace{0\,0\ldots0}^{p}\,\overbrace{1\,1\ldots1}^{p} \\ 1\,1\ldots1\,0\,0\ldots0 \end{array}\right]$ *and* $m \ge p \ge 1$

$$\mathrm{forb}(m, F) = \mathrm{forb}(m, F_1) = (m-1)p + 2.$$

Proof: We first show that $\mathrm{forb}(m, F) \le (m-1)p + 2$. Let A be a simple matrix of m rows not containing F. We construct a directed graph D using row numbers of A as vertices and adding the edges $i \to j$ for each i and j such that the number of $\left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]$ connections between row i and row j is less than p. As a result of forbidding F, we see that if i, j are not joined in D then row i over row j lacks the column $\begin{bmatrix}0\\0\end{bmatrix}$.

We use Lemma 2.1 to obtain a vertex ordering for D and a forest of indirected trees T consisting of at most $m-1$ edges. Rearrange the rows of A in accordance with the ordering. The number of columns of A with an entry $\begin{bmatrix}0\\1\end{bmatrix}$ in row i over row j for some edge $i \to j$ in our forest T is at most $(m-1)(p-1)$, using the definition of $i \to j$.

There are at most $m+1$ columns which for each edge $i \to j$ of T, do not have an entry $\begin{bmatrix}0\\1\end{bmatrix}$ in row i over row j. To see this note that apart from the column of 1's, each such column must have a highest row i with a 0, namely 1's in rows $1, 2, \ldots, i-1$ and a 0 in row i. Applying the lemma we consider any row j with $i < j \leq m$. If there is a directed path from i to j then all the entries in the rows corresponding to the vertices of the path are forced to be 0's. If there is a directed path from i to k and there is no edge in D from k to j then the entry in row k is forced to be 0 and then the entry in row j is forced to be 1 (by the lack of an edge from j to k in D). Thus there are at most $(m-1)(p-1) + (m+1) = (m-1)p + 2$ columns in A.

We complete our proof by showing $\text{forb}(m, F_1) \geq (m-1)p + 2$ using a construction. We form $(m-1)(p-1)$ columns C_{ij} for $1 \leq i \leq p-1$ and $1 \leq j < i$ and $i+1 < j \leq p$ with C_{ij} having 1's in rows $1, 2, \ldots, i-1$, 0 in row i, 1 in row $i+1$, and 0's in rows $i+1, i+3, \ldots, m$. We then add the $m+1$ columns containing no submatrix $\begin{bmatrix}0\\1\end{bmatrix}$ to obtain a simple $m \times ((m-1)p+2)$ matrix with no configuration F_1. ■

Some of the proof for the bounds for forbidden configurations proceed by what we call the *standard argument*, made in reference to a particular row of a simple $m \times n$ matrix. We first bring this row to the top of the matrix and then rearrange the columns to produce a matrix in the following form:

$$\begin{bmatrix} 11\ldots1 & 11\ldots1 & 00\ldots0 & 00\ldots0 \\ B_1 & B_2 & B_2 & B_3 \end{bmatrix}$$

where B_1, B_2 and B_3 are each matrices with $m-1$ rows, but B_1 and B_3 have no identical columns. The matrix

$$\begin{bmatrix} B_1 B_2 B_3 \end{bmatrix}$$

is a simple matrix. This process may be extended for multiple rows to obtain new results and simplified proofs of old results as in the following result for which a weaker linear bound (of $(2p-2)m$) is given as Theorem 2.2 in [4].

Theorem 2.3 *For* $F = \begin{bmatrix} 0 & \overbrace{1 \quad 1 \quad \ldots \quad 1}^{p} & 0 & 1 \\ 0 & 0 \quad 0 \quad \ldots \quad 0 & 1 & 1 \end{bmatrix}$, *forb*$(m, F) \leq (p - \frac{1}{2})m + 1$.

Proof: Let A be an $m \times n$ simple matrix with no configuration F. If a pair of rows has the configuration K_2, then that pair of rows has no configuration $F' = \begin{bmatrix} \overbrace{0 \; 0 \; \ldots \; 0}^{p} \\ 1 \; 1 \; \ldots \; 1 \end{bmatrix}$.

Let C be a maximal set of rows so that for every pair $i, j \in C$, the rows i, j do not have the configuration F'. We may assume $|C| > 1$ since if there are no pair of rows avoiding F' then A avoids K_2 and so $n \leq m+1$ by Theorem 1.2.

We may reorder the rows of A so that the k rows of C are first and then decompose A as

$$A = \begin{bmatrix} 0's & B & 1's \\ D & E & G \end{bmatrix} \tag{1}$$

where every column in the $k \times l$ matrix B contains the configuration $\left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]$ and so in fact contains at least $k-1$ $\left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]$'s. Now B has no configuration F' and so we deduce $(k-1)l \leq 2(p-1)\binom{k}{2}$ since there are at most $2(p-1)$ configurations $\left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]$. This yields $l \leq (p-1)k$.

We find that at most one column can be in common to both D and G since if there were two columns α, γ in both D and G then in some row t there is the configuration $[0\ 1]$ in both D and G. But then for every $i \in C$, there is a K_2 on rows i, t and hence no F' on rows i, t, which contradicts that C is a maximal set.

Thus we can delete from A the k rows of C and at most $(p-1)k+1$ columns to obtain a simple matrix A'. Thus

$$\mathrm{forb}(m, F) \leq \mathrm{forb}(m-k, F) + (p-1)k + 1.$$

For $k \geq 2$ we have $\frac{(p-1)k+1}{k} \leq p - \frac{1}{2}$ and so we obtain the bound. ■

The following constructions provide lower bounds on forb(m, F) for F as above.

Proposition 2.4 *For* $F = \begin{bmatrix} 0 & \overbrace{1 \quad 1 \quad \dots \quad 1}^{p} \\ 0 & 0 \quad 0 \quad \dots \quad 0 \end{bmatrix}$ *or* $\begin{bmatrix} 0 & \overbrace{1 \quad 1 \quad \dots \quad 1}^{p} \\ 1 & 0 \quad 0 \quad \dots \quad 0 \end{bmatrix}$ *and p odd and $p \geq 9$*

$$\mathit{forb}(m, F) \geq \left(\frac{p}{2} + \frac{3}{2} + \frac{2}{p+1}\right) m + \mathrm{O}(1),$$

while for p even and $p \geq 12$

$$\mathit{forb}(m, F) \geq \left(\frac{p}{2} + \frac{4}{3} + \frac{2(1-(p \pmod 6))}{p}\right) m + \mathrm{O}(1).$$

Proof: For each p, the construction below determines a number l and a simple l-rowed matrix K. Inductively we expand a simple m-rowed matrix A_m avoiding F to an $m+l$-rowed matrix A_{m+l} avoiding F as follows:

$$A_{m+l} = \begin{bmatrix} A_m & \text{1's} \\ \text{1's} & K \end{bmatrix}.$$

The new matrix has $|A_m| + |K|$ columns.
The matrix $K_l = \left[K_l^0, K_l^1, K_l^2, K_l^{l-2}, K_l^{l-1}\right]$ contains $l^2 + l + 1$ columns and $2l-2$ $\left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]$'s

and $2l-2$ $\begin{bmatrix}1\\0\end{bmatrix}$'s in each pair of rows.
For p odd, we choose $l=(p+1)/2$ and $K=K_l$. This matrix contains $(p^2+4p+7)/4$ columns and has exactly $p-1$ $\begin{bmatrix}0\\1\end{bmatrix}$'s and $\begin{bmatrix}1\\0\end{bmatrix}$'s in each pair of rows. This gives a growth rate of $\left(\frac{p^2+4p+7}{4}\right)/\left(\frac{p+1}{2}\right)=\left(\frac{p}{2}+\frac{3}{2}+\frac{2}{p+1}\right)$ per added row.
When p is even, we choose $l=p/2$. K_l now has $p-2$ $\begin{bmatrix}0\\1\end{bmatrix}$'s and $\begin{bmatrix}1\\0\end{bmatrix}$'s in each pair of rows. We produce K by augmenting K_l with $\lfloor\frac{p}{6}\rfloor$ columns from $K^3_{p/2}$ so that we add no more that an extra 1 to each row of the $p/2$ rows; thus, we still have fewer than $p-1$ $\begin{bmatrix}0\\1\end{bmatrix}$'s and $\begin{bmatrix}1\\0\end{bmatrix}$'s in each pair of rows of K.
K has $\frac{p^2}{4}+\frac{2p}{3}+1-(p \pmod 6)$ columns giving a growth rate of $\frac{p}{2}+\frac{4}{3}+\frac{2(1-(p \pmod 6))}{p}$ columns per added row. ■

The following result includes Theorem 2.5 of [4].

Corollary 2.5 *For any submatrix F_s of $F=\begin{bmatrix}1&1&1&0&0\\1&0&0&1&0\end{bmatrix}$ containing either the configuration $\begin{bmatrix}1&1&0\\0&0&1\end{bmatrix}$ or $\begin{bmatrix}0&0&0\end{bmatrix}$ or $\begin{bmatrix}1&1&1\end{bmatrix}$,*

$$\mathrm{forb}(m,F_s)=\left\lfloor\frac{3}{2}m\right\rfloor+1.$$

Proof: The bound $\mathrm{forb}(m,F)\le\lfloor\frac{3}{2}m\rfloor+1$ follows from Theorem 2.3. Since $\mathrm{forb}(m,F_s)\le\mathrm{forb}(m,F)$, we have only to prove $\mathrm{forb}(m,F_s)\ge\lfloor\frac{3}{2}m\rfloor+1$. Note that the matrices A_m constructed in Proposition 2.4 do not contain the configuration $\begin{bmatrix}1&1&0\\0&0&1\end{bmatrix}$. This gives the first case. For the second case, we construct the $m\times(\lfloor\frac{3}{2}m\rfloor+1)$ matrices C_m inductively starting with $C_1=\begin{bmatrix}0&1\end{bmatrix}$ and $C_2=\begin{bmatrix}0&1&0&1\\0&0&1&1\end{bmatrix}$ as before. C_m is constructed by adding a row of 1's under the matrix C_{m-1} and the $m\times 1$ column $\begin{bmatrix}1\\1\\\vdots\\1\\1\\0\end{bmatrix}$ for m odd, the $m\times 2$ matrix $\begin{bmatrix}1&1\\1&1\\\vdots&\vdots\\1&1\\1&0\\0&0\end{bmatrix}$ for m even. These matrices do not contain the configuration $\begin{bmatrix}0&0&0\end{bmatrix}$.
The complementary construction works for $\begin{bmatrix}1&1&1\end{bmatrix}$. ■

The following theorem gives a new exact bound using a similar method.

Theorem 2.6 *For* $F = \begin{bmatrix} 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}$ *and* $m \geq 3$,

$$\mathrm{forb}(m, F) = \left\lfloor \frac{7}{3} m \right\rfloor + 1.$$

Proof: We construct the matrices A_m inductively as follows. $A_3 = K_3$. A_4 and A_5 have the constant columns along with those with only one 0 or 1. For $m \geq 6$ we obtain A_m from A_{m-3} by adding 7 columns in the following way:

$$A_m = \begin{bmatrix} A_{m-3} & & & & \text{1's} & & & \\ & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ \text{0's} & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ & 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix}.$$

None of these simple matrices contains the configuration F, and each has $\lfloor \frac{7}{3} m \rfloor + 1$ columns, establishing the quantity on the right as a lower bound.

We now show that the upper bound using the proof ideas of Theorem 2.3. We use induction on m. If we can find a C as in Theorem 2.3 with $|C| = k \geq 3$ then $\frac{(p-1)k+1}{k} \leq \frac{7}{3}$ (using $p = 3$). Hence we would have $\mathrm{forb}(m, F) \leq \mathrm{forb}(m-k, F) + \frac{7}{3}k$ and so we obtain our bound. Hence we may assume that the largest C that can be found is of size 2 and that the $2 \times l$ matrix B in (1) has $l = 4$ since for $l \leq 3$ we could delete 2 rows and at most 4 columns yielding the desired bound by induction. This forces B to be the matrix $\begin{bmatrix} 0011 \\ 1100 \end{bmatrix}$ apart from column order. We may rearrange A to obtain

$$\mathrm{A} = \left[\begin{array}{ccc|ccc|cccc|ccc|ccc} 0 & \dots & 0 & 0 & \dots & 0 & 1 & 1 & 0 & 0 & 1 & \dots & 1 & 1 & \dots & 1 \\ 0 & \dots & 0 & 0 & \dots & 0 & 0 & 0 & 1 & 1 & 1 & \dots & 1 & 1 & \dots & 1 \\ & B_1 & & & B_2 & & & & E & & & B_2 & & & B_3 & \end{array}\right]$$

where $[B_1 B_2] = D$ and $[B_2 B_3] = G$ and B_2 are the columns in common. We find that B_2 has exactly one column. Our argument in Theorem 2.3 showed that there is at most one column and if there is no column then we could delete the first two rows and the 4 columns (containing E) and obtain a simple matrix so $\mathrm{forb}(m, F) \leq \mathrm{forb}(m-2, F) + 4$ which yields the bound.

We find a row such that the first 2 entries in E are different and rearrange A so the first three rows give:

$$\begin{bmatrix} \dots & 0 & 1 & 1 & 0 & 0 & 1 & \dots \\ \dots & 0 & 0 & 0 & 1 & 1 & 1 & \dots \\ \dots & a & 0 & 1 & b & c & a & \dots \end{bmatrix}$$

Because of the symmetry, we may assume the entry a to be 0. For b and c there are 3 cases to consider.

Case 1: $b = c = 0$. Rows 2 and 3 have the configuration $\begin{bmatrix} 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$ and therefore must not contain a configuration $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$. This forces all of the entries on the right for these 3 rows to be $\begin{bmatrix}1\\1\\0\end{bmatrix}$'s. If there are 2 or more of these, then rows 1 and 3 contain $\begin{bmatrix}0&1&1&1&1\\0&0&0&0&1\end{bmatrix}$. To avoid F on rows 1 and 3, there are at most two 1's in row 3 under the $\begin{bmatrix}0\\0\end{bmatrix}$'s. Now we could delete the first 3 rows of A and at most 7 columns (the 4 columns of E, the common column B_2, and the columns with 1's in row 3) to get a simple matrix and so

$$\text{forb}(m,F)\le\text{forb}(m-3,F)+7$$

and so the bound is proven.
Case 2: Assume $b=c=1$. But now the pairs of rows 1,3 and 2,3 both contain K_2 and this contradicts the choice of C.
Case 3: Assume $b=0$, $c=1$. In the first three rows we have

$$\begin{bmatrix}1&1&0&0\\0&0&1&1\\0&1&0&1\end{bmatrix}.$$

and so the pairs of rows 1,3 and 2,3 both contain K_2, contradicting the choice of C. ∎

Corollary 2.7 *For any submatrix of F_s of $F=\begin{bmatrix}0&1&1&1&0&1\\0&0&0&0&1&1\end{bmatrix}$ containing either $F_1=\begin{bmatrix}0&1&1&1\\0&0&0&0\end{bmatrix}$ or $F_2=\begin{bmatrix}0&1&1&1\\1&0&0&0\end{bmatrix}$ and $m\ge 3$*

$$\text{forb}(m,F_s)=\left\lfloor\frac{7}{3}m\right\rfloor+1.$$

Proof: The matrices A_m constructed above do not contain the configuration F_2. We modify this construction slightly, setting $A'_3=A_3$, $A'_4=A_4$, $A'_5=A_5$ and

$$A'_m=\begin{bmatrix}A'_{m-3} & \text{1's}\\ & \begin{matrix}0&0&0&1&1&1&1\end{matrix}\\ \text{1's} & \begin{matrix}0&1&1&0&0&1&1\end{matrix}\\ & \begin{matrix}1&0&1&0&1&0&1\end{matrix}\end{bmatrix}.$$

These matrices do not contain F_1. ∎

The following exact bound uses graph theory to aid the analysis.

Theorem 2.8 *For $F=\begin{bmatrix}1&1&0&0&0\\0&0&1&1&1\end{bmatrix}$ and $m\ge 3$,*

$$\textit{forb}(m,F)=\left\lfloor\frac{8}{3}m\right\rfloor.$$

Proof: We provide a construction for A_m, a simple matrix $m \times \left\lfloor \frac{8}{3}m \right\rfloor$ which avoids the configuration F. We take $A_1 = K_1$, $A_3 = K_3$ and

$$A_5 = \left[\begin{array}{cccccc} & 0 & 0 & 1 & 1 & 1 \\ A_3 & 0 & 0 & 1 & 1 & 1 \\ & 0 & 0 & 1 & 1 & 1 \\ 0's & 1 & 0 & 1 & 0 & 1 \\ & 0 & 1 & 0 & 1 & 1 \end{array}\right]$$

For $m \neq 1, 2, 3, 5$, we can construct A_m inductively using

$$A_m = \left[\begin{array}{ccccccccc} & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ A_{m-3} & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ \text{0's} & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \end{array}\right].$$

Hence, for $m \geq 3$, we have $\mathrm{forb}(m, F) \geq \left\lfloor \frac{8}{3}m \right\rfloor$. We now show the reverse inequality.

Assume the theorem true where the number of rows is strictly between 2 and m. Let A be a simple matrix of dimensions $m \times \mathrm{forb}(m, F)$ with $m > 3$. We construct a graph considering rows as vertices and having directed edges $i \rightarrow j$ if the number of $\left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]$'s in $\left[\begin{smallmatrix}R_i\\R_j\end{smallmatrix}\right]$ < 2 and edges $i \cdots\cdots j$ if the numbers of $\left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]$'s and $\left[\begin{smallmatrix}1\\0\end{smallmatrix}\right]$'s in $\left[\begin{smallmatrix}R_i\\R_j\end{smallmatrix}\right]$ are each < 3. We then have the following properties.

(i) *Each pair of rows is connected by at least one edge.* This is clear.

(ii) *The graph on the directed edges is transitive and contains no cycles.*
If $i \rightarrow j$ and $j \rightarrow k$, then we have the three possibilities

$$(a) \quad \begin{tikzcd}[row sep=small, column sep=small] & j \arrow[d] \\ i \arrow[ur] \arrow[r, dotted, no head] & k \end{tikzcd}, (b) \quad \begin{tikzcd}[row sep=small, column sep=small] & j \arrow[d] \\ i \arrow[ur] & k \arrow[l] \end{tikzcd}, \text{ and } (c) \quad \begin{tikzcd}[row sep=small, column sep=small] & j \arrow[d] \\ i \arrow[ur] \arrow[r] & k \end{tikzcd}.$$

For cases (a) and (b) we look at the possible entries for these three rows. The entries above the braces indicate the number of possible columns of these types.

$$\begin{array}{cccccccccc} & \multicolumn{2}{c}{\overbrace{\hphantom{0\cdots0\;0\cdots0}}^{\leq 1}} & \multicolumn{2}{c}{\overbrace{\hphantom{1\cdots1\;0\cdots0}}^{\leq 1}} & \multicolumn{2}{c}{\overbrace{\hphantom{1\cdots1\;1\cdots1}}^{\leq 2}} & & \\ i & 0\cdots0 & 0\cdots0 & 1\cdots1 & 0\cdots0 & 1\cdots1 & 1\cdots1 & 0\cdots0 & 1\cdots1 \\ j & 1\cdots1 & 1\cdots1 & 0\cdots0 & 0\cdots0 & 0\cdots0 & 1\cdots1 & 1\cdots1 & 1\cdots1 \\ k & 0\cdots0 & 1\cdots1 & 0\cdots0 & 1\cdots1 & 0\cdots0 & 0\cdots0 & 0\cdots0 & 1\cdots1 \end{array}.$$

We can eliminate rows i and j and at most 4 columns to produce a simple matrix A_{m-2}. Using the second construction we produce an m rowed simple matrix which does not

contain F but has more columns that A.
For A to be maximal we can have (c) only - the graph must be transitive and, as well, there are no cycles.

(iii) *Induction holds if* $\cdots$ *edges occur other than in cliques.*

If there are edges
$$\begin{tikzcd} & j \arrow[d, dotted, no head] \\ i \arrow[ur] \arrow[r, dotted, no head] & k \end{tikzcd}$$
then the possible entries for these rows are

$$\begin{array}{cl} i & \overbrace{0\cdots0\;0\cdots0}^{\le 1}\overbrace{1\cdots1\;0\cdots0}^{\le 2}\overbrace{1\cdots1\;1\cdots1}^{\le 2}\;0\cdots0\;1\cdots1 \\ j & 1\cdots1\;1\cdots1\,0\cdots0\;0\cdots0\,0\cdots0\;1\cdots1\;0\cdots0\;1\cdots1 \\ k & 0\cdots0\;1\cdots1\,0\cdots0\;1\cdots1\,0\cdots0\;0\cdots0\,0\cdots0\;;1\cdots1 \end{array}.$$

We eliminate rows j and k and at most 5 columns to produce a simple matrix A_{m-2}. But then
$$|A| \le |A_{m-2}| + 5 \le \left\lfloor \frac{8}{3}(m-2) \right\rfloor + 5 \le \left\lfloor \frac{8}{3}m \right\rfloor .$$

(iv) *The theorem is true if the* $\cdots$ *edges occur only in cliques.*
Consider a clique with k vertices. We wish to discover how many columns are possible which are not constant on these k rows.
For $k = 1$ there are no columns which are non-constant on this clique.
For $k > 1$, the presence of the $\cdots$ edges means we avoid the configuration $\begin{bmatrix} 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$ in the clique. If a column has both 0 and 1 entries on these k rows, it has at least $k-1$ of the $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ configurations in these k rows. More that $2k$ of such columns would give more than $2k(k-1)$ of the $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ configurations in $k(k-1)/2$ pairs of rows from this clique. One pair would then contain more than 4 and hence the configuration $\begin{bmatrix} 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$. Thus there can be no more than $2k$ columns which are not constant on these k rows. Let c_k be the number of cliques which have k vertices. The maximum number of columns of A which are not constant on cliques is then $2(m - c_1)$.

There is an ordering on the cliques by defining $C_1 < C_2$ if there is a directed edge $i \to j$ for $i \in C_1$ and $j \in C_2$. By properties (i) and (ii) this ordering is strictly linear. Reorder the rows of A to respect this linear order. The columns of A which are constant on cliques avoid the submatrix $\begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}$. If the number of cliques is m', then this is equivalent to avoiding this submatrix in an m' rowed matrix. There are a maximum of $m'-1$ distinct columns possible which contain a $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ and a maximum of $m'+1$ columns with no $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$.
Thus, A has a maximum of $2(m - c_1) + 2m'$ columns, but where there are cliques of size 2, we have overcounted. This is because on clique C_t of size two, say on rows $i, i+1$, we can distinguish only 2 nonconstant columns. To achieve the count of 4 we must have $\begin{bmatrix} 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \end{bmatrix}$ in rows $i, i+1$. Two of these columns could have 1's above and 0's below. The remaining two columns must either have 0's on a clique above C_t and then a 1 in

C_t or a 1's on a clique following C_t and a 0 in C_t. The count of $m'-1$ distinct columns possible which contain a $\left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]$ must be reduced by at least 1 for every clique of size 2 (the two columns can both contribute a $\left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]$ to the same clique pair).
We then have

$$|A| \le 2(m-c_1) + 2m' - c_2 = 2c_1 + 5c_2 + \sum_{k>2}(2k+2)c_k$$
$$\le \left\lfloor \frac{8}{3}(c_1 + 2c_2 + 3c_3 + \ldots) \right\rfloor = \left\lfloor \frac{8}{3}m \right\rfloor . \blacksquare$$

3 Quadratic Bounds

In [4], we classified which $2\times l$ F give rise to quadratic bounds. We are able to substantially strengthen the bounds obtaining some exact bounds or bounds with correct quadratic terms. Graph theory is remarkably useful.
In the following, we investigate forb(m,F) for

$$F \quad = \quad \left[\begin{array}{cccc} \overbrace{0\ldots0}^{r} & \overbrace{1\ldots1}^{p} & \overbrace{0\ldots0}^{p} & \overbrace{1\ldots1}^{s} \\ 0\ldots0 & 0\ldots0 & 1\ldots1 & 1\ldots1 \end{array}\right] \tag{2}$$

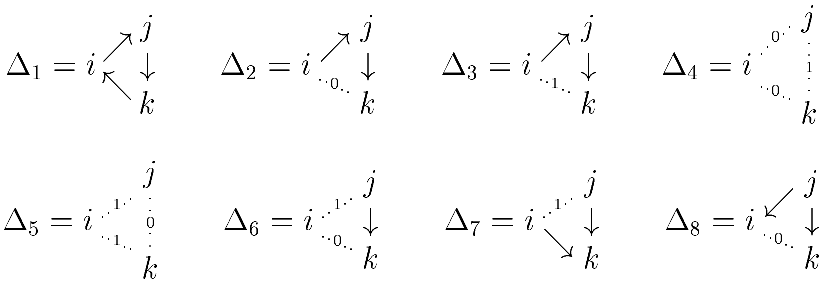
By symmetry we may assume $r \ge s \ge 1$. From [4], forb(m,F) is quadratic if $r \ge 2$ or $r = s = 1$ and $p \ge 2$. Certain general cases are known 'exactly'.

Theorem 3.1 ([3]) *Let $0_{2,r}$ denote the $2\times r$ matrix of 0's. For F as above with $r \ge 2, p, s$ or any submatrix F_s of F containing $0_{2,r}$, then forb(m,F) = forb(m,F_s) = forb$(m,0_{2,r})$ = $\frac{r+1}{6}m^2 + O(m)$.* ■

In what follows we will often consider cases with $p > r$. We will assume $p \ge 2$. In [4] it is shown that $F_0 = \left[\begin{smallmatrix}0&1&1&0&0&1\\0&0&0&1&1&1\end{smallmatrix}\right]$, forb$(m,F_0) \ge \lfloor \frac{m^2}{4} \rfloor + m + 1$. Let A be a simple $m\times n$ matrix which does not have the configuration F of (2). Let R_i refer to the ith row of A. We consider the rows as vertices in a graph with the following edges:
(a) The directed edge $i \to j$ if the number of $\left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]$'s in $\begin{bmatrix}R_i\\R_j\end{bmatrix}$ is less than p.
(b) $i \overset{0}{\cdots} j$ if the number of $\left[\begin{smallmatrix}0\\0\end{smallmatrix}\right]$'s in $\begin{bmatrix}R_i\\R_j\end{bmatrix}$ is less than r.
(c) $i \overset{1}{\cdots} j$ if the number of $\left[\begin{smallmatrix}1\\1\end{smallmatrix}\right]$'s in $\begin{bmatrix}R_i\\R_j\end{bmatrix}$ is less than s.

We then have the following:

Lemma 3.2 *If we are given forb$(m-1,F) \le a(m-1)^2 + b(m-1) + c$ (for $a,b > 0$ and $a \le b$), then either forb$(m,F) \le am^2 + bm + c$ or we may assume that for $m \ge \frac{3}{2a}\max\{p,r,s\}$ that each pair of rows of A is connected by exactly one edge and, in the graph arising from A, the following 8 triangles do not occur:*



Proof: We compute $am^2 + bm + c - (a(m-1)^2 + b(m-1) + c) = a(2m-1) + b \geq 2am$. Thus we need only show that one of the cases of the Lemma yield

$$\mathrm{forb}(m, F) \leq \mathrm{forb}(m-1, F) + 2am$$

(or $\mathrm{forb}(m, F) \leq \mathrm{forb}(m-2, F) + 2am$. Our choice of m ensures that $3\max\{p, r, s\} \leq 2am$ and this can be used to show the above inequality in what follows. In trying to prove $\mathrm{forb}(m, F) \leq am^2 + bm + c$ by induction, the base cases can be handled with large b and/or c. Since the results are not usually exact, this is reasonable.

For A not to have the configuration F, each pair of rows must be connected by an edge. Since a directed edge may go either way there are $\binom{4}{2} = 6$ possible combinations for rows i and j to be connected by two edges.
If, for example, we have the edges $i \to j$ and $i \overset{0}{\cdots} j$, then A has the configuration

$$\begin{array}{ll}
R_i & \overbrace{0\ldots0}^{\leq p-1}\ \overbrace{0\ldots0}^{\leq r-1}\ 1\ldots1\,1\ldots1 \\
R_j & 1\ldots1\ 0\ldots0\ 0\ldots0\,1\ldots1 \\
\vdots & \cdots\cdots\cdots\cdots\cdots
\end{array} .$$

Eliminating the first $\leq p + r - 2$ columns and R_i leaves a simple matrix, giving that

$$\mathrm{forb}(m, F) \leq \mathrm{forb}(m-1, F) + p + r - 2.$$

Using $p + r - 2 < 2am$ would enable us to prove $\mathrm{forb}(m, F) \leq am^2 + bm + c$ by induction. The proofs for the other 5 cases are similar.

For Δ_1, we deduce

$$\begin{array}{ll}
R_i & \overbrace{0\ldots0\,0\ldots0}^{\leq p-1}\ \overbrace{0\ldots0\,1\ldots1}^{\leq p-1}\ \overbrace{1\ldots1\,1\ldots1}^{\leq p-1}\ 0\ldots0\,1\ldots1 \\
R_j & 1\ldots1\,1\ldots1\ 0\ldots0\,0\ldots0\ 0\ldots0\,1\ldots1\ 0\ldots0\,1\ldots1 \\
R_k & 0\ldots0\,1\ldots1\ 1\ldots1\,1\ldots1\ 0\ldots0\,0\ldots0\ 0\ldots0\,1\ldots1 \\
\vdots & \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots
\end{array}$$

Eliminating the first $\leq 3p - 3$ columns, R_i leaves a simple matrix, so that $\mathrm{forb}(m, F) \leq$

forb$(m-1,F)+3p-3$. For Δ_2 we get

$$\begin{array}{llll}
 & \overbrace{\hphantom{0\ldots0\,0\ldots0\,0\ldots0}}^{\leq p+r-2} & \overbrace{\hphantom{0\ldots0\,1\ldots1}}^{\leq p-1} & \\
R_i & 0\ldots0\,0\ldots0\,0\ldots0 & 0\ldots0\,1\ldots1 & 1\ldots1\,1\ldots1\,1\ldots1 \\
R_j & 1\ldots1\,1\ldots1\,0\ldots0 & 0\ldots0\,0\ldots0 & 0\ldots0\,1\ldots1\,1\ldots1 \\
R_k & 0\ldots0\,1\ldots1\,0\ldots0 & 1\ldots1\,1\ldots1 & 0\ldots0\,0\ldots0\,1\ldots1 \\
\vdots & \ldots\ldots\ldots\ldots\ldots\ldots & \ldots\ldots\ldots\ldots & \ldots\ldots\ldots\ldots\ldots\ldots
\end{array}$$

Eliminating the first $\leq 2p+r-3$ columns and R_i leaves a simple matrix, so forb$(m,F) \leq$ forb$(m-1,F)+2p+r-3$. Symmetrically we get forb$(m,F) \leq$ forb$(m-1,F)+2p+s-3$ with the triangle Δ_3. For Δ_4 we have

$$\begin{array}{llll}
 & \overbrace{\hphantom{0\ldots0\,0\ldots0\,0\ldots0}}^{\leq 2r-2} & \overbrace{\hphantom{0\ldots0\,1\ldots1}}^{\leq s-1} & \\
R_i & 0\ldots0\,0\ldots0\,0\ldots0 & 0\ldots0\,1\ldots1 & 1\ldots1\,1\ldots1\,1\ldots1 \\
R_j & 0\ldots0\,1\ldots1\,0\ldots0 & 1\ldots1\,1\ldots1 & 0\ldots0\,0\ldots0\,1\ldots1 \\
R_k & 0\ldots0\,0\ldots0\,1\ldots1 & 1\ldots1\,1\ldots1 & 1\ldots1\,1\ldots1\,1\ldots1 \\
\vdots & \ldots\ldots\ldots\ldots\ldots\ldots & \ldots\ldots\ldots\ldots & \ldots\ldots\ldots\ldots\ldots\ldots
\end{array}.$$

Eliminating row i and the first set of $\leq 2r+s-3$ columns leaves a simple matrix, so that forb$(m,F) \leq$ forb$(m-1,F)+2r+s-3$ Symmetrically we get forb$(m,F) \leq$ forb$(m-1,F)+2s+r-3$. with the triangle Δ_5. For Δ_6 we have

$$\begin{array}{lllll}
 & \overbrace{\hphantom{1\ldots1\,1\ldots1}}^{\leq s-1} & \overbrace{\hphantom{0\ldots0\,0\ldots0}}^{\leq r-1} & \overbrace{\hphantom{0\ldots0\,1\ldots1}}^{\leq p-1} & \\
R_i & 1\ldots1\,1\ldots1 & 0\ldots0\,0\ldots0 & 0\ldots0\,1\ldots1 & 0\ldots0\,1\ldots1 \\
R_j & 1\ldots1\,1\ldots1 & 0\ldots0\,1\ldots1 & 0\ldots0\,0\ldots0 & 1\ldots1\,0\ldots0 \\
R_k & 0\ldots0\,1\ldots1 & 0\ldots0\,0\ldots0 & 1\ldots1\,1\ldots1 & 1\ldots1\,0\ldots0 \\
\vdots & \ldots\ldots\ldots\ldots & \ldots\ldots\ldots\ldots & \ldots\ldots\ldots\ldots & \ldots\ldots\ldots\ldots
\end{array}$$

Eliminating the first $\leq p+r+s-3$ columns and R_j leaves a simple matrix, so that forb$(m,F) \leq$ forb$(m-1,F)+p+r+s-3$. For Δ_7 we have

$$\begin{array}{llll}
 & \overbrace{\hphantom{1\ldots1\,1\ldots1}}^{\leq s-1} & \overbrace{\hphantom{0\ldots0\,0\ldots0\,1\ldots1}}^{\leq 2p-2} & \\
R_i & 1\ldots1\,1\ldots1 & 0\ldots0\,0\ldots0\,1\ldots1 & 0\ldots0\,1\ldots1\,0\ldots0 \\
R_j & 1\ldots1\,1\ldots1 & 0\ldots0\,1\ldots1\,0\ldots0 & 0\ldots0\,0\ldots0\,1\ldots1 \\
R_k & 0\ldots0\,1\ldots1 & 1\ldots1\,1\ldots1\,1\ldots1 & 0\ldots0\,0\ldots0\,0\ldots0 \\
\vdots & \ldots\ldots\ldots\ldots & \ldots\ldots\ldots\ldots\ldots\ldots & \ldots\ldots\ldots\ldots\ldots\ldots
\end{array}$$

Eliminating the first $\leq p+r+s-3$ columns and R_k leaves a simple matrix, so that forb$(m,F) \leq$ forb$(m-1,F)+2p+s-3$. Symmetrically we get forb$(m,F) \leq$ forb$(m-1,F)+2p+r-3$ with the triangle Δ_8. ■

Theorem 3.3 *For* $F = \left[\begin{array}{cccccccccc} 0 & \overbrace{0 \quad 0 \quad \cdots \quad 0}^{p} & \overbrace{1 \quad 1 \quad \cdots \quad 1}^{p} & 1 \\ 0 & 1 \quad 1 \quad \cdots \quad 1 & 0 \quad 0 \quad \cdots \quad 0 & 1 \end{array}\right]$, *there exists a c, so*

$$\frac{m^2}{4} + (p - 1\frac{1}{2} - \sqrt{p-1})m + O(p) \leq \mathit{forb}(m,F) \leq \frac{m^2}{4} + (p-1)(m-2) + c$$

where the lower bound holds for $p \geq 4$ and $m > 4(p-1)^{3/2}$.

In addition, there exists an m_0 so that for $m > m_0$ we may take $c = 0$ in the upper bound.

Proof: Let A be an m rowed matrix with no configuration F. Form a graph on the rows of A as above ($r = s = 1$) with edges $i \to j$, $i \overset{0}{\cdots} j$, $i \overset{1}{\cdots} j$. To prove the upper bound we use induction on m with $m \geq 6p$. We can choose c large enough (2^{6p} will do) so that $\mathrm{forb}(m-1, F) \leq \frac{(m-1)^2}{4} + (p-1)((m-1)-2) + c$. Applying Lemma 3.2, (using $6p \geq \frac{3}{2a}\max\{p, r, s\}$) we may assume A has no doubled edges nor the forbidden triangles.
Property 1: If we have $i \to j$ and $j \to k$, then $i \to k$ (transitivity). This follows from the three forbidden triangles $\Delta_1, \Delta_2, \Delta_3$.
Property 2: If $i \to k$ and $j \to k$ and yet i, j are not joined by a directed edge then there is an edge $i \overset{0}{\cdots} j$. This follows from the forbidden triangle Δ_7.
Property 3: If $i \to j$ and $i \to k$ and yet j, k are not joined then there is an edge $j \overset{1}{\cdots} k$. This follows from the forbidden triangle Δ_8.
Property 4: If $i \to j$ and $j \overset{0}{\cdots} k$ then $i \overset{0}{\cdots} k$. We need the forbidden triangles $\Delta_2, \Delta_6, \Delta_8$.
Property 5: Symmetrically if $i \to j$ and $i \overset{1}{\cdots} k$ then $j \overset{1}{\cdots} k$. We need the forbidden triangles $\Delta_3, \Delta_6, \Delta_7$.

We decompose the rows into two sets M_1, M_2 as follows. Identify a maximal clique C_1 of at least two rows such that every pair $i, j \in C_1$ are joined $i \overset{0}{\cdots} j$. Such a clique exists since otherwise our matrix A is avoiding the configuration F' obtained from F by deleting the column of 0's but that yields a linear bound of $(p-1)m + 2$. In addition assume C_1 has the property that there is no clique C' below C_1, namely no clique C' of (at least two) rows such that every pair $i, j \in C'$ are joined $i \overset{0}{\cdots} j$ and for every $k \in C$ there is a $l \in C'$ so that there is a directed walk from k to l. Let M_1 be the rows i for which there exists a $k \in C$ with a directed path from i to k. We deduce by properties 1,2,4,5 that there are no edges $i \overset{1}{\cdots} j$ for $i, j \in M_1$. Use the techniques of Lemma 2.1 with the edges $\overset{0}{\cdots}$ being considered non edges, to find a forest of indirected trees T_1 on the rows M_1 and a row ordering. Alternatively, we can obtain an acyclic ordering of M_1 since there are no directed cycles (because of the transitivity and lack of 2-cycles by Lemma 3.2). Then delete the edges implied by transitivity to obtain T_1, a forest of indirected trees whose roots are C_1.

Let M_2 be the remaining rows of A. We assert that for no edges $k \overset{0}{\cdots} l$ for $k, l \in M_2$. Assume the contrary that $k \overset{0}{\cdots} l$ exists. Then by the definition of M_1, there are no edges from k or l to any $i \in C_1$. Let $D \subseteq C_1$ be the $i \in C_1$ for which there is either an edge $i \to k$ or $i \to l$. Let $E \subseteq C_1$ be the $i \in C_1$ for which there is either an edge $i \overset{0}{\cdots} k$ and $i \overset{0}{\cdots} l$. If $D \cup E = C_1$ then this would contradict the choice of C_1 with $C' = E \cup \{k, l\}$ as a clique below C_1. Thus for some $i \in C_1$, we have $i \overset{1}{\cdots} k$ and $i \overset{1}{\cdots} l$ but then this is a forbidden triangle Δ_8 or we have $i \overset{1}{\cdots} k$ and $i \overset{0}{\cdots} l$ (or vice versa) but then this is a forbidden triangle Δ_7 . We conclude that there are no edges $k \overset{0}{\cdots} l$ for $k, l \in M_2$. As above, use Lemma 2.1 with $\overset{1}{\cdots}$ edges being considered non edges to form a forest of outdirected trees T_2 on the rows M_2.

Let $|M_1| = m_1, |M_2| = m_2$. Using the methods of Theorem 2.2, there are at most

m_1+1 columns on rows M_1, which do not have a $\begin{bmatrix}0\\1\end{bmatrix}$ on an edge of T_1. Similarly, there are at most m_2+1 colums on rows M_2 which do not have $\begin{bmatrix}0\\1\end{bmatrix}$ on an edge of T_2. Hence at most $(m_1+1)(m_2+1)$ columns in A do not have a $\begin{bmatrix}0\\1\end{bmatrix}$ on an edge of $T_1 \cup T_2$. All other columns of A have a $\begin{bmatrix}0\\1\end{bmatrix}$ on an edge of T_1 or T_2 and so by the standard pidgeonhole argument there are at most $(p-1)(m-2)$ such columns. We may maximize this bound $(m_1+1)(m_2+1)+(p-1)(m-2)$ by taking $m_1=\lfloor m/2\rfloor, m_2=\lceil m/2\rceil$ and we get the bound.

To reduce c to 0 for large m, consider $A=A_0$ as an $m\times n$ simple matrix with no configuration F. If A_0 has no triangle or doubled edge then our above argument gives the bound with $c=0$. Assume, A_0 has a triangle or doubled edge. Using the arguments of Lemma 3.2 we may delete one row and some columns to obtain an $(m-1)\times n_1$ simple matrix A_1 with no configuration F and $n_1>n-3p$. Now if A_1 has no triangle or doubled edge, then $n_1 \le \frac{(m-1)^2}{4}+(p-1)((m-1)-2)$ and then, for $m\ge 6p$ we have $n\le \frac{m^2}{4}+(p-1)(m-2)$. If A_1 has a triangle or a doubled edge, then we can delete one row and some columns to obtain a simple $(m-2)\times n_2$ matrix A_2, and we repeat as for A_1. Either we will have shown the bound for m with $c=0$ or we will obtain an $(m-k)\times n_k$ simple matrix A_k with no configuration F and $n_k>n-3kp$. Assume $n>\frac{m^2}{4}+(p-1)(m-2)$. We have for A_k that $n-3kp<n_k\le\frac{(m-k)^2}{4}+(p-1)((m-k)-2)+c$ by induction. We combine to get

$$m<2(2p+1+\frac{c}{k}+\frac{k}{4}).$$

Taking $k=2\sqrt{c}$ we get $m<2(\sqrt{c}+2p+1)=m_0$. Hence we can assume $c=0$ for $m\ge m_0$.

We construct a simple m-rowed matrix avoiding F as follows. Let $m_1=\left\lfloor\frac{m}{2}\right\rfloor, m_2=m-m_1, l_1=\left\lfloor\sqrt{p-1}\right\rfloor, l_2=\left\lfloor\frac{p-1}{l_1}\right\rfloor$. The graph on the upper m_1 rows of A where we eliminate directed edges implied by transitivity will consist of disjoint paths of length l_1 (and one of r_1) and on the lower m_2 rows, disjoint paths of length l_2 (and one of r_2). Define the $l\times l$ matrices

$$K_{1,l}=\begin{bmatrix}0 & & & \\ & 0 & \text{1's} & \\ & \text{0's} & \ddots & \\ & & & 0\end{bmatrix} \text{ and } K_{2,l}=\begin{bmatrix}1 & & & \\ & 1 & \text{1's} & \\ & \text{0's} & \ddots & \\ & & & 1\end{bmatrix}.$$

There are no $\begin{bmatrix}0\\1\end{bmatrix}$ submatrices in $K_{1,l}$ and $K_{2,l}$. There are no $\rightarrow$ edges between upper and lower rows of these matrices. We use these blocks to obtain the $m_1\times(m_1+1)$ and

$m_2 \times (m_2+1)$ matrices

$$K_u = \begin{bmatrix} K_{1,l_1} & & & & 1 \\ & K_{1,l_1} & \text{1's} & & 1 \\ & & \ddots & & \vdots \\ & \text{1's} & K_{1,l_1} & & 1 \\ & & & K_{1,r_1} & 1 \end{bmatrix}, \; K_d = \begin{bmatrix} K_{2,l_2} & & & & 0 \\ & K_{2,l_2} & \text{0's} & & 0 \\ & & \ddots & & \vdots \\ & \text{0's} & K_{2,l_2} & & 0 \\ & & & K_{2,r_2} & 0 \end{bmatrix},$$

where r_1 and r_2 are the integer remainders in dividing m_1 and m_2 by l_1 and l_2 respectively. Note that K_u has no $\left[\begin{smallmatrix}0\\0\end{smallmatrix}\right]$ and K_u has no $\left[\begin{smallmatrix}1\\1\end{smallmatrix}\right]$

In the product

$$A_1 = K_u \otimes K_d$$

we have $\rightarrow$ trees of length l_1 on top and trees of length l_2. The bottom rows of the $K_{1,l}$ blocks on top and the top rows of the $K_{2,l}$ blocks on the bottom become the roots of paths. The maximum number of $\left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]$'s in a pair of rows $\left[\begin{smallmatrix}R_i\\R_j\end{smallmatrix}\right]$ is $l_1 l_2 < p-1$ occurring when both i and j are roots of paths. When i and j are not roots the maximum number is $\max\{(l_1-1)l_2, l_1(l_2-1)\} \le p-3$ when $p > 4$.

Between two rows $i < j$ in the same path in the top (or in the bottom) there are no $\left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]$'s in $\left[\begin{smallmatrix}R_i\\R_j\end{smallmatrix}\right]$. For each edge $i \rightarrow (i+1)$ in these $\left\lceil \frac{m_1}{l_1} \right\rceil + \left\lceil \frac{m_2}{l_2} \right\rceil$ trees, we add $p-1$ columns to our matrix to produce the matrix A so that

$$\begin{aligned} |A| &= (m_1+1)(m_2+1) + (p-1)\left(m - \left\lceil \frac{m_1}{l_1} \right\rceil - \left\lceil \frac{m_2}{l_2} \right\rceil\right) \\ &\ge \frac{m^2}{4} + (p-1-\sqrt{p-1}-\frac{1}{2})m + O(p), \end{aligned}$$

since

$$\begin{aligned} (p-1)\left(\left\lceil \frac{m_1}{l_1} \right\rceil + \left\lceil \frac{m_2}{l_2} \right\rceil\right) &\le \frac{p-1}{l_1} m_1 + \frac{p-1}{l_2} m_2 + 2(p-1) \\ &\le (l_1+2)m_1 + (\sqrt{p-1})m_2 + 2(p-1) \\ &\le (\sqrt{p-1}+1)m + 2(p-1). \end{aligned}$$

Let U_1, U_2 be m_1 rowed matrices consisting of columns with a single O. U_1 includes the $m_1 - \left\lceil \frac{m_1}{l_1} \right\rceil$ columns not in K_u. U_2 consists of $p-1$ such columns chosen from K_u (here we need $m > 4(p-1)^{3/2}$). Similarly let D_1, D_2 be the corresponding m_2 rowed matrices of columns with a single 1 with D_1 being from those not in K_d and D_2 being $p-1$ chosen from K_d. We then have

$$A = [A, U_1 \otimes D_2, U_2 \otimes D_1].$$

With $p > 4$ and $m > 4(p-1)^{3/2}$, $\left\lceil \frac{m_2}{l_2} \right\rceil > \frac{m}{2(\sqrt{p-1}+2)} \ge p-1$, so that we obtain enough columns from K_u and K_d. We add no more that two $\left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]$'s to each $\left[\begin{smallmatrix}R_i\\R_j\end{smallmatrix}\right]$, where i, j are rows from the top and bottom, respectively, without both being roots. The count never exceeds $p-3+2=p-1$. ∎

An exact bound is possible for $p=2$, improving on a result in [4].

Theorem 3.4 *For* $F=\begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix}$

$$forb(m,F)=\lfloor\frac{m^2}{4}\rfloor+m+1$$

Proof: The construction can be found as Configuration 5 in [4]. We can tighten the proof found there by seeking copies of K_2. Let A be an $m\times(\lfloor\frac{m^2}{4}\rfloor+m+2)$ simple matrix with no F. Since $\mathrm{forb}(m,K_2)=m+1$, on the first $m+2$ columns we can find a copy of K_2, say on rows i,j, and deduce that either $\begin{array}{c} i \\ j \end{array}\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ or $\begin{array}{c} i \\ j \end{array}\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ occurs just once in A in rows i,j. Set aside the column of A that that has the single occurence and repeat on the remaining columns which can be done if there are at least $m+2$ columns left. Note that we can no longer find K_2 on rows i,j after the deletion. We will have set aside $\lfloor\frac{m^2}{4}\rfloor+1$ columns. Form a graph G on the rows by joining i,j if we selected a K_2 on those rows. Thus G has $\lfloor\frac{m^2}{4}\rfloor+1$ edges. Following the proof in [4] we can verify that the graph has no triangles. By Turán's bound, the number of edges is at most $\lfloor\frac{m^2}{4}\rfloor$ and this contradiction proves the bound. ∎

The following result shows some promise for obtaining exact bounds for all $2\times l$ F.

Theorem 3.5 *Let*

$$F_{pq}=\begin{bmatrix} \overbrace{0\cdots 0}^{r}\overbrace{0\,0\cdots 0}^{p}\overbrace{1\,1\cdots 1}^{q}\overbrace{1\,1\cdots 1}^{s} \\ 0\cdots 0\,1\,1\cdots 1\,0\,0\cdots 0\,1\,1\cdots 1 \end{bmatrix}$$

with $p<q$. If $forb(m,F_{pp})\le am^2+bm+c$ for $0<a,b$, then for $m>q/a+a/2$, there exists a $c'\ge c$ so that

$$forb(m,F_{pq})\le am^2+bm+c'.$$

There exists an m_0 so that for $m>m_0$ we can replace c' by c.

Proof: Following in the spirit of Lemma 3.2 we show that $\mathrm{forb}(m,F_{pq})\le am^2+bm+c'$ by induction on m. We establish the base cases for $m\le q/a+a/2$ by choosing c' large enough. If A is a simple $m\times am^2+bm+c'+1$ matrix with no F_{pq}, then A contains a configuration F_{pp} say on rows 1,2. We deduce that A has at most $2(q-1)$ columns with the configuration $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ in rows 1,2 and so we could delete row 1 and at most $2(q-1)$ columns to obtain a simple matrix A' with no F_{pq}. Applying induction and the fact that

$am^2+bm+c'-(a(m-1)^2+b(m-1)+c'>2(q-1)$ for $m>q/a+a/2$, we deduce that A' has F_{pq}, a contradiction. Thus for all m

$$\mathrm{forb}(m,F_{pq})\leq am^2+bm+c'.$$

To establish $c'=c$ for m large, we assume $c'>c$ and let A be an $m\times am^2+bm+c+1$ simple matrix with no F_{pq}. Thus A contains F_{pp}, say on rows 1,2. As above, we can delete row 1 and at most $2(q-1)$ columns to obtain a simple matrix A'. But, $am^2+bm+c+1-2(q-1)>a(m-1)^2+b(m-1)+c'$ for $m>(2q-3+a^2+c'-c-b)/2a=m_0$. Hence for $m>m_0$, A' has F_{pq}, a contradiction. Hence we may choose $c'=c$ for $m>m_0$. ■

Theorem 3.6 *Let*

$$F=\begin{bmatrix}\overbrace{0\ldots0}^{r}&\overbrace{1\ldots1}^{p}&\overbrace{0\ldots0}^{p}&\overbrace{1\ldots1}^{s}\\0\ldots0&0\ldots0&1\ldots1&1\ldots1\end{bmatrix}\ \textit{with } r,p,s\geq2,\ r\geq s$$

then forb$(m,F)\leq(r-1)m^2/2+O(m)$. *Moreover for* $m>p$ *and* $r,s\geq3$, *we have* $\lim_{p\to\infty}$*forb*$(m,F)/m^2=(r-1)/2$.

Proof: We use induction on r. We first consider the case $r=2$ and follow the proof technique for Theorem 3.3. Let A be a matrix with no configuration F. Obtain the decomposition M_1 and M_2 as well as forests T_1 and T_2. As before, with $|M_1|=m_1$, $|M_2|=m_2$, we find that there are at most $(m_1+1)(m_2+1)$ columns with no $\begin{matrix}i\\j\end{matrix}\begin{bmatrix}0\\1\end{bmatrix}$ on an edge $i\to j$ of T_1 or T_2, or $\begin{bmatrix}0\\0\end{bmatrix}$ on an edge $i\overset{0}{\cdots}j$ with $i,j\in M_1$ or $\begin{bmatrix}1\\1\end{bmatrix}$ on an edge $i\overset{1}{\cdots}j$ with $i,j\in M_2$. There are at most $(p-1)(m_1-1+m_2-1)+\binom{m_1}{2}+\binom{m_2}{2}$ other columns before we violate the definition of our edges. Note that $(m_1+1)(m_2+1)+\binom{m_1}{2}+\binom{m_2}{2}=\frac{m^2}{2}+O(m)$ using $m_1+m_2=m$. The result now follows.

For general $r>3$, consider the smaller matrix

$$F'=\begin{bmatrix}\overbrace{0\ldots0}^{r-1}&\overbrace{1\ldots1}^{p}&\overbrace{0\ldots0}^{p}&\overbrace{1\ldots1}^{r-1}\\0\ldots0&0\ldots0&1\ldots1&1\ldots1\end{bmatrix}.$$

In the first $\mathrm{forb}(m,F')$ columns we can find a copy of F', say in rows i,j. Then there are either at most $r-1$ $\begin{bmatrix}0\\0\end{bmatrix}$'s or at most $r-1$ $\begin{bmatrix}1\\1\end{bmatrix}$'s in rows i,j. In the former case, remove a column with $\begin{bmatrix}0\\0\end{bmatrix}$ in rows i,j and in the latter case a column with $\begin{bmatrix}1\\1\end{bmatrix}$. Now rows i,j do not contain F'. Thus after $\binom{m}{2}$ deletions we must stop. Hence

$$\mathrm{forb}(m,F)\leq\mathrm{forb}(m,F')+\binom{m}{2}.$$

The result follows by induction.

A construction of an m-rowed matrix proves the second conclusion. We start with a division of the matrix into a set of m_1 rows on top and a disjoint set of m_2 rows on the bottom. We start with $[K_{m_1}^{m_1}, K_{m_1}^{m_1-1}]\otimes[K_{m_2}^0, K_{m_2}^1]$ and then add the columns $K_{m_1}^{m_1-2}\otimes K_{m_2}^0$ and $K_{m_1}^{m_1}\otimes K_{m_2}^{m_2-2}$ for a total of

$$(m_1+1)(m_2+1)+\binom{m_1}{2}+\binom{m_2}{2}=\frac{m^2+m+2}{2}$$

distinct columns. The trees at the top and bottom will consist of single vertices. At this stage there is a single $\begin{bmatrix}0\\1\end{bmatrix}$ in each pair of rows $\begin{bmatrix}R_i\\R_j\end{bmatrix}$ where R_i is a row at the top, R_2 is a row at the bottom. We now add more columns containing $r-2$ copies of $K_{m_1}^{m_1-2}$ at the top but with entries at the bottom chosen from I_{m_2} so the columns are distinct and we get no more than $p-1$ $\begin{bmatrix}0\\1\end{bmatrix}$'s in each pair of rows one from the top and one from the bottom. This gives $r-1$ $\begin{bmatrix}0\\0\end{bmatrix}$'s in each pair of rows from the top and hence avoids F.
For a column with 0's in the top rows i and j (in the top m_1 rows) in the kth additional copy of $K_{m_1}^{m_1-2}$ we choose the bottom 0's apart from a 1 in the $i+j+(k-1)m_1(\mathrm{mod} m_2)$-th entry. With $m_2=\left\lceil\frac{(r-2)m}{p-2}\right\rceil$ and $m_1<m$, this ensures that we do not get more than $p-2$ additional $\begin{bmatrix}0\\1\end{bmatrix}$'s between a row at the top and a row at the bottom.
We now add columns which have 1's in the top rows and different entries chosen from $K_{m_2}^3$ so that we add the maximum number of columns without obtaining more than $s-1$ $\begin{bmatrix}1\\1\end{bmatrix}$'s in each pair of rows from the bottom. This is a design problem; the number of extra columns of this form is approximately $\frac{s-2}{3}\binom{m_2}{2}$.
The number of columns in this matrix is then approximately

$$\frac{m^2+m+2}{2}+(r-2)\binom{m_1}{2}+\frac{s-2}{3}\binom{m_2}{2}.$$

As p increases, $m_1=m-m_2=\lfloor\frac{(p-r)m}{p-2}\rfloor$ approaches m, so that the coefficient of m^2 in this expression approaches $(r-1)m^2/2$. ■

4 A new bound for copies of the identity

For an $m\times n$ matrix A, let $t\cdot A$ denote the $m\times tn$ matrix obtained from t copies of A (this notation is ambiguous but in this paper we will not be multiplying a matrix by a constant). Let I_k denote the $k\times k$ identity matrix. Before this we only had $\mathrm{forb}(m,p\cdot I_k)$ to be $\mathrm{O}(m^k)$ by Theorem 1.1.

Theorem 4.1 *We have $forb(m,p\cdot I_k)=O(m^{k-1+\frac{1}{k}})$ and $forb(m,p\cdot I_k)=\Omega(m^{k-1})$.*

Proof: We use induction on p noting that Theorem 1.2 gives $\mathrm{forb}(m,I_k)=O(m^{k-1})$. Let A be an $m\times n$ simple matrix with no configuration $p\cdot I_k$. Assume $n>\mathrm{forb}(m,(p-1)\cdot I_k)$

so that A has $(p-1)\cdot I_k$ on rows $\{i_1, i_2, \ldots, i_k\}$. Since A doesn't have $p \cdot I_k$, there is some column of I_k occurring exactly $p-1$ times in A. Without loss of generality we may assume that the column has 0's in rows $i_1, i_2, \ldots, i_{k-1}$ and a 1 in row i_k. We call the set of $p-1$ columns of A with 0's in rows $i_1, i_2, \ldots, i_{k-1}$ and a 1 in row i_k to be a *special set* of columns on $\{i_1, i_2, \ldots, i_k\}$.

If we assume

$$n > \mathrm{forb}(m, (p-1)\cdot I_k) + (p-1)m^{k-1+\frac{1}{k}}$$

we can choose $m^{k-1+\frac{1}{k}}$ disjoint special sets of $p-1$ columns by removing them one set at a time. Use the notation $\binom{X}{t}$ to be all subsets of X of size t. Let $[m] = \{1, 2, \ldots, m\}$. For $S = \{i_1, i_2, \ldots i_{k-1}\} \subseteq \binom{[m]}{k-1}$, let a_S denote the number of disjoint special sets of columns on rows $i_1, i_2, \ldots, i_{k-1}, j$ for any j i.e. special sets with 0's in rows of S. Thus

$$\sum_{S \in \binom{[m]}{k-1}} a_S \geq m^{k-1+\frac{1}{k}}$$

Given a_S column disjoint special sets on rows $i_1, i_2, \ldots, i_{k-1}, j$, for $j \in \{j_1, j_2, \ldots, j_{a_S}\}$, then the columns will have 0's in all rows $j \in \{j_1, j_2, \ldots, j_{a_S}\}$ apart from one row, because the special sets are disjoint. Thus we get $\binom{a_S}{k}$ copies of $(p-1)\cdot I_k$ inside those $(p-1)a_S$ columns and a_S rows $\{j_1, j_2, \ldots, j_{a_S}\}$ and so we get $\binom{a_S}{k}$ new special sets columns contained in the $(p-1)a_S$ columns. Thus

$$\sum_{S \in \binom{[m]}{k-1}} \left(a_S + \binom{a_S}{k}\right) \leq \binom{m}{k}.$$

This yields a contradiction since the sum $\sum_{S \in \binom{[m]}{k-1}} \left(a_S + \binom{a_S}{k}\right)$ will ba minimized (by convexity) with $a_S = m^{k-1+\frac{1}{k}} / \binom{m}{k}$ and so $\binom{a_S}{k} > m$. Now $\sum_{S \in \binom{[m]}{k-1}} \binom{a_S}{k} > m\binom{m}{k-1} > \binom{m}{k}$. This is a contradiction and so $n \leq \mathrm{forb}(m, (p-1)\cdot I_k) + (p-1)m^{k-1+\frac{1}{k}}$. Use induction on p to get $\mathrm{forb}(m, p\cdot I_k) = O(m^{k-1+\frac{1}{k}})$.

A construction to show $\mathrm{forb}(m, p\cdot I_k) = \Omega(m^{k-1})$ is due to Füredi and Quinn [8]. ∎

We believe the truth is $\mathrm{forb}(m, F) = \Theta(m^{k-1})$.

References

[1] R.P. Anstee, General forbidden configuration theorems, *J. Combin. Th. (A)* **40**, (1985), 108-124.

[2] R.P. Anstee and Z. Füredi, Forbidden Submatrices, *Discrete Math.* **62**,(1986), 225-243.

[3] R.P. Anstee, A forbidden configuration theorem of Alon, *J. Combin.Th. (A)* **47**(1988), 16-27.

[4] R.P. Anstee, J.R. Griggs, and A. Sali, Small Forbidden Configurations, *Graphs and Combin.* **13**(1997), 91-118.

[5] P. Erdős and Stone, On the structure of linear graphs, *Bull. Amer. Math. Soc.* **52**,(1946), 1087-1091.

[6] P. Frankl, On the trace of finite sets, *J. Combin. Th. (A)* **34**, (1983), 41-45.

[7] P. Frankl, Z. Füredi and J. Pach, Bounding One-Way Differences, *Graphs and Combin.* **3**, (1987), 341-347.266

[8] Z. Füredi and F. Quinn, Traces of finite sets, *Ars Combin.* **18**, (1983), 195-200.

[9] Z. Füredi, *Private communication.*

[10] N. Sauer, On the density of families of sets, *J. Combin. Th. (A)* **13**(1973) 145-147.

[11] L. Rédei, Ein kombinatorischer Satz, Acta Litt. Szeged **7**(1934),39-43.

[12] S. Shelah, A combinatorial problem: Stability and order for models and theories in infinitary languages, *Pac. J. Math.* **4**(1972), 247-261.

[13] V.N. Vapnik and A.Ya. Chervonenkis, On the uniform convergence of relative frequencies of events to their probabilities, *Th. Prob. and Applics.* **16**(1971), 264-280.

5 Appendix: Some Basic Linear Bounds

This appendix contains some new basic bounds for completeness. Careful counting is employed.

Theorem 5.1 *Let* $F = [\overbrace{0\ 0\ \ \ldots\ \ 0}^{p}]$. *Then for* $m \geq p-1 \geq 1$,

$$\mathrm{forb}(m, F) = 1 + \left\lfloor \frac{pm}{2} \right\rfloor.$$

Proof: The proof uses a counting argument. We use, successively, the columns with the smallest numbers of 0's. The m row matrix containing the column of 1's together with the m columns containing a single 0 has a single 0 in each row. Adding k columns containg two 0's each adds an average of $\frac{2k}{m}$ 0's in each row. We must have

$$1 + \frac{2k}{m} \leq p - 1,$$

or,

$$k \leq \frac{(p-2)m}{2},$$

for the matrix not to contain F. The number of columns is

$$1+m+k \leq 1+\left\lfloor \frac{pm}{2} \right\rfloor .$$

With $m \geq p-1$ there are a sufficient number of distinct columns with only two 0's to give a simple $m \times (1+\lfloor \frac{pm}{2} \rfloor)$ with no configuration F. ■

Corollary 5.2 *Let F be as above and $F_1 = \left[\overbrace{0\ 0\ \ldots\ 0}^{p}\ \overbrace{1\ 1\ \ldots\ 1}^{q} \right]$ with $q \leq p$. Then for $m \geq p-1 \geq 1$,*

$$\mathrm{forb}(m, F_1) = \mathrm{forb}(m, F) = 1 + \left\lfloor \frac{pm}{2} \right\rfloor .$$

Proof: Let A be a simple $m \times n$ matrix with no configuration F_1. Each row of A has either fewer than p 0's or fewer than q 1's. If we take the (0,1)-complements of the rows which have fewer than q 1's, we obtain a simple matrix which has fewer than p 0's in every row. Thus

$$\mathrm{forb}(m, F_1) \leq \mathrm{forb}(m, F).$$

The reverse inequality is clear. ■

We use a counting argument as well to establish the following theorem.

Theorem 5.3 *Let $F = \left[\overbrace{\begin{matrix} 0 & 0 & \ldots & 0 \\ 1 & 1 & \ldots & 1 \end{matrix}}^{p} \right]$. Then, for $m \geq \left\lceil \frac{p+1}{2} \right\rceil$ and $p \geq 3$,*

$$2 + \left\lfloor \frac{(p+1)m}{2} \right\rfloor \leq \mathrm{forb}(m, F) \leq 2 + \left\lfloor \frac{(p+1)m}{2} + \frac{p-3}{2} + \frac{p-3}{m-2} \right\rfloor .$$

Proof: We construct a matrix of m rows starting first with the 0's column, the 1's column, the identity and the reverse of the identity. With these we have the configuration $\left[\begin{smallmatrix} 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \end{smallmatrix} \right]$ in each pair of rows, with the other entries either $\left[\begin{smallmatrix} 0 \\ 0 \end{smallmatrix} \right]$'s or $\left[\begin{smallmatrix} 1 \\ 1 \end{smallmatrix} \right]$'s. We can add the configuration $F' =$

$\left[\begin{matrix} \overbrace{\begin{matrix} 0 & \ldots & 0 \\ 1 & \ldots & 1 \end{matrix}}^{p-3} & \overbrace{\begin{matrix} 1 & \ldots & 1 \\ 0 & \ldots & 0 \end{matrix}}^{p-3} \end{matrix} \right]$ to each pair of rows without obtaining the configuration F. Each of the columns containing either exactly two 0's or exactly two 1's will contribute $2(m-2)$ $\left[\begin{smallmatrix} 0 \\ 1 \end{smallmatrix} \right]$ configurations to the $\binom{m}{2}$ pairs of rows, with columns containing more that two 1's and more than two 0's contributing more. If we add k additional distinct columns without obtaining F, we must therefore have

$$2(m-2)k \leq 2(p-3)\binom{m}{2} .$$

The number of columns in our matrix is then

$$2+2m+k \leq 2+2m+\binom{p-3}{m-2}\binom{m(m-1)}{2} = 2+\frac{(p+1)m}{2}+\frac{p-3}{2}+\frac{p-3}{m-2},$$

giving the inequality on the right.

To prove the first inequality, we construct a matrix as before, starting with the 0's column, the 1's column, the identity and the reverse of the identity.
Case 1 ($m \geq p-2$): We add $\lfloor (p-3)m/2 \rfloor$ distinct columns with exactly two 0's without adding more than $p-3$ zeros to each row and hence no more than the allowed matrix F'. Thus

$$2+2m+\left\lfloor \frac{(p-3)}{2} \right\rfloor = 2+\left\lfloor \frac{(p+1)}{2} \right\rfloor \leq \text{forb}(m,F).$$

Case 2 ($(p+1)/2 \leq m < p-2$): We first add all of the $\binom{m}{2}$ distinct columns containing two 0's. These add $m-2$ $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ configurations to each pair of rows. We can then add $\lfloor (p-m-1)m/2 \rfloor$ columns with two 1's without adding more than $(p-3-(m-2))$ 1's to each row. These $(p-m-1)m/2+m(m-1)/2=(p-2)m/2$ added columns will not have the configuration F'. Thus, for this case,

$$2+2m+\left\lfloor \frac{(p-2)}{2} \right\rfloor = 2+\left\lfloor \frac{(p+2)}{2} \right\rfloor \leq \text{forb}(m,F),$$

which is actually a better lower bound. ■

Corollary 5.4 *For F as above and $m = p-1$ or $m = (p+1)/2$ with p odd the upper bound is achieved.*

Proof: In the first case the columns which contain exactly two 0's have precisely the configuration F' with other entries being $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$'s or $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$'s in each pair of rows. In the second case, the columns containing exactly two 0's along with their complements have the same composition for each pair of rows. ■