Forbidden Configurations and Product Constructions

R.P. Anstee*[1], Christina Koch†[1], Miguel Raggi‡[1], A. Sali§[2]

[1] Mathematics Department, The University of British Columbia, Vancouver, B.C. Canada V6T 1Z2

[2] Alfréd Rényi Institute, Budapest, Hungary

Abstract. A *simple* matrix is a (0,1)-matrix with no repeated columns. For a (0,1)-matrix F, we define that a (0,1)-matrix A has F as a *configuration* if there is a submatrix of A which is a row and column permutation of F (*trace* is the set system version of a *configuration*). Let $\|A\|$ denote the number of columns of A. We define

$\text{forb}(m, F) = \max\{\|A\| : A \text{ is } m\text{-rowed simple matrix and has no configuration } F\}$.

We extend this to a family $\mathcal{F} = \{F_1, F_2, \ldots, F_t\}$ and define

$\text{forb}(m, \mathcal{F}) = \max\{\|A\| : A \text{ is } m\text{-rowed simple matrix and has no configuration } F \in \mathcal{F}\}$.

We consider products of matrices. Given an $m_1 \times n_1$ matrix A and an $m_2 \times n_2$ matrix B, we define the *product* $A \times B$ as the $(m_1 + m_2) \times n_1 n_2$ matrix columns consist of all possible combinations obtained from placing a column of A on top of a column of B. Let I_k denote the $k \times k$ identity matrix, let I_k^c denote the (0,1)-complement of I_k and let T_k denote the $k \times k$ upper triangular (0,1)-matrix with a 1 in position i, j if and only if $i \leq j$. We show $\text{forb}(m, \{I_2 \times I_2, T_2 \times T_2\})$ is $\Theta(m^{3/2})$ while obtaining a linear bound when forbidding all 2-fold products of all 2×2 (0,1)-simple matrices. For two matrices F, P, where P is m-rowed, let $f(F, P) = \max_A\{\|A\| : A \text{ is } m\text{-rowed submatrix of } P \text{ with no configuration } F\}$. We establish $f(I_2 \times I_2, I_{m/2} \times I_{m/2})$ is $\Theta(m^{3/2})$ whereas $f(I_2 \times T_2, I_{m/2} \times T_{m/2})$ and $f(T_2 \times T_2, T_{m/2} \times T_{m/2})$ are both $\Theta(m)$. Additional results are obtained. We use the results on patterns due to Marcus and Tardos and generalizations due to Klazar and Marcus, Balogh, Bollobás and Morris.

Key words. VC-dimension, forbidden configurations, trace, patterns, products

1. Introduction

The investigations into the extremal problem of the maximum number of edges in an n vertex graph with no subgraph H originated with Erdős and Stone [11] and Erdős and Simonovits [10] . There is a large and illustrious literature. There are several ways to generalize to the hypergraph setting. Typically we consider *simple hypergraphs*, namely those with no repeated edges. One can consider a r-uniform hypergaph H and forbid a given subhypergraph H', itself a r-uniform hypergraph. One can consider a r-uniform

* Research supported in part by NSERC and Hungarian National Research Fund (OTKA) grant no. NK 78439

† Research supported in part by NSERC of first author

‡ Research supported in part by NSERC of first author

§ Research was supported in part by Hungarian National Research Fund (OTKA) grant no. NK 78439

hypergraph H and forbid a given *trace*. Or one can extend to general hypergraphs and forbid a given *trace*. This latter problem in the language of matrices is our focus. We say a matrix is *simple* if it is a (0,1)-matrix and there are no repeated columns. Given a (0,1)-matrix F, we say a matrix A has F as a *configuration* denoted $F \in A$, if there is a submatrix of A which is a row and column permutation of F. Let $\|A\|$ denote the number of columns in A. We define

$\text{forb}(m, F) = \max\{\|A\| : A \text{ is } m\text{-rowed simple matrix with no configuration } F\}$.

We recall an important conjecture from [5]. Let I_k denote the $k \times k$ identity matrix, I_k^c denotes the (0,1)-complement of I_k, and T_k denotes the $k \times k$ upper triangular matrix whose ith column has 1's in rows $1, 2, \ldots, i$ and 0's in the remaining rows. For p matrices $m_1 \times n_1$ matrix A_1, an $m_2 \times n_2$ matrix $A_2, \ldots$, an $m_p \times n_p$ matrix A_p we define $A_1 \times A_2 \times \cdots \times A_p$ as the $(m_1 + \cdots + m_p) \times n_1 n_2 \cdots n_p$ matrix whose columns consist of all possible combinations obtained from placing a column of A_1 on top of a column of A_2 on top of a column of A_3 etc. For example the vertex-edge incidence matrix of the complete bipartite graph $K_{m/2,m/2}$ is $I_{m/2} \times I_{m/2}$. Define $\mathbf{1}_k$ to be the $k \times 1$ column of 1's and $\mathbf{0}_\ell$ to be the $\ell \times 1$ columns of 0's. We can define $\mathbf{1}_k\mathbf{0}_\ell$ to be the $(k+\ell) \times 1$ column $\mathbf{1}_k \times \mathbf{0}_\ell$.

Conjecture 1. [5] Let F be a $k \times \ell$ matrix with $F \neq \mathbf{1}_1\mathbf{0}_1$. Let $X(F)$ denote the largest p such that there are choices $A_1, A_2, \ldots, A_p \in \{I_{m/p}, I^c_{m/p}, T_{m/p}\}$ so that $F \notin A_1 \times A_2 \times \cdots \times A_p$, then $\text{forb}(m, F) = \Theta(m^{X(F)})$.

We are assuming p divides m which does not affect asymptotic bounds. We obtain evidence that supports the conjecture while also indicating some potential difficulties. We will be considering F that are products of 2-rowed matrices. The following are the maximal 2-rowed simple submatrices of the matrices I, T, I^c of the conjecture. Let

$$E_1 = \begin{bmatrix} 0\,1\,0 \\ 0\,0\,1 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 0\,1\,1 \\ 0\,0\,1 \end{bmatrix}, \quad E_3 = \begin{bmatrix} 1\,0\,1 \\ 0\,1\,1 \end{bmatrix}$$

For an m-rowed matrix P, we define

$$f(F, P) = \max\{\|A\| : A \text{ is } m\text{-rowed submatrix of } P \text{ with no configuration } F\}.$$

Let K_m denote the $m \times 2^m$ simple matrix consisting of all possible different columns. Then $\text{forb}(m, F) = f(F, K_m)$.

Theorem 1. *$f(E_1 \times E_1, I_{m/2} \times I_{m/2})$ is $\Theta(m^{3/2})$.*

Theorem 2. *$f(E_1 \times E_2, I_{m/2} \times T_{m/2}) \leq 2m$.*

Theorem 3. *$f(E_2 \times E_2, T_{m/2} \times T_{m/2}) \leq 2m$.*

The bound of Theorem 1 is perhaps unexpected in view of Conjecture 1 but it is not a counterexample. The remaining three cases ($E_1 \times E_3$ in $I_{m/2} \times I^c_{m/2}$, $E_2 \times E_3$ in $T_{m/2} \times I^c_{m/2}$ and $E_3 \times E_3$ in $I^c_{m/2} \times I^c_{m/2}$) follow by taking appropriate (0,1)-complements. The proof of Theorem 3 is in Section 3, the proof of Theorem 1 is in Section 4 and the proof of Theorem 2 is in Section 5. Related results such as $f(E_1 \times E_2 \times E_2, I_{m/3} \times T_{m/3} \times T_{m/3})$ is $\Theta(m^2)$ being $\Theta(m^2)$ (Lemma 6) are proved in Section 3,Section 4,Section 5.

A central idea to many of our proofs is to encode columns of a p-fold product $A_1 \times A_2 \times \cdots \times A_p$ as 1's in a p dimensional (0,1)-array whose ith coordinate is indexed by the columns of A_i. In Section 2 we use results about *patterns* including the fundamental result of Marcus and Tardos [16]) and generalizations of Klazar and Marcus [14], Balogh, Bollobás and Morris [8]. We establish some exact bounds for some small cases. We relate patterns to forbidden configuration results in Section 3. The following basic result is proven in Section 3.

Proposition 1. *Let p, q, r, u, v, w be given positive integers. Define $x^+ = \max\{0, x\}$. The configuration given by the product $F = \overbrace{E_1 \times \cdots \times E_1}^{u} \times \overbrace{E_2 \times \cdots \times E_2}^{v} \times \overbrace{E_3 \times \cdots \times E_3}^{w}$ is contained in the a-fold product $\overbrace{I_{m/a} \times \cdots \times I_{m/a}}^{p} \times \overbrace{T_{m/a} \times \cdots \times T_{m/a}}^{q} \times \overbrace{I^c_{m/a} \times \cdots \times I^c_{m/a}}^{r}$ (where we have set $a = p + q + r$) if and only if*

$$2\left((u-p)^+ + (v-q)^+ + (w-r)^+\right) \le (p-u)^+ + (q-v)^+ + (r-w)^+.$$

For example with $u = 2, q = 3$ and the rest being 0, Proposition 1 yields that $E_1 \times E_1 \notin T \times T \times T$ and hence $\text{forb}(m, E_1 \times E_1)$ is $\Omega(m^3)$.

We now consider forbidden families of configurations. We have noted (in [3]) that $\text{forb}(m, \{E_1, E_2, E_3\}) = 2$. Balogh and Bollobás [7] have the much more general result that for a given k, there is a constant c_k such that $\text{forb}(m, \{I_k, T_k, I^c_k\}) = c_k$.

Let $\{E_1, E_2, E_3\} \times \{E_1, E_2, E_3\}$ denote the 6 possible 2-fold products whose terms are chosen from $\{E_1, E_2, E_3\}$. We would like to compute $\text{forb}(m, \{E_1, E_2, E_3\} \times \{E_1, E_2, E_3\})$ but in the interest of a more tractable proof we consider I_2 as a replacement for both E_1 and E_3 (I^c_2 is the same configuration as I_2) and T_2 as replacement for E_2. We note $\text{forb}(m, \{I_2, T_2\}) = 2$ and the bounds of Theorem 1, Theorem 2 and Theorem 3 apply. In Section 6 we prove:

Theorem 4. *forb$(m, \{I_2 \times I_2, T_2 \times T_2\})$ is $\Theta(m^{3/2})$.*

$$I_2 \times I_2 = \begin{bmatrix} 1\,1\,0\,0 \\ 0\,0\,1\,1 \\ 1\,0\,1\,0 \\ 0\,1\,0\,1 \end{bmatrix}, \quad T_2 \times T_2 = \begin{bmatrix} 1\,1\,1\,1 \\ 0\,0\,1\,1 \\ 1\,1\,1\,1 \\ 0\,1\,0\,1 \end{bmatrix}. \tag{1}$$

We make novel use of our *standard decomposition* (4) that has been useful studying forbidden configurations in the past. We also prove the following exact bound in Section 7, which contrasts with Theorem 4. The following four matrices are all 2×2 simple matrices (up to row and column permutations). Let

$$I_2 = \begin{bmatrix} 1\,0 \\ 0\,1 \end{bmatrix} \quad T_2 = \begin{bmatrix} 1\,1 \\ 0\,1 \end{bmatrix} \quad U_2 = \begin{bmatrix} 0\,0 \\ 0\,1 \end{bmatrix} \quad V_2 = \begin{bmatrix} 1\,0 \\ 1\,0 \end{bmatrix}.$$

We note $\text{forb}(m, \{I_2, T_2, U_2, V_2\}) = 1$. Define $\{I_2, T_2, U_2, V_2\} \times \{I_2, T_2, U_2, V_2\} = \{X \times Y \,:\, X, Y \in \{I_2, T_2, U_2, V_2\}\}$ as the 10 possible products of these matrices.

Theorem 5. *We have forb$(m, \{I_2, T_2, U_2, V_2\} \times \{I_2, T_2, U_2, V_2\}) = m + 3$.*

The following definitions are used. Let $[n] = \{1, 2, \ldots, n\}$. We discuss d-dimensional (0,1)-arrays B. We define $\sigma_1(B)$ to be the number of 1's in B. Thinking of the positions in B as elements of $[n]^d$, we let the coordinates of B be $x_1, x_2, \ldots, x_d$ and for a position $\mathbf{y} \in [n]^d$ we define $x_i(\mathbf{y})$ to be the value of coordinate x_i in $\mathbf{y}$. Let $\text{proj}_i(B)$ denote the

$[n]^{d-1}$ $(d-1)$-dimensional (0,1)-array obtained from B by projecting in the direction e_i (where e_i is the d-dimensional (0,1)-vector with a single 1 in coordinate x_i). For each position $\mathbf{y} \in [n]^d$ we form $\mathbf{y}_{\bar{i}}$ in $[n]^{d-1}$ by deleting the ith coordinate of $\mathbf{y}$. If B has a 1 in position $\mathbf{z}$, then we place a 1 in position $\mathbf{z}_{\bar{i}}$ in $\text{proj}_i(B)$. We repeat for all entries of B. All other entries in $\text{proj}_i(B)$ are 0. For example if B is a $2 \times 2 \times 2$ 3-dimensional array with 1's in positions $(1,1,2),(1,2,1),(2,2,1)$ and then $\text{proj}_1(B)$ has two 1's in positions $(1,2),(2,1)$ while $\text{proj}_i 2(B)$ has three 1's in positions $(1,2),(1,1),(2,1)$.

2. Splits

We are going to consider the maximum number of 1's in a $n_1 \times n_2$ (0,1)-matrix A subject to some property. The problems in this section are close relatives of Zarankiewicz' problem [15],[12] and indeed the investigations of *patterns* [13],[16],[17],[14] and have the slightly geometric flavour of points in space.

For any subset $R(1) \subset [n_1]$ and $R(2) \subset [n_2]$ we define $A|_{(R(1),R(2))}$ as the submatrix of A formed of the entries contained in the rows $R(1)$ and in the columns $R(2)$. In this section we will be considering cases where both $R(1)$ and $R(2)$ consist of consecutive integers. Let p_1, p_2 be given with $1 \le p_j \le n_j$ for $j = 1,2$. Assume we are given $I(j) = \{r_1(j), r_2(j), \ldots, r_{p-1}(j)\}$ with $0 < r_1(j) < r_2(j) < \cdots < r_{p_j-1}(j) < n_j$ for $j = 1,2$. Define $R_1(j) = \{1,2,\ldots,r_1(j)\}$, $R_i(j) = \{r_{i-1}(j)+1, r_{i-1}(j)+2, \ldots, r_i(j)\}$ for $1 < i < p_j$ and $R_{p_j}(j) = \{r_{p_j-1}(j)+1, r_{p_j-1}(j)+2, \ldots, n_j\}$. We observe that $\cup_{i=1}^{p_j} R_i(j) = [n_j]$. We say that A has a p_1, p_2 *split* if there are choices $I(1), I(2)$ and hence $R_i(1)$ for $1 \le i \le p_1$ and $R_j(2)$ for $1 \le j \le p_2$ so that $A|_{(R_i(1),R_j(2))}$ is a non zero matrix for all choices $1 \le i \le p_1$ and $1 \le j \le p_2$. Let $g(m,n;k,k)$ denote the maximum number of 1's in a $m \times n$ (0,1)-matrix that does not have a k,k split. Below is an example of a 3,3 split where a 1 from each block is indicated

$$
\begin{array}{c}
\\ R_1(1) \\ \\ R_2(1) \\ \\ R_3(1) \\ \\
\end{array}
\begin{array}{ccc}
R_1(2) & R_2(2) & R_3(2) \\
\left[\begin{array}{c|c|c}
 & 1 & \\
1 & & 1 \\
\hline
1 & 1 & 1 \\
 & & \\
\hline
1 & & \\
 & 1 & 1
\end{array}\right]
\end{array}
$$

A 3,3 split

Theorem 6. *Marcus and Tardos [16]. Let k be given. Then there exists a constant c_k such that $g(n,n;k,k) \le c_k n$.*

The result in [16] involving forbidden permutation patterns implies the above result by choosing the permutation appropriately. Moreover the proof directly extends to the above result. While the constants involved in [16] are not optimal, we can produce best possible constants for small values:

Theorem 7. *Let m,n be given with $m,n \ge 2$. Then $g(m,n;2,2) = m+n-1$ and $g(m,n;3,3) = 2m+2n-4$.*

Proof: An $m \times n$ matrix B with $(p_1-1)m + (p_2-1)n - (p_1-1)(p_2-1)$ 1's can be constructed with 1's in the first $p_1 - 1$ rows and the first $p_2 - 1$ columns. Then B has no p_1, p_2 split.

The following graph theory argument proves the upper bound for $p_1 = 2$ and $p_2 = 2$. Consider the bipartite graph on $m+n$ vertices $r_1, r_2, \ldots, r_m, c_1, c_2, \ldots, c_n$ given by the matrix A where a 1 in entry i, j joins vertex r_i to c_j. If $\sigma_1(A) \geq m+n$, then the bipartite graph has a cycle. If we consider the smallest column index c_1 of a vertex of the cycle which joins two vertices $r_1 < r_2$, then setting $I(1) = \{r_1\}$ and $I(2) = \{c_1\}$ yield a 2,2 split of A.

Now consider $p_1 = p_2 = 3$. Let A be a given $m \times n$ (0,1)-matrix with $\sigma_1(A) > 2m+2n-4$. Now create a new matrix A' from A by deleting, if possible, the topmost 1 and bottommost 1 in each column and so $\sigma_1(A') \geq \sigma_1(A) - 2n$. Note that we use deleting a 1 to refer to replacing a 1 by a 0. Now create a new matrix A'' from A' by deleting, if possible, the two remaining rightmost 1's in each row and so $\sigma_1(A'') \geq \sigma_1(A') - (2m-4)$ where we note that A' has no 1's in the first or last row. By hypothesis, $\sigma_1(A'') > 0$. Say there is a 1 in A'' in position r_2, c_1. By construction there are two entries in A' to the right of that 1, say in positions r_2, c_2 and r_2, c_3. By construction for each of these three entries there are 1's above and below in A in columns c_1, c_2, c_3. We now identify a $3, 3$ split in A by setting $I(1) = \{r_2 - 1, r_2\}$ and $I(2) = \{c_1, c_2\}$. We conclude that $g(m, n; 3, 3) \leq 2m + 2n - 4$. ■

This proof technique was introduced to the authors by Jozsef Solymosi as a curling technique (the winter sport of curling uses a strategy called 'peeling').

The papers [14],[8] consider Theorem 6 generalized to d-dimensional arrays. The following is our notation. Given integers $n_1, n_2, \ldots, n_d$ we can consider the positions $\prod_{i=1}^{d}[n_i]$ in an $n_1 \times n_2 \times \cdots \times n_d$ (0,1)-array A. Our main interest is in the case $n_1 = n_2 = \cdots = n_d$. Let $p_1, p_2, \ldots, p_d \geq 2$ be given. Assume we have d sets of indices $I(j) = \{r_1(j), r_2(j), \ldots, r_{p_j-1}(j)\}$ for coordinate j, for $j = 1, 2, \ldots, d$. We can form d sets $R_1(j), R_2(j), \ldots, R_{p_j}(j)$ with $\cup_{i=1}^{p_j} R_i(j) = [n_i]$ as follows: $R_1(j) = \{1, 2, \ldots, r_1(j)\}$, $R_2(j) = \{r_1(j)+1, r_1(j)+2, \ldots, r_2(j)\}$,..., $R_{p_j}(j) = \{r_{p_j-1}(j)+1, r_{p_j-1}(j)+2, \ldots, n_j\}$. We say A has a $p_1, p_2, \ldots, p_d$ split if we can choose the sets as above and for each $j \in [d]$ and for each possible choice $t \in [p_j]$ with $R(j) = R_t(j)$, the $\prod_{i=1}^{d} p_j$ block $A|_{(R(1),R(2),\ldots,R(d))}$ contains at least one 1. Let $g(n_1, n_2, \ldots, n_d; p_1, p_2, \ldots, p_d)$ be the maximum number of 1's in $n_1 \times n_2 \times \cdots \times n_d$ (0,1)-array that has no $p_1, p_2, \ldots, p_d$ split. The following yields the asymptotics.

Theorem 8. *Klazar and Marcus [14], Balogh, Bollobás and Morris [8]. Let k, d be given. Then there exists a constant $c_{k,d}$ so that $g(\overbrace{n, \ldots, n}^{d}; \overbrace{k, \ldots, k}^{d}) \leq c_{k,d} n^{d-1}$.*

The easy way to have many 1's in a d-dimensional array and have no $p_1, p_2, \ldots, p_d$ split is to place 1's in all positions which, for some choice $i \in [d]$ has the ith coordinate less than p_i. There will be $\prod_{i=1}^{d} p_i$ blocks and the block $(R_{p_1}(1), R_{p_2}(2), \ldots, R_{p_d}(d))$ will have a 1 and its coordinate in the ith direction will be at least p_i. In general this is not optimal. We may extend the argument in Theorem 7 to $\overbrace{3, 3, \ldots, 3}^{d-1}, q$ splits of d-dimensional arrays. It is surprising that we get exact results here yet do not have a reasonable bound for $g(m, n; 4, 4)$ but can conjecture $g(m, m; 4, 4) = 7m - 13$.

Theorem 9. *Let B be the $\prod_{i=1}^{d}[m_i]$ (0,1)-array with 1's in entries whose j coordinate is 1 or 2 for some $j = 1, 2, \ldots, d-1$ or whose dth coordinate is $1, 2, \ldots$ or $q-1$. The matrix*

B has no $\overbrace{3,3,\ldots,3}^{d-1},q$ split. Let A be any $\prod_{i=1}^{d}[m_i]$ (0,1)-array with no $\overbrace{3,3,\ldots,3}^{d-1},q$ split. Then $\sigma_1(A)\le\sigma_1(B)$, and hence for $q=3$, $g(m_1,m_2,\ldots,m_d;\overbrace{3,3,\ldots,3}^{d-1},q)=\sigma_1(B)$.

Proof: Let A be any $\prod_{i=1}^{d}[m_i]$ (0,1)-array with no $\overbrace{3,3,\ldots,3}^{d-1},q$ split. For each direction e_i with $i=1,2,\ldots,d-1$, delete from A, if possible, the two 1's with the smallest and largest coordinate value x_i. Finally in the direction e_d, delete from A, if possible, the largest $q-1$ points in each line in the direction e_d. Let A' be the resulting (0,1)-array.

Now if A' has a 1 in position $(y_1,y_2,\ldots,y_d)$ then we note that A has 1's in positions $(y_1,y_2,\ldots,y_d(j))$ for $j\in[q]$ where $y_d=y_d(1)$ and $y_d(1)<y_d(2)<\cdots<y_d(q)$. It is now straightforward to show that choosing indices $I(1)=\{y_1-1,y_1\}$, $I(2)=\{y_2-1,y_2\}$, ..., $I(d-1)=\{y_{d-1}-1,y_d\}$ and $I(d)=\{y_d(1),y_d(2),\ldots,y_d(q-1)\}$ yields a $\overbrace{3,3,\ldots,3}^{d-1},q$ split. We can show that $\sigma_1(B)\le\sigma_1(A)-\sigma_1(A')$, hence if $\sigma_1(A)>\sigma_1(B)$, then A would have the desired split. ■

3. Submatrices of $T\times T$

We show how to exploit the results about splits in the context of forbidden configurations but begin with the following elementary argument.

Proof of Proposition 1: We note that any row from E_i contains [0 1] and we define $K_2=[0\ 1]\times[0\ 1]$. None of our 2-rowed product terms I,T,I^c contain K_2. Two rows of F chosen from two different terms of the $(u+v+w)$-fold product, will necessarily contain K_2. This implies that if F is contained in the a-fold product $\overbrace{I_{m/a}\times\cdots\times I_{m/a}}^{p}$ $\times\overbrace{T_{m/a}\times\cdots\times T_{m/a}}^{q}\times\overbrace{I^c_{m/a}\times\cdots\times I^c_{m/a}}^{r}$, then each product term $I_{m/a}$, $T_{m/a}$, $I^c_{m/a}$ has at most 2 rows of F and if it has two rows then they come from the same 2-rowed product term E_i of F. Of the three matrices $I_{m/a},T,I^c$, we note that we can find E_1 only in $I_{m/a}$, E_2 only in $T_{m/a}$ and E_3 only in $I^c_{m/a}$. ■

Proof of Theorem 3 that $f(E_2\times E_2,T_{m/2}\times T_{m/2})\le 2m$. Let $F=E_2\times E_2$. Recall the ith column of T_k is the column with 1's in rows $1,2,\ldots,i$ and 0's in the remaining rows. Let A be an m-rowed submatrix of $T_{m/2}\times T_{m/2}$. We can create an $m/2\times m/2$ (0,1)-matrix B from A by placing a 1 in position (r,c) if A contains the column obtained from the rth column of $T_{m/2}$ placed on top of the cth column of $T_{m/2}$ namely the column with 1's only in rows $1,2,\ldots r$ and $m+1,m+2,\ldots,m+c$. We note that $\|A\|$ is $\sigma_1(B)$.

We claim that A has F as a configuration if and only if B has a $3,3$ split. The only way for a submatrix of $T_{m/2}\times T_{m/2}$ to be a row and column permutation of F is to lie in rows $r_1,r_2,m/2+c_1,m/2+c_2$ for some choices $2\le r_1<r_2\le m/2$ and $2\le c_1<c_2\le m/2$ (using the argument of Proposition 1 for $E_2\times E_2$ and noting that first row of $T_{m/2}$ is 1's). We have that any two rows of the upper triangular matrix $T_{m/2}$ (not including the first) have a copy of E_2. We note that the t th column of $T_{m/2}$ on rows r_1,r_2 (with $r_1<r_2$) has

$$\begin{array}{c} t \\ \begin{array}{c} r_1 \\ r_2 \end{array}\begin{bmatrix}0\\0\end{bmatrix} \end{array} \text{ for } 1\le t<r_1,\quad \begin{array}{c} t \\ \begin{array}{c} r_1 \\ r_2 \end{array}\begin{bmatrix}1\\0\end{bmatrix} \end{array} \text{ for } r_1\le t<r_2,\quad \text{and } \begin{array}{c} t \\ \begin{array}{c} r_1 \\ r_2 \end{array}\begin{bmatrix}1\\1\end{bmatrix} \end{array} \text{ for } r_2\le t. \qquad (2)$$

Assume A has a copy of F in the 4 rows $r_1, r_2, m/2+c_1, m/2+c_2$. We discover that the nine columns of F would correspond to nine 1's, one 1 in each of the nine blocks in the $3,3$ split of B given by $I(1)=\{r_1-1, r_2-1\}$ and $I(2)=\{c_1-1, c_2-1\}$ (notation from Section 2). Similarly a $3,3$ split of B yields a copy of F in A. We now appeal to the bound in Theorem 7. ■

An immediate generalization is the following.

Lemma 1. *$f(\overbrace{E_2\times E_2\times\cdots\times E_2}^{d}, \overbrace{T_{m/d}\times T_{m/d}\times\cdots\times T_{m/d}}^{d})$ is $\Theta(m^{d-1})$.*

Proof: Let A be an m-rowed submatrix of $\overbrace{T_{m/d}\times T_{m/d}\times\cdots\times T_{m/d}}^{d})$. We generalize the proof of Theorem 7 and encode A as an d-dimension (0,1)-array B where we place a 1 in B in position $(a_1, a_2, \ldots, a_d)$ if A has a column consisting of the a_1st column of $T_{m/d}$ on top of the a_2nd column of $T_{m/d}$ etc on top of the a_dth column of $T_{m/d}$. As before, $\|A\| = \sigma_1(B)$. We verify that A will have the configuration of the d-fold product $E_2\times E_2\times\cdots\times E_2$ if and only if B has a $\overbrace{3,3,\ldots,3}^{d}$ split. We have an exact bound from Theorem 9 if needed. ■

A further generalization considers the matrix $E_2(k) = [\mathbf{0}_k \,|\, T_k]$ (the columns of k-rowed submatrices of T_m).

Lemma 2. *We have that*

$$f(\overbrace{E_2(k)\times E_2(k)\times\cdots\times E_2(k)}^{d}, \overbrace{T_{m/d}\times T_{m/d}\times\cdots\times T_{m/d}}^{d})$$

is equal to

$$g(m/d, m/d, \ldots, m/d; k+1, k+1, \ldots, k+1),$$

and so is $\Theta(m^{d-1})$.

Proof: We use the d-dimensional generalization of splits Klazar,Marcus [14] and Balogh, Bollabás, Morris [8] where the d-fold product $E_2(k)\times E_2(k)\times\cdots\times E_2(k)$ will correspond to a $\overbrace{k+1, k+1, \ldots, k+1}^{d}$ split. ■

A rather interesting version of Theorem 3 and Lemma 1 that uses the idea of 'peeling' from Theorem 7 is the following.

Lemma 3. *Let $p \geq 3$. Then $f(E_2\times E_2, T_{m/p}\times T_{m/p}\times\cdots\times T_{m/p}) \leq \frac{m}{p}\cdot 4^{p-1}$.*

Proof: Let $F = E_2\times E_2$. We will show that $f(E_2\times E_2, T_{\hat{m}}\times T_{\hat{m}}\times\cdots\times T_{\hat{m}}) \leq 4^{p-1}\hat{m}$. We consider A as an $\hat{m}\times\hat{m}\cdots\times\hat{m})$ p-dimensional (0,1)-array B as follows. Let $x_1, x_2, \ldots, x_p$ be the p coordinate directions in B. The entries in coordinate direction x_i are indexed by the columns of $T_{m/p}$ in the given order. We note that $\|A\| = \sigma_1(B)$.

We first handle the case $p=3$. By Theorem 3, we have that for $i=1,2,3$, $\sigma_1(\mathrm{proj}_i(B))$ is at most $4\hat{m}$. In fact if $\sigma_1(\mathrm{proj}_i(B)) > 4\hat{m}$ then we have a $3,3$ split in $\mathrm{proj}_i(B)$ and that yields F in A where no rows of F come from the ith term $T_{\hat{m}}$ of the product and 2 rows of F come from another $T_{\hat{m}}$ of the product and the other 2 rows of F come from remaining part $T_{\hat{m}}$.

Now proceed to form a matrix B' from B by deleting from B in turn the top 1 in each line in the direction x_3 and then deleting the bottom 1 in each line in the direction x_2 and finally deleting the top two entries in each line in the direction x_1. We have

$$\sigma_1(B) \le \sigma_1(B') + \sigma_1(\text{proj}_3(B)) + \sigma_1(\text{proj}_2(B)) + 2\sigma_1(\text{proj}_1(B)) \le \sigma_1(B') + 4 \cdot 4\hat{m},$$

where we are using the size of the projections to upper bound the number of deleted 1's. Let $\mathbf{y}_1$ be a 1 of B'. Then, by our construction, there are 2 further 1's of B in positions $\mathbf{y}_2, \mathbf{y}_3$ with $x_1(\mathbf{y}_1) < x_1(\mathbf{y}_2) < x_1(\mathbf{y}_3)$, $x_2(\mathbf{y}_1) = x_2(\mathbf{y}_2) = x_2(\mathbf{y}_3)$ and $x_3(\mathbf{y}_1) = x_3(\mathbf{y}_2) = x_3(\mathbf{y}_3)$. For each $\mathbf{y}_j$ for $j = 1, 2, 3$, we will have two 1's in positions $\mathbf{y}'_j, \mathbf{y}''_j$ of B where $\mathbf{y}'_j$ agrees with $\mathbf{y}_j$ except in coordinate x_2 where $x_2(\mathbf{y}'_j) < x_2(\mathbf{y}_j)$ and $\mathbf{y}''_j$ agrees with $\mathbf{y}_j$ except in coordinate x_3 where $x_3(\mathbf{y}_j) < x_3(\mathbf{y}''_j)$. Then these 9 1's in B correspond to a copy of F in A as follows. Note that column t of $T_{\hat{m}}$ has a 0 in row r if and only if $t < r$. We choose two values $a = x_1(\mathbf{y}_2)$ and $b = a+1$ for coordinate x_1 so that when we consider the columns of A corresponding to $\mathbf{y}_1$ (and $\mathbf{y}'_1, \mathbf{y}''_1$ respectively), $\mathbf{y}_2$ (and $\mathbf{y}'_2, \mathbf{y}''_2$ resp.), $\mathbf{y}_3$ (and $\mathbf{y}'_3, \mathbf{y}''_3$ resp.) we have

$$\begin{matrix} & \mathbf{y}_1 & \mathbf{y}_2 & \mathbf{y}_3 \\ \begin{matrix} a \\ a+1 \end{matrix} & \begin{bmatrix} 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \end{bmatrix} \end{matrix}, \qquad \begin{matrix} & \mathbf{y}'_1 & \mathbf{y}'_2 & \mathbf{y}'_3 \\ \begin{matrix} a \\ b \end{matrix} & \begin{bmatrix} 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \end{bmatrix} \end{matrix}, \qquad \begin{matrix} & \mathbf{y}''_1 & \mathbf{y}''_2 & \mathbf{y}''_3 \\ \begin{matrix} a \\ b \end{matrix} & \begin{bmatrix} 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \end{bmatrix} \end{matrix}.$$

Note that for each $j = 1, 2, 3$, we have $x_2(\mathbf{y}'_j) < x_2(\mathbf{y}_j) = x_2(\mathbf{y}''_j)$ and $x_3(\mathbf{y}'_j) = x_3(\mathbf{y}_j) < x_3(\mathbf{y}''_j)$. We can choose a value $c = x_2(\mathbf{y}_1)$ for x_2 and a value $d = x_3(\mathbf{y}_1) + 1$ for x_3 (independent of j) so that in A, the columns corresponding to $\mathbf{y}_j, \mathbf{y}'_j, \mathbf{y}''_j$ have

$$\begin{matrix} & \mathbf{y}'_1 & \mathbf{y}_1 & \mathbf{y}''_1 \\ \begin{matrix} c+\hat{m} \\ d+2\hat{m} \end{matrix} & \begin{bmatrix} 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \end{bmatrix} \end{matrix}, \quad \begin{matrix} & \mathbf{y}'_2 & \mathbf{y}_2 & \mathbf{y}''_2 \\ \begin{matrix} c+\hat{m} \\ d+2\hat{m} \end{matrix} & \begin{bmatrix} 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \end{bmatrix} \end{matrix}, \quad \begin{matrix} & \mathbf{y}'_3 & \mathbf{y}_3 & \mathbf{y}''_3 \\ \begin{matrix} c+\hat{m} \\ d+2\hat{m} \end{matrix} & \begin{bmatrix} 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \end{bmatrix} \end{matrix}.$$

This yields a copy of F in A in rows $a, b, c+\hat{m}, d+2\hat{m}$, a contradiction. We deduce that $\sigma_1(B') = 0$ and hence $\sigma_1(B) \le 4^2\hat{m}$, concluding the proof for $p = 3$.

For $p \ge 4$, we proceed in a somewhat similar fashion but focussing on the first 4 coordinates. By induction on p, $\sigma_1(\text{proj}_i(B))$ is at most $4^{p-2}\hat{m}$. We form a matrix B' from B by deleting from B, if possible, the top 1 in each line in the direction x_4 and then deleting the bottom 1 in each line in the direction x_3 and then deleting the top 1 in each line in the direction x_2 finally deleting the bottom 1 in each line in the direction x_1. We have

$$\sigma_1(B) \le \sigma_1(B') + \sum_{i=1}^{4} \sigma_1(\text{proj}_i(B)) \le \sigma_1(B') + 4 \cdot 4^{p-2}\hat{m} = 4^{p-1}\hat{m},$$

using the fact that $\text{proj}_i(B)$ is a (p-1)-dimensional array and induction on p. Let $\mathbf{y}_1$ be an 1 of B'. Then, by our construction, there are 2 further 1's of B in positions $\mathbf{y}_2, \mathbf{y}_3$ with $x_1(\mathbf{y}_2) < x_1(\mathbf{y}_1)$ and $x_i(\mathbf{y}_2) = x_i(\mathbf{y}_1)$ for $i \ne 1$, and $x_2(\mathbf{y}_1) < x_2(\mathbf{y}_3)$ and $x_i(\mathbf{y}_1) = x_i(\mathbf{y}_3)$ for $i \ne 2$. For each $\mathbf{y}_j$ we will have two 1's in positions $\mathbf{y}'_j, \mathbf{y}''_j$ of B where $\mathbf{y}'_j$ agrees with $\mathbf{y}_j$ except in coordinate x_3 where $x_3(\mathbf{y}'_j) < x_3(\mathbf{y}_j)$ and $\mathbf{y}''_j$ agrees with $\mathbf{y}_j$ except in coordinate x_4 where $x_4(\mathbf{y}_j) < x_4(\mathbf{y}''_j)$. Then these 9 1's in B correspond to a copy of F. In particular we can choose values $a = x_1(\mathbf{y}_1)$, $b = x_2(\mathbf{y}_1) + 1$ so that in A the columns contain

$$\begin{matrix} & \mathbf{y}_2 & \mathbf{y}_1 & \mathbf{y}_3 \\ \begin{matrix} a \\ b+\hat{m} \end{matrix} & \begin{bmatrix} 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \end{bmatrix} \end{matrix}, \quad \begin{matrix} & \mathbf{y}'_2 & \mathbf{y}'_1 & \mathbf{y}'_3 \\ \begin{matrix} a \\ b+\hat{m} \end{matrix} & \begin{bmatrix} 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \end{bmatrix} \end{matrix}, \quad \begin{matrix} & \mathbf{y}''_2 & \mathbf{y}''_1 & \mathbf{y}''_3 \\ \begin{matrix} a \\ b+\hat{m} \end{matrix} & \begin{bmatrix} 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \end{bmatrix} \end{matrix}.$$

As above we can choose values $c = x_3(\mathbf{y}_1)$ and $d = x_4(\mathbf{y}_1) + 1$ and obtain a copy of F in rows $a, b + \hat{m}, c + 2\hat{m}, d + 3\hat{m}$ of A from the 9 columns of A given by the 9 1's of B , a contradiction.

We deduce that $\sigma_1(B') = 0$ and hence $\sigma_1(B) \leq 4^{p-1}\hat{m}$ ■

Some growth in the bound with respect to p is to be expected since forb$(m, E_2 \times E_2)$ is $\Theta(m^3)$.

4. Submatrices of $I \times I$

Proof of Theorem 1 that $f(E_1 \times E_1, I_{m/2} \times I_{m/2})$ is $\Theta(m^{3/2})$. Let $F = E_1 \times E_1$. Let A be a submatrix of $I_{m/2} \times I_{m/2}$ with no configuration F. We consider A as an $(m/2) \times (m/2)$ (0,1)-matrix B whose rows are indexed by the columns of $I_{m/2}$ and whose columns are indexed by the columns of $I_{m/2}$. Then $\|A\| = \sigma_1(B)$. Note that the rth row of $I_{m/2}$ has a 1 in column r of $I_{m/2}$ and 0's in all other columns. Thus a 2×2 submatrix of 4 1's in B in rows r_1, r_2 and columns c_1, c_2 will yield $I_2 \times I_2$ in the associated 4 columns of A (we let (r, c) label the column of A associated with the 1 in position (r, c) of B):

$$\begin{array}{c} \\ r_1 \\ r_2 \\ c_1 \\ c_2 \end{array} \begin{array}{cccc} (r_1,c_1) & (r_1,c_2) & (r_2,c_1) & (r_2,c_2) \\ \left[\begin{array}{c} 1 \\ 0 \\ 1 \\ 0 \end{array}\right. & \begin{array}{c} 1 \\ 0 \\ 0 \\ 1 \end{array} & \begin{array}{c} 0 \\ 1 \\ 1 \\ 0 \end{array} & \left.\begin{array}{c} 0 \\ 1 \\ 0 \\ 1 \end{array}\right] \end{array}$$

The remaining 5 columns of $E_1 \times E_1$ are less structured. Four rows of A contains F if and only if 2 rows of A (say r_1, r_2 chosen from the first $m/2$ rows of A) contain the first two rows of F corresponding to a copy of E_1 in $E_1 \times E_1$ and 2 rows of A (say $c_1 + m/2, c_2 + m/2$ chosen from the last $m/2$) contain the third and fourth rows of F (and the other E_1 in $E_1 \times E_1$). Now A having 9 columns containing F correspond to B having 9 1's as follows: a 2×2 submatrix of 1's in rows r_1, r_2 and columns c_1, c_2 and one more 1 in each row of the 2×2 submatrix and one more 1 in each column of the submatrix and one more 1 in neither of the two chosen rows or two chosen columns. To see this consider a 1 in position r_1, c of B with $c \neq c_1, c_2$. Then column (r_1, c) of A will have $\left[\begin{smallmatrix}1\\0\end{smallmatrix}\right]$ in rows r_1, r_2 and $\left[\begin{smallmatrix}0\\0\end{smallmatrix}\right]$ in rows $c_1 + m/2, c_2 + m/2$. A 1 in position r, c of B with $r \neq r_1, r_2$ and $c \neq c_1, c_2$ will yield a column (r, c) in A that is all 0's on row $r_1, r_2, c_1 + m/2, c_2 + m/2$.

Assume $\|A\| = \sigma_1(B) \geq \left(\frac{m}{2}\right)^{3/2} + 2.5m$. We initially process B by deleting any row or column with at most two 1's (and hence up to $2 \cdot 2 \cdot \frac{m}{2}$ 1's) repeating the deletion process if necessary so that the resulting matrix $\overline{B}$ has row and column sums at least 3. We note that $\sigma_1(B) \leq \sigma_1(\overline{B}) + 2m$. Then $\sigma_1(\overline{B}) \geq \left(\frac{m}{2}\right)^{3/2} + m/2$. We now appeal to Kővari, Sós and Turán [15] for a solution of Zarankiewicz' problem and deduce that $\overline{B}$ has a 2×2 block of 1's and then $\overline{B}$ has the configuration of 9 1's yielding F in A. Thus $\|A\| < \left(\frac{m}{2}\right)^{3/2} + 2.5m$.

A construction using projective planes [15] establishes $f(E_1 \times E_1, I_{m/2} \times I_{m/2})$ is $\Omega(m^{3/2})$. ■

Problem 1. Determine $f(E_1 \times E_1 \times E_1, I_{m/3} \times I_{m/3} \times I_{m/3})$.

The crux of this problem would be determining the maximum number of 1's in a 3-dimensional $(m/3) \times (m/3) \times (m/3)$ (0,1)-array which has no $2 \times 2 \times 2$ subarray of 8 1's.

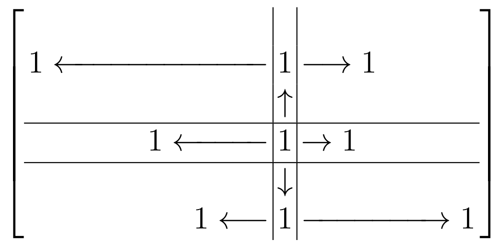
Erdős [9] has obtained a bound $O(m^{11/4})$ for this but only a $\Omega(m^{5/2}$ construction. Note the sharp contrast with results such as Theorem 3, Lemma 1, Lemma 3.

5. Submatrices of $\boldsymbol{I} \times \boldsymbol{T}$

It is useful to state a result about arrangements of 1's in a 2-dimensional array.

Lemma 4. *Let B be an $\hat{m} \times \hat{m}$ matrix with $\sigma_1(B) > 4\hat{m}$. Then there are 9 1's in B in the following configuration consisting of three lines of three 1's:*

Special 3, 3 *split:*



(3)

If the central 1 is in position (r_1, c_1) then we have a 3,3 split using $I(1) = \{r_1 - 1, r_1\}$ and $I(2) = \{c_1 - 1, c_1\}$ (notation from Section 2) ■

Proof: We may argue that $\sigma_1(B) \le 4\hat{m}$ as follows. Form a matrix B' from B by deleting from B, if possible, the top and bottom 1 in each row (a line in the direction x_2) and then deleting, if possible, the top and bottom 1 in each column (a line in the direction x_1). We have that $\sigma_1(B) \le \sigma_1(B') + 4\hat{m}$.

If $\sigma_1(B') > 0$, then select a 1 in B' in position $\mathbf{y}_1$. Then there are two 1's in B in positions $\mathbf{y}_2, \mathbf{y}_3$ in the same column as $\mathbf{y}_1$, $\mathbf{y}_2$ lying below $\mathbf{y}_1$ and $\mathbf{y}_3$ lying above. Then for each $\mathbf{y}_i$ there are two additional 1's in B positions $\mathbf{y}_i', \mathbf{y}_i''$ lying to the left and to the right of $\mathbf{y}_i$ in the same row as $\mathbf{y}_i$.

Now $\mathbf{y}_1 = (r_1, c_1)$ is the central 1 and this yields a 3,3 split as described. ■

Proof of Theorem 2 that $f(E_1 \times E_2, I_{m/2} \times T_{m/2}) \le 2m$. Let $F = E_1 \times E_2$. Let A be an m-rowed submatrix of $I_{m/2} \times T_{m/2}$ with no F. We consider A as an $(m/2) \times (m/2)$ (0,1)-matrix B whose rows are indexed by the columns of $I_{m/2}$ and whose columns are indexed by the columns of $T_{m/2}$ in the usual order. We note that $\|A\| = \sigma_1(B)$.

By Lemma 4, if $\sigma_1(B) > 2m$, we can find an arrangement of 9 1's as in (3). Let $\mathbf{y}_1$ denote the central 1 with $\mathbf{y}_2$ denoting the 1 below $\mathbf{y}_1$ and $\mathbf{y}_3$ denoting the 1 above. Let the position of the three 1's be $\mathbf{y}_i = (r_i, c_1)$ for $i = 1, 2, 3$. Now $r_i = x_1(\mathbf{y}_i) = x_1(\mathbf{y}_i') = x_1(\mathbf{y}_i'')$ for $i = 1, 2, 3$ and so in the columns of A corresponding to the 9 1's we find

$$\begin{array}{c}\\ r_1\\ r_2\\ r_3\end{array}\begin{array}{c}\mathbf{y}_1\\ \begin{bmatrix}1\\0\\0\end{bmatrix}\end{array}\begin{array}{c}\mathbf{y}_1'\\ \begin{bmatrix}1\\0\\0\end{bmatrix}\end{array}\begin{array}{c}\mathbf{y}_1''\\ \begin{bmatrix}1\\0\\0\end{bmatrix}\end{array}\begin{array}{c}\mathbf{y}_2\\ \begin{bmatrix}0\\1\\0\end{bmatrix}\end{array}\begin{array}{c}\mathbf{y}_2'\\ \begin{bmatrix}0\\1\\0\end{bmatrix}\end{array}\begin{array}{c}\mathbf{y}_2''\\ \begin{bmatrix}0\\1\\0\end{bmatrix}\end{array}\begin{array}{c}\mathbf{y}_3\\ \begin{bmatrix}0\\0\\1\end{bmatrix}\end{array}\begin{array}{c}\mathbf{y}_3'\\ \begin{bmatrix}0\\0\\1\end{bmatrix}\end{array}\begin{array}{c}\mathbf{y}_3''\\ \begin{bmatrix}0\\0\\1\end{bmatrix}\end{array}$$

We now use the 3,3 split idea. Let $c_1 = x_2(\mathbf{y}_1) = x_2(\mathbf{y}_2) = x_2(\mathbf{y}_3)$. Column c_1 of $T_{m/2}$ has $\begin{smallmatrix}c_1\\c_1+1\end{smallmatrix}\begin{bmatrix}1\\0\end{bmatrix}$. Also any column a of $T_{m/2}$ with $a < c_1$ has $\begin{smallmatrix}c_1\\c_1+1\end{smallmatrix}\begin{bmatrix}0\\0\end{bmatrix}$ and any column b of $T_{m/2}$ with $c_1 < b$ has $\begin{smallmatrix}c_1\\c_1+1\end{smallmatrix}\begin{bmatrix}1\\1\end{bmatrix}$. Note that $x_2(\mathbf{y}_j') < c_1 = x_2(\mathbf{y}_j) < x_2(\mathbf{y}_j'')$ for $j = 1, 2, 3$. Then we find in A in the 9 columns:

$$\begin{array}{r} \\ c_1+m/2 \\ c_1+1+m/2 \end{array}\begin{array}{ccccccccc} \mathbf{y}_1' & \mathbf{y}_1 & \mathbf{y}_1'' & \mathbf{y}_2' & \mathbf{y}_2 & \mathbf{y}_2'' & \mathbf{y}_3' & \mathbf{y}_3 & \mathbf{y}_3'' \\ \begin{bmatrix}0\\0\end{bmatrix} & \begin{bmatrix}1\\0\end{bmatrix} & \begin{bmatrix}1\\1\end{bmatrix} & \begin{bmatrix}0\\0\end{bmatrix} & \begin{bmatrix}1\\0\end{bmatrix} & \begin{bmatrix}1\\1\end{bmatrix} & \begin{bmatrix}0\\0\end{bmatrix} & \begin{bmatrix}1\\0\end{bmatrix} & \begin{bmatrix}1\\1\end{bmatrix} \end{array}$$

Thus in rows $r_1, r_2, c_1 + m/1, c_1 + 1 + m/2$ of A we find $E_1 \times E_2$. This contradiction establishes the bound.

The construction $A = I_{m/2} \times \mathbf{1}_{m/2}$ avoids F and has $\Omega(m)$ columns. ■

It is useful to give one result about arrangements of 1's in a 3-dimensional array.

Lemma 5. *Let C be an $\hat{m} \times \hat{m} \times \hat{m}$ 3-dimensional (0,1)-array with more than $6\hat{m}^2$ 1's. Then there are 27 1's as follows. There are three values a, b, c for x_1 coordinate such that the three planes $x_1 = a$, $x_1 = b$ and $x_1 = c$ of C each contains 9 points. The 9 points in each plane form a special 3,3 split as in (3) with the central 1 in each of the three planes having the same x_2, x_3 coordinates.*

Proof: Form a matrix C' from C by deleting from C, if possible, the top and bottom 1 in each line in direction x_3 then the top and bottom 1 in each line in the direction x_2 and then the top two 1's in each line in the direction x_1. We obtain

$$\sigma_1(C) \le \sigma_1(C') + 6\hat{m}^2.$$

If C' has a 1 in position $\mathbf{y}_1$, we can find 27 1's yielding a special 3,3,3 split as follows. There are 2 1's of C in positions $\mathbf{y}_2, \mathbf{y}_3$ with $x_1(\mathbf{y}_1) < x_1(\mathbf{y}_2) < x_1(\mathbf{y}_3)$, $x_2(\mathbf{y}_1) = x_2(\mathbf{y}_2) = x_2(\mathbf{y}_3)$ and $x_3(\mathbf{y}_1) = x_3(\mathbf{y}_2) = x_3(\mathbf{y}_3)$. Then there are 1's of C in positions $\mathbf{x}_j$, $\mathbf{z}_j$ for $j = 1, 2, 3$, where $x_2(\mathbf{x}_j) < x_2(\mathbf{y}_j) < x_2(\mathbf{z}_j)$ and $x_1(\mathbf{x}_j) = x_1(\mathbf{y}_j) = x_1(\mathbf{z}_j)$, $x_3(\mathbf{x}_j) = x_3(\mathbf{y}_j) = x_3(\mathbf{z}_j)$. Now for each choice $\mathbf{v} \in \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$, we obtain positions $\mathbf{v}_j', \mathbf{v}_j''$ for $j = 1, 2, 3$ with $x_3(\mathbf{v}_j') < x_3(\mathbf{v}_j) < x_3(\mathbf{v}_j'')$ and $x_1(\mathbf{v}_j') = x_1(\mathbf{v}_j) = x_1(\mathbf{v}_j'')$, $x_2(\mathbf{v}_j') = x_2(\mathbf{v}_j) = x_2(\mathbf{v}_j'')$. In particular there are three planes $x_1 = a$, $x_1 = b$, $x_1 = c$ each with 9 1's and each plane has a special 3,3 split as in (3) with the central 1's of each plane (namely $\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3$) having the same x_2, x_3 coordinates. The horizontal direction in (3) corresponds to the x_3 direction. ■

Lemma 6. *$f(E_1 \times E_2 \times E_2, I_{m/3} \times T_{m/3} \times T_{m/3})$ is $\Theta(m^2)$.*

Proof: Let A be an m-rowed submatrix of $I_{m/3} \times T_{m/3} \times T_{m/3}$ with no configuration $E_1 \times E_2 \times E_2$. As above, we translate A into a 3-dimensional array B with $\|A\| = \sigma_1(B)$.

Now by Lemma 5, if $\sigma_1(B) > 6(m/3)^2$ there will be 27 1's in B as described and we use the notation of the proof. This will yield a copy of $E_1 \times E_2 \times E_2$ in A. Let the central 1 in each plane be in position $\mathbf{y}_i$ for $i = 1, 2, 3$ with $x_1(\mathbf{y}_1) = r_1 < x_1(\mathbf{y}_2) = r_2 < x_1(\mathbf{y}_3) = r_3$. Then there are 1's of B in positions $\mathbf{x}_j$, $\mathbf{z}_j$ for $j = 1, 2, 3$, where $x_2(\mathbf{x}_j) < x_2(\mathbf{y}_j) < x_2(\mathbf{z}_j)$ and $x_1(\mathbf{x}_j) = x_1(\mathbf{y}_j) = x_1(\mathbf{z}_j)$, $x_3(\mathbf{x}_j) = x_3(\mathbf{y}_j) = x_3(\mathbf{z}_j)$. Now for each choice $\mathbf{v} \in \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$, we obtain positions $\mathbf{v}_j', \mathbf{v}_j''$ for $j = 1, 2, 3$ with $x_3(\mathbf{v}_j') < x_3(\mathbf{v}_j) < x_3(\mathbf{v}_j'')$ and $x_1(\mathbf{v}_j') = x_1(\mathbf{v}_j) = x_1(\mathbf{v}_j'')$, $x_2(\mathbf{v}_j') = x_2(\mathbf{v}_j) = x_2(\mathbf{v}_j'')$. Now $r_i = x_1(\mathbf{v}_i') = x_1(\mathbf{v}_i) = x_1(\mathbf{v}_i'')$ for $i = 1, 2, 3$ and also all choices $\mathbf{v} \in \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$. In the columns of A corresponding to the 27 1's we find

$$\begin{array}{r} \\ r_1 \\ r_2 \\ r_3 \end{array}\begin{array}{ccccccccc} \mathbf{v}_1 & \mathbf{v}_1' & \mathbf{v}_1'' & \mathbf{v}_2 & \mathbf{v}_2' & \mathbf{v}_2'' & \mathbf{v}_3 & \mathbf{v}_3' & \mathbf{v}_3'' \\ \begin{bmatrix}1\\0\\0\end{bmatrix} & \begin{bmatrix}1\\0\\0\end{bmatrix} & \begin{bmatrix}1\\0\\0\end{bmatrix} & \begin{bmatrix}0\\1\\0\end{bmatrix} & \begin{bmatrix}0\\1\\0\end{bmatrix} & \begin{bmatrix}0\\1\\0\end{bmatrix} & \begin{bmatrix}0\\0\\1\end{bmatrix} & \begin{bmatrix}0\\0\\1\end{bmatrix} & \begin{bmatrix}0\\0\\1\end{bmatrix} \end{array}$$

Note that in rows r_1, r_2, we see copies of E_1. Now let $c_2 = x_2(\mathbf{y}_1), c_3 = x_3(\mathbf{y}_1)$. Recall that $x_2(\mathbf{y}_1) = x_2(\mathbf{y}_2) = x_2(\mathbf{y}_3) = x_2(\mathbf{y}'_j) = x_2(\mathbf{y}''_j)$ for all $j = 1, 2, 3$. These 9 1's lie on the plane with x_2 coordinate c_2. The same need not be true if we replace $\mathbf{y}$ by $\mathbf{x}$ or $\mathbf{z}$ but, for each choice of $j \in \{1, 2, 3\}$, we have $x_2(\mathbf{x}_j) = x_2(\mathbf{x}'_j) = x_2(\mathbf{x}''_j) < c_2$ and $c_2 < x_2(\mathbf{z}_j) = x_2(\mathbf{z}'_j) = x_2(\mathbf{z}''_j)$. Column c_2 of $T_{m/2}$ has $\begin{smallmatrix} c_2 \\ c_2+1 \end{smallmatrix} \left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right]$. Also any column a of $T_{m/2}$ with $a < c_2$ has $\begin{smallmatrix} c_2 \\ c_2+1 \end{smallmatrix} \left[\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right]$ and any column b of $T_{m/2}$ with $c_2 < b$ has $\begin{smallmatrix} c_2 \\ c_2+1 \end{smallmatrix} \left[\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right]$. Recalling that $x_2(\mathbf{x}_j) < c_2 = x_2(\mathbf{y}_j) < x_2(\mathbf{z}_j)$, we have for each $j = 1, 2, 3$ in A in rows $c_2 + m/1, c_2 + 1 + m/3$ the following in the columns of corresponding to the 27 1's while letting $j = 1, 2, 3$:

$$\begin{array}{r} \\ c_2 + m/3 \\ c_2 + 1 + m/3 \end{array} \begin{array}{ccccccccc} \mathbf{x}_j & \mathbf{y}_j & \mathbf{z}_j & \mathbf{x}'_j & \mathbf{y}'_j & \mathbf{z}'_j & \mathbf{x}''_j & \mathbf{y}''_j & \mathbf{z}''_j \\ \begin{bmatrix} 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \end{bmatrix} \end{array}$$

We use the same ideas noting that $x_3(\mathbf{v}') < x_3(\mathbf{v}) < x_3(\mathbf{v}'')$ for each $j = 1, 2, 3$ and each choice $\mathbf{v} \in \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$. In this case there are 9 1's with the same x_3 coordinate as $\mathbf{y}_1$ (the positions $\mathbf{v}_j$ for all choices $j = 1, 2, 3$ and $\mathbf{v} \in \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$). With $c_3 = x_3(\mathbf{y}_1)$ we have the following in the columns corresponding to the 27 1's while letting $j = 1, 2, 3$:

$$\begin{array}{r} \\ c_3 + 2m/3 \\ c_3 + 1 + 2m/3 \end{array} \begin{array}{ccccccccc} \mathbf{x}'_j & \mathbf{x}_j & \mathbf{x}''_j & \mathbf{y}'_j & \mathbf{y}_j & \mathbf{y}''_j & \mathbf{z}'_j & \mathbf{z}_j & \mathbf{z}''_j \\ \begin{bmatrix} 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \end{bmatrix} \end{array}.$$

We now have a copy of $E_1 \times E_2 \times E_2$ in A, a contradiction and so $\|A\| \le 6(m/3)^2$.

Let α being first column of $I_{m/3}$. Using Proposition 1, the construction $\alpha \times T_{m/3} \times T_{m/3}$ avoids F and has $\Theta(m^2)$ columns. ■

Using an analogous argument one obtains

Lemma 7. $f(E_1 \times \overbrace{E_2 \times \cdots \times E_2}^{p-1}, I_{m/p} \times \overbrace{T_{m/p} \times \cdots \times T_{m/p}}^{p-1})$ *is* $\Theta(m^{p-1})$. ■

6. Proof of the unexpected bound

Proof of Theorem 4. Let A be an m-rowed simple matrix with no configurations $\{I_2 \times I_2, T_2 \times T_2\}$. Our *standard decomposition* on row r considers deleting row r from A and reordering the columns as below with C_r consisting of the columns which are repeated in the matrix obtained from A by deleting row r.

$$A = \begin{array}{c} r \rightarrow \\ \\ \end{array} \begin{bmatrix} 00\cdots0 & 11\cdots1 \\ B_r \quad C_r & C_r \quad D_r \end{bmatrix} \tag{4}$$

Our typical use of this is to note that $[B_rC_rD_r]$ is simple with no configurations $\{I_2 \times I_2, T_2 \times T_2\}$ and C_r is simple with no configurations F_4, F_5:

$$F_4 = \begin{bmatrix} 1\,1 \\ 1\,1 \\ 0\,1 \end{bmatrix}, \quad F_5 = \begin{bmatrix} 1\,1\,0\,0 \\ 1\,0\,1\,0 \\ 0\,1\,0\,1 \end{bmatrix}. \tag{5}$$

In particular if C_r has F_4, then A has $T_2 \times T_2$, and if C_r has F_5, then A has $I_2 \times I_2$, both forbidden configurations. We deduce by induction (on m) that $\|A\| = \|[B_rC_rD_r]\| +$

$\|C_r\| \le \text{forb}(m-1, \{I_2 \times I_2, T_2 \times T_2\}) + \text{forb}(m-1, \{F_4, F_5\})$. We would be done if we could show that $\|C_r\| \le 20m^{1/2}$ for some r. The choice of 20 is an artifact of our proof. We may now assume $\|C_r\| \ge 20m^{1/2}$ for all r and will arrive at a contradiction. Our proof will show that for each $r \in [m]$, we can choose a set of rows $S(r) \subseteq [m]\backslash r$ where $|S(r)| \ge \|C_r\|/4 \ge 5m^{1/2}$. Note that $S(r)$ is a set so $|S(r)|$ denotes *cardinality* while $\|C_r\|$ denotes the *number of columns*. Choose one row r. We show that there is a subset $K \subseteq S(r)$, of size $|S(r)|/3$, every pair of rows $r_i, r_j \in K$ satisfying

$$|S(r_i) \cap S(r_j)| \le 5. \tag{6}$$

Then if we let $t = m^{1/2}$ we can choose $r_1, r_2, \dots, r_t \in K \subseteq S(r)$ and obtain t disjoint sets

$$S(r_1), S(r_2)\backslash S(r_1), S(r_3)\backslash(S(r_1) \cup S(r_2)), \dots, S(r_t)\backslash(S(r_1) \cup S(r_2) \cup \dots \cup S(r_{t-1})).$$

This yields that $S(r_1) \cup S(r_2) \cup S(r_3) \cdots \cup S(r_t)$ is of size at least

$$5m^{1/2} + (5m^{1/2} - 5) + (5m^{1/2} - 10) + \cdots \quad > m,$$

a contradiction given that we have m rows.

Consider the following operation on C_r. Delete as many rows as we can while preserving simplicity of the remaining matrix. Doing this may involve some choices. The remaining set of rows is denoted $R(r)$ and so $C_r|_{R(r)}$ is simple and $\|C_r\| = \|C_r|_{R(r)}\|$. Now by Lemma 8 found below, $\|C_r|_{R(r)}\| \le 2|R(r)|$ and so $|R(r)| \ge \|C_r\|/2$.

We consider the standard decomposition applied to $C_r|_{R(r)}$

$$C_r|_{R(r)} = \begin{matrix} s \rightarrow \\ \\ \end{matrix} \begin{bmatrix} 00\cdots0 & 11\cdots1 \\ E_s \quad G_s & G_s \quad H_s \end{bmatrix} \tag{7}$$

Given our choice for $R(r)$, we deduce that $1 \le \|G_s\|$. As in the proof of Lemma 8, we note that G_s does not have the configurations $\left[\begin{smallmatrix}1\\1\end{smallmatrix}\right], I_2$. We deduce that $1 \le \|G_s\| \le 2$ where G_s will either consist of a column of 0's or a column of sum 1 or both. If G_s has a column of 0's then $C_r|_{R(r)}$ has a column of sum 1 with a 1 in row s. If G_s has a column of sum 1 with a 1 in row t, then $C_r|_{R(r)}$ has a column of sum 2 with 1's in rows s, t and also a column of sum 1 with a 1 in row t. In this latter case record a directed arc $s \rightarrow t$ and in this way form a directed graph D on the rows $R(r)$ with at most $|R(r)|$ arcs. We now indicate how to find a set $S(r) \subseteq R(r)$ so that $C_r|_{S(r)}$ has $I_{|S(r)|}$ and $|S(r)| \ge |R(r)|/2$. First let T denote the set of rows $t \in R(r)$ for which $C_r|_{R(r)}$ has a column of sum 1 with a 1 in row t. Let $U = R(r)\backslash T$. Then for each $u \in U$, there is exactly one arc $u \rightarrow v$ in D and, by our choice of T, $v \in T$. Let $V = \{v \in T : \text{ there is a } u \in U \text{ with } u \rightarrow v\}$. We may now form $S(r) = U \cup (T\backslash V)$. We see that $C_r|_{S(r)}$ has $I_{|S(r)|}$ since for each $t \in T\backslash V$, we take the column of sum 1 with a single 1 in row t and for $u \in U$, we take the column with sum 2 with 1's in rows u, v. In this latter case we have $u \rightarrow v$ and so $v \in V$ hence $v \notin S(r)$. We may verify $|S(r)| \ge |R(r)|/2$ by noting that $|S(r)| = |R(r)| - |V|$ and $|V| \le |U|$.

We need more detailed information and begin by computing what happens on quadruples of rows $\{i, j, k, \ell\}$ of A in order to avoid the two 4×4 configurations $I_2 \times I_2$ and $T_2 \times T_2$. There are 5 cases $Q_0, Q_1, \dots, Q_4$. These cases were computed using a C++ program (can be downloaded at[2]) that had many test runs checking correctness and was also independently checked by a program written in sage (public code that uses Python). In each case one may easily check that, if the case is satisfied, indeed our three matrices of (1) aren't present as configurations. But checking completeness of the list without a computer would require an enormous amount of work. In general we would write

$$\begin{array}{c} \\ i \\ j \\ k \\ \ell \end{array}\begin{array}{c} no \\ \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} \end{array},$$

to denote that in A we do *not* have the specified vector $(a,b,c,d)^T$ on rows $\{i,j,k,\ell\}$ in that order. In what follows, the row order is not specified but is the same for each column for a given case Q_i. In each case, either 3 or 4 or 5 or 6 columns must not be present on a quadruple of rows in order to not contain the configurations $\{I_2 \times I_2, T_2 \times T_2\}$.

$$Q_0 = \overset{no}{\begin{bmatrix}0\\1\\0\\1\end{bmatrix}} \overset{no}{\begin{bmatrix}0\\0\\1\\1\end{bmatrix}} \overset{no}{\begin{bmatrix}1\\1\\1\\1\end{bmatrix}}, \quad Q_1 = \overset{no}{\begin{bmatrix}0\\1\\0\\1\end{bmatrix}} \overset{no}{\begin{bmatrix}1\\1\\0\\1\end{bmatrix}} \overset{no}{\begin{bmatrix}0\\0\\1\\1\end{bmatrix}} \overset{no}{\begin{bmatrix}1\\0\\1\\1\end{bmatrix}} \overset{no}{\begin{bmatrix}0\\1\\1\\1\end{bmatrix}}, \quad Q_2 = \overset{no}{\begin{bmatrix}1\\1\\1\\0\end{bmatrix}} \overset{no}{\begin{bmatrix}0\\1\\0\\1\end{bmatrix}} \overset{no}{\begin{bmatrix}0\\0\\1\\1\end{bmatrix}} \overset{no}{\begin{bmatrix}1\\0\\1\\1\end{bmatrix}},$$

$$Q_3 = \overset{no}{\begin{bmatrix}1\\1\\1\\0\end{bmatrix}} \overset{no}{\begin{bmatrix}1\\0\\0\\1\end{bmatrix}} \overset{no}{\begin{bmatrix}0\\1\\0\\1\end{bmatrix}} \overset{no}{\begin{bmatrix}0\\0\\1\\1\end{bmatrix}}, \quad Q_4 = \overset{no}{\begin{bmatrix}1\\1\\0\\0\end{bmatrix}} \overset{no}{\begin{bmatrix}1\\0\\1\\0\end{bmatrix}} \overset{no}{\begin{bmatrix}0\\1\\1\\0\end{bmatrix}} \overset{no}{\begin{bmatrix}1\\0\\0\\1\end{bmatrix}} \overset{no}{\begin{bmatrix}0\\1\\0\\1\end{bmatrix}} \overset{no}{\begin{bmatrix}0\\0\\1\\1\end{bmatrix}}.$$

In what follows we analyze closely each of the 5 cases above to deduce that the columns of $[B_r\, C_r]|_{S(r)}$ form a laminar family in order to help prove (6).

If we are missing $(a,b,c,d)^T$ on the quadruple of rows r,i,j,k, then we are missing $(b,c,d)^T$ on the triple of rows i,j,k in C_r else A has both $(0,b,c,d)^T$ and $(1,b,c,d)^T$ on rows r,i,j,k, a contradiction. Thus it is possible to determine what is missing in $C_r|_{R(r)}$ on the triple of rows $i,j,k \in R(r)$ by considering what is missing on the quadruple of rows r,i,j,k. We also obtain a contradiction, based on the choice of $R(r)$, if we find a copy of 'K_2' in what is missing, namely if on the triple i,j,k there is a pair of rows i,j with all 4 columns of K_2 appearing as follows:

$$\begin{array}{c} \\ i \\ j \\ k \end{array} \overset{\text{no}}{\begin{bmatrix}0\\0\\a\end{bmatrix}} \overset{\text{no}}{\begin{bmatrix}0\\1\\b\end{bmatrix}} \overset{\text{no}}{\begin{bmatrix}1\\0\\c\end{bmatrix}} \overset{\text{no}}{\begin{bmatrix}1\\1\\d\end{bmatrix}},$$

where $a,b,c,d \in \{0,1\}$. Perhaps other columns are missing on rows i,j,k. Note that we could delete row k from $C_r|_{R(r)}$ and preserve simplicity, contradicting our choice of $R(r)$. The reason for this is that on the three rows i,j,k, the columns present would possibly be

$$\begin{array}{c} i \\ j \\ k \end{array} \begin{bmatrix}0\\0\\\overline{a}\end{bmatrix} \quad \begin{bmatrix}0\\1\\\overline{b}\end{bmatrix} \quad \begin{bmatrix}1\\0\\\overline{c}\end{bmatrix} \quad \begin{bmatrix}1\\1\\\overline{d}\end{bmatrix},$$

where $\overline{x}$ denotes the (0,1)-complement of x. We can see that deleting row k will not result in repeated columns assuming $C_r|_{R(r)}$ has no repeated columns.

Consider the quadruple of rows r, i, j, k. We note that Q_3, Q_4 each contain 3 rows (rows $1, 2, 3$ in each case), each pair of rows of which 'has a K_2' in what is missing and so any choice for r, will leave that on the remaining triple of rows in C_r 'has a K_2'. For $\{i, j, k\} \subseteq R(r)$, the quadruple of rows r, i, j, k cannot be in cases Q_3 or Q_4. Case Q_0 'has no K_2' and so if Q_0 applies to a quadruple of rows, then apparently any row of Q_0 as given could be row r. Case Q_1 has 'has a K_2' on rows 1,2 and also on rows 1,3 and so row r would have to be row 1 of Q_1 as given. Case Q_2 'has a K_2' on rows 1,2 and so row r must either be row 1 or row 2 of Q_2 as given. This produces the following cases. We use P_i to denote a triple arising from the quadruple Q_i in these ways.

$$P_0 = \overset{no}{\begin{bmatrix}1\\0\\1\end{bmatrix}} \overset{no}{\begin{bmatrix}0\\1\\1\end{bmatrix}} \overset{no}{\begin{bmatrix}1\\1\\1\end{bmatrix}} \text{ or } \overset{no}{\begin{bmatrix}0\\0\\1\end{bmatrix}} \overset{no}{\begin{bmatrix}0\\1\\1\end{bmatrix}} \overset{no}{\begin{bmatrix}1\\1\\1\end{bmatrix}} \text{ or } \overset{no}{\begin{bmatrix}0\\1\\1\end{bmatrix}} \overset{no}{\begin{bmatrix}0\\0\\1\end{bmatrix}} \overset{no}{\begin{bmatrix}1\\1\\1\end{bmatrix}} \text{ or } \overset{no}{\begin{bmatrix}0\\1\\0\end{bmatrix}} \overset{no}{\begin{bmatrix}0\\0\\1\end{bmatrix}} \overset{no}{\begin{bmatrix}1\\1\\1\end{bmatrix}},$$

$$P_1 = \overset{no}{\begin{bmatrix}1\\0\\1\end{bmatrix}} \overset{no}{\begin{bmatrix}1\\0\\1\end{bmatrix}} \overset{no}{\begin{bmatrix}0\\1\\1\end{bmatrix}} \overset{no}{\begin{bmatrix}0\\1\\1\end{bmatrix}} \overset{no}{\begin{bmatrix}1\\1\\1\end{bmatrix}}, \quad P_2 = \overset{no}{\begin{bmatrix}1\\1\\0\end{bmatrix}} \overset{no}{\begin{bmatrix}1\\0\\1\end{bmatrix}} \overset{no}{\begin{bmatrix}0\\1\\1\end{bmatrix}} \text{ or } \overset{no}{\begin{bmatrix}1\\1\\0\end{bmatrix}} \overset{no}{\begin{bmatrix}0\\1\\1\end{bmatrix}} \overset{no}{\begin{bmatrix}1\\1\\1\end{bmatrix}}.$$

These then, are the only cases we need consider for what is missing on a triple of rows of $R(r)$ in C_r. Because we know that the columns of $C_r|_{S(r)}$ contain an identity matrix $I_{|S(r)|}$ on the rows $S(r)$, we know that the final three cases of P_0 and the second case of P_2 cannot be what is missing on a triple of rows in $S(r)$. Thus r would have to be the first row in Q_0, Q_1 or Q_2 as given.

Among the remaining options for what is missing (P_1, the first case for P_2 and the first case for P_0), by looking at Q_0, Q_1 and Q_2, we see that below the zeroes in row r, all triple of rows in $S(r)$ are missing two columns of sum 2. This means that the columns in $[B_rC_r]|_{S(r)}$ must form a laminar family. We will use this fact in what follows.

Let $r_j \in S(r)$. The columns of C_{r_j} must correspond to columns of A which appear with a 1 and 0 in row r_j and are the same elsewhere, in particular, on the rows $(r \cup S(r)) \setminus r_j$. We thus have pairs of columns in $A|_{S(r)\cup r}$ from C_{r_j} as follows:

$$\begin{matrix} r_j \\ S(r)\setminus r_j\{ \\ r \end{matrix} \begin{bmatrix} 0 & 1 \\ \alpha & \alpha \\ a & a \end{bmatrix}. \tag{8}$$

We begin with the case where $a = 0$ by considering $[B_rC_r]|_{S(r)}$. Suppose there exist two non-zero choices for α, say $\beta \neq \gamma$. We have the following situation for columns in $A|_{S(r)\cup r}$ from C_{r_j}:

$$\begin{matrix} r_j \\ S(r)\backslash r_j\,\{ \\ r \end{matrix} \begin{bmatrix} 0 & 1 & 0 & 1 \\ \beta & \beta & \gamma & \gamma \\ 0 & 0 & 0 & 0 \end{bmatrix}. \tag{9}$$

We know that the columns of $[B_rC_r]|_{S(r)}$ form a laminar family. Thus the columns from (9) must form a laminar family. Columns 2 and 4 have 1's in common on row r_j and so the columns as sets are not disjoint so one must be contained in the other. We deduce without loss of generality that $\beta \leq \gamma$. Now, considering columns $2, 3$, and the fact that $\beta \neq \mathbf{0}$, we violate the laminar property, a contradiction. Thus there is only one non-zero choice for α in (8) when $a = 0$ but since $C_{r_j}|_{S(r)}$ need not be simple, the column with α could be repeated. If α has column sum at least 2 with 1's on rows $a, b \in S(r)$, and it is repeated, it creates a $T_2 \times T_2$ as follows. This happens because A is a simple matrix and so on some row r' ($r' \notin S(r)$) the columns must differ, yielding the following, with $T_2 \times T_2$ on rows r_j, a, b, r'.

$$\begin{array}{r} r_j \\ a \\ b \\ S(r) \setminus \{r_j, a, b\} \\ r' \end{array} \begin{bmatrix} 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ \alpha' & \alpha' & \alpha' & \alpha' \\ 0 & 1 & 0 & 1 \end{bmatrix}.$$

Thus if any α is repeated, it must have column sum 1. We now consider the second case when $a = 1$ in (8). This case will be simpler since columns with $a = 1$ will automatically yield a row of ones on row r, and $T_2 \times T_2$ has a row of 1's. Suppose there are two non-zero columns β, γ in $C_{r_j}|_{S(r)\setminus r_j}$. Since $C_{r_j}|_{S(r)\setminus r_j}$ need not be simple, we do not require $\beta \neq \gamma$. In some columns of $A|_{S(r)\cup r}$ from C_{r_j} we have

$$\begin{array}{r} r_j \\ S(r)\setminus r_j \{ \\ r \end{array} \begin{bmatrix} 0\,0\,1\,1 \\ \beta\,\gamma\,\beta\,\gamma \\ 1\,1\,1\,1 \end{bmatrix}.$$

Now suppose β and γ both have entry 1 on some row $r' \in S(r) \setminus r_j$. Also, as above, because A is a simple matrix, there is some row r'' (r'' need not be in $S(r)$) where the columns differ. The situation is then as follows:

$$\begin{array}{r} r_j \\ r' \\ S(r) \setminus \{r_j, r', r''\} \\ r'' \\ r \end{array} \begin{bmatrix} 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ \beta' & \gamma' & \beta' & \gamma' \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}.$$

This creates $T_2 \times T_2$ in A. Therefore for any pair of non-zero columns α in $C_{r_j}|_{S(r)\setminus r_j}$, with 1's in row r, there cannot be a row in $S(r)\setminus r_j$ where both are 1's.

We have shown that for the columns in $C_{r_p}|_{S(r)\cup\{r\}}$ which are 0 in row r , then those columns can be $\mathbf{0}$ and/or a single non-zero column α. If α is of column sum 1 it may be repeated, otherwise there is only one copy in $C_{r_p}|_{S(r)\cup\{r\}}$. We have also shown that for the columns in $C_{r_p}|_{S(r)\cup\{r\}}$ which are 1 in row r, that there is no row in $S(r)\setminus r_p$ where more than one of the columns have 1's.

Let $r_p \in S(r)$ be given. We say that row $r_s \in S(r)\setminus r_p$ is a *bad* row for r_p if *more than two* columns in $C_{r_p}|_{S(r)}$ have 1's in row r_s. By our case analysis above, we see that there is at most one *bad* row for a given row r_p which will arise from columns which are 0's in row r and have a column α of sum 1 repeated plus possibly with a column which is 1 in row r. For all rows $r_q \in S(r) \setminus r_p$, except possibly a single *bad* row, there is at most one column of the columns of C_{r_p} with a 0 in row r, which also has a 1 in row r_q and there is at most one column of the columns of C_{r_p} with a 1 in row r, which also has a 1 in row r_q. Thus for all rows in $S(r) \setminus r_p$ with the exception of the previously described *bad* row, the remaining rows will have at most two 1's on the columns of $C_{r_p}|_{S(r)\setminus r_p}$.

Now given a pair r_p, r_q, for which neither is a *bad* row for the other, we can now show that $|S(r_p) \cap S(r_q)| \leq 5$. Assume the contrary, that $|S(r_p) \cap S(r_q)| \geq 6$. We decompose A first using r_p and then using r_q. In each case we only show those column of A arising from C_{r_p} (respectively from C_{r_q}) which yield the identity matrix on the rows $S(r_p) \cap S(r_q)$. Let $t = |S(r_p) \cap S(r_q)| \geq 6$.

$$\begin{array}{r} \\ r_p \\ r_q \\ S(r_p) \cap S(r_q)\{ \end{array} \begin{array}{c} \begin{array}{cc} C_{r_p} & C_{r_p} \end{array} \\ \begin{bmatrix} 1 & 0 \\ \beta^T & \beta^T \\ I_t & I_t \end{bmatrix} \end{array} \qquad \begin{array}{r} \\ r_p \\ r_q \\ S(r_p) \cap S(r_q)\{ \end{array} \begin{array}{c} \begin{array}{cc} C_{r_q} & C_{r_q} \end{array} \\ \begin{bmatrix} \gamma^T & \gamma^T \\ 1 & 0 \\ I_t & I_t \end{bmatrix} \end{array}$$

Let $\beta = (b_1, b_2, \ldots, b_t)^T$ and $\gamma = (g_1, g_2, \ldots, g_t)^T$ and let $J = \{j : b_j = g_j = 0\}$. By assumption, β^T and γ^T each can contain at most two 1's and so $|J| \geq t - 4$. We now consider the columns indexed by J and the rows indexed by $\{1, 2\} \cup J'$ where $J' = \{j + 2 : j \in J\}$ for each of the following $(t + 2) \times t$ matrices:

$$\begin{array}{r} \\ r_p \\ r_q \\ S(r_p) \cap S(r_q)\{ \end{array} \begin{array}{c} C_{r_p} \\ \begin{bmatrix} 1 \\ \beta^T \\ I_t \end{bmatrix} \end{array} \qquad \begin{array}{r} \\ r_p \\ r_q \\ S(r_p) \cap S(r_q)\{ \end{array} \begin{array}{c} C_{r_q} \\ \begin{bmatrix} \gamma^T \\ 1 \\ I_t \end{bmatrix} \end{array}$$

This yields $I_2 \times I_{t-4}$ and hence $I_2 \times I_2$ in A, a contradiction. Therefore $|S(r_p) \cap S(r_q)| \leq 5$.

It remains to show that we can choose sufficient rows from $S(r)$ such that each row r_q is not a *bad* row for all other rows chosen. However, this is straightforward when interpreted as a graph theory problem. We create a graph with $|S(r)|$ vertices and assign each row of $S(r)$ to a vertex. We assign an arc $r_i \to r_j$ if r_j is a *bad* row for r_i. Because for each row r_i there can only be one *bad* row r_j, the out-degree of any vertex r_i is at most one. A set of rows that has our desired property is equivalent to choosing an independent set in our graph. By Lemma 9 we know that it is possible to choose such a set that is at least a third of the remaining rows. Thus we can choose at least $|S(r)|/3$ rows where for any chosen rows r_p, r_q, neither row is a *bad* row for the other and so $|S(r_p) \cap S(r_q)| < 6$. We then generate the series of disjoint sets described at the beginning of this proof and arrive at our contradiction. ■

Lemma 8. *We have that forb*$(m, \{F_4, F_5\}) \leq 2m$.

Proof: Let A be an m-rowed simple matrix with no configurations F_4, F_5. We apply the standard decomposition of (4) and note that C_r does not contain configurations $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ or I_2 else A has F_4 and F_5 respectively. But then $\|C_r\| \leq 2$ and can either have no columns or a column of 0's or a column of sum 1 or two columns consisting of a column of 0's and a column of sum 1. Then by induction, $\|A\| = \|[B_r C_r D_r]\| + \|C_r\| \leq \text{forb}(m-1, \{F_4, F_4\}) + 2$, which yields the bound. ■

Lemma 9. *Given a directed graph $D = (V, A)$ where for all $v \in V$ the maximum out-degree is 1, it is always possible to colour the vertices of D with 3 colours so that if $u \to v \in A$ then u and v have different colours.*

Proof: We use induction on the number of vertices. Suppose there exists a $v' \in V$ such that v' has in-degree 0. Let D' be the induced subgraph of $D \setminus v'$. Directed graph D' maintains our property of maximum out-degree 1, so we colour D' with 3 colors and then add v' back with an appropriate colour. Suppose there is no vertex $v \in V$ with in-degree 0. In this case all vertices have in-degree 1 and out-degree 1, and so the edges of D form the union of disjoint cycles, which we can colour with 3 colours. ■

It is likely (but unknown) that $\text{forb}(m, \{E_1, E_2, E_3\} \times \{E_1, E_2, E_3\})$ is $O(m^{3/2})$. One might ask the relationship of Theorem 4 to Conjecture 1. The Conjecture (which applies only for a single forbidden configuration) says that only product constructions are needed for best possible asymptotics, but in this case $\{I_2 \times I_2, T_2 \times T_2\}$ are simultaneously missing from all 1-fold products. In particular $I \times I$ avoids $T_2 \times T_2$ but does not avoid $I_2 \times I_2$ (Proposition 1). Surprisingly there is an $O(m^{3/2})$ construction contained in $I \times I$ and yet avoiding $I_2 \times I_2$ (Theorem 1). The other 2-fold products $I_2 \times T_2$ (Theorem 2) and $T_2 \times T_2$ (Theorem 3) behave as the conjecture might suggest.

7. A bound for 10 4 × 4 Forbidden Configurations

Proof of Theorem 5: Use the notation $\mathcal{F} = \{I_2, T_2, U_2, V_2\} \times \{I_2, T_2, U_2, V_2\}$. We establish the lower bound by construction. Let $\alpha = \mathbf{1}_{m-1}\mathbf{0}_1$. We construct the $m \times (m+3)$ matrix A consisting of $[\mathbf{0}_m \, I \, \alpha \, \mathbf{1}_m]$. Not that each of $\{I_2, T_2, U_2\}$ has one column of sum 1 and hence each matrix $\{I_2, T_2, U_2\} \times \{I_2, T_2, U_2\}$ has a column of sum 2 but no column of A thas at least 2 1's and 2 0's. The same observation holds for $T_2 \times V_2$, $U_2 \times V_2$, and $V_2 \times V_2$. Now $I_2 \times V_2$ has its first two rows containing $[I_2 \,|\, I_2]$ yet no two rows of A have $[I_2 \,|\, I_2]$ and so A has no $I_2 \times V_2$. Hence A avoids all configurations in $\mathcal{F}$, thus $\text{forb}(m, \mathcal{F}) \geq m + 3$.

We use induction on m for the upper bound. Begin by verifying $\text{forb}(4, \mathcal{F}) = 7$ using a C++ program which exhaustively considered all subsets of 8 columns of K_4 and verified that in all of them F is contained as a configuration. To prove the bound for $m \geq 5$, we will proceed by induction on m. For an m-rowed matrix A that doesn't contain any configuration in $\mathcal{F}$ it suffices by induction to show there exists a row r for which $\|C_r\| \leq 1$, using the standard decomposition as in (7). If this were so, we could delete row r and perhaps one column (one instance of the column forming C_r) from A, keeping the remaining matrix simple. This would yield $\text{forb}(m, \mathcal{F}) \leq 1 + \text{forb}(m-1, \mathcal{F}) = 1 + (m-1) + 3 = m + 3$ as desired.

Let us proceed by contradiction. Suppose then that for every row r, $|C_r(A)| \geq 2$. We then have at least two columns α and β in $C_1(A)$. The matrix A would look like this

$$1 \begin{bmatrix} 0 \cdots & 0 \; 0 \; 1 \; 1 & \cdots 1 \\ & \alpha \; \beta \; \alpha \; \beta & \end{bmatrix}.$$

But α and β must differ in some row. Without loss of generality, assume they differ on row 2, and suppose $\alpha_2 = 0$ and $\beta_2 = 1$. We will prove that α and β must be (0,1)-complements of each other. Suppose otherwise and suppose they had something in common, say in row 3. The first four rows of A would look like this (where we are not requiring $b \neq c$):

$$\begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{bmatrix} 0 \cdots & 0 \; 0 \; 1 \; 1 & \cdots 1 \\ & 0 \; 1 \; 0 \; 1 & \\ & a \; a \; a \; a & \\ & b \; c \; b \; c & \end{bmatrix}.$$

for some values of a, b, c (we are using the fact that the matrix has at least 4 rows). Then in rows 1 and 3 and rows 2 and 4 we get that this matrix contains

$$\begin{array}{c}\begin{matrix}1\\3\end{matrix}\begin{bmatrix}0&1\\a&a\end{bmatrix}\\ \times \\ \begin{matrix}2\\4\end{matrix}\begin{bmatrix}0&1\\b&c\end{bmatrix}\end{array}.$$

which is a configuration of $\mathcal{F}$ (for any a, b, c), so we conclude $\alpha = \overline{\beta}$.

Now $C_2(A)$ must have two repeated columns, say γ and δ. As argued above, they must be (0,1)-complements. Here is part of the matrix A:

$$\begin{bmatrix}0&0&0&0&1&1&1&1\\0&0&1&1&0&0&1&1\\\alpha&\gamma&\alpha&\overline{\gamma}&\overline{\alpha}&\gamma&\overline{\alpha}&\overline{\gamma}\end{bmatrix}.$$

Since α and $\overline{\gamma}$ have to differ somewhere, we can assume $\alpha_3 = a$, and $\gamma_3 = a$. Since α and γ must differ somewhere, we can assume $\alpha_4 = b$ and $\gamma_4 = \overline{b}$. Furthermore, since we have at least 5 rows, we can then write the selected columns of A where the columns are given labels below to indicate the source of the column.

$$\begin{array}{c}\begin{bmatrix}0&0&0&0&1&1&1&1\\0&0&1&1&0&0&1&1\\a&a&a&\overline{a}&\overline{a}&a&\overline{a}&\overline{a}\\b&\overline{b}&b&b&\overline{b}&\overline{b}&\overline{b}&b\\d&c&d&\overline{c}&\overline{d}&c&\overline{d}&\overline{c}\end{bmatrix}\\ \begin{matrix}\alpha&\gamma&\alpha&\overline{\gamma}&\overline{\alpha}&\gamma&\overline{\alpha}&\overline{\gamma}\end{matrix}\end{array}.$$

There are two cases. Either $d = c$ or $d = \overline{c}$. So we either have

$$(d=c):\ \begin{array}{c}\begin{bmatrix}0&0&0&0&1&1&1&1\\0&0&1&1&0&0&1&1\\a&a&a&\overline{a}&\overline{a}&a&\overline{a}&\overline{a}\\b&\overline{b}&b&b&\overline{b}&\overline{b}&\overline{b}&b\\d&d&d&\overline{d}&\overline{d}&d&\overline{d}&\overline{d}\end{bmatrix}\\ \begin{matrix}\alpha&\gamma&\alpha&\overline{\gamma}&\overline{\alpha}&\gamma&\overline{\alpha}&\overline{\gamma}\end{matrix}\end{array} \quad \text{or} \quad (d=\overline{c}):\ \begin{array}{c}\begin{bmatrix}0&0&0&0&1&1&1&1\\0&0&1&1&0&0&1&1\\a&a&a&\overline{a}&\overline{a}&a&\overline{a}&\overline{a}\\b&\overline{b}&b&b&\overline{b}&\overline{b}&\overline{b}&b\\d&\overline{d}&d&d&\overline{d}&\overline{d}&\overline{d}&d\end{bmatrix}\\ \begin{matrix}\alpha&\gamma&\alpha&\overline{\gamma}&\overline{\alpha}&\gamma&\overline{\alpha}&\overline{\gamma}\end{matrix}\end{array}.$$

These yield the following configurations in $\mathcal{F}$ respectively:

$$(d=c):\ \begin{array}{c}\begin{matrix}2\\4\end{matrix}\begin{bmatrix}0&1\\\overline{b}&b\end{bmatrix}\\ \times \\ \begin{matrix}3\\5\end{matrix}\begin{bmatrix}a&\overline{a}\\d&\overline{d}\end{bmatrix}\end{array}, \qquad (d=\overline{c}):\ \begin{array}{c}\begin{matrix}2\\3\end{matrix}\begin{bmatrix}0&1\\a&\overline{a}\end{bmatrix}\\ \times \\ \begin{matrix}4\\5\end{matrix}\begin{bmatrix}b&\overline{b}\\d&\overline{d}\end{bmatrix}\end{array}.$$

This is a contradiction to $\|C_r\| \geq 2$ and hence for $m \geq 5$, there must be some row r for which $|C_r(A)| \leq 1$ which then yields the bound by induction. ■

References

1. R.P. Anstee, A Survey of Forbidden Configurations results, http://www.math.ubc.ca/∼ anstee.
2. R.P. Anstee, A Survey of Forbidden Configurations results, http://www.math.ubc.ca/∼ anstee/FCConfThesisVersion.tar.gz
3. R.P. Anstee, Balin Fleming, Two refinements of the bound of Sauer, Perles and Shelah and Vapnik and Chervonenkis, *Discrete Math.*, to appear.
4. R.P. Anstee, J.R. Griggs, A. Sali, Small Forbidden Configurations, *Graphs and Combinatorics* **13**(1997), 97-118.
5. R.P. Anstee, A. Sali, Small Forbidden Configurations IV, *Combinatorica* **25**(2005), 503-518.
6. Anstee, R. P., B. Fleming, Z. Füredi, and A. Sali, Color critical hypergraphs and forbidden configurations, proceedings of EuroComb 2005, Berlin, Germany. *Discrete mathematics and Theoretical Computer Science*, 2005, 117-122.
7. J. Balogh, B. Bollobás, Unavoidable Traces of Set Systems, *Combinatorica*, **25** (2005), 633-643.
8. J. Balogh, B. Bollobás, R. Morris, Hereditary properties of partitions, ordered graphs and ordered hypergraphs, *Eur. J. Combin.*, 8 (2006), 1263-1281.
9. P. Erdős, On Extremal Problems of Graphs and Generalized Graphs, Israel J. Math., **2**(1964), 183-190.
10. P. Erdős and M. Simonovits, A limit theorem in graph theory. *Studia Sci. Math. Hungar* **1** (1966), 51–57.
11. P. Erdős, A.H. Stone, On the Structure of Linear Graphs, *Bull. A.M.S.*, **52**(1946), 1089-1091.
12. Z. Füredi, An upper bound on Zarankiewicz problem, *Combinatorics, Probability and Computing* **5**(1996), 29-33.
13. Z. Füredi, P. Hajnal, Davenport-Schinzel theory of matrices, *Discrete Math.* **103**(1992), 233-251.
14. M. Klazar, A. Marcus, Extensions of the linear bound in the Füredi-Hajnal conjecture, *Adv. in Appl. Math.*, **38** (2007), 258-266.
15. Kővari, V. Sós, P. Turán, On a problem of K. Zarankiewicz, *Colloq. Math* **3** (1954), 50-57.
16. A. Marcus, G. Tardos, Excluded permutation matrices and the Stanley Wilf Conjecture, *J. Combin. Th. Ser. A* **107** (2004), 153-160.
17. G. Tardos, On 0-1 matrices and small excluded submatrices, *J. Combin. Th. Ser. A* **111** (2005), 266-288.