Multivalued Matrices and Forbidden Configurations

R.P. Anstee*, Jeffrey Dawson †
Mathematics Department
The University of British Columbia
Vancouver, B.C. Canada V6T 1Z2

Linyuan Lu
University of South Carolina
Columbia SC, USA

Attila Sali
Rényi Institute
Budapest, Hungary

April 17, 2018

Abstract

An r-matrix is a matrix with symbols in $\{0, 1, \ldots, r-1\}$. A matrix is simple if it has no repeated columns. Let $\mathcal{F}$ be a finite set of r-matrices. Let forb$(m, r, \mathcal{F})$ denote the maximum number of columns possible in a simple r-matrix A that has no submatrix which is a row and column permutation of any $F \in \mathcal{F}$. Many investigations have involved $r = 2$. For general r, forb$(m, r, \mathcal{F})$ is polynomial in m if and only if for every pair $i, j \in \{0, 1, \ldots, r-1\}$ there is a matrix in $\mathcal{F}$ whose entries are only i or j. Let $\mathcal{T}_\ell(r)$ denote the following r-matrices. For a pair $i, j \in \{0, 1, \ldots, r-1\}$ we form four $\ell \times \ell$ matrices namely the matrix with i's on the diagonal and j's off the diagonal and the matrix with i's on and above the diagonal and j's below the diagonal and the two matrices with the roles of i, j reversed. Anstee and Lu determined that forb$(m, r, \mathcal{T}_\ell(r))$ is a constant. Let $\mathcal{F}$ be a finite set of 2-matrices. We ask if forb$(m, r, \mathcal{T}_\ell(3) \backslash \mathcal{T}_\ell(2) \cup \mathcal{F})$ is $\Theta($forb$(m, 2, \mathcal{F}))$ and settle this in the affirmative for some cases including most 2-columned F.

Keywords: extremal set theory, (0,1)-matrices, multivalued matrices, forbidden configurations, trace, Ramsey Theory.

*Research supported in part by NSERC
†Research supported in part by NSERC USRA

1 Introduction

We define a matrix to be *simple* if it has no repeated columns. A (0,1)-matrix that is simple is the matrix analogue of a set system (or simple hypergraph) thinking of the matrix as the element-set incidence matrix. We generalize to allow more entries in our matrices and define an r*-matrix* be a matrix whose entries are in $\{0, 1, \ldots, r-1\}$. We can think of this as an r-coloured matrix. For $r = 2$, r-matrices are (0,1)-matrices and for $r = 3$, r-matrices are (0,1,2)-matrices. We examine extremal problems and let $\|A\|$ denote the number of columns in A.

We will use the language of matrices in this paper rather than sets. For two matrices F and A, we write $F \prec A$, and say that A has F as a *configuration*, if there is a submatrix of A which is a row and column permutation of F. Row and column order matter to submatrices but not to configurations. Let $\mathcal{F}$ denote a finite set of matrices. Let

$$\text{Avoid}(m, r, \mathcal{F}) = \{A \,:\, A \text{ is } m\text{-rowed and simple } r\text{-matrix}, F \not\prec A \text{ for } F \in \mathcal{F}\}.$$

Our extremal function of interest is

$$\text{forb}(m, r, \mathcal{F}) = \max_A\{\|A\| \,:\, A \in \text{Avoid}(m, r, \mathcal{F})\}.$$

In the case $r = 2$, we are considering (0,1)-matrices and then we drop r from the notation to write $\text{Avoid}(m, 2, \mathcal{F}) = \text{Avoid}(m, \mathcal{F})$ and $\text{forb}(m, 2, \mathcal{F}) = \text{forb}(m, \mathcal{F})$. We define

$$\text{forbmax}(m, r, \mathcal{F}) = \max_{m' \le m} \text{forb}(m', r, \mathcal{F}).$$

It has been conjectured by Anstee and Raggi [9] that $\text{forbmax}(m, 2, \mathcal{F}) = \text{forb}(m, 2, \mathcal{F})$ for large m (which is a type of monotonicity). For many $\mathcal{F}$ this is readily proven.

The following dichotomy between polynomial and exponential bounds is striking. Denote an (i, j)-matrix as a matrix whose entries are i or j.

Theorem 1.1 *(Füredi and Sali [8]) Let $\mathcal{F}$ be a family of r-matrices. If for every pair $i, j \in \{0, 1, \ldots, r-1\}$, there is an (i, j)-matrix in $\mathcal{F}$ then for some k, forb$(m, r, \mathcal{F})$ is $O(m^k)$. If there is some pair $i, j \in \{0, 1, \ldots, r-1\}$ so that $\mathcal{F}$ has no (i, j)-matrix then forb$(m, r, \mathcal{F})$ is $\Omega(2^m)$.*

It would be of interest to have more examples of forbidden families of configurations where we can determine the asymptotics of $\text{forb}(m, r, \mathcal{F})$. There are known examples given in [8]. There is a generalization of a result of Balogh and Bollobás [6] for (0,1)-matrices to r-matrices. Define the generalized identity matrix $I_\ell(a, b)$ as the $\ell \times \ell$ r-matrix with a's on the diagonal and b's elsewhere. The standard identity matrix is $I_\ell(1, 0)$. Define the generalized triangular matrix $T_\ell(a, b)$ as the $\ell \times \ell$ r-matrix with a's below the diagonal and b's elsewhere. The standard upper triangular matrix is $T_\ell(0, 1)$. Let

$$\mathcal{T}_\ell(r) = \{I_\ell(a, b) \,:\, a, b \in \{0, 1, \cdots, r-1\}, a \neq b\}$$

$$\bigcup \{T_\ell(a,b) \,:\, a,b \in \{0,1,\cdots,r-1\}, a \neq b\}\,. \tag{1}$$

By Theorem 1.1, forb$(m,r,\mathcal{T}_\ell(r))$ is bounded by a polynomial but much more is true.

Theorem 1.2 *[3] Given r,ℓ, there is a constant $c(r,\ell)$ so that forb$(m,r,\mathcal{T}_\ell(r)) \le c(r,\ell)$.*

We will use the constant $c(r,\ell)$ repeatedly in this paper. This is a kind of Ramsey Theorem, a particular structured configuration appears in any r-matrix of a suitably large number of distinct columns. An important result is that $c(r,\ell)$ is $O(2^{c_r\ell^2})$ for some constant c_r. Not unexpectedly, Ramsey Theory shows up in the proof. Section 2 contains a number of proofs using Ramsey theory.

$\mathcal{T}_\ell(2)$ consists of (0,1)-matrices (i.e. 2-matrices). This paper considers forbidding the matrices $\mathcal{T}_\ell(r)\backslash\mathcal{T}_\ell(2)$. Note that any (0,1)-matrix $A \in$ Avoid$(m,r,\mathcal{T}_\ell(r)\backslash\mathcal{T}_\ell(2))$ and so forb$(m,r,\mathcal{T}_\ell(r)\backslash\mathcal{T}_\ell(2)) = \Omega(2^m)$. Forbidding $\mathcal{T}_\ell(r)\backslash\mathcal{T}_\ell(2)$ may be somewhat like asking the matrices to be (0,1)-matrices.

Theorem 1.3 *Let r,ℓ be given. Then forb$(m,r,\mathcal{T}_\ell(r)\backslash\mathcal{T}_\ell(2))$ is $\Theta(2^m)$.*

Proof: A construction in Avoid$(m,r,\mathcal{T}_\ell(r)\backslash\mathcal{T}_\ell(2))$ is to take all 2-columns on m rows.

Take any matrix $A \in$ Avoid$(m,r,\mathcal{T}_\ell(r)\backslash\mathcal{T}_\ell(2))$ and replace all entries $2,3,\ldots,r-1$ by 1's to obtain the 2-matrix A', not necessarily simple. The number of different columns in A' is at most 2^m.

Let α be a column of A'. Let B denote the submatrix of A consisting of all columns of A that map to α under the replacements. Let B' be the simple submatrix of B consisting of the rows of B where α has 1's. Then $\|B\| = \|B'\| \le c(r-1,\ell)$ else we have a configuration in $\mathcal{T}_\ell(r)\backslash\mathcal{T}_\ell(2)$ in B' (using Theorem 1.2) with symbols chosen from $\{1,2,\ldots,r-1\}$.

Combining these two observations yields the desired bound. ■

By the same argument we can show forb$(m,r,\mathcal{T}_\ell(r)\backslash\mathcal{T}_\ell(s))$ is $\Theta(s^m)$ but the focus is on $s=2$ in this paper. In this paper we will also take $r=3$. Note that Lemma 1.6 provides a justification for this restriction. Define the matrices $T_\ell(a,b,c)$ as the $\ell\times\ell$ matrix with a's below the diagonal, b's on the diagonal and c's above the diagonal. In our problems we can require $a \neq b$. These appear in the proof of Theorem 1.2 but, for $a \neq b \neq c$, are not matrices of just two entries which are referred to in Theorem 1.1. One general result in this direction is the following.

Theorem 1.4 *[4] Let $\mathcal{F}$ be a finite family of (0,1)-matrices. Then forb$(m,3,\mathcal{T}_\ell(3)\backslash\mathcal{T}_\ell(2)\cup T_\ell(0,2,1)\cup\mathcal{F})$ is $O($forbmax$(m,\mathcal{F}))$.*

Another version of Theorem 1.4 with restricted column sums (*column sum* will refer in this setting to the number of 1's) is given in Section 2 with the analogous proof. We are not pleased with the inclusion of $T_\ell(0,2,1)$ in Theorem 1.4 and think it can be avoided.

Problem 1.5 *Let F be a (0,1)-matrix. Is it true that forb$(m, 3, \mathcal{T}_\ell(3)\backslash\mathcal{T}_\ell(2) \cup F)$ is $\Theta($forbmax$(m, F))$?*

Obviously the configuration $T_\ell(0, 2, 1)$ will be problematical. We will let ℓ take on large but constant values. Some results given below support a yes answer. For example if we forbid nothing in the (0,1)-world then the maximum number of possible distinct (0,1)-columns is 2^m. One could say that "forbmax$(m, \emptyset) = 2^m$ ". Now using Theorem 1.3, we see that Problem 1.5 is true in this case.

Given $s = 2$, one can show it suffices to consider $r = 3$ in Problem 1.5. The argument is similar to Theorem 1.3 and uses Ramsey Theory. The proof is given in Section 2.

Lemma 1.6 *Let $r > 2$ and ℓ be given. Then there is a constant $bd(\ell)$ so that forb$(m, r, \mathcal{T}_\ell(r)\backslash\mathcal{T}_\ell(2) \cup F)$ is $O($forb$(m, 3, \mathcal{T}_{bd(\ell)}(3)\backslash\mathcal{T}_{bd(\ell)}(2) \cup F))$.*

Given the answer 'yes' to Problem 1.5, this yields a justification for restricting to $r = 3$. The argument could also be extended to $\mathcal{T}_\ell(r)\backslash\mathcal{T}_\ell(s)$ but the focus is on $s = 2$. Many configurations F can be handled by Theorem 1.7 and in particular configurations with more than two columns.

Theorem 1.7 *Let $F \prec T_{\ell/2}(0, 1)$. Then forb$(m, 3, \mathcal{T}_\ell(3)\backslash\mathcal{T}_\ell(2)\cup F)$ is $\Theta($forbmax$(m, F))$.*

Proof: Note that $T_{\ell/2}(0, 1) \prec T_\ell(0, 2, 1)$ by considering the submatrix of $T_\ell(0, 2, 1)$ consisting of the even indexed columns and the odd indexed rows. Thus if $F \not\prec A$, then $T_\ell(0, 2, 1) \not\prec A$. Apply Theorem 1.4. ■

One important corollary is the following.

Corollary 1.8 *forb$(m, 3, \mathcal{T}_\ell(3)\backslash\mathcal{T}_\ell(2) \cup [0\,1])$ is $\Theta(1)$.*

Proof: $[0\,1] \prec T_{\ell/2}(0, 1)$ for $\ell \geq 4$. ■

This paper provides a number more results in this direction mostly involving configurations of two columns. Define $F_{a,b,c,d}$ to be the $(a+b+c+d) \times 2$ configuration with a rows $[1\,1]$, b rows $[1\,0]$, c rows $[0\,1]$, and d rows $[0\,0]$. The asymptotics of forb$(m, F_{a,b,c,d})$ have been completely determined by Anstee and Keevash [1]. Note that we can assume $a \geq d$ since otherwise we can tale the (0,1)-complement $F^c_{a,b,c,d} = F_{d,c,b,a}$. Also we may assume $b \geq c$ since as configurations $F_{a,b,c,d} = F_{a,c,b,d}$. We note that forb$(m, F_{a,b,0,0})$ is $\Omega(m^{a+b-1})$ by taking all columns of column sum $a+b$ and a different construction shows forb$(m, F_{0,b,b,0})$ is $\Omega(m^b)$. The important upper bounds are for $a \geq 1$, forb$(m, F_{a,b,b,a})$ is $\Theta(m^{a+b-1})$ [1] and forb$(m, F_{0,b+1,b,0})$ is $\Theta(m^b)$ [1]. Note that $I_2 = F_{0,1,1,0}$. This is the first result not covered by Theorem 1.7.

Theorem 1.9 *forb$(m, 3, \mathcal{T}_\ell(3)\backslash\mathcal{T}_\ell(2) \cup I_2)$ is $\Theta($forbmax$(m, I_2))$.*

Theorem 1.10 *Let $a \geq 0$ and $b \geq 2$ be given. Then forb$(m, 3, \mathcal{T}_\ell(3)\backslash\mathcal{T}_\ell(2) \cup F_{a,b,b,a})$ is $\Theta($forbmax$(m, F_{a,b,b,a}))$.*

We give the proofs in Section 4. Note the subtlety that forbmax$(m, F_{0,b,b,0})$ is $\Theta(m^b)$ where as, for $a \geq 1$, forbmax$(m, F_{a,b,b,a})$ is $\Theta(m^{a+b-1})$. The proofs use results for two columned forbidden configurations from [1]. The other critical two columned result concerns $F = F_{0,b+1,b,0}$ for which we don't know the answer for Problem 1.5.

Define $t \cdot F = [F\, F\, \cdots F]$ to be the concatenation of t copies of F.

Theorem 1.11 *Let F be a given $k \times p$ (0,1)-matrix. Then forb$(m, 3, \mathcal{T}_\ell(3)\backslash\mathcal{T}_\ell(2) \cup t \cdot F)$ is $O(\max\{m^k, \text{forb}(m, 3, \mathcal{T}_\ell(3)\backslash\mathcal{T}_\ell(2) \cup F)\})$.*

Proof: Let $A \in \text{Avoid}(m, 3, \mathcal{T}_\ell(3)\backslash\mathcal{T}_\ell(2))$ with

$$\|A\| > (t-1)p\binom{m}{k} + \text{forb}(m, 3, \mathcal{T}_\ell(3)\backslash\mathcal{T}_\ell(2) \cup F) + 1.$$

Then $F \prec A$. Remove from A the p columns containing a copy of F and repeat. We will generate at least $(t-1)\binom{m}{k} + 1$ copies of F and hence at least t column disjoint copies of F in the same set of k rows and so $t \cdot F \prec A$. ■

To apply this, we need to know forb$(m, 3, \mathcal{T}_\ell(3)\backslash\mathcal{T}_\ell(2) \cup F)$. The following is established in Section 4.

Theorem 1.12 *Let*

$$H = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{2}$$

Then forb$(m, 3, \mathcal{T}_\ell(3)\backslash\mathcal{T}_\ell(2) \cup H)$ is $\Theta(m)$.

Corollary 1.13 *Given H in (2), we have forb$(m, 3, \mathcal{T}_\ell(3)\backslash\mathcal{T}_\ell(2) \cup t \cdot H)$ is $\Theta(m^2)$.*

Proof: We apply Theorem 1.11 and Theorem 1.12. ■

Theorem 1.12 and Corollary 1.13 are yes instances of Problem 1.5 since forb(m, H) is $\Theta(m)$ and forb$(m, t \cdot H)$ is $\Theta(m^2)$ [5].

Given an $m_1 \times n_1$ matrix A and a $m_2 \times n_2$ matrix B, defne the product of two matrices $A \times B$ as the $(m_1 + m_2) \times n_1 n_2$ matrix obtained from placing each column of A on top of each column of B for all possible pairs of columns. Let F be given with

$$0 \times 1 \times F = \begin{bmatrix} 0\,0 \cdots 0 \\ 1\,1 \cdots 1 \\ F \end{bmatrix}$$

In [5], we establish that forb$(m, 0 \times 1 \times F)$ is $O(m \cdot \text{forb}(m, F))$. We establish this version of the Problem 1.5 in Section 3.

Theorem 1.14 *forb$(m, 3, \mathcal{T}_\ell(3)\backslash\mathcal{T}_\ell(2) \cup 0 \times 1 \times F)$ is $O(m \cdot \text{forb}(m, \mathcal{T}_\ell(3)\backslash\mathcal{T}_\ell(2) \cup F))$*

This result extends results for $F_{a,b,b,a}$ to $F_{a+1,b,b,a+1}$ and can be used in other instances such as H above. We finish the paper with some open problems.

2 Results using Ramsey Theory

We apply Ramsey Theory to help us find configurations in $\mathcal{T}_\ell(3)\backslash\mathcal{T}_\ell(2)$ etc. We use the p colour Ramsey number $R_p(t_1, t_2, \ldots, t_p)$ as the smallest number n such that for every edge colouring of K_n with p colours there is some colour i so that there is a clique of size t_i with all edges of colour i. Typical notation is that for $t_1 = t_2 = \cdots = t_p = t$, we write $R_p(t_1, t_2, \ldots, t_r) = R_p(t^p)$. While these numbers can be large, we can for example bound $R_p(t^p) \leq 2^{pt}$.

Let r, s be given integers with $r > s \geq 2$. Let us define a set $\mathcal{P}_t^x(r)$ of $t \times t$ matrices by the following template which will have choices $x, y_1, y_2, \ldots, y_t \in \{1, 2, \ldots, r-1\}$ where we require $y_j \neq x$ for $j \in [t]$. The entries marked * may be given entries in $\{0, 1, \ldots, r-1\}$ in any possible way.

$$\mathcal{P}_t^x : \quad \begin{bmatrix} y_1 & & & & & \\ x & y_2 & & & * & \\ x & x & y_3 & & & \\ \vdots & & & \ddots & & \\ x & x & x & & y_{t-1} & \\ x & x & x & \cdots & x & y_t \end{bmatrix} \tag{3}$$

Lemma 2.1 *Let ℓ, r, s be given with $r > s \geq 2$. Let $t = (r-1)(R_r((2\ell)^r) - 1) - 1$. Assume A is an m-rowed simple r-matrix. Assume there is some $G \in \mathcal{P}_t^x$ with $G \prec A$ and such that if $x \in \{0, 1, \ldots, s-1\}$ then $y_j \in \{s, s+1, \ldots, r-1\}$ for all $j \in [t]$. Then there is some F with $F \prec G$ and*

$$F \in (\mathcal{T}_\ell(r)\backslash\mathcal{T}_\ell(s)) \bigcup \{T_\ell(x, z, u)\,;\, x, u \in \{0, 1, \ldots, s-1\}, x \neq u, z \notin \{0, 1, \ldots, s-1\}\}.$$

.

Proof: Assume there is some $G \in \mathcal{P}_t^x$ with $G \prec A$ and such that if $x \in \{0, 1, \ldots, s-1\}$ then $y_i \in \{s, s+1, \ldots, r-1\}$ for all $i \in [t]$.

First assume $x \notin \{0, 1, \ldots, s-1\}$. There are $r-1$ choices for each y_j and hence there is some choice $z \in \{0, 1, \ldots, s-1\}\backslash x$ which appears at least $R_r((2\ell)^r)$ times on the diagonal. Now form a graph whose vertices are the rows i with $y_j = z$ and we colour edge a, b for $a < b$ by the entry in the a, b location of G (above the diagonal). There will be at least $R_r((2\ell)^r)$ vertices and there will be at most r colours and so by the Ramsey number there will be a clique of size 2ℓ of all edges of the same colour, say colour u. If $u = z$ we have $T_{2\ell}(x, z) \prec A$. If $u = x$ we have $I_{2\ell}(x, z) \prec A$. If $u \neq x, z$ then we consider the configuration $T_{2\ell}(x, u, z)$ of size 2ℓ induced by the clique and the even columns and the odd rows to show $T_\ell(x, y) \prec A$. All three cases yield a configuration in $\mathcal{T}_\ell(r)\backslash\mathcal{T}_\ell(s)$.

Now assume $x \in \{0, 1, \ldots, s-1\}$ then there are $r-s-1$ choices for each y_j and hence there is some choice $z \notin \{0, 1, \ldots, s-1\}$ which appears at least $R_r((2\ell)^r)$ times on the diagonal. Now we proceed as above to obtain a configuration $T_{2\ell}(x,z,u)$. If $u \in \{s, s+1, \ldots, r-1\}$ then we obtain a configuration in $\mathcal{T}_\ell(r)\backslash\mathcal{T}_\ell(s)$. If $u = x$ we obtain a configuration $I_{2\ell}(x,z)$ which is in $\mathcal{T}_\ell(r)\backslash\mathcal{T}_\ell(s)$. If $u \in \{0,1,\ldots,s-1\}$ with $u \neq x$, then we obtain a configuraton $T_{2\ell}(x,z,u)$ with $x, u \in \{0,1,\ldots,s-1\}$ and $x \neq u$ (which does not yield a configuration in $\mathcal{T}_\ell(r)\backslash\mathcal{T}_\ell(s)$). ■

Our application of the Lemma 2.1 to Theorem 1.4 will be in the case $r = 3$ and $s = 2$ and then $\{T_\ell(x,z,u)\,;\, x,u \in \{0,1,\ldots,s-1\}, x \neq u, z \notin \{0,1,\ldots,s-1\}$ is the single configuration $T_\ell(0,2,1)$. We prove in greater generality.

Proof of Theorem 1.4: The idea of the proof is to use the induction to generate configurations corresponding to matrices in $\mathcal{P}_t^x$ that enable us to apply the proof of Lemma 2.1 and obtain matrices in $\mathcal{T}_\ell(r)\backslash\mathcal{T}_\ell(s)$.

We use the following function f in our proof. Let f be determined by the recurrence

$$f(p_0, p_1, \ldots, p_{r-1}) = \sum_{i=0}^{r-1} f(p_0, p_1, \ldots, p_i - 1, \ldots, p_{r-1}), \tag{4}$$

and the base cases that $f(p_0, p_1, \ldots, p_{r-1}) = 1$ if $p_i = 1$ for any $i \in \{0,1,\ldots,r-1\}$. Solving this exactly seems difficult but since f satisfies the same recurrence as the multinomial coefficients, with smaller base cases, we obtain

$$f(p_0, p_1, \ldots, p_{r-1}) \leq \frac{(p_0 + p_1 + \cdot + p_{r-1} - r)!}{(p_0 - 1)!(p_1 - 1)! \cdots (p_{r-1} - 1)!} \tag{5}$$

Let $g(p_0, p_1, \ldots, p_{r-1}) = f(p_0, p_1, \ldots, p_{r-1}) \cdot \text{forbmax}(m, \mathcal{F})$.

We will establish for fixed m but by induction on $\sum_i p_i$, that if A is an n-rowed simple r-matrix with $n \leq m$ and $\|A\| > g(p_0, p_1, \ldots, p_{r-1})$ then for some $i \in \{0,1,\ldots,r-1\}$, A will contain configuration $F \in \mathcal{F}$ or a configuration in $\mathcal{P}_{p_i}^i$ satisfying the condition that if $i \in \{0,1,\ldots,s-1\}$, then $y_j \in \{s, s+1, \ldots, r-1\}$ for $j \in [P_i]$. We use forbmax so that $\text{forbmax}(m, s, \mathcal{F}) \geq \text{forb}(n, s, \mathcal{F})$.

If $p_i = 1$, then an element of $\mathcal{P}_{p_i}^i$ is a 1×1 matrix. For $i \in \{0,1,\ldots,s-1\}$, then the entry in the 1×1 matrix must not be in $\{0,1,\ldots,s-1\}$ and if $i \notin \{0,1,\ldots,s-1\}$, then the entry in the 1×1 matrix must not be i. In the former case, we require the matrix to have some entry not in $\{0,1,\ldots,s-1\}$ which would only be difficult if A was an s-matrix. In that case $\|A\| \leq \text{forb}(n, s, \mathcal{F}) \leq \text{forbmax}(m, s, \mathcal{F})$ and we note that $f(p_0, p_1, \ldots, p_{r-1}) = 1$ for $p_i = 1$. In the latter case we are merely requiring that the matrix A has at least two different entries which would only not occur for $\|A\| = 1$. In either case we are able to obtain an instance of $\mathcal{P}_1^i$ in A if $\|A\| > g(p_0, p_1, \ldots, p_{r-1})$. This establishes the required base cases for the induction.

Assume $p_i \geq 2$ or all $i \in \{0,1,\ldots,r-1\}$. Consider a matrix $A \in \text{Avoid}(n, r, \mathcal{P}_{p_0}^0 \cup \mathcal{P}_{p_1}^1 \cup \cdots \cup \mathcal{P}_{p_{r-1}}^{r-1} \cup \mathcal{F})$ with $n \leq m$ and $\|A\| > g(p_0, p_1, \ldots, p_{r-1})$. We wish to obtain a contradiction.

Choose a row w of A which has at least two different entries one of which is not in $\{0,1,\ldots,s-1\}$. If there is no such row then either $\|A\|=1$ or A is an s-matrix. In the latter case, we have $\|A\|>g(p_1,p_2,\ldots p_{r-1})\geq \text{forbmax}(m,s,\mathcal{F})\geq \text{forb}(n,s,\mathcal{F})$ and so $F\prec A$, a contradiction. We may assume a row w of A, which has at least two different entries one of which is not in $\{0,1,\ldots,s-1\}$, exists.

Decompose A as follows by permuting rows and columns

$$A=\begin{array}{c}\text{row } w\to\\ \\ \end{array}\left[\begin{array}{c|c|c|c|c} 0\,0\cdots 0 & 1\,1\cdots 1 & 2\,2\cdots 2 & \cdots & r-1\,r-1\cdots r-1\\ G_0 & G_1 & G_2 & \cdots & G_{r-1}\end{array}\right] \tag{6}$$

Each G_i is simple. Now

$$\|A\|=\sum_{i=0}^{r-1}\|G_i\|>g(p_0,p_1,\ldots,p_{r-1})=f(p_0,p_1,\ldots,p_{r-1})\cdot\text{forbmax}(m,s,\mathcal{F})$$

$$=\left(\sum_{i=0}^{r-1}f(p_0,p_1,\ldots,p_i-1,\ldots,p_{r-1})\right)\cdot\text{forbmax}(m,s,\mathcal{F}).$$

From the recurrence (4), there is some i with

$$\|G_i\|>g(p_0,p_1,\ldots,p_i-1,\ldots,p_{r-2},p_{r-1}).$$

Certainly $G_i\prec A$ and $G\in\text{Avoid}(n-1,3,\mathcal{P}^0_{p_0}\cup\mathcal{P}^1_{p_1}\cup\cdots\cup\mathcal{P}^{r-1}_{p_{r-1}}\cup\mathcal{F})$. Then by induction on $\sum_i p_i$, we can assume G_i has a copy of $\mathcal{P}^i_{p_i-1}$ using the template (3) with $x=i$ and if $x=i\in\{0,1,\ldots,s-1\}$ then $y_j\in\{s,s+1,\ldots,r-1\}$ for all $j=1,2,\ldots,p_i-1$. We can extend to a copy of $\mathcal{P}^i_{p_i}$ in A by adding row w to extend by a row of i's and then extend by a column from some G_j with $j\neq i$. If $i\in\{0,1,\ldots,s-1\}$, then we can extend to a copy of $\mathcal{P}^i_{p_i}$ in A by adding row w to extend by a row of i's and then extend by a column from some G_h with $h\in\{s,s+1,\ldots,r-1\}$. This is possible since we have assumed that row w has at least two different entries one of which is not in $\{0,1,\ldots,s-1\}$. Now some matrix G in the family $\mathcal{P}^i_{p_i}$ has $G\prec A$.

Specializing to $p_0=p_1=\cdots=p_{r-1}=(r-1)(R_r((2\ell)^r)-1$ and applying Lemma 2.1 yields that G contains a configuration in $(\mathcal{T}_\ell(r)\backslash\mathcal{T}_\ell(s))\bigcup\{T_\ell(x,z,u)\,;\,x,u\in\{0,1,\ldots,s-1\},x\neq u,z\notin\{0,1,\ldots,s-1\}\}$ and then specializing to $r=3$ and $s=2$ yields the result. ■

It was convenient to consider general r,s but we will focus on $r=3$ and $s=2$. The proof of Theorem 1.4 can be adapted to considering fixed column sum i.e. columns with a fixed number of 1's. In the case of 3-matrices, we define the *column sum* of a 3-column α to be the number of 1's present. When there are no 2's in α, this is the usual column sum. Define

$$\text{forb}_k(m,3,\mathcal{F})=\max\{\|A\|\,:\,A\in\text{Avoid}(m,3,\mathcal{F}),\ \text{all columns in } A \text{ have } k\ 1's\},$$

and define forbmax_k similarly. There are $\mathcal{F}$ for which we can exploit information about $\text{forb}_k(m,3,\mathcal{F})$, deducing some information from $\text{forb}_k(m,\mathcal{F})$.

Theorem 2.2 *Let $\mathcal{F}$ be a finite set of $(0,1)$-matrices. Let ℓ be given. Then there exists a constant d so that*

$$\text{forb}_k(m,3,\mathcal{T}_\ell(3)\backslash\mathcal{T}_\ell(2)\cup T_\ell(0,2,1)\cup\mathcal{F}) \text{ is } O\left(\sum_{j=k-d}^{k}\text{forbmax}_j(m,\mathcal{F})\right) \tag{7}$$

Proof: We will follow the proof of Theorem 1.4 but note how columns sums are affected. Let $g_k(p_0,p_1,p_2)=f(p_0,p_1,p_2)\cdot\text{forbmax}_k(m,F))$.

Consider a matrix $A\in\text{Avoid}_k(n,3,\mathcal{P}^0_{p_0}\cup\mathcal{P}^1_{p_1}\cup\mathcal{P}^2_{p_2}\cup\mathcal{F})$ with $n\le m$ and $k>p+2R_3((2\ell)^3)$ and $\|A\|>g_k(p_0,p_1,p_2)$. We wish to obtain a contradiction.

It is convenient to interpret the proof of Theorem 1.4 as growing a tree where each node is associated with a matrix with three associated parameters (p_0,p_1,p_2) and has some fixed column sum s. We begin with a root node corresponding to a matrix A with parameters (p,p,p) where $p=2R_3(2\ell,2\ell,2\ell)$. Then the matrices G_0,G_1,G_2 can be viewed as the children. Our recursive growth of the tree begins with a node corresponding matrix B for which we decompose by some row w with at least two entries one of which is 2. If we can't decompose then either $\|B\|=1$ or B is an (0,1)-matrix.

Assume each column of B has s 1's. Decompose B as follows by permuting rows and columns

$$B=\begin{array}{c}\text{row } w\rightarrow\\ \\\end{array}\left[\begin{array}{c|c|c}0\,0\cdots0 & 1\,1\cdots1 & 2\,2\cdots2\\ H_0 & H_1 & H_2\end{array}\right] \tag{8}$$

Each H_i is simple. Given that each column in B has s 1's then for each column in H_0 and H_2 has s 1's and each column in H_1 has $s-1$ 1's. Thus the nodes of our tree correspond to matrices with fixed column sum.

We also need to keep track of the current triple (q_0,q_1,q_2) for each node. Thus if B has the triple (q_0,q_1,q_2) then G_0 has triple (q_0-1,q_1,q_2), G_1 has triple (q_0,q_1-1,q_2) and G_2 has triple (q_0,q_1,q_2-1). We do not decompose B if $q_0=1$ or $q_1=1$ of $q_2=1$. Otherwise the node corresponding to B has children G_0,G_1,G_2 with the possibility that $\|G_0\|=0$ or $\|G_1\|=0$ in which case B would only have two children.

Given the decomposition (8), then $\|A\|$ is the sum of $\|B\|$ over all leaves B of the tree. The leaves of the tree which cannot be further decomposed correspond to matrices B with $\|B\|=1$ or B is a (0,1)-matrix or B where the three parameters (q_0,q_1,q_2) have either $q_0=1$ or $q_1=1$ of $q_2=1$.

We deduce that the depth of the tree is at most $d=3p=6R_3(2\ell,2\ell,2\ell)$ with a branching factor of 3 and so there are at most 3^d nodes in the tree which is a constant. Also we have that each node corresponds to a matrix with constant column sum $s\in\{k-d,k-d+1,\ldots,k\}$ which is a constant cardinality set.

Now continue growing the tree until no further growth is possible. If the process generates a node B with $q_0=1$ or $q_1=1$ of $q_2=1$, then by the arguments of Theorem 1.4, there will be some configuration in $\mathcal{T}_\ell(3)\backslash\mathcal{T}_\ell(2)\cup T_\ell(0,2,1)$ in B and hence in A. A leaf node is one which corresponds to some (0,1)-matrix B with constant column sum $s\in\{k-d,k-d+1,\ldots,k\}$ for which we deduce that $\|B\|\le\text{forbmax}_s(m,F)$.

The bound (7) now follows with the inclusion of some large constants. ∎

We will apply this result to 2-columned F.

Proof of Lemma 1.6: We readily note that $\text{forb}(m,r,\mathcal{T}_\ell(r)\backslash\mathcal{T}_\ell(2)\cup F)\geq \text{forb}(m,3,\mathcal{T}_\ell(r)\backslash\mathcal{T}_\ell(2)\cup F)$ since $\text{Avoid}(m,3,\mathcal{T}_\ell(3)\backslash\mathcal{T}_\ell(2)\cup F)\subseteq \text{Avoid}(m,r,\mathcal{T}_\ell(r)\backslash\mathcal{T}_\ell(2)\cup F)$.

Let $bd(\ell)=R_{(r-2)(r-2)}((2\ell)^{(r-2)(r-2)})$ where we assume $bd(\ell)>(r-2)\ell$. Let $A\in\text{Avoid}(m,r,\mathcal{T}_\ell(r)\backslash\mathcal{T}_\ell(2)\cup F)$. Replace all entries $3,4,\ldots,r-1$ by 2's to obtain A'. The number of different columns in A' is at most $\text{forb}(m,3,\mathcal{T}_{bd(\ell)}(3)\backslash\mathcal{T}_{bd(\ell)}(2)\cup F)$ for the following reason. If $F\prec A'$, then $F\prec A$ so we may assume $F\not\prec A'$. Let A'' be the matrix obtained from A' by keeping exactly one copy of each column. If $\|A''\|>\text{forb}(m,3,\mathcal{T}_{bd(\ell)}(3)\backslash\mathcal{T}_{bd(\ell)}(2)\cup F)$ then there is configuration $G\prec A''$ with $G\in\mathcal{T}_{bd(\ell)}(3)\backslash\mathcal{T}_{bd(\ell)}(2)$. There are several cases.

If G is a generalized identity matrix say $I_{bd(\ell)}(1,2)$, then in A, we have a configuration which has entries in $\{2,3,\ldots,r-1\}$ on the diagonal and 1's off the diagonal. Then there is some entry $q\in\{2,3,\ldots,r-1\}$ appearing $\lceil bd(\ell)/(r-2)\rceil\geq\ell$ times (using $bd(\ell)>(r-2)\ell$) and we obtain a principal submatrix of G (row and column indices given by the diagonal entries q) in $\mathcal{T}_{\lceil bd(\ell)/(r-2)\rceil}(r)\backslash\mathcal{T}_{\lceil bd(\ell)/(r-2)\rceil}(2)$ in A.

If G is a generalized identity matrix say $I_{bd(\ell)}(2,1)$, then in A, we have a configuration which has entries in $\{2,3,\ldots,r-1\}$ off the diagonal and 1's on the diagonal. Now apply Ramsey Theory by colouring a graph on $bd(\ell)$ vertices with the colour of edge (i,j) for $i<j$ being the 2-tuple $a_{i,j},a_{j,i}$. There are $(r-2)^2$ colours and so if $bd(\ell)>R_{(r-2)(r-2)}((2\ell)^{(r-2)(r-2)})$, then there is a clique of colour p,q of size 2ℓ and so $2\ell\times 2\ell$ configuration whose entries on the diagonal are 1's and above the diagonal are p and whose entries below the diagonal are q. If $p=q$, we have a configuration in $\mathcal{T}_{2\ell}(r)\backslash\mathcal{T}_{2\ell}(2)$. If $p\neq q$, then we form an $\ell\times\ell$ configuration with p's above the diagonal and q's below the diagonal (by taking even indexed columns and odd indexed rows) which is a configuration in $\mathcal{T}_\ell(r)\backslash\mathcal{T}_\ell(2)$. Similar arguments handle the remaining cases.

To determine the maximum number of columns of A that map into a given $(0,1,2)$-column α in A', let α have t 2's and then the columns mapping into α correspond to a t-rowed simple matrix with entries in $\{2,3,\ldots,r-1\}$. If the number of columns is bigger that $c(r-2,\ell)$, then those columns contain a configuration in $\mathcal{T}_\ell(r)$ whose entries are in $\{2,3,\ldots,r-1\}$ and so the configuration is in $\mathcal{T}_\ell(r)\backslash\mathcal{T}_\ell(2)$. We now deduce that $\|A\|\leq c(r-2,\ell)\times forb(m,3,\mathcal{T}_{bd(\ell)}(3)\backslash\mathcal{T}_{bd(\ell)}(2)\cup F)$ yielding our bound. ∎

3 $0\times 1\times F$

Let $A\in\text{Avoid}(m,3,\mathcal{T}_\ell(3)\backslash\mathcal{T}_\ell(2)\cup 0\times 1\times F)$. If we can choose a pair of rows i,j so that there are $\text{forb}(m-2,3,\mathcal{T}_\ell(3)\backslash\mathcal{T}_\ell(2)\cup F)+1$ columns of A which have 0's on row i and 1's in row j, then we have $F\prec A$, a contradiction.

Lemma 3.1 *Let $\epsilon>0$ be given. Let A be an m-rowed simple 3-matrix with each column* *having both a 0 and a 1 and at least ϵm entries either 0 or 1. Assume*

$$\|A\| > 2 \cdot forbmax(m-2, 3, \mathcal{T}_\ell(3) \backslash \mathcal{T}_\ell(2) \cup F) \frac{\binom{m}{2}}{\epsilon m - 1}. \tag{9}$$

Then $0 \times 1 \times F \prec A$.

Proof: We note that a column of m rows that has p 0's and q 1's will have pq pairs of rows i, j containing the configuration $\left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]$. For a given p, q, the minimum number of configurations $\left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]$ is $p + q - 1$ when for example there is one 1 and $p + q - 1$ 0's. An m-rowed column with at least one 0 and at least one 1 and at least ϵm entries that are 0 or 1 will have at least $\epsilon m - 1$ configurations $\left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]$. There are $2\binom{m}{2}$ choices for i, j when considered as an ordered pair.

If (9) is valid then then there will be a pair of rows i, j with more than $2 \cdot \text{forbmax}(m, 3, \mathcal{T}_\ell(3) \backslash \mathcal{T}_\ell(2) \cup F)$ columns with the configuration $\left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]$. Thus there will be a pair of rows i, j with at least $\text{forb}(m, \mathcal{T}_\ell(3) \backslash \mathcal{T}_\ell(2) \cup F) + 1$ columns all with the submatrix $\left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]$ (or all the reverse). Then we can form an $(m-2) \times (\text{forbmax}(m, 3, \mathcal{T}_\ell(3) \backslash \mathcal{T}_\ell(2) \cup F) + 1)$ simple matrix A' that when extends by a row of 0's and a row of 1's is contained in A. SInce $A' \in \text{Avoid}(m, \mathcal{T}_\ell(3) \backslash \mathcal{T}_\ell(2))$, we deduce that $F \prec A'$ and then $0 \times 1 \times F \prec A$, as desired. ∎

Proof of Theorem 1.14: If we have many columns with few 0's and 1's then we will show we are able to find in A a $c \times c$ configuration G in $\mathcal{P}_c^2$ of A as in (11) and then can use Lemma 2.1.

Let $A \in \text{Avoid}(m, 3, \mathcal{T}_\ell(3) \backslash \mathcal{T}_\ell(2) \cup F)$. There are at most $c(2, \ell)$ (0,2)-columns and at most $c(2, \ell)$ (1,2)-columns. Let A' be the matrix obtained from A by deleting (0,2)-columns and (1,2)-columns.

Now each column in A' has at least one 0 and one 1. Let

$$\epsilon = \frac{1}{4R(2\ell, 2\ell, 2\ell)}. \tag{10}$$

Delete from A' any rows entirely of 2's to obtain a simple matrix $A'' \in \text{Avoid}(t, 3, \mathcal{T}_\ell(3) \backslash \mathcal{T}_\ell(2) \cup F)$ where $t \leq m$. Let A_2 denote those columns of A'' with at most ϵt 0's and 1's and let A_{01} denote those columns of A'' with more than ϵt 0's and 1's .

We select columns of A_2 in turn to form the pattern $\mathcal{P}_c^2$ in (3) with $c = 2 \cdot R((2\ell)^3)$. We can begin with a column on $(1-\epsilon)t$ 2's. At the kth stage we have k columns (selected in the order displayed) with

$$\begin{array}{l}\overbrace{\begin{array}{ccccc} \neq 2 & & & & \\ 2 & \neq 2 & & & \\ 2 & 2 & \neq 2 & & \\ 2 & 2 & \cdots & \neq 2 & \\ 2 & 2 & \cdots & 2 & \neq 2 \end{array}}^{k} \\ \left.\begin{array}{ccccc} 2 & 2 & \cdots & 2 & 2 \\ 2 & 2 & \cdots & 2 & 2 \\ 2 & 2 & \cdots & 2 & 2 \end{array}\right\} \geq (1-k\epsilon)t \end{array} \tag{11}$$

where the final block of 2's in rows S has $|S| \geq (1-k\epsilon)t$. Any column of A_2 not already chosen has 2's in at least $(1-(k+1)\epsilon)t$ rows of S. To proceed we need that $\|A_2\| \geq c = 2R_3(2\ell, 2\ell, 2\ell)$ and we require that $(1-k\epsilon)t \geq 1$ for $k+1 \leq c = 2R(2\ell, 2\ell, 2\ell)$. Our choice of ϵ (10) ensures this. A_2 has no rows of 2's and so a column with a 0 or 1 in rows S can be used to extend (11) to the situation with $k+1$ columns. We repeat until we have $c = 2R_3(2\ell, 2\ell, 2\ell)$ columns. Applying Lemma 2.1, we obtain a matrix $F \in \mathcal{T}_\ell(3)$ with $F \prec A$ that has 2's below the diagonal and so we have obtained a configuration in $\mathcal{T}_\ell(3)\backslash\mathcal{T}_\ell(2)$, a contradiction. Thus

$$\|A_2\| \leq 2 \cdot R_3(2\ell, 2\ell, 2\ell). \tag{12}$$

If

$$\|A_{01}\| > 2 \cdot \text{forbmax}(m-2, 3, \mathcal{T}_\ell(3)\backslash\mathcal{T}_\ell(2) \cup F)\frac{\binom{t}{2}}{\epsilon t - 1}$$

then by Lemma 3.1, $0 \times 1 \times F \prec A_{01}$. Thus

$$\|A_{01}\| \leq 2 \cdot \text{forb}(m-2, \mathcal{T}_\ell(3)\backslash\mathcal{T}_\ell(2) \cup F)\frac{\binom{t}{2}}{\epsilon t - 1} \leq m \cdot \text{forbmax}(m-2, \mathcal{T}_\ell(3)\backslash\mathcal{T}_\ell(2) \cup F). \tag{13}$$

Using $\|A\| = 2 \cdot c(2, \ell) + \|A_2\| + \|A_{01}\|$, we obtain our desired bound. ■

4 Two-columned matrices

The main result of this section is the following. The proof is given after Lemma 4.5 and Lemma 4.6.

Theorem 4.1 *forb*$(m, 3, \mathcal{T}_\ell(3)\backslash\mathcal{T}_\ell(2) \cup F)$ *is* $O\left(\sum_{k=0}^{m} \text{forbmax}_k(m, F)\right)$ *for all two-columned matrices* F.

We have some useful results for two-columned F. Theorem 1.2 in [1], gives us insight into $F_{0,b,b,0}$ with a strong stability result. For our purposes we only need the following.

Lemma 4.2 *[1] Let k, b be given with $b \geq 1$. Then $forb_k(m, F_{0,b,b,0})$ is $O(m^{b-1})$.*

Lemma 5.4 in [1], repeated below, gives us insight into $F_{0,b,b,1}$ which helps us consider $F_{1,b,b,1}$ for $b \geq 2$..

Lemma 4.3 *Suppose $r \geq 1$ and $\mathcal{F}$ is a k-uniform family of subsets of $[m]$, with $k \geq r+2$, so that every pair $A, B \in \mathcal{F}$ is either disjoint or intersects in at least $k-r$ points, and for every $A \in \mathcal{F}$ we have $1 \notin A$. Then $|\mathcal{F}|$ is $O(m^r)$.*

Translated in our language it says $\mathrm{forb}_k(m, F_{0,r+1,r+1,1})$ is $\Theta(m^r)$ for $k \neq r+1$. Note that $\mathrm{forb}_{r+1}(m, F_{0,r+1,r+1,1})$ is $\Theta(m^{r+1})$ by taking all columns of $r+1$ 1's. By taking (0,1)-complements where $F^c_{0,r+1,r+1,1} = F_{1,r+1,r+1,0}$. Thus for $k \neq m-r-1$, $\mathrm{forb}_k(m, F_{1,r+1,r+1,0})$ is $\Theta(m^r)$ and $\mathrm{forb}_{r+1}(m, F_{1,r+1,r+1,0})$ is $\Theta(m^{r+1})$ by taking all columns of $r+1$ 1's.

Corollary 4.4 *Let $r \geq 1$. Let m be given. For $k \neq r+1, r+2, m-r-2, m-r-1$, we have $forb_k(m, F_{1,r+1,r+1,1})$ is $\Theta(m^r)$. For $k = r+1, r+2, m-r-2$ or $m-r-1$ we have $forb_k(m, F_{1,r+1,r+1,1})$ is $\Theta(m^{r+1})$.*

Proof: Assume $k \neq r+1, r+2, m-r-2, m-r-1$ and $\mathrm{forb}_k(m, F_{1,r+1,r+1,0}) \leq cm^r$. Let $A \in \mathrm{Avoid}(m, F_{1,r+1,r+1,1})$. Consider row 1. The number of 0's plus the number of 1's in row 1 is $\|A\|$. Let B be the submatrix of A formed by the columns with a 1 in row 1 and rows $2, 3, \ldots m$. Then $B \in \mathrm{Avoid}_{k-1}(m-1, F_{0,r+1,r+1,1})$ and so $\|B\| \leq \mathrm{forb}_{k-1}(m-1, F_{0,r+1,r+1,1})$. Note that $k-1 \neq r+1$. Let C be the submatrix of A formed by the columns with a 0 in row 1 and rows $2, 3, \ldots m$. Then $C \in \mathrm{Avoid}_k(m-1, F_{1,r+1,r+1,0})$. Now $F_{1,r+1,r+1,0}$ is the (0,1)-complement of $F_{0,r+1,r+1,1}$ and so $\mathrm{forb}_k(m-1, F_{1,r+1,r+1,0}) = \mathrm{forb}_{m-k-1}(m-1, F_{0,r+1,r+1,1})$. Note that $k \neq (m-1)-r$. We deduce that $\|A\| = \|B\| + \|C\| \leq 2cm^r$ for $k \neq r+1, r+2, m-r-2, m-r-1$. For the remaining cases $k = r+1, r+2, m-r-2$ or $m-r-1$, we deduce that $\|A\|$ is $O(m^{r+1})$. ■

Let a two-columned F and an $A \in \mathrm{Avoid}(m, 3, \mathcal{T}_\ell(3)\backslash\mathcal{T}_\ell(2) \cup F)$ with fixed column sum k be given. Note that by Theorem 2.2, given L, there exists a constant t such that after $t\left(\sum_{i=k-t}^{k} \mathrm{forbmax}_i(m, F)\right)$ columns, we either find in A a configuration of $\mathcal{T}_L(3)\backslash\mathcal{T}_L(2)$, or $T_L(0,2,1)$ or F. If $L \geq \ell$, the only object on this list not forbidden is $T_L(0,2,1)$, so we may assume we find this configuration. Note that $T_{L/2}(1,0) \prec T_L(0,2,1)$, so $T_{L/2}(1,0)$ must appear. Reorder the columns so that the 1's are above the diagonal in $T_L(0,1)$. Now, using the previous notation for two-columned matrices, let $F = F_{a,b,c,d}$. Notice that if we delete the first a columns of $T_L(0,1)$, every pair of columns has a copies of $[1\,1]$; if we delete the last d columns, every pair of columns has d copies of $[0\,0]$; and if we take every cth column of what remains, every pair of columns has c copies of $[0\,1]$.

Let A' be the submatrix of A obtained by taking the selected columns from $T_L(0,1)$ and deleting the rows from $T_L(0,1)$. Note that in the deleted rows, every pair of columns has a copies of $[1\,1]$, c copies of $[0\,1]$, and d copies of $[0\,0]$, so if any pair of columns of A' have b copies of $[1\,0]$, A contains F. Also, since A has fixed column sum, and the column sums of $T_L(0,1)$ increase from left to right, the column sums of A' decrease from left to right. To use these facts we need the following lemma.

Lemma 4.5 *Let M be an m-rowed matrix such that:*

(i) M does not contain $\begin{bmatrix} \mathbf{1}_b & \mathbf{0}_b \end{bmatrix}$ as a $b \times 2$ submatrix for some b.

(ii) If $i < j$, column i of M has more 1's than column j

(iii) M avoids $\mathcal{T}_\ell(3)\backslash\mathcal{T}_\ell(2)$

Then for every r, there exists a constant c_r (dependent on k) such that if M has more than c_r columns, it has an $r \times r$ configuration in $\mathcal{P}_r^2$ (3) with 1's on the diagonal and 2's below the diagonal.

Proof: We proceed by induction. When $r = 1$, the desired object is just a single 0, so the lemma is trivial. Suppose the lemma holds for r. We claim that the lemma holds for $r+1$ with $c_{r+1} = R_5(\ell, \ell, \ell, c_r, b+1) + b$. Suppose M satisfies the hypotheses of the lemma and has c_{r+1} columns. Define M' to be the restriction of M to the rows with a 1 in the first column. Since the column sums of M strictly decrease from left to right, the $(b+1)$th column of M has at least b fewer 1's than the first, which implies that there must be at least b non-1 entries in the $(b+1)$th column of M'. At most $b-1$ of these entries are 0 by condition (i), so there is at least one 2. Pick one. The $(b+2)$th column of M' has at least two 2's, at least one of which is in a different row than the one already chosen. Pick one such 2. Similarly the $(b+3)$th column of M' has a 2 in a different row than the 2's already selected, and so on; continuing in this way, we find a diagonal of 2's of length $\|M\| - b = R_5(\ell, \ell, \ell, c_r, b+1)$. Let the square submatrix of A induced by the row and column indices of the chosen diagonal be M''.

We now produce a colouring of the complete graph on $\|M''\|$ vertices as follows. Given $i < j$, if $M''_{ij}, M''_{ji} \neq 0$, colour edge $\{i, j\}$ with the ordered pair (M''_{ij}, M''_{ji}); if $M''_{ij} = 0$ or $M''_{ji} = 0$, colour $\{i, j\}$ with 0. Now there are five colours: $(1,1), (1,2), (2,1), (2,2)$, and 0. By Ramsey Theory, we have a clique of size ℓ of colour $(1,1), (1,2)$ or $(2,1)$ or a clique of size $b+1$ of colour 0 or a clique of size c_r of colour $(2,2)$. In the first case, all three colours give rise to a member of $\mathcal{T}_\ell(3)\backslash\mathcal{T}_\ell(2)$. In the second case, we have a column with b 0's opposite the 1's in the first column of M, contradicting condition (i). Hence the only allowed case is the third, which corresponds to a block of 2's. In particular, there is a row x of M with c_r 2's and a 1. Look under the 2's; the resulting matrix has c_r columns and satisfy the hypotheses of the lemma, so by induction there is an $r \times r$ configuration with 1's on the diagonal and 2's above. Adding in row x gives an $(r+1) \times (r+1)$ configuration of the desired type. ■

Lemma 4.6 *Let M satisfy the hypotheses of Lemma 4.5 with c_r defined there. Then $\|M\| < c_{R_3(2\ell,\ell,\ell)}$.*

Proof: We use the notation c_r from the statement of Lemma 4.5. Suppose $\|M\| \geq c_{R_3(2\ell,\ell,\ell)}$. By Lemma 4.5, there exists a configuration $N \prec M$ in $\mathcal{P}^2_{R_3(2\ell,\ell,\ell)}$ with 1's on the diagonal and 2's below the diagonal. We colour a complete graph $K_{R_3(2\ell,\ell,\ell)}$ as follows: for $i < j$, colour edge (i,j) with N_{ji} (note that $N_{ij} = 2$). By the definition of $R_3(2\ell,\ell,\ell)$, there is a monochromatic clique. Three cases are possible. If there is a clique of size 2ℓ of colour 0, we get $T_{2\ell}(0,1,2)$, which contains $T_\ell(0,2)$. If there is a clique of size ℓ of colour 1, we have $T_\ell(1,2)$, and a clique of size ℓ of colour 2 yields $I_\ell(1,2)$. This contradicts our assumption that $\|M\| \geq c_{R(2\ell,\ell,\ell)}$. ■

Proof of Theorem 4.1: Let $A \in \text{Avoid}(m,3,\mathcal{T}_\ell(3)\backslash\mathcal{T}_\ell(2) \cup F)$ be given, with fixed column sum k. By Theorem 2.2, there exist constants C and d independent of k such with more than $C(\sum_{i=k-d}^{k}\text{forbmax}_i(m,F))$ columns, we either have one of the forbidden objects or a very large triangular matrix $T_t(0,2,1)$ with $t = c_{R_3(2\ell,\ell,\ell)}$. This yields a matrix M satisfying the hypotheses of Lemma 4.5 with $\|M\| > c_{R_3(2\ell,\ell,\ell)}$. Then Lemma 4.6 yields a contradiction. Hence, $||A|| \leq C(\sum_{i=k-d}^{k}\text{forbmax}_i(m,F))$, so $\text{forbmax}_k(m,3,\mathcal{T}_\ell(3)\backslash\mathcal{T}_\ell(2) \cup F) \leq C(\sum_{i=k-d}^{k}\text{forbmax}_i(m,F))$. Summing over k gives the desired result. ■

This result can be used to give bounds for many 2-columned matrices.

Proof of Theorem 1.10: For $F_{0,b,b,0}$ we use Lemma 4.2 which yields that $\text{forbmax}_k(m,3,\mathcal{T}_\ell(3)\backslash\mathcal{T}_\ell(2) \cup F_{0,b,b,0})$ is $O(m^{b-1})$. Then Theorem 4.1 yields that $\text{forb}(m,3,\mathcal{T}_\ell(3)\backslash\mathcal{T}_\ell(2) \cup F_{0,b,b,0})$ is $O(m^b)$. From [1], $\text{forb}(m,F_{0,b,b,0})$ is $\Theta(m^b)$.

By Corollary 4.4, $\text{forbmax}_k(m,F_{1,b,b,1})$ is $O(m^{b-1})$ for $b \geq 2$ and $k \neq r+1, r+2, m-r-2, m-r-1$. For $k = r+1, r+2, m-r-2$ or $m-r-1$, the bound os $O(m^b)$. By Theorem 4.1, $\text{forb}(m,3,\mathcal{T}_\ell(3)\backslash\mathcal{T}_\ell(2) \cup F_{1,b,b,1})$ is $O(m^b)$. From Theorem 1.14 we may extend this to obtain $\text{forb}(m,3,\mathcal{T}_\ell(3)\backslash\mathcal{T}_\ell(2) \cup F_{a,b,b,a})$ is $O(m^{a+b-1})$. This is the correct bound by [1]. ■

Proof of Theorem 1.9: Use Lemma 4.2 with $b = 1$ which by Theorem 4.1, yields that $\text{forb}(m,3,\mathcal{T}_\ell(3)\backslash\mathcal{T}_\ell(2) \cup F_{0,1,1,0})$ is $O(m)$. ■

We do not know how to do solve for $F = F_{1,1,1,1}$ for which $\text{forb}(m,F_{1,1,1,1})$ and $\text{forb}_k(m,F_{1,1,1,1})$ are both $\Theta(m)$. Similarly, the case $F = F_{a,1,1,a}$ for $a \geq 2$ is not solved. We have $\text{forb}(m,F_{a,1,1,a})$ is $\Omega(m^a)$. The following results give bounds which must be close to the correct bounds.

Theorem 4.7 *$\text{forb}(m,3,\mathcal{T}_\ell(3)\backslash\mathcal{T}_\ell(2) \cup F_{1,1,1,1})$ is $O(m\log m)$*

Proof: Let $A \in \text{Avoid}(m,F_{1,1,1,1})$ with column sum k be given. If $k \leq m/2$, then every pair of columns has a $[0\,0]$. Since the column sum is fixed, every pair of columns has an I_2. Hence there must be no [1 1] in any pair of columns. This means the 1's must all appear on disjoint rows, so there are at most $\frac{m}{k}$ columns. If $k > m/2$, take the 0-1 complement to get a similar result. Applying Theorem 4.1 and summing over k gives the result. ■

Applying Theorem 1.14 gives the following corollary.

Corollary 4.8 *forb*$(m, 3, \mathcal{T}_\ell(3)\backslash\mathcal{T}_\ell(2) \cup F_{a,1,1,a})$ *is* $O(m^a \log m)$

Of course, the extra factor of $\log m$ is undesirable. However, given that all known forbidden families have a polynomial bound, this strongly suggests that the actual bound for $F_{a,1,1,a}$ is $O(m^a)$.

5 An example with 3 columns

Define the useful notation $A|_S$ to denote the submatrix of A given by the rows S. In order to prove Theorem 1.12 for H given in (2), we find the following lemma usful. A standard decomposition applied to 3-matrices considers deleting a row i from a simple 3-matrix A. The resulting matrix might not be simple. Let $C_{a,b}(i)$ be the simple 3-matrix that consists of the repeated columns of the matrix that is obtained when deleting row r from A that lie under both symbol a and b in row i. In particular $[a\,b] \times C_{a,b}(i) \prec A$. Let $B(i)$ denote the $(m-1)$-rowed simple 3-matrix obtained from A by deleting row i and any repeats of columns so that

$$\|A\| \leq \|B(i)\| + \|C_{0,1}(i)\| + \|C_{1,2}(i)\| + \|C_{0,2}(i)\|.$$

The inequality arises from columns that are repeated three times in the matrix obtained from A by deleting row r but get counted four times on the right hand side. This bound on $\|A\|$ is often amenable to induction on the number of rows. If $K_2 = [0\,1] \times [0\,1] \not\prec A$, or in our case $H \not\prec A$, we deduce that $\|C_{0,1}(i)\|$ is $O(1)$, namely the constant bound for forb$(m, 3, \mathcal{T}_\ell(3)\backslash\mathcal{T}_\ell(2) \cup [0\,1])$ using Corollary 1.8. The following lemma could help with $\|C_{1,2}(i)\|$ and $\|C_{0,2}(i)\|$.

Lemma 5.1 *Let* $A \in \text{Avoid}(m, 3, \mathcal{T}_\ell(3)\backslash\mathcal{T}_\ell(2))$. *Assume for some set of rows* S *we have* $[\mathbf{0} \mid I_{|S|}] \prec A|_S$ *and for each pair of rows* $i, j \in S$, *we have no* $\left[\begin{smallmatrix}1\\1\end{smallmatrix}\right]$ *in* A. *If* $|S| > 3\ell \cdot c(2, \ell)$, *then there is some row* $i \in S$ *for which* $\|C_{1,2}(i)\| = \|C_{0,2}(i)\| = 0$.

Proof: We will show that $\|C_{1,2}(i)\| > 0$ for only a few choices $i \in S$ and similarly show that $\|C_{0,2}(i)\| > 0$ for only a few choices $i \in S$. Then for S large enough, there will be some $i \in S$ with $\|C_{1,2}(i)\| = \|C_{0,2}(i)\| = 0$.

Let U denote the rows $i \in S$ for which $\|C_{1,2}(i)\| > 0$. Assume $|U| \geq \ell \cdot c(2,\ell)$. When $\|C_{1,2}(i)\| > 0$, we have (at least) two columns in A differing only in row i, one with a 1 and one with a 2. Choose one such pair of columns γ,δ as shown:

$$\begin{array}{c} i \\ U\backslash i \\ [m]\backslash U \end{array} \left[\begin{array}{cc} 1 & 2 \\ \alpha & \alpha \\ \beta & \beta \end{array}\right] \prec A.$$

It is possible that for many i, the same second column might be chosen. By the property of A that A has no $\left[\begin{smallmatrix}1\\1\end{smallmatrix}\right]$ on rows of S and hence U, we deduce $\delta|_U = \left[\begin{smallmatrix}2\\\alpha\end{smallmatrix}\right]$ is a (0,2)-vector. By Theorem 1.2 (in this case due to [6]), we have that there are at most $c(2,\ell)$ choices for $\left[\begin{smallmatrix}2\\\alpha\end{smallmatrix}\right]$. Now there are $|U|$ choices for i and so, given our bound on $|U|$, there are ℓ choices for $i \in U$ which have the same $\left[\begin{smallmatrix}2\\\alpha\end{smallmatrix}\right]$. Now considering the ℓ columns $\left[\begin{smallmatrix}1\\\alpha\end{smallmatrix}\right]$ yields an $\ell \times \ell$ matrix in $A|_U$ with 1's on diagonal and 2's off the diagonal namely $I_\ell(2,1) \in \mathcal{T}_\ell(3)\backslash\mathcal{T}_\ell(2)$, a contradiction. Thus $\|C_{1,2}(i)\| > 0$ for less than $\ell \cdot c(2,\ell)$ choices i.

Assume $\|C_{0,2}(i)\| > 0$ for $2\ell \cdot c(2,\ell)$ choices i. Denote the choices by V. Then we have the following

$$\begin{array}{c} i \\ V\backslash i \\ [m]\backslash V \end{array} \left[\begin{array}{cc} 0 & 2 \\ \alpha & \alpha \\ \beta & \beta \end{array}\right] \prec A \tag{14}$$

This case is a little more complicated because α may have up to one 1. We choose a subset $W \subseteq V$ of the rows where α has no 1's. This can be done as follows. Choose some row $i_1 \in V$ and assume the corresponding choice of columns yields an α with a 1 in row $j_1 \in V$ and if not let $j_1 = i_1$. Now choose a row $i_2 \in V\backslash\{i_1, j_1\}$ and assume the corresponding α has a 1 in row $j_2 \in V$ and if not $j_2 = i_2$. Now choose a row $i_3 \in V\backslash\{i_1, j_1, i_2, j_2\}$ and assume the corresponding α has a 1 in row $j_3 \in V$ and if not $j_3 = i_3$. Continue in this way to form $W = \{i_1, i_2, \ldots, i_{\ell\cdot c(2,\ell)}\}$ using the fact $|V| \geq 2\ell \cdot c(2,\ell)$. $|W| \geq \ell \cdot c(2,\ell))$ and for each $i \in W$ we have $\|C_{0,2}(i)\| > 0$ where we have on pair of cols in A as in (14) with α having no 1's. Now repeat the above argument for the (1,2)-case to obtain an $\ell \times \ell$ matrix in A with 0's on diagonal and 2's off the diagonal, namely $I_\ell(2,0) \in \mathcal{T}_\ell(3)\backslash\mathcal{T}_\ell(2)$, a contradiction. Thus $\|C_{0,2}(i)\| > 0$ for less than $2\ell \cdot c(2,\ell)$ choices i.

We deduce that for $|S| > 3\ell \cdot c(2,\ell)$, there exists a row i with $|C_{1,2}(i)| = |C_{0,2}(i)| = 0$. ■

Proof of Theorem 1.12: Let $A \in \text{Avoid}(m, 3, \mathcal{T}_\ell(3)\backslash\mathcal{T}_\ell(2) \cup H)$ be given. If A contains a large identity (or its complement) then by Lemma 5.1 there exists a row i with $C_{1,2}(i) = C_{0,2}(i) = \emptyset$. Note that $\|C_{0,1}(i)\|$ is $O(1)$ by Corollary 1.8, since $C_{0,1}(i)$ avoids $[0\,1]$. Thus, we can delete row i and at most $O(1)$ columns and obtain a simple matrix. Then induction on m would yield the desired $O(m)$ bound. Our goal is to show that a large identity must occur.

Let A_k be the submatrix of A with column sum k. If, for any L, $\|A_k\| > c(3,L)$ then either $I_L(0,1) \prec A_k$, $I_L(1,0) \prec A_k$, or $T_L(0,1) \prec A_k$. We will take L to be large.

We note that $H \prec I_L(0,1)$ and so the first case does not occur. In the second case $I_L(1,0) \prec A_k$ we have that A_k and indeed A does not have $\left[\begin{smallmatrix}0\\0\end{smallmatrix}\right]$ on the L rows containing $I_L(1,0)$ else $H \prec A$. Then apply Lemma 5.1, using the (0,1)-complement and note that $\|C_{1,0}(i)\|$ is $O(1)$ by Corollary 1.8 since $H \prec [0\,1] \times [0\,1]$ and hence $[0\,1] \not\prec C_{0,1}(i)$. This yields that $\|A\|$ is $O(m)$ by induction on m.

In the third case, with the triangular matrix $T_L(0,1) \prec A_k$, let A'_k be the matrix consisting of the columns from A_k containing $T_L(0,1)$. Assume the columns of A'_k are ordered consistent with $T_L(0,1)$. Let A''_k be the submatrix obtained from A'_k by deleting the L rows containing $T_L(0,1)$. Then the column sums of A''_k are decreasing from left to right. Let S be the rows containing 1's in the first column of A''_k. Every triple of columns in $T_L(0,1)$ has the submatrix $[0\,0\,1]$, so A''_k does not contain any submatrix $[1\,0\,0]$ else $H \prec A''_k \prec A$. Thus $A''_k|_S$ does not contain $[0\,0]$. Also A''_k has decreasing column sums from left to right. We proceed in a manner similar to the proof of Lemma 4.5. We first find a diagonal of entries either 0 or 2. By the pigeonhole principle, there is a long diagonal of 2's or a long diagonal of 0's. If there is a long diagonal of 2's we apply Ramsey Theory as before. Large cliques involving 0's are not allowed since $[0\,0]$ is forbidden, and hence we are forced to have a block of 2's. This yields a submatrix $[1\,2\,2\cdots 2]$ and so we proceed, as in proof of Lemma 4.5, considering the columns containing the 2's. If we continue doing this for enough rows, we find a forbidden object. Hence, there must be a point where no sufficiently long diagonal of 2's exists, so there is a long diagonal of 0's. In this case, apply Ramsey Theory again. Recalling that we have no submatrix $[0\,0]$, the only configurations that result are either in $\mathcal{T}_\ell(3)\backslash\mathcal{T}_\ell(2)$ or an identity complement $I_t(1,0)$ for some large t. Given that $H \not\prec A$, we have that there is no $\left[\begin{smallmatrix}0\\0\end{smallmatrix}\right]$ on any pair of the t rows. This allows us to use Lemma 5.1.

If $\|A_k\|$ is bounded by a constant for all k then $\|A\|$ is $O(m)$. If $\|A_k\|$ is a big enough constant then we obtain an $I_t(1,0) \prec A_k$ for some appropriately large t. By Lemma 5.1 we find some $i \in [m]$ with $\|C_{1,2}(i)\| = \|C_{0,2}(i)\| = 0$. As noted above, $\|C_{0,1}(i)\|$ is $O(1)$. Thus we can delete row i of A and at most $O(1)$ columns from A to obtain a simple matrix in Avoid$(m-1, 3, \mathcal{T}_\ell(3)\backslash\mathcal{T}_\ell(2) \cup H)$ and then apply induction. ■

6 Open problems

Some small examples of F for which we have not handled forb$(m, 3, \mathcal{T}_\ell(3)\backslash\mathcal{T}_\ell(2) \cup F)$ include:

$$K_2 = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix}, \text{ bound should be } O(m)$$

$$F_{0,2,1,0} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, \text{ bound should be } O(m)$$

$$F_{1,1,1,0} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, \text{ bound should be } O(m)$$

We would particularly like to have a general result that $\mathrm{forb}(m, 3, \mathcal{T}_\ell(3)\backslash\mathcal{T}_\ell(2)\cup([0\,1]\times F))$ is $O(m \times \mathrm{forb}(m, 3, \mathcal{T}_\ell(3)\backslash\mathcal{T}_\ell(2) \cup F))$ matching the standard induction results for (0,1)-forbidden configurations.

Given a $(0, 1)$-column α, we might consider a 3-matrix $A \in \mathrm{Avoid}(m, 3, \mathcal{T}_\ell(3)\backslash\mathcal{T}_\ell(2))$ such that each column of A arises from α by setting certain entries to 2. We deduce that $[0\,1] \nprec A$ and so by Corollary 1.8, we have the interesting fact that $\|A\|$ is $O(1)$. In some sense the columns of A are a 3-matrix replacement for α. We were unable to exploit this for Problem 1.5.

References

[1] R.P. Anstee and P. Keevash, Pairwise intersections and forbidden configurations, *European J. of Combin.* **27** (2006), 1235–1248.

[2] R.P. Anstee and C.L. Koch, Forbidden Families of Configurations, Australasian J. of Combinatorics **59**(2014), 361-377.

[3] R.P. Anstee and Linyuan Lu, Repeated columns and an old chestnut, *Elec. J. of Combinatorics* **20** (2013), P2, 11pp.

[4] R.P. Anstee and Linyuan Lu, Multicoloured Families of Configurations, arXiv 1409.4123, 16pp.

[5] R.P. Anstee, A Survey of forbidden configurations results, *Elec. J. of Combinatorics* **20** (2013), DS20, 56pp.

[6] J. Balogh, B. Bollobás, Unavoidable Traces of Set Systems, *Combinatorica*, **25** (2005), 633–643.

[7] J. Balogh, P. Keevash, and B. Sudakov, Disjoint representability of sets and their complements, *Journal of Combinatorial Theory Ser. B* **95** (2005), 12–28.

[8] Z. Füredi, A. Sali, Optimal Multivalued Shattering, *SIAM J. on Discrete Mathematics* **26**(2012), 737–744.

[9] M. Raggi, *Forbidden configurations*, Ph.D. Thesis, University of British Columbia, 2011.