

# Exponential multivalued forbidden configurations

Travis Dillon<sup>1\*</sup>Attila Sali<sup>2,3†</sup><sup>1</sup> Lawrence University, WI, USA<sup>2</sup> Alfréd Rényi Institute of Mathematics, Budapest, Hungary<sup>3</sup> Department of Computer Science, Budapest University of Technology and Economics, Budapest, Hungaryreceived 2<sup>nd</sup> July 2020, revised 2<sup>nd</sup> Mar. 2021, accepted 8<sup>th</sup> Mar. 2021.

The *forbidden number*  $\text{forb}(m, F)$ , which denotes the maximum number of unique columns in an  $m$ -rowed  $(0, 1)$ -matrix with no submatrix that is a row and column permutation of  $F$ , has been widely studied in extremal set theory. Recently, this function was extended to  $r$ -matrices, whose entries lie in  $\{0, 1, \dots, r-1\}$ . The combinatorics of the generalized forbidden number is less well-studied. In this paper, we provide exact bounds for many  $(0, 1)$ -matrices  $F$ , including all 2-rowed matrices when  $r > 3$ . We also prove a stability result for the  $2 \times 2$  identity matrix. Along the way, we expose some interesting qualitative differences between the cases  $r = 2$ ,  $r = 3$ , and  $r > 3$ .

**Keywords:** forbidden configurations,  $(0, 1)$ -matrices, extremal set theory

## 1 Introduction

We call a matrix *simple* if it has no repeated columns. Every set system (or simple hypergraph) corresponds to a simple  $(0, 1)$ -matrix via its element-set incidence matrix, and such matrices provide a convenient language for extremal set theory. We generalize this situation to  $r$ -matrices, which have entries in  $\{0, 1, \dots, r-1\}$ . Such matrices can be thought of as  $r$ -coloured set systems or as representations of collections of functions from a given finite set into  $\{0, 1, \dots, r-1\}$ .

For two matrices  $F$  and  $A$ , we say that  $F$  is a *configuration* of  $A$ , denoted  $F \prec A$ , if  $A$  contains a submatrix which is a row and column permutation of  $F$ . If  $F \not\prec A$ , we say that  $A$  *avoids*  $F$ . Configurations of simple  $(0, 1)$ -matrices correspond to traces of set systems or hypergraphs. For a given finite collection  $\mathcal{F}$  of matrices, we denote by  $\text{Avoid}(m, r, \mathcal{F})$  the collection of  $m$ -rowed, simple  $r$ -matrices that avoid

\*Research conducted under the auspices of the Budapest Semesters in Mathematics program.

†This author's work is partially supported by the National Research, Development and Innovation Office (NKFIH) [grant K-116769 and K-132696]; the National Research, Development and Innovation Fund [TUDFO/51757/2019-ITM, Thematic Excellence Program]; the BME NC TKP2020 grant of NKFIH Hungary; and the BME-Artificial Intelligence FIKP grant of EMMI [BME FIKP-MI/SC]. It is also connected to the "Development of quality-oriented and harmonized R+D+I strategy and functional model at BME" project supported by the New Hungary Development Plan [Project ID: TÁMOP-4.2.1/B-09/1/KMR-2010-0002].

every matrix  $F \in \mathcal{F}$ . We let  $|A|$  denote the number of columns of  $A$ . The main extremal function in the study of forbidden configurations is

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

When  $r = 2$  we usually write  $\text{forb}(m, \mathcal{F})$  in place of  $\text{forb}(m, 2, \mathcal{F})$ . We also use  $\text{forb}(m, r, F)$  instead of the more cumbersome  $\text{forb}(m, r, \{F\})$ . We will use several simple properties of this function. For example, if  $F \prec F'$  then  $\text{forb}(m, r, F) \leq \text{forb}(m, r, F')$ . Also, we let  $F^c$  denote the *complement* of the  $(0, 1)$ -matrix  $F$ , where each 0 is replaced by a 1 and vice versa; then  $\text{forb}(m, r, F) = \text{forb}(m, r, F^c)$ .

The foundational result in the theory of forbidden configurations is Sauer's theorem (proven in [10], also by Perles and Shelah [11] and Vapnik and Chervonenkis [12]). Let  $K_k$  denote the complete  $k \times 2^k$  simple  $(0, 1)$ -configuration (corresponding to the power set of a  $k$ -element set).

**Theorem 1.1.** *For every positive integer  $m$ ,*

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

Alon [1] gave a generalization for complete  $r$ -matrices, but the forbidden number is exponential when  $r > 2$ . This is a special case of a more general phenomenon proved by Füredi and Sali [9].

**Theorem 1.2.** *Let  $\mathcal{F}$  be a family of  $r$ -matrices. If for every pair  $i, j \in \{0, 1, \dots, r-1\}$  there is an  $(i, j)$ -matrix in  $\mathcal{F}$ , then  $\text{forb}(m, r, \mathcal{F}) = O(m^k)$  for some positive integer  $k$ . If  $\mathcal{F}$  has no  $(i, j)$ -matrix for some pair  $i, j \in \{0, 1, \dots, r-1\}$ , then  $\text{forb}(m, r, \mathcal{F}) = \Omega(2^m)$ .*

Extensive investigations have been undertaken for forbidden configurations of simple  $(0, 1)$ -matrices; see, for example, the excellent dynamic survey of Anstee [3]. On the other hand, the more general case of  $r$ -matrices is not so well-explored. Previous papers mainly focus on providing bounds on the forbidden number for special classes of sets in the polynomial case [5, 6]. In this paper, we dive into exponential forbidden numbers and provide exact bounds when  $(0, 1)$ -configurations of  $r$ -matrices are forbidden. This is similar in flavour to a recent paper of Füredi, Kostochka, and Luo [7], which proves several minimum-degree conditions that guarantee cycles in hypergraphs; by dropping the assumption of uniformity, their bounds jump from polynomial to exponential.

The structure of the paper is as follows. Section 2 provides a method to transfer bounds for  $r = 2$  to larger values of  $r$ . The following three sections calculate forbidden numbers of specific classes of matrices. We obtain exact results when  $r > 3$  and bounds for  $r = 3$  that differ from the forbidden number by an additive constant. We also prove a stability result for the identity configuration. Our work culminates in Section 6, which provides exact forbidden numbers for all two-rowed  $(0, 1)$ -configurations for every  $r > 3$  and a large class of two-rowed  $(0, 1)$  configurations when  $r = 3$ . The main tool in both cases a reduction lemma. Finally, Section 7 applies the method of Section 2 to obtain a nearly complete classification of  $(0, 1)$ -configurations of size  $3 \times 2$  and  $3 \times 3$ .

## 2 General bounds

For a given configuration  $A$ , let  $\bar{A}$  denote its underlying simple configuration. If  $A$  has  $m$  columns and  $S \subseteq [m]$ , then we let  $A|_S$  be the restriction of  $A$  to the rows with indices in  $S$ . By convention, we set  $\text{forb}(0, F) = 1$  for all  $F$ . In general, if  $F$  has  $t$  rows, then  $\text{forb}(k, F) = 2^k$  when  $0 \leq k < t$ .

**Lemma 2.1.** *If  $F$  is a  $(0, 1)$ -matrix and  $r \geq 3$ , then*

$$\text{forb}(m, r, F) \leq \sum_{k=0}^m \binom{m}{k} (r-2)^{m-k} \text{forb}(k, F). \quad (2.1)$$

**Proof:** Let  $A \in \text{Avoid}(m, r, F)$ , and let  $X$  be a  $k$ -element subset of the rows. Consider the matrix  $C$  obtained by taking all columns of  $A$  that have 0's and 1's in exactly the rows in  $X$ , and let  $C' = \overline{C|_X}$ . We know that  $|C'| \leq \text{forb}(k, F)$ . Each column in  $C'$  appears with multiplicity at most  $(r-2)^{m-k}$  in  $C|_X$ , so  $|C| \leq (r-2)^{m-k} \text{forb}(k, F)$ . To finish the proof, we sum over all subsets of the rows.  $\square$

The bound given by this lemma may be quite bad, especially if  $F$  is not simple. However, for simple matrices, we have the following lower bound.

**Lemma 2.2.** *Let  $F$  be a simple  $(0, 1)$ -matrix with  $n$  rows and fix  $r \geq 3$ . Suppose that  $(A_k)_{k=1}^\infty$  is a sequence of  $(0, 1)$ -matrices that avoids  $F$ , where  $A_k$  has  $k$  rows, such that  $\overline{A_k|_S} \subseteq A_n$  for every  $k \geq n$  and  $S \in \binom{[k]}{n}$ . If we set  $|A_0| = 1$ , then*

$$\text{forb}(m, r, F) \geq \sum_{k=0}^m \binom{m}{k} (r-2)^{m-k} |A_k|. \quad (2.2)$$

**Proof:** We construct a configuration that avoids  $F$  as follows. Let  $k \in [m]$ . For each  $k$ -set  $X$  of rows, we choose the  $(r-2)^{m-k}$  columns that contain a copy of  $A_k$  in the rows of  $X$  and have elements of  $\{2, \dots, r-1\}$  in every other position. Let  $A$  be the configuration that contains all such columns. If  $F \prec A$ , then  $F \prec A|_S$  for some  $n$ -set of rows  $S$ . But every column in  $A|_S$  appears in  $A_n$ , so  $F \prec A_n$ , a contradiction.  $\square$

The condition that  $F$  is simple is absolutely essential. For simple matrices, however, this lemma can easily extend bounds from the classical case to the generalized one. In particular, combining Lemmas 2.1 and 2.2 proves the following.

**Lemma 2.3.** *Let  $F$  be a simple  $n$ -rowed  $(0, 1)$ -matrix. If there exists a sequence  $(A_k)_{k=1}^\infty$  of  $(0, 1)$ -matrices, each of which avoids  $F$ , such that*

- $A_k$  has  $k$  rows,
- $|A_k| = \text{forb}(k, F)$ , and
- $\overline{A_k|_S}$  is contained in  $A_n$  for every  $k \geq n$  and  $n$ -set  $S \subseteq [k]$ , then

$$\text{forb}(m, r, F) = \sum_{k=0}^m \binom{m}{k} (r-2)^{m-k} \text{forb}(k, F). \quad (2.3)$$

### 3 Complete configurations

**Proposition 3.1.** *We have  $\text{forb}(m, r, K_k) = \sum_{i=0}^{k-1} \binom{m}{i} (r-1)^{m-i}$ . If  $(r-1)^{m-k} \geq p-1$ , then  $\text{forb}(m, r, p \cdot K_k) = \sum_{i=0}^{k-1} \binom{m}{i} (r-1)^{m-i} + (p-1) \binom{m}{k}$ .*

**Proof:** We first prove that  $\text{forb}(m, r, K_k) = \sum_{i=0}^{k-1} \binom{m}{i} (r-1)^{m-i}$ . Let  $A_n$  denote the  $n$ -rowed configuration that contains every column with at most  $k-1$  zeros. Then  $(A_n)$  satisfies the conditions of Lemma 2.3, so Theorem 1.1 implies that

$$\begin{aligned} \text{forb}(m, r, K_k) &= \sum_{n=0}^m \binom{m}{n} (r-2)^{m-n} \sum_{i=0}^{k-1} \binom{n}{i} \\ &= \sum_{i=0}^{k-1} \binom{m}{i} \sum_{n=0}^m \binom{m-i}{n-i} (r-2)^{m-n} \\ &= \sum_{i=0}^{k-1} \binom{m}{i} (r-1)^{m-i}. \end{aligned}$$

Now we prove the forbidden number for all  $p$ . The configuration that contains every column with at most  $k-1$  zeros avoids  $K_k$ . If  $(r-1)^{m-k} \geq p-1$ , for each  $k$ -set of rows, we may append  $p-1$  columns to this matrix that have zeros in that  $k$ -set and nowhere else. The resulting configuration avoids  $p \cdot K_k$  and has  $\sum_{i=0}^{k-1} \binom{m}{i} (r-1)^{m-i} + (p-1) \binom{m}{k}$  columns.

Now suppose that  $A \in \text{Avoid}(m, r, p \cdot K_k)$ . For each  $k$ -set  $X$  of rows, there is a column of  $K_k$  that appears at most  $p-1$  times in  $A|_X$ . Let  $A'$  be the configuration obtained by deleting the corresponding columns of  $A$  for all  $k$ -sets. Since  $K_k$  is symmetric, no row-permutation of  $K_k$  is a subset of  $A'|_X$ , so  $K_k \not\prec A'|_X$  for every  $k$ -set  $X$ . Therefore  $K_k \not\prec A'$ , which implies that

$$|A| \leq |A'| + (p-1) \binom{m}{k} \leq \sum_{n=0}^{k-1} \binom{m}{n} (r-1)^{m-n} + (p-1) \binom{m}{k},$$

as claimed.  $\square$

Proposition 3.1 is enough to determine the logarithmic growth rate of  $\text{forb}(m, r, F)$  asymptotically for every  $(0, 1)$ -configuration  $F$ .

**Corollary 3.2.** *The asymptotic formula  $\log \text{forb}(m, r, F) \sim m \log(r-1)$  holds as  $m \rightarrow \infty$  for every fixed  $(0, 1)$ -configuration  $F$  and  $r \geq 3$ .*

**Proof:** Since  $F \prec p \cdot K_k$  for some  $p$  and  $k$ , Proposition 3.1 guarantees a constant  $C > 0$  so that  $\text{forb}(m, r, F) \leq Cm^{k-1}(r-1)^m$  for every  $m$  and  $r$ . We may assume by complementation that  $F$  contains at least one 0, in which case the configuration that contains every column with no 0's avoids  $F$ ; this implies that  $\text{forb}(m, r, F) \geq (r-1)^m$  for every  $m, r \in \mathbb{N}$ . If  $r \geq 3$  is fixed, then the logarithmic growth rates of the lower and upper bounds are asymptotically equal as  $m \rightarrow \infty$ .  $\square$

The trivial bound  $\text{forb}(m, r, F) \leq r^m$  combined with the lower bound  $\text{forb}(m, r, F) \geq (r-1)^m$  shows that  $\text{forb}(m, r, F) = \Theta(r^m)$  if  $m$  is fixed and  $\text{forb}(m, r, F)$  is regarded as a function of  $r$ .

Going back to exact results, let  $K_k^s$  denote the  $k \times \binom{k}{s}$  configuration of zeros and ones in which every column contains  $s$  ones, called the *complete uniform configuration of weight  $s$* . Füredi and Quinn proved in [8] that  $\text{forb}(m, K_k^s) = \sum_{i=0}^{k-1} \binom{m}{i}$ . The configuration where  $s$  ones never appear above  $k-s$  zeros provides the lower bound; since  $K_k^s \prec K_k$ , Sauer's theorem provides the upper bound. The construction easily extends, yielding the following result.

**Proposition 3.3.** *If  $s \leq k$ , then  $\text{forb}(m, r, K_k^s) = \sum_{i=0}^{k-1} \binom{m}{i} (r-1)^{m-i}$ . If  $(r-2)^{m-k} \geq p-1$ , then  $\text{forb}(m, r, p \cdot K_k^s) = \text{forb}(m, r, K_k^s) + (p-1) \binom{m}{k}$ .*

**Proof:** Let  $A_n$  be the  $n$ -rowed configuration that contains every column in which  $s$  ones do not appear above  $k-s$  zeros. The sequence  $(A_n)$  satisfies the conditions of Lemma 2.3, and an identical calculation to the one in the proof of Proposition 3.1 proves the first statement.

The proof of the upper bound for the second statement is identical to the one in Proposition 3.1. For the lower bound, let  $A$  be the configuration that contains every column where  $s$  ones never appear above  $k-s$  zeros; this configuration avoids  $K_k^s$ . If  $(r-2)^{m-k} \geq p-1$ , then for each  $X \in \binom{[m]}{k}$  we can append  $p-1$  columns to  $A$  that have  $s$  ones above  $k-s$  zeros in the rows of  $X$  and non-binary digits elsewhere. The resulting configuration avoids  $p \cdot K_k^s$  and has  $\text{forb}(m, r, p \cdot K_k^s) = \text{forb}(m, r, K_k^s) + (p-1) \binom{m}{k}$  columns.  $\square$

A matrix is called  $p$ -simple if each column has multiplicity at most  $p$ .

**Corollary 3.4.** *Assume that  $F$  is a  $k$ -rowed  $p$ -simple matrix such that  $p \cdot K_k^s \prec F$  for some  $0 \leq s \leq k$ . If  $(r-2)^{m-k} \geq p-1$ , then*

$$\text{forb}(m, r, F) = \sum_{n=0}^{k-1} \binom{m}{n} (r-1)^{m-n} + (p-1) \binom{m}{k}.$$

*If  $F$  is simple and  $K_k^s \prec F$ , then  $\text{forb}(m, r, F) = \text{forb}(m, r, K_k^s) = \text{forb}(m, r, K_k)$  for all  $m \in \mathbb{N}$  and  $r \geq 2$ .*

**Proof:** Since  $p \cdot K_k^s \prec F \prec p \cdot K_k$ , the statement follows from Propositions 3.1 and 3.3.  $\square$

The result for non-simple matrices in Proposition 3.3 is only applicable when  $r > 3$ . The argument can be modified to show that  $\text{forb}(m, 3, p \cdot K_k^s)$  is at most an additive constant away from  $\sum_{i=0}^{k-1} \binom{m}{i} 2^{m-i} + (p-1) \binom{m}{k}$ .

**Proposition 3.5.** *Suppose  $p > 1$  and  $a = \lceil \log_2(p-1) \rceil$ . Then*

$$\text{forb}(m, 3, p \cdot K_k^s) \geq \sum_{i=0}^{k-1} \binom{m}{i} 2^{m-i} + (p-1) \left( \binom{m}{k} - \binom{k+a-1}{k} \right).$$

**Proof:** Let  $A$  be the configuration with all columns that do not contain  $s$  ones above  $k-s$  zeros. For every  $k$ -set  $X$  with elements  $i_1 < i_2 < \dots < i_k$  and  $i_s + (m - i_{s+1}) - k \geq \log_2(p-1)$ , we may append  $p-1$  columns to  $A$  with entries  $c_i$  given by

$$\begin{cases} c_i = 1 & \text{if } i \in X \text{ and } i \leq i_s \\ c_i \in \{0, 2\} & \text{if } i \notin X \text{ and } i \leq i_s \\ c_i = 2 & \text{if } i_s < i < i_{s+1} \\ c_i = 0 & \text{if } i \in X \text{ and } i \geq i_{s+1} \\ c_i \in \{1, 2\} & \text{if } i \notin X \text{ and } i \geq i_{s+1}. \end{cases}$$

For each such column  $\mathbf{c}$ , there is exactly one  $k$ -set  $S$  (namely  $S = X$ ) so that  $\mathbf{c}|_S$  is  $s$  ones above  $k - s$  zeros. Therefore, the resulting configuration  $A'$  avoids  $p \cdot K_k^s$ .

To determine the number of columns added to  $A$ , we count the number of choices of  $X$  with  $i_s + (m - i_{s+1}) - k < \log_2(p - 1)$ . The number of choices with  $i_s + (m - i_{s+1}) - k = b$  is  $\binom{k-1+b}{k-1}$ , so the number of choices of  $X$  not covered in our strategy is

$$\sum_{b=0}^{a-1} \binom{k-1+b}{k-1} = \binom{k+a-1}{k}.$$

In total, then  $A'$  contains  $(p - 1) \left( \binom{m}{k} - \binom{k+a-1}{k} \right)$  more columns than  $A$ .  $\square$

**Corollary 3.6.**  $\text{forb}(m, 3, 2 \cdot K_k^s) = \sum_{i=0}^{k-1} \binom{m}{i} (r - 1)^{m-i} + \binom{m}{k}$ .

**Proof:** Applying Proposition 3.5 with  $p = 2$  gives the lower bound, and the upper bound follows from Proposition 3.1 together with  $2 \cdot K_k^s \prec 2 \cdot K_k$ .  $\square$

## 4 Identity matrices

Noting that  $I_k = K_k^{-1}$  yields the following corollary of Proposition 3.3.

**Corollary 4.1.** *If  $r > 3$ , then  $\text{forb}(m, r, p \cdot I_k) = \sum_{i=0}^{k-1} \binom{m}{i} (r - 1)^{m-i} + (p - 1) \binom{m}{k}$  for all  $m$  such that  $(r - 2)^{m-k} \geq p - 1$ .*

The main result of this section is a stability theorem for  $I_2$ . It would be interesting to see similar stability theorems for other complete uniform configurations.

With each configuration  $A \in \text{Avoid}(m, r, I_2)$  we can associate a tournament on  $m$  vertices. Direct an edge from  $i$  to  $j$  if there is no column in which 0 appears in row  $i$  and 1 appears in row  $j$ . If both  $ij$  and  $ji$  are possible edges, choose just one. Since  $A$  avoids  $I_2$ , there must be an edge between each pair of vertices, so this construction gives a tournament  $T_A$  on  $m$  vertices.

**Proposition 4.2.** *Let  $r \geq 2$  and  $A \in \text{Avoid}(m, r, I_2)$  such that  $T_A$  is not transitive. Then  $|A| \leq m(r - 1)^{m-1} + (r - 1)^m - 2(r - 1)^{m-3}$ .*

**Proof:** We first prove the case  $r = 2$ : If  $A \in \text{Avoid}(k, 2, I_2)$  such that  $T_A$  is not transitive, then  $|A| \leq m - 1$ . Since  $T_A$  is not transitive, it contains a 3-cycle  $ijk$ . The only possible columns in  $A|_{\{i,j,k\}}$  are  $\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$  and  $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ . If we delete rows  $i$  and  $j$ , then the resulting configuration  $A'$  is simple and avoids  $I_2$ , so  $|A| = |A'| \leq \text{forb}(m - 2, I_2) = m - 1$ .

We now proceed with the general case. Suppose that  $A \in \text{Avoid}(m, r, I_2)$  with  $T_A$  not transitive. As before, there is a 3-cycle  $ijk$  in  $T_A$ . Applying the argument used in the proof of Lemma 2.1 and splitting the sum over sets that do or do not contain  $\{i, j, k\}$  gives the bound

$$\text{forb}(m, r, I_2) \leq \sum_{\substack{X \subseteq [m] \\ \{i,j,k\} \not\subseteq X}} (r - 2)^{m-|X|} (|X| + 1) + \sum_{\substack{X \subseteq [m] \\ \{i,j,k\} \subseteq X}} (r - 2)^{m-|X|} (|X| - 1)$$

$$= \sum_{X \subseteq [m]} (r-2)^{m-|X|} (|X|+1) - 2 \sum_{\substack{X \subseteq [m] \\ \{i,j,k\} \subseteq X}} (r-2)^{m-|X|}.$$

The left sum simplifies to

$$\sum_{k=0}^m \binom{m}{k} (r-2)^{m-k} (k+1) = m(r-1)^{m-1} + (r-1)^m,$$

and the right sum is

$$\sum_{k=3}^m \binom{m-3}{k-3} (r-2)^{m-k} = \sum_{k=0}^{m-3} \binom{m-3}{k} (r-2)^{m-3-k} = (r-1)^{m-3}.$$

Combining the two evaluations completes the proof.  $\square$

**Theorem 4.3.** *For each integer  $r \geq 2$ , there is a unique extremal  $r$ -configuration with  $m$  rows that avoids  $I_2$ .*

**Proof:** By Proposition 4.2, if  $A$  is extremal, then  $T_A$  is transitive. Therefore there is an ordering  $i_1, \dots, i_m$  of  $[m]$  so that  $i_s i_t$  is an edge of  $T$  if and only if  $s < t$ . After permuting the rows of  $A$  according to this order, no 0 appears above a 1. There are  $m(r-1)^{m-1}$  such columns that contain a 0 and  $(r-1)^m$  columns with no 0. Since  $A$  is extremal, it contains all these columns. Up to row and column permutation, therefore,  $A$  is unique.  $\square$

Thus, there is a gap between the unique extremal configuration that avoids  $I_2$  and any other configuration that avoids  $I_2$  but is not a subconfiguration of the extremal one.

In another direction, Propositions 3.1 and 3.3 and Corollary 3.6 show that  $\text{forb}(m, 3, p \cdot I_2) = \text{forb}(m, 3, p \cdot K_2)$  when  $p = 1$  or  $p = 2$ . However, equality does not hold for higher values of  $p$ . The following exact evaluation of  $\text{forb}(m, r, 3 \cdot I_2)$  shows that  $\text{forb}(m, 3, p \cdot I_k) \neq \text{forb}(m, 3, p \cdot K_k)$  in general. In contrast, Corollary 4.1 states that  $\text{forb}(m, r, p \cdot I_k) = \text{forb}(m, r, p \cdot K_k)$  for every  $p \geq 1$  when  $r > 3$ .

**Proposition 4.4.** *If  $m \geq 4$ , then  $\text{forb}(m, 3, 3 \cdot I_2) = \text{forb}(m, 3, 3 \cdot K_2) - 1$ .*

**Proof:** Let  $A$  be the configuration constructed in the proof of Proposition 3.5 with  $\text{forb}(m, 3, 3 \cdot K_2) - 2$  columns that avoids  $3 \cdot I_2$ . Appending the column  $\mathbf{c}$  with  $\mathbf{c}_1 = 1$ ,  $\mathbf{c}_m = 0$ , and  $\mathbf{c}_i = 2$  for every  $1 < i < m$  creates a configuration with  $\text{forb}(m, 3, 3 \cdot K_2) - 1$  columns that avoids  $3 \cdot I_2$ .

We now show that any 3-configuration that avoids  $3 \cdot I_2$  has at most  $\text{forb}(m, 3, 3 \cdot K_2) - 1$  columns. In each pair of rows, either  $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$  or  $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$  appears at most twice. Permuting the corresponding columns to the right end of the configuration  $A$ , we create a decomposition  $A = [BC]$  where  $|C| \leq 2 \binom{m}{2}$  and  $B$  avoids  $I_2$ . If  $B$  is not the unique extremal configuration that avoids  $I_2$ , then

$$|A| = |B| + |C| \leq \text{forb}(m, 3, K_2) - 1 + 2 \binom{m}{2} = \text{forb}(m, 3, 3 \cdot K_2) - 1.$$

Otherwise, Theorem 4.3 shows that we may permute the rows of  $B$  so it contains every column where no 0 appears above a 1. Since  $B$  has at least four rows,  $B|_{\{i,j\}}$  contains at least four columns of the form  $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$  for every  $i, j \in [m]$  with  $i < j$ .

We mark the pair  $i < j$  for each time that 0 appears in row  $i$  and 1 appears in row  $j$  in the configuration  $C$ . Since  $A$  avoids  $3 \cdot I_2$  and  $B|_{\{i,j\}}$  already contains four columns of the form  $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ , each pair has at most two marks. Each column of  $C$  contributes at least one mark. If the pair  $(1, m)$  has at most one mark, then there are at most  $2^{\binom{m}{2}} - 1$  columns in  $C$ . If  $(1, m)$  has two marks, then there is a column  $\mathbf{c}$  in  $C$  with  $c_1 = 0$ ,  $c_m = 1$ , and  $c_s \neq 2$  for some  $1 < s < m$ . In this case the column  $\mathbf{c}$  contributes at least two marks: one for  $(1, m)$ , and one for either  $(1, s)$  or  $(s, m)$ . So in this case, too, there are at most  $2^{\binom{m}{2}} - 1$  columns in  $C$ . In either case,

$$|A| = |B| + |C| \leq \text{forb}(m, 3, 3 \cdot K_2) - 1,$$

proving the lower bound.  $\square$

The upper bound in this argument shows that  $\text{forb}(m, 3, p \cdot K_2) < \text{forb}(m, 3, p \cdot I_2)$  for all  $p \geq 3$ . Indeed, by following this mark argument, one can calculate exact forbidden numbers for larger  $p$ . It's not too hard to show, for example, that  $\text{forb}(m, 3, 4 \cdot I_2) = \text{forb}(m, 3, 4 \cdot K_2) - 2$  and  $\text{forb}(m, 3, 5 \cdot I_2) = \text{forb}(m, 3, 5 \cdot K_2) - 5$ . The computations, however, rapidly become rather case-heavy as  $p$  increases. In general, the mark argument can be extended to show that the difference between  $\text{forb}(m, 3, p \cdot I_2)$  and  $\text{forb}(m, 3, p \cdot K_2)$  is superlinear in  $p$ ; for example,

$$\text{forb}(m, 3, p \cdot I_2) \leq \text{forb}(m, 3, p \cdot K_2) - \frac{1}{4}(p-1) \log_2(p-1) (\log_2(p-1) - 1), \quad (4.1)$$

although this is not sharp.

## 5 Block matrices

**Proposition 5.1.** *If  $(r-2)^{m-a-b} \geq p-1$ , then*

$$\begin{aligned} \text{forb}\left(m, r, \begin{bmatrix} \mathbf{0}_{a \times p} \\ \mathbf{1}_{b \times p} \end{bmatrix}\right) &= \sum_{\ell=0}^{a-1} \binom{m}{\ell} (r-1)^{m-\ell} + \sum_{k=0}^{b-1} \binom{m}{k} (r-1)^{m-k} \\ &\quad - \sum_{\ell=0}^{a-1} \sum_{k=0}^{b-1} \binom{m}{\ell} \binom{m-\ell}{k} (r-2)^{m-\ell-k} + (p-1) \binom{m}{a} \binom{m-a}{b}. \end{aligned} \quad (5.1)$$

**Proof:** Any maximal matrix that avoids  $F := \begin{bmatrix} \mathbf{0}_{a \times p} \\ \mathbf{1}_{b \times p} \end{bmatrix}$  contains all columns that have fewer than  $a$  zeros or fewer than  $b$  ones. This accounts for the first three terms of (5.1). Thus we need only bound the number of columns that contain at least  $a$  zeros and at least  $b$  ones. There are  $(r-2)^{m-a-b}$  columns that contain exactly  $a$  zeros and  $b$  ones for a fixed  $a$ -set  $X$  and  $b$ -set  $Y$  of rows. If  $(r-2)^{m-a-b} \geq p-1$ , then for each disjoint  $X, Y \subseteq [m]$  with  $|X| = a$  and  $|Y| = b$ , we may take  $p-1$  columns with 0's in the rows in  $X$  and 1's in the rows of  $Y$  and entries in  $\{2, \dots, r-1\}$  elsewhere. This is  $(p-1) \binom{m}{a} \binom{m-a}{b}$  columns, which provides the lower bound.

For the upper bound, we again use a mark argument. Consider the set of ordered pairs  $(X, Y)$  where  $X, Y \subseteq [m]$  are disjoint,  $|X| = a$ , and  $|Y| = b$ . Given a matrix  $A$ , we place a mark on the pair  $(X, Y)$  for every column  $\mathbf{c} \in A$  such that  $\mathbf{c}|_X$  contains only zeros and  $\mathbf{c}|_Y$  contains only ones. There can be at most  $(p-1) \binom{m}{a} \binom{m-a}{b}$  marks in total if the matrix  $A$  avoids  $F$ . Every column that contains at least  $a$  zeros and



$b$  ones contributes at least one mark, so there are at most  $(p-1)\binom{m}{a}\binom{m-a}{b}$  such columns, which gives the upper bound.  $\square$

**Corollary 5.2.** *If  $(r-2)^{m-2} \geq p-1$ , then*

$$\text{forb}(m, r, p \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix}) = 2(r-1)^m - (r-2)^m + (p-1)m(m-1).$$

## 6 Forbidden configurations with 2 rows

We define the general 2-rowed  $(0, 1)$ -forbidden configuration

$$F(a, b, c, d) = \begin{bmatrix} 0 & \cdots & 01 & \cdots & 10 & \cdots & 01 & \cdots & 1 \\ 0 & \cdots & 00 & \cdots & 01 & \cdots & 11 & \cdots & 1 \end{bmatrix}. \quad (6.1)$$

$\underbrace{\hspace{10em}}_a \quad \underbrace{\hspace{10em}}_b \quad \underbrace{\hspace{10em}}_c \quad \underbrace{\hspace{10em}}_d$

Our main tools will be two reduction lemmas.

**Lemma 6.1** (Reduction Lemma for  $r > 3$ ). *Suppose  $b, c \geq 1$  and set  $b' = \min\{b, c\}$ . If  $(r-2)^{m-2} \geq 2(\max\{b, c\} - 1)$ , then*

$$\text{forb}(m, r, F(a, b, c, d)) = \text{forb}(m, r, F(a, b', b', d)).$$

**Proof:** If  $b = c$  the statement is trivial, so suppose without loss of generality that  $b < c$ . We set  $F := F(a, b, c, d)$  and  $F' = F(a, b, b, d)$ . It follows from  $F' \prec F$  that  $\text{forb}(m, r, F') \leq \text{forb}(m, r, F)$ . To prove the reverse inequality, we want to show that  $|A| \leq \text{forb}(m, r, F')$  for every  $A \in \text{Avoid}(m, r, F)$ . This is true if  $A$  avoids  $F'$ , so suppose instead that  $F' \prec A$ . By permuting the rows of  $A$ , we may assume that some instance of  $F'$  appears in its first two rows. We write  $A$  in the block form

$$A = \begin{bmatrix} 0 & 0 & 1 & \cdots & r-1 \\ 0 & 1 & 0 & \cdots & r-1 \\ A_{0,0} & A_{0,1} & A_{1,0} & \cdots & A_{r-1,r-1} \end{bmatrix}. \quad (6.2)$$

Because  $F'$  appears in the first two rows, we know that  $|A_{0,0}| \geq a$ , that  $|A_{0,1}|, |A_{1,0}| \geq b$ , and that  $|A_{1,1}| \geq d$ . If either of  $A_{0,1}$  or  $A_{1,0}$  contains at least  $c$  columns, then  $A$  contains  $F$  in the first two rows. But  $A$  avoids  $F$ , so  $|A_{1,0}|, |A_{0,1}| < c$ . We assumed that  $(r-2)^{m-2} \geq 2(c-1)$ , so it is possible to delete the columns with  $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$  or  $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$  in the first two rows and append  $|A_{0,1}| + |A_{1,0}|$  distinct columns  $c$  to  $A$  with  $c_1 = 0, c_2 = 1$ , and  $c_i \notin \{0, 1\}$  for  $i > 2$ . The resulting configuration does not contain  $I_2$  in its first two rows, so it does not contain  $F'$  in the first two rows, either. Moreover, this operation does not create a new instance of  $F'$  in  $A$ .

Iterating this process for every appearance of  $F'$  in  $A$  produces a matrix with the same number of columns as  $A$  that avoids  $F'$ . Thus  $|A| \leq \text{forb}(m, r, F')$ , as desired.  $\square$

**Theorem 6.2** (Forbidden numbers for 2-rowed  $(0, 1)$ -matrices with  $r > 3$ ). *Let  $F = F(a, b, c, d)$  and  $\alpha = \max\{a, d, \min\{b, c\}\}$ , and suppose  $(r-2)^{m-2} \geq 2 \max\{a, b, c, d\}$ . If  $\alpha > 0$ , then*

$$\text{forb}(m, r, F(a, b, c, d)) = m(r-1)^{m-1} + (r-1)^m + (\alpha-1) \binom{m}{2}. \quad (6.3)$$

Otherwise,  $F = p \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix}$  and

$$\text{forb}(m, r, p \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix}) = 2(r-1)^m - (r-2)^m + (p-1)m(m-1). \quad (6.4)$$

**Proof:** The case  $F = p \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix}$  is given by Corollary 5.2. We prove the statement for  $\alpha > 0$  in cases.

*Case 1:*  $\alpha = a$  and  $b, c \geq 1$ . By Lemma 6.1, we may assume that  $b = c$ . Then  $\mathbf{0}_{2 \times a} \prec F$  and  $F$  is  $a$ -simple, so the statement follows from Corollary 3.4. Taking the  $(0, 1)$ -complement of  $F$  handles the case  $\alpha = d$  with  $b, c \geq 1$ .

*Case 2:*  $\alpha = \min\{b, c\}$ . This implies  $b, c \geq 1$ , so by Lemma 6.1, we may assume  $b = c$ . Then  $c \cdot I_2 \prec F \prec c \cdot K_2$ , so the upper bound follows from Proposition 3.1 and the lower bound from Corollary 4.1 with  $k = 2$ .

*Case 3:*  $b = 0$  or  $c = 0$ . By possibly taking the complement, we may assume that  $b = 0$ . Since the arguments are symmetric, suppose  $a \geq d$ , which implies that  $\alpha = a$ . Then  $F \prec F(a, 1, \max\{1, c\}, d)$ , so by Lemma 6.1,

$$\text{forb}(m, r, a \cdot \mathbf{0}_2) \leq \text{forb}(m, r, F) \leq \text{forb}(m, r, F(a, 1, 1, d)).$$

The lower and upper bounds are equal by Corollary 3.4 and Case 1.  $\square$

Proving a reduction lemma for  $r = 3$  requires a different approach.

**Lemma 6.3** (Reduction Lemma for  $r = 3$ ). *Let  $b' = \min\{b, c\}$ . If  $2^{m-2} \geq (\max\{a, b, c, d\} - 1)m^2$  and  $b' \geq 1$ , then*

$$\text{forb}(m, 3, F(a, b, c, d)) = \text{forb}(m, 3, F(a, b', b', d)).$$

**Proof:** Let  $p = \max\{a, b, c, d\}$ , so that  $F := F(a, b, c, d)$  is  $p$ -simple, and set  $F' = F(a, b', b', d)$ . As above,  $\text{forb}(m, r, F') \leq \text{forb}(m, r, F)$  follows from the observation that  $F' \prec F$ .

Now let  $A \in \text{Avoid}(m, 3, F)$ . We want to show that  $|A| \leq \text{forb}(m, r, F')$ . If  $A$  does not contain  $F'$ , this is clear, so we assume that  $F' \prec A$ . We write  $A$  in block form as

$$A = \left[ \begin{array}{cc|cccccc} 0 & 1 & 0 & 1 & 0 & 1 & 2 & 2 & 2 \\ 1 & 0 & 0 & 1 & 2 & 2 & 0 & 1 & 2 \\ A_{0,1} & A_{1,0} & B & C & D & E & F & G & H \end{array} \right]. \quad (6.5)$$

By possibly taking the complement of  $F$ , we may assume that  $b \leq c$ . Moreover, since the statement is trivial if  $b = c$ , we assume that strict inequality holds. Since  $F' \prec A$  but  $F \not\prec A$ , we have that  $b \leq |A_{0,1}|, |A_{1,0}| < c$ . If there is a column in  $B$  that is not in  $D$ , then we may delete the column  $\begin{bmatrix} 0 \\ 0 \\ v \end{bmatrix}$  and insert the column  $\begin{bmatrix} 0 \\ 2 \\ v \end{bmatrix}$  without introducing  $F$  as a configuration. By replacing binary digits in the first two rows with 2's in this manner, we may assume that

$$\begin{array}{l} B \subseteq D \subseteq H \\ B \subseteq F \subseteq H \end{array} \quad \text{and} \quad \begin{array}{l} C \subseteq E \subseteq H \\ C \subseteq G \subseteq H. \end{array}$$

The matrices  $\begin{bmatrix} 0 & 1 & 2 \\ D & E & H \end{bmatrix}$  and  $\begin{bmatrix} 0 & 1 & 2 \\ F & G & H \end{bmatrix}$  both avoid  $F$ , so  $|DEH| + |FGH| \leq 2 \text{forb}(m-1, 3, F)$ . Also,  $B \cup C \subseteq H$ . From inclusion-exclusion,  $|B| + |C| - |H| \leq |B \cap C|$ . Because  $A$  avoids  $p \cdot K_2$ , we know that  $B \cap C$  avoids  $p \cdot K_1$ , so  $|B \cap C| \leq 2^{m-2} + (p-1)(m-2)$ . Therefore

$$|A| \leq 2(c-1) + 2 \text{forb}(m-1, 3, F) + 2^{m-2} + (p-1)(m-2).$$

If  $2^{m-2} \geq (p-1)m^2$ , then Proposition 3.1 implies

$$\begin{aligned} |A| &\leq 2(p-1) + 2 \operatorname{forb}(m-1, 3, p \cdot K_2) + 2^{m-2} + (p-1)(m-2) \\ &= (m-1)2^{m-1} + 2^m + 2^{m-2} + (p-1)m(m-1) + (p-1)m \\ &\leq m2^{m-1} + 2^m = \operatorname{forb}(m, 3, I_2). \end{aligned}$$

Since  $I_2 \prec F'$ , this shows that  $|A| \leq \operatorname{forb}(m, 3, F')$ , completing the proof.  $\square$

**Theorem 6.4.** *Suppose  $\max\{a, d\} \geq \min\{b, c\}$ . If  $2^{m-2} \geq (\max\{a, b, c, d\} - 1)m^2$ , then*

$$\operatorname{forb}(m, 3, F(a, b, c, d)) = (r-1)^m + m(r-1)^{m-1} + (\max\{a, d\} - 1) \binom{m}{2}. \quad (6.6)$$

**Proof:** Since all arguments are symmetric, we assume that  $a \leq d$  and  $b \leq c$ . If  $b \geq 1$ , then

$$d \cdot \mathbf{1}_2 \prec F(a, b, b, d) \prec d \cdot K_2.$$

The forbidden numbers of both bounding configurations are equal by Corollary 3.4, and applying Lemma 6.3 shows that  $F(a, b, b, d)$  and  $F(a, b, c, d)$  have the same forbidden number. If  $b = 0$  then  $F(a, 0, c, d) \prec F(a, 1, \max\{1, c\}, d)$ . We have  $d \cdot \mathbf{1}_2 \prec F(a, 0, c, d)$ , and by Lemma 6.3,

$$\operatorname{forb}(m, 3, F(a, 1, \max\{1, c\}, d)) = \operatorname{forb}(m, 3, F(a, 1, 1, d)) \leq \operatorname{forb}(m, 3, d \cdot K_2).$$

Again the upper and lower bounds are equal by Corollary 3.4.  $\square$

The remaining question is to evaluate  $\operatorname{forb}(m, 3, F(a, b, c, d))$  when  $\min\{b, c\} > \max\{a, d\}$ . By the Reduction Lemma, we need only consider the case  $b = c$ . The smallest 2-rowed  $(0, 1)$ -matrix whose forbidden number is not known when  $r = 3$  is

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

## 7 Forbidden configurations with 3 rows

Lemma 2.3 provides a handful of results on 3-rowed forbidden matrices for free.

**Corollary 7.1.** *The following forbidden numbers are exact for all  $r \geq 2$  when  $m \geq 3$ .*

$F$	$\operatorname{forb}(m, r, F)$
$\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$ or $\begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$	$m(r-1)^{m-1} + 2(r-1)^m - (r-2)^m - m(r-2)^{m-1}$
$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$ or $\begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$	$2m(r-1)^{m-1} + (r-2)^m.$

**Proof:** Theorem 3.2 of [4] proves that the configuration  $A_m = [0_m I_m \mathbf{1}_m]$  is extremal for the second matrix in the first row of the table when  $r = 2$ , and it is not too hard to see that  $A_m$  is extremal for the

first matrix, as well. The sequence  $(A_m)$  satisfies the conditions of Lemma 2.3, and  $|A_m| = m + 1$  if  $m \in \{0, 1\}$  and  $|A_m| = m + 2$  otherwise, so for either matrix  $F$ , we have

$$\text{forb}(m, r, F) = \sum_{k=0}^m \binom{m}{k} (r-2)^{m-k} \text{forb}(k, F);$$

simplifying this sum yields the expression in the first row of the table.

Let  $U_m$  denote the  $m \times m$  upper-triangular matrix with 1's on and above the diagonal and 0's elsewhere. Theorem 3.3 of [4] proves that  $A_m = U_m \cup I_m^c \cup \mathbf{0}_m$  is an extremal configuration (when  $r = 2$ ) for both matrices in the second row of the table. Both  $U_m$  and  $I_m^c$  have a column with exactly one 0 and are otherwise disjoint, so  $|A_m| = 2m$  for  $m \geq 1$ , and  $|A_0| = 1$  by convention. As before,  $(A_m)$  satisfies the conditions of Lemma 2.3, so for either forbidden matrix  $F$ ,

$$\begin{aligned} \text{forb}(m, r, F) &= \sum_{k=0}^m \binom{m}{k} (r-2)^{m-k} \text{forb}(k, F) \\ &= \sum_{k=0}^m \binom{m}{k} (r-2)^{m-k} 2k + (r-2)^m \\ &= 2m(r-1)^{m-1} + (r-2)^m, \end{aligned}$$

as claimed.  $\square$

The matrices  $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$  and  $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$  are sandwiched between the two matrices in the second row of Corollary 7.1, so they have the same forbidden number. Our results, together with  $(0, 1)$ -complementation, evaluate the exact forbidden number for all  $3 \times 2$  and  $3 \times 3$  simple matrices for  $r \geq 3$  except  $F = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}$ . We can, however, bound  $\text{forb}(m, r, F)$  using our previous results. Theorem 3.3 in [2] states that  $\text{forb}(m, F) \leq \frac{3}{2}m + 1$ ; applications of our Lemma 2.1 and Proposition 3.3 show that

$$m(r-1)^{m-1} + (r-1)^m \leq \text{forb}(m, r, F) \leq \frac{3}{2}m(r-1)^{m-1} + (r-1)^m.$$

## References

- [1] Noga Alon, *On the density of sets of vectors*, Discrete Mathematics **46** (1983), no. 2, 199–202.
- [2] Richard Anstee, *Some problems concerning forbidden configurations*, preprint.
- [3] Richard Anstee, *A survey of forbidden configuration results*, The Electronic Journal of Combinatorics **20** (2013), no. 1.
- [4] Richard Anstee, Jerrold Griggs, and Attila Sali, *Small forbidden configurations*, Graphs and Combinatorics **13** (1997), 97–118.
- [5] Richard Anstee and Linyuan Lu, *Unavoidable multicoloured families of configurations*, 2014, arXiv:1409.4123.

- [6] Keaton Ellis, Baian Liu, and Attila Sali, *Multi-symbol forbidden configurations*, *Discrete Applied Mathematics* **276** (2020), 24 – 36.
- [7] Zoltán Füredi, Alexander Kostochka, and Ruth Luo, *Berge cycles in non-uniform hypergraphs*, 2020, arXiv:2002.01597.
- [8] Zoltán Füredi and F. Quinn, *Traces of finite sets*, *Ars Combinatoria* **18** (1984), 195–200.
- [9] Zoltán Füredi and Attila Sali, *Optimal multivalued shattering*, *SIAM Journal on Discrete Mathematics* **26** (2012), 737–744.
- [10] Norbert Sauer, *On the density of families of sets*, *Journal of Combinatorial Theory, Series A* **13** (1972), 145–147.
- [11] Saharon Shelah, *A combinatorial problem: Stability and order for models and theories in infinitary language*, *Pacific Journal of Mathematics* **41** (1972), 247–261.
- [12] Vladimir Vapnik and Alexey Chervonenkis, *On the uniform convergence of relative frequencies of events to their probabilities*, *Theory of Probability and Its Applications* **16** (1971), 264–280.