

Families of Forbidden Configurations

by

Christina Louise Koch

B.A., St. Olaf College, 2005

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

The Faculty of Graduate Studies
(Mathematics)

THE UNIVERSITY OF BRITISH COLUMBIA
(Vancouver)

April 2013

© Christina Louise Koch 2013

Abstract

The forbidden configuration problem arises from a question in extremal set theory. The question seeks a bound on the maximum number of subsets of an m -set given that some trace is forbidden. In terms of hypergraphs, we seek the maximum number of edges on a simple hypergraph of m vertices such that this hypergraph does not contain a forbidden sub-hypergraph.

We will use the notation of matrices to describe the problem as follows. We call a $(0, 1)$ -matrix simple if it has no repeated columns; this will be the analogue of a simple hypergraph. Let F be a given matrix. We say that a $(0,1)$ -matrix A contains F as a configuration if there is some submatrix of A that is a row and column permutation of F . This is equivalent to a hypergraph containing some sub-hypergraph. F need not be simple. We define $\text{forb}(m, F)$ as the maximum number of columns possible for a simple, m -rowed, $(0,1)$ -matrix that does not contain F as a configuration.

A variation of the forbidden configuration problem forbids a family of configurations $\mathcal{F} = \{F_1, F_2, \dots, F_n\}$ instead of a single configuration F . Thus, $\text{forb}(m, \mathcal{F})$ becomes the maximum number of columns possible for a simple, m -rowed, $(0, 1)$ -matrix that does not contain any one of the $F_i \in \mathcal{F}$ as a configuration.

We will present a series of results organized by the character of $\text{forb}(m, \mathcal{F})$. These include a classification of families \mathcal{F} such that $\text{forb}(m, \mathcal{F})$ is a constant and the unexpected result that for a certain family, $\text{forb}(m, \mathcal{F})$ is on the order of $m^{3/2}$, where previous results typically had $\text{forb}(m, F)$ on the order of integer powers of m . We will conclude with a case study of families of forbidden configurations and suggestions for future work.

Preface

The question of forbidden configurations was introduced to me by my supervisor Dr. Richard Anstee, and the choice to explore families of forbidden configurations came at both his suggestion and that of his previous PhD student Dr. Miguel Raggi.

There are two sections of the paper that contain work done by other people. The first is a result of Balin Fleming included in Section 2.4, which is used with his permission and clearly identified as his work. The second is the proof of the result in Chapter 4, which borrows from the proof for a very similar bound originally included in a paper by Dr. Anstee, Dr. Attila Sali, and Dr. Raggi. However, the necessary modifications to the proof for the new result are the original work of Dr. Anstee and myself.

All other results contained in this paper were produced by myself, some jointly with Dr. Anstee.

Contents

Abstract	ii
Preface	iii
Table of Contents	iv
List of Tables	vi
Acknowledgements	vii
1 Introduction	1
1.1 Extremal problems and hypergraphs	1
1.2 Definitions and notation	4
1.2.1 (0,1)-matrices	4
1.2.2 Forbidden configurations and bounds	5
1.2.3 Matrix constructions	6
1.2.4 Set system terminology	9
1.3 History of results	9
1.3.1 Single forbidden configurations	10
1.3.2 Families of forbidden configurations	11
1.3.3 Elementary remarks and lemmas	13
1.4 Basic methods	14
1.4.1 Standard induction	14
1.4.2 What is missing	15
1.4.3 Linear algebra reduction	16
2 Families With a Constant Bound	18
2.1 A classification for families of constant bound	18

2.1.1	A constant construction	19
2.2	Failed monotonicity of the constant bound	20
2.3	Exact bound for blocks of 0's and 1's	21
2.3.1	Column bound	22
2.3.2	A sharp boundary	23
2.3.3	Other bounds and constructions	27
2.4	Constant bound for another family	28
3	Families With a Linear Bound	31
3.1	Families of minimal quadratic configurations	31
3.1.1	Families proved using basic methods	33
3.1.2	Families proved using a structure argument	36
3.2	Other linear results	39
4	Families With a Super-linear Bound	40
4.1	Unexpected bound for product configurations	40
4.1.1	Introduction	40
4.1.2	Theorem and proof	41
4.2	A similar bound for larger product configurations	51
5	A Case Study: Families of 2 x 2 Configurations	55
5.1	Constant families	56
5.2	Linear families	58
5.3	Table of results	60
5.4	Critical subfamilies	60
6	Conclusion	62
6.1	Summary of results	62
6.2	Future work	63
	Bibliography	65

List of Tables

3.1	Minimal Quadratic Configurations	32
3.2	Pairs of Minimal Quadratic Configurations	33
5.1	List of 2 x 2 Configurations	55
5.2	Summary of Results	60

Acknowledgements

First, an enormous thank you goes to my supervisor, Richard Anstee. I am deeply grateful for his generosity which includes but is not limited to: large blocks of his time, collaboration on proofs, brief lectures on linear programming, advice on teaching, digressions into antique collecting, gentle reminders to keep writing up results, and of course, numerous cups of tea and coffee.

An additional thanks to Miguel Raggi and Attila Sali for agreeing to modify a previously submitted paper to include one of my results. Also, thanks to Balin Fleming for permission to include one of his unpublished results. A special thanks to Jozsef Solymosi for serving as second reader for this thesis.

This degree wouldn't have been nearly as enjoyable as it was without the faculty, staff, and students of the math department. Thanks to all my professors, the office staff, and especially my graduate colleagues with whom I've facilitated workshops, tutored students, graded exams, written a wiki, and discussed many an assignment.

Profound thanks to the friends both near and far who have supported me over the past two years: the Songbirds, the Bon House girls, everyone at the Hope, my house mates, my community group, other Regent and St. Olaf friends, my First Lutheran, First Trinity and Fairview families . . . it's truly a wonderful thing to have so many people who believe in me.

And finally, thank you to my family.

Chapter 1

Introduction

In this chapter we provide an introduction to the topic of this paper: the question of forbidden configurations. We begin with the origin of the research question and its statement. We then provide a list of all necessary notation and definitions for understanding the content of the paper, including a re-statement of the research question. This leads to a section summarizing a history of relevant results. We conclude with an overview of the basic methods that will be used to prove results.

1.1 Extremal problems and hypergraphs

At its heart, combinatorics is a field of mathematics that asks “how many?” This initial question is often modified as follows: given some condition to be met, what is the maximum (or minimum) family such that the condition holds, or conversely, that the condition fails? In extremal combinatorics we are not always able to obtain exact maximum or minimum bounds, and in these cases seek an asymptotic bound. Our combinatorial question for this paper relates to similar questions from graph theory, so we will begin with the following basic definitions.

Definition 1.1.1. *A simple graph G is an ordered pair $G = (V, E)$ where V stands for a set of vertices and E stands for a set of edges, where edges are unique 2-element subsets of V .*

Definition 1.1.2. *The complete graph on n vertices, denoted K_n , is a graph on n vertices such that every possible edge is included in E .*

Forbidden configurations, the subject of this paper, relate to theorems in extremal graph theory. These theorems address the following combinatorial question: supposing that a graph G on n vertices has no subgraph isomorphic to a graph H , what is the maximum number of edges permissible for G ? One can also ask the dual question: given n vertices what is the minimum number of edges in G such that you are guaranteed a copy of the subgraph H ?

The first theorem, by Turán (and Mantel for $k = 3$), answers this question for H as the complete graph K_k as follows:

Theorem 1.1.3. [17] *Let $G = (V, E)$ be a simple graph on n vertices with the property that there is no triangle. Then $|E| \leq \frac{n^2}{4}$.*

Theorem 1.1.4. [22] *Let G be any simple graph on n vertices. Then the minimum number of edges in G such that G must contain a subgraph isomorphic to K_k is bounded above by*

$$\frac{k-2}{k-1} \cdot \frac{n^2}{2} \text{ or } \left(1 - \frac{1}{k-1}\right) \cdot \frac{n^2}{2}$$

Erdős and Simonovits extended a previous result of Erdős and Stone to present a limit bound when forbidding any subgraph H . It uses the chromatic number of the subgraph H .

Definition 1.1.5. *The chromatic number of a graph G , or $\chi(G)$, is the smallest natural number c such that the vertices of G can be colored with c colors and no two vertices of the same color share an edge.*

Theorem 1.1.6. [12, 13] *Let n, H be given. We define $ex(n, H)$ as the maximum number of edges in an n vertex (simple) graph that has no subgraph isomorphic to H . Then for all $\epsilon > 0$, there exists some N such that for all $n > N$:*

$$\frac{n^2}{2} \left(1 - \frac{1}{1 - \chi(H)} - \epsilon\right) \leq ex(n, H) \leq \frac{n^2}{2} \left(1 - \frac{1}{1 - \chi(H)} + \epsilon\right).$$

Not only does this theorem provide an asymptotic bound for all forbidden subgraphs, H , it also connects the idea of a forbidden subgraph with the chromatic

number of that subgraph. We will conjecture a similar notion is helpful when looking at forbidden configurations.

We will now generalize this problem of a forbidden subgraph to hypergraphs. We begin with our original definition of a graph G as an ordered pair $G = (V, E)$ and generalize this definition to a hypergraph.

Definition 1.1.7. *Let X be a finite set, with 2^X denoting the set of all subsets of X , or the power set of X .*

Definition 1.1.8. *A hypergraph is an ordered pair $H = (X, \mathcal{E})$, where X is a finite set of elements called vertices, \mathcal{E} is a multiset of sets in 2^X , that is, for each $E \in \mathcal{E}$, we have $E \in 2^X$. The elements of \mathcal{E} are called edges. The hypergraph H is called simple if the elements $E \in \mathcal{E}$ are unique. i.e. \mathcal{E} is a set.*

Definition 1.1.9. *A set system \mathcal{A} on X is a family of subsets of X , that is, $\mathcal{A} \subseteq 2^X$.*

It should be clear that the set system \mathcal{A} is equivalent to the simple hypergraph $H = (X, \mathcal{E})$ if \mathcal{A} is defined on X and for all $E \in \mathcal{E}$, $E \in \mathcal{A}$. A simple graph as defined in Definition 1.1.1 is equivalent to a simple hypergraph where all edges have size 2.

There is some disagreement over whether the edge set \mathcal{E} can include the empty edge, where the empty edge is the set $\emptyset \in 2^X$. In the understanding of a hypergraph as a set system, inclusion of the empty set is sensible; however, there is value to avoiding the empty edge when considering the transversal of a hypergraph. In the next section, when we recast the forbidden subgraph problem in matrix notation, we will include the empty edge as a viable edge for the hypergraph.

Definition 1.1.10. *A sub-hypergraph is an ordered pair (S, \mathcal{E}_S) , where, given a hypergraph $H = (X, \mathcal{E})$, $S \subseteq X$ and \mathcal{E}_S consists of all sets $E \in \mathcal{E}$ restricted to the elements in S .*

Note that the sub-hypergraph of a simple hypergraph may not itself be simple; two edges that both contain the same subset of vertices will be the same if they are restricted to that subset. Also, as in the case of the hypergraph, it is possible for the empty edge to be included in the edge set of the sub-hypergraph.

Hypergraphs are a generalization of the initial notion of a graph; it is thus natural to ask if the combinatorial question motivating Theorem 1.1.4 and Theorem 1.1.6 can be applied to hypergraphs. For example, given a hypergraph $H = (X, \mathcal{E})$, where

$|X| = m$, what is the maximum possible size of \mathcal{E} so that H does not contain a sub-hypergraph F ?

This is precisely the question founding the study of forbidden configurations. In the rest of this paper, we will approach this question using the language of matrices.

1.2 Definitions and notation

1.2.1 (0,1)-matrices

Recall that a simple hypergraph $H = (X, \mathcal{E})$ can be thought as a set system \mathcal{A} on X . We will now translate a set system \mathcal{A} to a $(0, 1)$ -matrix A .

Let \mathcal{A} be a set system defined on the set X . Let $|X| = m$ and $|\mathcal{A}| = n$. For each set $B \in \mathcal{A}$, we generate an incidence vector β , where β is a single column with m entries. As an incidence vector, β has a 1 in a row i if $i \in B$ and 0 otherwise. We can thus describe all our sets in \mathcal{A} as incidence vectors. We join these vectors together side by side to form the $m \times n$ $(0, 1)$ -matrix A . The elements of X index the rows of A and the sets $B \in \mathcal{A}$ index the columns. Note that row and column order is not so important to the combinatorial object and so when we refer to such a matrix A we are actually thinking of the equivalence class of matrices that could be obtained from A by row and column permutations.

We say a matrix A is *simple* if it is a $(0, 1)$ -matrix with no repeated columns. A simple matrix A can be written as a set system \mathcal{A} by reversing the process described above: if the rows of A are indexed by the elements of the set X , each column β can be interpreted as a set $B \in \mathcal{A} \subseteq 2^X$, where the elements of X included in B are those with a 1 in the corresponding row in β .

We use the notation $[m]$ to mean the set $\{1, 2, 3, \dots, m\}$. We will often use $[m]$ to index the rows of A .

We use the notation $\|A\|$ to describe the number of columns in a matrix A .

Let a set S be a subset of the rows of a matrix A . We use the notation $A|_S$ to describe the matrix A *restricted to* S , that is, the submatrix of A consisting of the rows indexed by elements of S .

Given a $(0, 1)$ -matrix A , the $(0, 1)$ -*complement* of the matrix, A^c , is the matrix that results from replacing all the 1's in A with 0's and all the 0's with 1's.

1.2.2 Forbidden configurations and bounds

In our matrix interpretation of hypergraphs, a sub-hypergraph becomes a configuration.

We define that A *contains* F as a *configuration*, written $F \prec A$, if there is a submatrix of A that is a row and column permutation of F . For example, let

$$F = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \text{ and } A = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}.$$

Then $F \prec A$. We see this by looking at the 2×2 submatrix of A given by rows 1 and 3 and columns 3 and 4.

On the other hand, we say that A *avoids* F , or $F \not\prec A$ if there is no submatrix in A that is a row and column permutation of F . For example, let

$$F = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \text{ and } A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Then $F \not\prec A$.

If A is a simple $(0,1)$ -matrix with m rows that avoids F , we use the following notation: $A \in \text{Avoid}(m, F)$. So:

$$\text{Avoid}(m, F) = \{A : A \text{ simple, } m\text{-rowed matrix; } F \not\prec A\}.$$

The primary question of forbidden configurations concerns the following function:

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

In other words, $\text{forb}(m, F)$ is the largest value n such that there exists an $m \times n$ simple $(0,1)$ -matrix A where $A \in \text{Avoid}(m, F)$. This is simply a restatement of the question from Section 1.1 translated into a matrix representation: given a simple $(0,1)$ -matrix A on m rows, what is the maximum number of columns in A so that A does not contain a configuration F ?

For this paper, we intend to modify the forbidden configuration question slightly by forbidding a family of configurations instead of a single configuration.

A *forbidden family* is a finite set of forbidden configurations, denoted $\mathcal{F} = \{F_1, F_2, \dots, F_t\}$.

We define

$$\text{Avoid}(m, \mathcal{F}) = \{A : A \text{ simple, } m\text{-rowed matrix; } F_i \not\prec A \text{ for } i = 1, 2, \dots, t\}.$$

Our previous definition of $\text{forb}(m, F)$ naturally generalizes to:

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

So $\text{forb}(m, \mathcal{F})$ is the largest value n such that there exists an $m \times n$ simple $(0, 1)$ -matrix A where $A \in \text{Avoid}(m, \mathcal{F})$.

In order to find $\text{forb}(m, \mathcal{F})$, there are often separate arguments to establish a lower bound and upper bound. The lower bound is frequently established by a construction of appropriate size that avoids \mathcal{F} , while the upper bound is proved separately. Certain standard techniques for establishing an upper bound will be discussed in Section 1.4.

Often it is not possible to determine exact bounds for the columns of A . In this case, we must be satisfied with determining asymptotic upper and lower bounds, expressed in big-O notation.

- $\text{forb}(m, \mathcal{F})$ is $O(m^k)$ if there exists $c \in \mathbb{R}$ such that for sufficiently large m $\text{forb}(m, \mathcal{F}) \leq cm^k$.
- $\text{forb}(m, \mathcal{F})$ is $\Omega(m^k)$ if there exists $c \in \mathbb{R}$ such that for sufficiently large m $\text{forb}(m, \mathcal{F}) \geq cm^k$.
- $\text{forb}(m, \mathcal{F})$ is $\Theta(m^k)$ if there exists $c_1, c_2 \in \mathbb{R}$ such that for sufficiently large m , we have $c_1 m^k \leq \text{forb}(m, \mathcal{F}) \leq c_2 m^k$.

1.2.3 Matrix constructions

We will use the following notation to describe various matrix configurations and constructions.

We define $\mathbf{1}_k$ to be the column of k 1's and $\mathbf{0}_k$ to be the column of k 0's.

We say that a $(0, 1)$ -column α has *column sum* s when it contains s 1's.

We define K_k as the *complete object* on k rows, that is, the $k \times 2^k$ simple matrix of all possible columns on k rows. We define K_k^s as the $k \times \binom{k}{s}$ simple matrix of all

columns of column sum s on a matrix of k rows. For example:

$$K_4^2 = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix}.$$

A *cycle of length k* is a graph on k vertices with edges of size 2, such that we can order the vertices x_1, x_2, \dots, x_k and we have $k-1$ edges (x_j, x_{j+1}) for $j = 1, 2, \dots, k-1$ and the edge (x_k, x_1) . We will define C_k to be the vertex-edge incidence matrix of the cycle of length k , where the vertices of the cycle index the rows and the columns indicate which vertices are connected by an edge. Note that the matrix C_k will have column sum 2 for all columns and row sum 2 for all rows. For example,

$$C_4 = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}.$$

In our configurations and constructions, we will frequently see the following three matrices:

The *identity matrix* on m rows, or I_m or K_m^1 , consists of an $m \times m$ matrix with 1's on the diagonal and 0's everywhere else. It corresponds to the set system of singleton sets on m elements. For example,

$$I_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

The *identity complement matrix* on m rows, or I_m^c or K_m^{m-1} , is the $(0, 1)$ -complement of I_m , or an $m \times m$ matrix with 0's on the diagonal and 1's everywhere else. It cor-

responds to a set system of sets of size $m - 1$ on m elements. For example,

$$I_4^c = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}.$$

The *upper triangular matrix* on m rows, or T_m , consists of an $m \times m$ matrix with columns of strictly increasing column sum, with 1's justified towards the top of the matrix. It corresponds to the set system $\{\{1\}, \{1, 2\}, \{1, 2, 3\}, \dots, \{1, 2, \dots, m\}\}$. For example,

$$T_4 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

We will also refer to this matrix as a *tower* matrix.

We use the notation $[A|B]$ to denote the matrix obtained from concatenating the two matrices A and B . We use the notation $k \cdot A$ to denote the matrix $[A|A|\dots|A]$ consisting of k copies of A concatenated together. We give precedence to the operation \cdot (multiplication) over concatenation so that for example $[2 \cdot A|B]$ is the matrix consisting of the concatenation of B with the concatenation of two copies of A .

We will use the notation A^T to denote the *matrix transpose* of A .

If A is an $m_1 \times n_1$ simple matrix and B an $m_2 \times n_2$ simple matrix, we define the *product* $A \times B$ as the $(m_1 + m_2) \times (n_1 n_2)$ matrix of all possible combinations of columns of A over columns of B . Note that this is a different definition than the conventional notion of matrix multiplication. An example of $I_2 \times I_2$ is shown below.

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

It is interesting to note that after some row and column permutation, $I_2 \times I_2$ is the same matrix as the C_4 shown earlier.

1.2.4 Set system terminology

The following definitions appropriate terminology used for set systems (comparable, disjoint, laminar) and recast them in our matrix notation.

Let α, β be two columns of a $(0, 1)$ -matrix, with the entry in the i th row denoted α_i, β_i respectively. We say that α and β are *comparable*, with $\alpha \leq \beta$ (or $\alpha \geq \beta$) if for all rows i , $\alpha_i \leq \beta_i$ (or $\alpha_i \geq \beta_i$). We say that α and β are *disjoint* if on every row i , α_i and β_i are not both 1. For example, let

$$\alpha = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \beta = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \gamma = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

Then α and β are comparable, with $\beta \leq \alpha$; β and γ are disjoint; and α and γ are neither comparable nor disjoint.

We define a *laminar family* as the following: given a set system \mathcal{A} , for all $B, C \in \mathcal{A}$, $B \subseteq C$, $C \subseteq B$ or $B \cap C = \emptyset$. A *laminar matrix* is defined similarly: given a $(0, 1)$ -matrix A , for all columns $\alpha, \beta \in A$, α and β are comparable or α and β are disjoint. Another way to define a laminar matrix is by forbidding a configuration or observing what columns are missing on any triple of rows in a matrix A (see Section 1.4 for more detail on “what is missing”). Let

$$F_L = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad L_1 = \begin{matrix} i \\ j \\ k \end{matrix} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad L_2 = \begin{matrix} i \\ j \\ k \end{matrix} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

Remark 1.2.1. *Let A be a $(0, 1)$ -matrix with no configuration F_L . Then A is laminar.*

Remark 1.2.2. *Let A be a $(0, 1)$ -matrix such that on every triple of rows $S = \{a, b, c\}$ there is a bijection $\sigma : S \rightarrow \{i, j, k\}$ such that $A|_S$ has no L_1 and L_2 as indexed. Then $F_L \not\prec A$ and so A is laminar.*

1.3 History of results

Here we present a history of results for the forbidden configuration problem. Many of these results will be used to motivate or prove results given in the research chapters

of this paper.

1.3.1 Single forbidden configurations

The primary result in the study of forbidden configurations is the following theorem, proved independently by Sauer and Perles, Shelah, and Vapnik and Chervonenkis.

Theorem 1.3.1. [20, 21, 23] *We have that*

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

This implies that $\text{forb}(m, K_k)$ is $\Theta(m^{k-1})$.

Corollary 1.3.2. *For any simple configuration F on k rows, $\text{forb}(m, F)$ is $O(m^{k-1})$.*

Anstee and Füredi proved a generalization of Theorem 1.3.1 as follows:

Theorem 1.3.3. [5] *For fixed k, t , as $m \rightarrow \infty$,*

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

Corollary 1.3.4. *For any k -rowed $(0, 1)$ -matrix F , we have that $\text{forb}(m, F)$ is $O(m^k)$.*

These results establish a maximum upper bound for all single forbidden configurations: in the case of a simple k -rowed configuration, Corollary 1.3.2 gives a bound of at most $O(m^{k-1})$, while for all other k -rowed configurations, Theorem 1.3.3 gives a bound of at most $O(m^k)$.

When generating constructions to achieve a lower bound for $\text{forb}(m, F)$, three matrices frequently occur. These are the I (identity), I^c (identity complement), and T (upper triangular) matrix, or products of the three. The prevalence of these three matrices gave rise to the following conjecture of Anstee and Sali. They have established the conjecture for all 3-rowed F .

Conjecture 1.3.5. [8] *Let $X(F)$ be the smallest p such that $F \prec A_1 \times A_2 \times \cdots \times A_p$ for every A_i as $I_{m/p}, I_{m/p}^c$ or $T_{m/p}$. Then $\text{forb}(m, F) = \Theta(m^{X(F)-1})$.*

The notation $X(F)$ is a direct allusion to Theorem 1.1.6 which linked chromatic number of a graph (denoted $\chi(G)$) to bounds of this kind.

Here, Anstee and Sali conjecture that matrix products of the identity, identity complement, and triangular matrices provide sufficient constructions to achieve the asymptotic bounds for $\text{forb}(m, F)$ for any forbidden configuration F . The three matrices of Conjecture 1.3.5 will appear shortly in a result regarding forbidden families, indicating that while Conjecture 1.3.5 remains unproven, the use of the identity, identity complement and triangular matrix is somehow fundamental to solving questions about forbidden configurations.

1.3.2 Families of forbidden configurations

Because the focus of this paper will be on families of forbidden configurations, we now shift our attention to results that pertain specifically to forbidden families. For our first result, we present a theorem of Balogh and Bollobás

Theorem 1.3.6. [9] *Let k be given. Then $\text{forb}(m, \{I_k, I_k^c, T_k\})$ is $O(1)$.*

In other words, when we forbid all three matrices used for constructions in Conjecture 1.3.5, the resulting matrices $A \in \text{Avoid}(m, \{I_k, I_k^c, T_k\})$ all have a constant upper bound. We should point out that the bound increases dramatically with k and the asymptotics as a function of k are not fully understood. However, the following remark gives an easy lower bound.

Proposition 1.3.7. [11] *Let k be given. Then $\text{forb}(m, \{I_k, I_k^c, T_k\}) \geq \binom{2k-2}{k-1}$.*

We establish this result in Proposition 2.1.3.

We have two further results that provide more exact bounds for specific instances of families for which Theorem 1.3.6 would apply. The first set of bounds comes from Anstee and Dunwoody, and are exact bounds for the families of Theorem 1.3.6 for $k = 1$, $k = 2$ and $k = 3$.

Theorem 1.3.8. [11] *We have $\text{forb}(m, \{I_1, I_1^c, T_1\}) = 0$, $\text{forb}(m, \{I_2, I_2^c, T_2\}) = 2$, and $\text{forb}(m, \{I_3, I_3^c, T_3\}) = 6$.*

A result of Fleming provides a more precise (if not exact) upper bound for a family in which the matrices are all sub-matrices of those used in Theorem 1.3.6

Theorem 1.3.9. [14] Let $F_a = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} | t \cdot \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, $F_b = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} | t \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $F_c = t \cdot \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$.
If $t \geq 2$, $\text{forb}(m, \{F_a, F_b, F_c\}) \leq 6t - 6$.

Note that $F_a \prec I_{t+2}$, $F_b \prec I_{t+2}^c$ and $F_c \prec T_{3t}$

We will take up this result in greater detail in Section 2.4, while discussing forbidden families with a constant bound.

Once we forbid only two of the three matrices of Conjecture 1.3.5, a non-constant bound results, as the following three theorems, proved by Balogh, Keevash and Sudakov, indicate.

Theorem 1.3.10. [10] Let $k \geq 2$. Then $\text{forb}(m, \{I_k, I_k^c\})$ is $\Theta(m^{k-1})$.

Theorem 1.3.11. [10] Let $k \geq 2$. Then $\text{forb}(m, \{I_k^c, T_k\})$ is $\Theta(m^{k-1})$.

Theorem 1.3.12. [10] Let $k \geq 2$. Then $\text{forb}(m, \{I_k, T_k\})$ is $\Theta(m^{k-2})$.

Forbidding families of forbidden configurations also results in asymptotic bounds with non-integer exponents.

Theorem 1.3.13. [1] We have $\text{forb}(m, \{C_4, \mathbf{1}_3\})$ is $\Theta(m^{3/2})$.

This theorem is linked to another family with the same bound.

Theorem 1.3.14. [7] We have $\text{forb}(m, \{I_2 \times I_2, T_2 \times T_2, I_2 \times T_2\})$ is $\Theta(m^{3/2})$.

The previous result was in a preliminary version of [7] and in Section 4.1 we will prove a stronger result, Theorem 4.1.3.

Finally, we have a result regarding families of balanced and totally balanced matrices. The cycle-incidence matrices of the family $\{C_3, C_4, C_5, \dots\}$ are also called *balanced* matrices; the matrices of the family $\{C_3, C_5, C_7, \dots\}$ are called *totally balanced*. It turns out that the bound on these infinite families is dependent simply on the bound for C_3 , i.e.

Theorem 1.3.15. [2] We have

$$\text{forb}(m, C_3) = \text{forb}(m, \{C_3, C_5, C_7, \dots\}) = \binom{m}{2} + \binom{m}{1} + \binom{m}{0}.$$

Thus we have a series of results characterizing forbidden families that contain the three matrices of Conjecture 1.3.5, as well as some more precise constant bounds, and finally, a set of results relating to balanced and totally balanced matrices.

1.3.3 Elementary remarks and lemmas

In addition to the results described above, one can easily make a series of helpful remarks and lemmas regarding $\text{forb}(m, F)$ and $\text{forb}(m, \mathcal{F})$ simply by using the definitions from Section 1.2.

Remark 1.3.16. (Complementarity) *Assume $\mathcal{F} = \{F_1, F_2, \dots, F_t\}$ and $\mathcal{F}^c = \{F_1^c, F_2^c, \dots, F_t^c\}$. Then $\text{forb}(m, F_i) = \text{forb}(m, F_i^c)$ and $\text{forb}(m, \mathcal{F}) = \text{forb}(m, \mathcal{F}^c)$.*

Lemma 1.3.17. (Monotonicity) *Suppose F', F are configurations. If $F' \prec F$, then $\text{forb}(m, F') \leq \text{forb}(m, F)$.*

PROOF. Any matrix that contains F will also contain F' , and any matrix that avoids F' will also avoid F . The lemma follows. ■

Lemma 1.3.18. (Induced lower bound) *If $\mathcal{F}' \subseteq \mathcal{F}$, then $\text{forb}(m, \mathcal{F}') \geq \text{forb}(m, \mathcal{F})$.*

PROOF. If A avoids \mathcal{F} it will also avoid \mathcal{F}' , however there may be a better (larger) construction for \mathcal{F}' that does not avoid the configurations in $\mathcal{F} \setminus \mathcal{F}'$. ■

Note that this lemma is superficially the exact opposite of the previous lemma. In Lemma 1.3.17, a smaller configuration F' means a smaller bound $\text{forb}(m, F)$, while in Lemma 1.3.18 a smaller family \mathcal{F} means a larger bound $\text{forb}(m, \mathcal{F})$.

Lemma 1.3.19. (Induced upper bound) *If $\mathcal{F} = \{F_1, F_2, \dots, F_t\}$, then $\text{forb}(m, \mathcal{F}) \leq \min_{i \in [t]}(\text{forb}(m, F_i))$. Furthermore, $\text{forb}(m, \mathcal{F}) \leq \min_{\mathcal{F}' \subseteq \mathcal{F}}\{\text{forb}(m, \mathcal{F}')\}$.*

PROOF. Because each $F \in \mathcal{F}$ must be avoided, $\text{forb}(m, \mathcal{F})$ can be no larger than the smallest $\text{forb}(m, F)$ for all $F \in \mathcal{F}$ or the smallest $\text{forb}(m, \mathcal{F}')$ for all $\mathcal{F}' \subseteq \mathcal{F}$. ■

We conclude with a final lemma that, while easy to state and simple to prove, will be remarkably powerful, particularly in conjunction with Theorem 1.3.6 regarding forbidden families with a constant bound.

Lemma 1.3.20. *Let $\mathcal{F} = \{F_1, F_2, \dots, F_t\}$. Suppose \mathcal{F}' is a family such that for each $j \in [t]$, there exists $F' \in \mathcal{F}'$ such that $F' \prec F_j$. Then $\text{forb}(m, \mathcal{F}') \leq \text{forb}(m, \mathcal{F})$.*

PROOF. Let \mathcal{F} and \mathcal{F}' be as stated above. Suppose $\text{forb}(m, \mathcal{F}') > \text{forb}(m, \mathcal{F})$. This implies that there is a construction A of size $m \times \text{forb}(m, \mathcal{F}')$ that avoids \mathcal{F}' . For any $F_i \in \mathcal{F}$, if there is $F' \prec \mathcal{F}'$, such that $F' \prec F_i$, then $F_i \not\prec A$. However, by our hypothesis, this is true for all $F_i \in \mathcal{F}$. Thus A avoids \mathcal{F} and $\text{forb}(m, \mathcal{F}) \geq \text{forb}(m, \mathcal{F}')$, a contradiction.

1.4 Basic methods

1.4.1 Standard induction

One way to establish an upper bound on $\text{forb}(m, \mathcal{F})$ is by inducting on m , the number of rows. The name for this process is *standard induction*, and it is a common proof technique used in the field of forbidden configurations. As we will use it in several results, it is worth describing in some detail here.

Let $A \in \text{Avoid}(m, \mathcal{F})$, with $\|A\| = \text{forb}(m, \mathcal{F})$ for some forbidden family \mathcal{F} . Standard induction is based on the following decomposition: Choose a row r in A . Sort the columns of A based on their entry in column r . Now suppose we delete r . Certain columns in $A_{[m] \setminus r}$ may now be repeated. These repeated columns form the sub-matrix C_r . The non-repeated columns with a 0 in row r form the submatrix B_r . Non-repeated columns with a 1 in row r form the submatrix D_r . After this sorting of rows and columns, the decomposition appears as:

$$A = \begin{array}{c} r \\ [m] \setminus r \end{array} \begin{bmatrix} 0 \dots 0 & 1 \dots 1 \\ B_r & C_r & C_r & D_r \end{bmatrix}. \quad (1.1)$$

Note that B_r, C_r , and D_r are all simple matrices in their own right and the matrix $[B_r|C_r|D_r]$ is also simple. We also know that because $\mathcal{F} \not\prec A$, $\mathcal{F} \not\prec [B_r|C_r|D_r]$, so $[B_r|C_r|D_r] \in \text{Avoid}(m-1, \mathcal{F})$. However $\|A\| = \|C_r\| + \|[B_r|C_r|D_r]\|$. Thus if we can find an upper bound for $\|C_r\|$, the upper bound for $\|A\|$ arises from induction on the rows of A as follows:

$$\text{forb}(m, \mathcal{F}) \leq \|C_r\| + \text{forb}(m-1, \mathcal{F}). \quad (1.2)$$

In order to establish an upper bound on $\|C_r\|$, we will frequently use what we have termed *inductive children*, that is, sub-configurations that must be avoided in

C_r in order that $F_i \not\prec [01] \times C_r$ for all $F_i \in \mathcal{F}$. Consider the following example:

$$\mathcal{F} = \left\{ \left[\begin{array}{cccc} 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{array} \right], \left[\begin{array}{cc} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{array} \right] \right\}.$$

Inductive Children of F : $\left[\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \end{array} \right], \left[\begin{array}{ccc} 0 & 1 & 1 \\ 0 & 0 & 1 \end{array} \right], \left[\begin{array}{ccc} 1 & 0 & 1 \\ 0 & 1 & 1 \end{array} \right], \left[\begin{array}{c} 1 \\ 1 \\ 1 \end{array} \right], \left[\begin{array}{cc} 1 & 1 \\ 1 & 1 \end{array} \right].$

If either of the last two inductive children are present in C_r , the second configuration of our forbidden family will appear in the original matrix. Similarly, if any of the first three inductive children are present in C_r , we will have the first configuration of our forbidden family. If the forbidden family of inductive children is “simpler” than the original forbidden family, it will sometimes be easier to construct a bound or apply a previous result. Then by induction, a bound can be found for the original family.

1.4.2 What is missing

Another proof technique is that of considering “what is missing” or “in short supply”. For this technique, we have a configuration of k rows. On any k -tuple of rows we consider which k -rowed columns must be absent (missing), so that the configuration is avoided, or which k -rowed columns must have a maximum number of copies to avoid the configuration. For example, if

$$F = \left[\begin{array}{cccccc} 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{array} \right],$$

on any 2-tuple of rows, i, j either there are ‘absent’ columns:

$$\begin{array}{c} i \\ j \end{array} \begin{array}{c} no \\ \left[\begin{array}{c} 0 \\ 0 \end{array} \right] \end{array} \quad \text{or} \quad \begin{array}{c} i \\ j \end{array} \begin{array}{c} no \\ \left[\begin{array}{c} 1 \\ 1 \end{array} \right] \end{array},$$

or one column appears ‘in short supply,’ which in this case, appears at most once, namely either

$$\begin{matrix} i \\ j \end{matrix} \begin{matrix} \leq 1 \\ \left[\begin{matrix} 1 \\ 0 \end{matrix} \right] \end{matrix} .$$

or the same with the roles of i, j reversed.

If this is true for all pairs of rows in A then we can see that indeed $A \in \text{Avoid}(m, F)$. In general for a set of rows $S \in [m]$ with $|S| = k$ and for a $k \times 1$ $(0, 1)$ -column α , we say that α is in *short supply* on S (in A) if there is a small bound on the number of times α appears in $A|_S$. Note that here we are considering α as a submatrix of $A|_S$ not as a configuration; order matters. Also when α appears no times in $A|_S$ then we say α is *absent* on S .

One basic result that follows from a “what is missing?” analysis is as follows: given a matrix A , if for each k -tuple of rows from A , some $k \times 1$ column with k entries is missing, then A has no K_k and the bound from Corollary 1.3.2 applies, making $\|A\|$ be $O(m^{k-1})$.

1.4.3 Linear algebra reduction

In the previous section, we noted that a matrix with at least one $k \times 1$ column missing on each k -tuple of rows contains no K_k and is thus of size at most $O(m^{k-1})$. In this section we will present a method of Anstee and Fleming that reduces a matrix to this scenario by eliminating $O(m^{k-1})$ columns, making the bound for the total matrix also $O(m^{k-1})$ [4]. This method presupposes that we have already performed an analysis of what is missing or in short supply, and all k -tuples of rows are either missing a column or have at least two columns in short supply.

Theorem 1.4.1. *Let t be given. Let A be an m -rowed simple matrix such that for each $S \in \binom{[m]}{k}$, either $A|_S$ has one column absent or has two columns α_S, β_S which are in short supply, say appearing at most t times. Then $\|A\|$ is $O(m^{k-1})$.*

PROOF. Let \mathcal{S} be the set of all k -tuples of rows S such that there are at least two columns in short supply on S . Let $S \in \mathcal{S}$ with columns α and β its columns in short supply.

We will first define a multi-linear indicator polynomial g_S such that, for all columns γ in the matrix A , $g_S(\gamma)$ is non-zero only when $\gamma|_S = \alpha$ or $\gamma|_S = \beta$. We first define a

multi-linear expression that tracks each α and β on a k -tuple of rows, S . If each row r in A is assigned a variable x_r , this will be:

$$f_{S,\alpha} = \prod_{r \in S} (x_r - (\alpha)_r).$$

And similarly for β . We combine these expressions into our multi-linear polynomial:

$$g_S = f_{S,\alpha} - f_{S,\beta}$$

Note that the combination of leading terms for each $f_{S,\alpha}, f_{S,\beta}$ cancels out, leaving g_S with degree $k - 1$.

We choose pairs $S_i \in \binom{[m]}{k}$ and γ_i (a column of A) for $i = 1, 2, \dots, p$ such that $\gamma_i|_{S_i}$ is either α or β . In addition we assume that for all pairs $i < j$, that $\gamma_j|_{S_i} \neq \alpha, \beta$ and for any column γ of A we have that there is an i with $1 \leq i \leq p$, such that $\gamma|_{S_i} = \alpha$ or β . We can achieve this by greedily seeking such pairs S_i, γ_i until no more are possible.

We form indicator polynomials for each i namely $g_{S_i} = f_{S_i,\alpha} - f_{S_i,\beta}$ which are nonzero for any column γ that has $\gamma|_{S_i} = \alpha$ or β . We recall that the highest degree terms in $f_{S_i,\delta}$ are $\prod_{j \in S_i} x_j$ and so $g_{S_i} = f_{S_i,\alpha} - f_{S_i,\beta}$ has degree at most $k - 1$.

Our choice of pairs has for $i < j$, that $g_{S_i}(\gamma_j) = 0$ and $g_{S_j}(\gamma_j) \neq 0$. Thus the p polynomials g_{S_i} are linearly independent. But the polynomials lie in a space of dimension $\binom{m}{k-1} + \binom{m}{k-2} + \dots + \binom{m}{0}$ and so we have the wonderful conclusion that p is at most this dimension and so is $O(m^{k-1})$.

We now remove from A all columns γ which have for some i , $\gamma|_{S_i} = \alpha$ or β . There can be at most $2tp$ such columns which is $O(m^{k-1})$. Our resulting matrix A' has the property that for every $S \in \binom{[m]}{k}$, we have that $A'|_S$ has (at least) one absent column and so A' has no K_k . Thus $\|A'\|$ is $O(m^{k-1})$. The number of deleted columns is also $O(m^{k-1})$ and so $\|A\|$ is $O(m^{k-1})$. ■

Chapter 2

Families With a Constant Bound

In this chapter we will discuss results pertaining to forbidden families that have a constant bound. We will use a previously cited result of Balogh and Bollobás to classify all forbidden families that have a constant upper bound. We will also show that the complement of this set is all families with at least a linear lower bound. We will then briefly discuss how the constant bound for these forbidden families is not monotonic, that is, the constant bound is not weakly decreasing as the number of rows in the matrix A increases. Finally we will discuss exact bounds for two forbidden families of constant bound: both an exact bound for a specific forbidden family and a maximal upper bound for another forbidden family.

2.1 A classification for families of constant bound

In Section 1.2, we stated an important result of Balogh and Bollobás regarding forbidden families, Theorem 1.3.6. This theorem states that for each $k \in \mathbb{N}$, forbidding the family $\mathcal{F} = \{I_k, I_k^c, T_k\}$ will yield a constant bound on $\text{forb}(m, \mathcal{F})$. We will now prove that this theorem in fact characterizes all forbidden families of constant bound.

Theorem 2.1.1. *Let \mathcal{F} be a finite forbidden family. Then $\text{forb}(m, \mathcal{F})$ is $O(1)$ if and only if there exists $F_i, F_j, F_\ell \in \mathcal{F}$ (not necessarily distinct configurations) such that for some value $k \in \mathbb{N}$, $F_i \prec I_k$, $F_j \prec I_k^c$, and $F_\ell \prec T_k$.*

PROOF. We prove the forward direction by the contrapositive. Let \mathcal{F} be a forbidden family, where $F_i \in \mathcal{F}$ is an $m_i \times n_i$ configuration. We choose d to be the maximum of $\{\{m_i\}, \{n_i\}\}$ over i . Now suppose there is at least one matrix of I_{2d}, I_{2d}^c , or T_{2d}

(without loss of generality, we will assume I_{2d}) where for all $F_i \in \mathcal{F}$, $F_i \not\prec I_{2d}$. Then for all $F_i \in \mathcal{F}$, $F_i \not\prec I_m$ for $m \geq 2d$ and so there exists a linear construction $A \in \text{Avoid}(m, \mathcal{F})$, namely, $A = I_m$. Therefore $\text{forb}(m, \mathcal{F})$ is $\Omega(m)$ and cannot have a constant upper bound.

The reverse direction is straightforward to prove using Lemma 1.3.20. As stated in the hypothesis, suppose there exists F_i, F_j, F_ℓ (not necessarily unique configurations) in our family \mathcal{F} such that for some value $k \in \mathbb{N}$, $F_i \prec I_k$, $F_j \prec I_k^c$, and $F_\ell \prec T_k$. Then by Lemma 1.3.20, $\text{forb}(m, \mathcal{F}) \leq \text{forb}(m, \{I_k, I_k^c, T_k\})$. But by Theorem 1.3.6 $\text{forb}(m, \{I_k, I_k^c, T_k\})$ is $O(1)$, so $\text{forb}(m, \mathcal{F})$ is also $O(1)$. ■

The proof of this classification suggests the following proposition.

Proposition 2.1.2. *For all forbidden families \mathcal{F} with a non-constant upper bound, $\text{forb}(m, \mathcal{F})$ is $\Omega(m)$.*

PROOF. By Theorem 2.1.1, a family has a non-constant upper bound if there is some matrix from I_k, I_k^c, T_k such that for all $F_i \in \mathcal{F}$, none of the F_i are contained in this matrix. Without loss of generality, let this matrix be I_k . Then for $m \geq k$ the linear matrix $I_m \in \text{Avoid}(m, \mathcal{F})$ and so there is at least a linear construction in $\text{Avoid}(m, \mathcal{F})$. Thus $\text{forb}(m, \mathcal{F})$ is $\Omega(m)$. ■

These two results together tell us that there is no family \mathcal{F} such that $\text{forb}(m, \mathcal{F})$ is $\Theta(m^\ell)$ where $0 < \ell < 1$.

2.1.1 A constant construction

Dunwoody describes the following construction that avoids the family $\{I_k, T_k, I_k^c\}$.

Proposition 2.1.3. *[11] We select for A all columns from the $(k-1)$ -fold product $[\mathbf{0}_{k-1}|T_{k-1}] \times [\mathbf{0}_{k-1}|T_{k-1}] \times \cdots \times [\mathbf{0}_{k-1}|T_{k-1}]$ that have column sum at most $k-1$. A is a $(k-1)^2 \times \binom{2k-2}{k-1}$ simple matrix in $\text{Avoid}((k-1)^2, \{I_k, I_k^c, T_k\})$.*

PROOF. We readily see that $T_k \not\prec A$ since A has no column with k 1's. We also see that $I_k \not\prec A$ since if $I_k \prec A$, then it must use two rows from one of the products T_{k-1} but every two rows of I_k contains I_2 and no two rows of T_{k-1} contain I_2 . Similarly $I_k^c \not\prec A$ So $A \in \text{Avoid}((k-1)^2, \{I_k, I_k^c, T_k\})$.

We note that $[\mathbf{0}_{k-1}|T_{k-1}]$ contains exactly one column of sum i for each i with $0 \leq i \leq k-1$. We create a bijection between the elements of $\binom{2k-2}{k-1}$ and the columns as

follows. We consider $1, 2, \dots, 2k-2$ and imagine choosing $k-1$ entries $a_1 < a_2 < \dots < a_{k-1}$. Then we choose the column of sum $a_1 - 1$ from the first T_{k-1} and the column of sum $a_i - a_{i-1} - 1$ from the i th T_{k-1} (from the $k-1$ -fold product) for $i = 2, 3, \dots, k-1$. We verify that sum of column sums is at most $k-1$ ($a_1 - 1 + \sum_{i=2}^{k-1} (a_i - a_{i-1} - 1) = a_{k-1} - k + 1$ and we use $a_{k-1} \leq 2k - 2$) and no column sum chosen from a T_{k-1} exceeds $k-1$ (By our choice of a_i 's we have $i \leq a_i \leq 2k - 2 - (k - 1 - i)$ and so $a_i \leq k + i - 1$ and $a_{i-1} \geq i - 1$ and so $0 \leq a_i - a_{i-1} - 1 \leq k - 1$). ■

This construction is interesting for two reasons. First, it establishes a lower bound for $\text{forb}(m, \{I_k, T_k, I_k^c\})$, as seen in Proposition 1.3.7. Second, in contrast to the bounds that we will present in Section 2.3 and Section 2.4, this is quite a substantial constant bound, demonstrating that even though the asymptotic bound of the family is constant by Theorem 1.3.6, there is still room for a large construction that will avoid the family.

2.2 Failed monotonicity of the constant bound

When considering $\text{forb}(m, F)$ and $\text{forb}(m, \mathcal{F})$, a natural question arises about whether the bound is monotonic with respect to the number of rows, i.e. as m increases, is $\text{forb}(m, F)$ or $\text{forb}(m, \mathcal{F})$ weakly decreasing?

We have one result of Raggi that states that $\text{forb}(m, F)$ is monotonic for single configurations of a certain kind.

Proposition 2.2.1. *[18] Let F be a $(0, 1)$ -matrix that does not contain a row of 0's, a row of 1's, or a repeated row. Then for all m , $\text{forb}(m, F) \leq \text{forb}(m + 1, F)$.*

While we have this result for single configurations of the type described above, when considering $\text{forb}(m, \mathcal{F})$, particularly for families with a constant bound, the monotonicity conjecture fails almost immediately. There is a simple family with a constant bound where $\text{forb}(m, \mathcal{F})$ does not satisfy the inequality $\text{forb}(m, \mathcal{F}) \leq \text{forb}(m + 1, \mathcal{F})$, namely the 2×2 block of 0's and the 2×2 block of 1's

Theorem 2.2.2. *We have*

$$\text{forb} \left(m, \left\{ \left[\begin{array}{cc} 0 & 0 \\ 0 & 0 \end{array} \right], \left[\begin{array}{cc} 1 & 1 \\ 1 & 1 \end{array} \right] \right\} \right) = \begin{cases} 2 & \text{if } m = 1, m \geq 7 \\ 4 & \text{if } m = 2, 5, 6 \\ 6 & \text{if } m = 3, 4 \end{cases} .$$

PROOF. The proof for $m = 1, 2$ is trivial; on one row the complete matrix $[01]$ avoids the family and on two rows the complete matrix K_2 (with four columns) also avoids the family.

For $m = 3$ and $m = 4$ we use Proposition 2.3.10 to prove the upper bound of $6(\ell - 1) = 6(1) = 6$ and then the construction K_4^2 achieves the bound (for $m = 3$, simply delete one of the rows of the matrix construction for $m = 4$).

For $m = 5$ and $m = 6$ we have a construction with four columns, which is $(K_4^2)^T$. However, we must show that the upper bound for these rows is no greater than 4. Suppose there is a matrix of five columns on five or six rows that avoids our family. We analyze this matrix by rows. In the first row there must be at least three 1's or three 0's. Without loss of generality, we assume there are at least three 1's. We now restrict our attention to the remaining rows in the columns underneath these 1's. These rows must have at least two 0's, otherwise we generate the block of 1's. However, we cannot repeat any of these rows with two 0's because that will generate the block of 0's. There are only 3 possible ways to order the two 0's and one 1 in these three columns before creating a repeat. Thus it is not possible to go beyond four rows, and $\text{forb}(m, \mathcal{F})$ for $m = 5, m = 6$ must be less than 5.

Finally, the bound and construction for $m \geq 7$ follows from both Theorem 2.3.4 and Theorem 2.3.6. ■

The pattern of bounds in this proof suggests that while the bound on a forbidden family need not be always monotonic, $\text{forb}(m, \mathcal{F})$ may be eventually monotonic, for m sufficiently large.

2.3 Exact bound for blocks of 0's and 1's

In this section, we offer a refinement of the constant bound of Theorem 1.3.6 by giving an exact bound for forbidden families consisting of a block of zeros and block of ones,

of which the family in Theorem 2.2.2 is an example.

Definition 2.3.1. Let k, ℓ, p, q be positive integers. We define $0_{j\ell}$ as a $k \times \ell$ block of 0's and J_{pq} as a $p \times q$ block of 1's.

These two configurations, $0_{j\ell}$ and J_{pq} , are interesting both for their non-monotonic bound, and for their frequent appearance in other configurations, particularly I_k, I_k^c and T_k .

2.3.1 Column bound

Our first theorem will establish an exact bound for $\text{forb}(m, \{0_{k\ell}, J_{pq}\})$ for m sufficiently large. We will first prove a series of lemmas.

Lemma 2.3.2. Let k, ℓ, p, q be given. Let $A \in \text{Avoid}(m, \{0_{k\ell}, J_{pq}\})$, with $\|A\| = n$. Also let a_r denote the number of 0's in row r of A , and b_r the number of 1's in row r so that $a_r + b_r = n$. Then:

$$\sum_{r=1}^m \left(\binom{a_r}{\ell} + \binom{b_r}{q} \right) \leq (k-1) \binom{n}{\ell} + (p-1) \binom{n}{q}. \quad (2.1)$$

PROOF. We consider the columns of A . We take all ℓ sized subsets of the columns and call them 0-buckets. Similarly, we take all q sized subsets of the columns as 1-buckets. We will have $\binom{n}{\ell}$ 0-buckets and $\binom{n}{q}$ 1-buckets. We then process the rows of A one by one, considering all possible ℓ sized and q sized subsets of columns on that row. If one of these subsets contains all 0's or all 1's, it makes a contribution to the appropriate 0-bucket or 1-bucket. Thus if there are a 0's in a row, and b 1's (where $a + b = n$), then the row will make contributions to $\binom{a}{\ell}$ 0-buckets and $\binom{b}{q}$ 1-buckets. The left side of (2.1) is thus the total number of contributions over the rows of A . Each of our $\binom{n}{\ell}$ 0-buckets can have a maximum of $k-1$ contributions, and similarly, our $\binom{n}{q}$ 1-buckets can have a maximum of $p-1$ contributions, which produces the right side of the inequality. ■

Lemma 2.3.3. Let m, ℓ, q be given with $m \geq \binom{\ell+q-2}{q-1}$. Let $A = [K_{\ell+q-2}^{q-1} | j \cdot \gamma]^T$, where γ is a column in $K_{\ell+q-2}^{q-1}$ and $j = m - \binom{\ell+q-2}{q-1}$. Then A is an $m \times (\ell + q - 2)$ simple matrix and $A \in \text{Avoid}(m, \{0_{1\ell}, J_{1q}\})$.

PROOF. By definition A is simple, with m rows and $\ell + q - 2$ columns. Furthermore, each row has $\ell - 1$ zeroes and $q - 1$ ones, so $A \in \text{Avoid}(m, \{0_{1\ell}, J_{1q}\})$. ■

With these two lemmas, we can now establish an upper bound and constructive lower bound for $\text{forb}(m, \{0_{k\ell}, J_{pq}\})$.

Theorem 2.3.4. *Let k, ℓ, p, q be given. Then there exists some constant $c_{k\ell pq}$ such that for $m \geq c_{k\ell pq}$, we have $\text{forb}(m, \{0_{k\ell}, J_{pq}\}) = \ell + q - 2$*

PROOF. We let $d = \max\{k, \ell, p, q\}$. We can then say $0_{k\ell} \prec T_{2d}, 0_{k\ell} \prec I_{2d}$ and $J_{pq} \prec T_{2d}, J_{pq} \prec I_{2d}^c$. Thus by Theorem 2.1.1, $\text{forb}(m, \{0_{k\ell}, J_{pq}\})$ is $O(1)$.

We wish to show that $\text{forb}(m, \{0_{k\ell}, J_{pq}\}) = \ell + q - 2$. From Lemma 2.3.2 we know that the right side of (2.1) is constant based on n, k, ℓ, p and q . In order to maintain the inequality in (2.1), the terms of the left side must be zero for most rows r , which means $a_r < \ell$ and $b_r < q$. Because $a + b = n$, this requires $n \leq \ell + q - 2$. So for sufficiently large m , $\text{forb}(m, \{0_{k\ell}, J_{pq}\}) \leq \ell + q - 2$.

It remains to show we have a construction that achieves the bound. Let $A = [K_{\ell+q-2}^{q-1} | j \cdot \gamma]^T$ as described in Lemma 2.3.3. By Lemma 2.3.3 $A \in \text{Avoid}(m, \{0_{1\ell}, J_{1q}\})$, and so we know $A \in \text{Avoid}(m, \{0_{k\ell}, J_{pq}\})$. Furthermore, $\|A\| = \ell + q - 2$. Thus A is a construction that achieves the upper bound and so $\text{forb}(m, \{0_{k\ell}, J_{pq}\}) = \ell + q - 2$. ■

2.3.2 A sharp boundary

We have now shown that $\text{forb}(m, \{0_{k\ell}, J_{pq}\}) = \ell + q - 2$ for m larger than some constant $c_{k\ell pq}$. We can say something more specific about $c_{k\ell pq}$ when $\ell = q$ or $p = k$. In fact, in these two cases, we can show that there is a sharp boundary for $\text{forb}(m, \{0_{k\ell}, J_{pq}\})$; for $m \leq c_{k\ell pq}$, $\text{forb}(m, \{0_{k\ell}, J_{pq}\}) > \ell + q - 2$ and for $m > c_{k\ell pq}$, the bound proved in Theorem 2.3.4 applies.

First we consider the case where $\ell = q$.

Lemma 2.3.5. *Let $\ell = q$, $\|A\| = n$ and $a + b = n$. Then the terms of the summand in (2.1) are bounded from below as follows:*

$$\binom{\lfloor \frac{n}{2} \rfloor}{\ell} + \binom{\lceil \frac{n}{2} \rceil}{\ell} \leq \binom{a}{\ell} + \binom{b}{\ell}. \quad (2.2)$$

PROOF. Note that if a is written as $\lfloor \frac{n}{2} \rfloor - i$ for some $i \in \mathbb{Z}$, then $b = n - a = n - \lfloor \frac{n}{2} \rfloor + i = \lceil \frac{n}{2} \rceil + i$. We proceed by induction on i .

Let $a = \lfloor \frac{n}{2} \rfloor$, and $b = \lceil \frac{n}{2} \rceil$. Then $\binom{a}{\ell} + \binom{b}{\ell} \geq \binom{\lfloor \frac{n}{2} \rfloor}{\ell} + \binom{\lceil \frac{n}{2} \rceil}{\ell}$ is true trivially.

Now assume for the inequality is true for $a = \lfloor \frac{n}{2} \rfloor - i$, $b = \lceil \frac{n}{2} \rceil + i$, for some $i \in \mathbb{Z}$.

We must show that

$$\binom{\lfloor \frac{n}{2} \rfloor - (i+1)}{\ell} + \binom{\lceil \frac{n}{2} \rceil + (i+1)}{\ell} \geq \binom{\lfloor \frac{n}{2} \rfloor}{\ell} + \binom{\lceil \frac{n}{2} \rceil}{\ell}.$$

Using the combinatorial identity $\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$, we can rewrite the left side of the previous expression as

$$\binom{\lfloor \frac{n}{2} \rfloor - i}{\ell} - \binom{\lfloor \frac{n}{2} \rfloor - i - 1}{\ell - 1} + \binom{\lceil \frac{n}{2} \rceil + i}{\ell - 1} + \binom{\lceil \frac{n}{2} \rceil + i}{\ell}.$$

By induction we know that $\binom{\lfloor \frac{n}{2} \rfloor - i}{\ell} + \binom{\lceil \frac{n}{2} \rceil + i}{\ell} \geq \binom{\lfloor \frac{n}{2} \rfloor}{\ell} + \binom{\lceil \frac{n}{2} \rceil}{\ell}$ and by properties of the binomial coefficient the difference $\binom{\lceil \frac{n}{2} \rceil + i}{\ell - 1} - \binom{\lfloor \frac{n}{2} \rfloor - i - 1}{\ell - 1} \geq 0$. Thus

$$\binom{\lfloor \frac{n}{2} \rfloor - (i+1)}{\ell} + \binom{\lceil \frac{n}{2} \rceil + (i+1)}{\ell} \geq \binom{\lfloor \frac{n}{2} \rfloor}{\ell} + \binom{\lceil \frac{n}{2} \rceil}{\ell},$$

as desired. ■

Proposition 2.3.6. *Let p, k be given, and $l = q$. Let $A \in \text{Avoid}(m, \{0_{k\ell}, J_{p\ell}\})$, with $\|A\| = n$. Then the number of rows m in A is subject to the following bound:*

$$m \leq \frac{(k+p-2)\binom{n}{\ell}}{\binom{\lfloor \frac{n}{2} \rfloor}{\ell} + \binom{\lceil \frac{n}{2} \rceil}{\ell}} \quad (2.3)$$

PROOF. Because Lemma 2.3.5 gives a lower bound for the summand from the left side of (2.1), we know the following is true:

$$m \left(\binom{\lfloor \frac{n}{2} \rfloor}{\ell} + \binom{\lceil \frac{n}{2} \rceil}{\ell} \right) \leq \sum_{r=1}^m \binom{\alpha_r}{\ell} + \binom{\beta_r}{q}$$

We know from Lemma 2.3.2 that the summand above is bounded by

$$(k-1)\binom{n}{\ell} + (p-1)\binom{n}{q}$$

Combining these two inequalities, gives

$$m \left(\binom{\lfloor \frac{n}{2} \rfloor}{\ell} + \binom{\lceil \frac{n}{2} \rceil}{\ell} \right) \leq (k-1)\binom{n}{\ell} + (p-1)\binom{n}{\ell},$$

resulting in the bound of (2.3). ■

We can now prove that when $\ell = q$, the column bound has a sharp boundary where the column bound changes. In particular, the following results also show lack of monotonicity for this constant bound.

Theorem 2.3.7. *Let p, k be given, and $l = q$. Then if $m \leq (k-1)\binom{2\ell-1}{\ell} + (p-1)\binom{2\ell-1}{\ell}$, then there exists a matrix $A \in \text{Avoid}(m, \{O_{k\ell}, J_{p\ell}\})$ such that $\|A\| > 2\ell - 2$. If $m > (k-1)\binom{2\ell-1}{\ell} + (p-1)\binom{2\ell-1}{\ell}$ then $\text{forb}(m, \{O_{k\ell}, J_{p\ell}\}) = 2\ell - 2$.*

PROOF. Let $l = q$ and $n = 2\ell - 1$. The row bound described by (2.3) is as follows:

$$\begin{aligned} m &\leq \frac{(k+p-2)\binom{2\ell-1}{\ell}}{\binom{\ell-1}{\ell} + \binom{\ell}{\ell}} \\ &= (k+p-2)\binom{2\ell}{\ell} \\ &= (k-1)\binom{2\ell-1}{\ell} + (p-1)\binom{2\ell-1}{\ell}. \end{aligned}$$

Let $A = [(k-1) \cdot K_{2\ell-1}^{\ell-1} | (p-1) \cdot K_{2\ell-1}^{\ell}]^T$. Then A has $(k-1)\binom{2\ell-1}{\ell} + (p-1)\binom{2\ell-1}{\ell}$ rows and $2\ell - 1$ columns. Furthermore, the configuration $O_{1\ell}$ will appear exactly $k-1$ times on any subset of ℓ columns and the configuration $J_{1\ell}$ will appear exactly $p-1$ times on any subset of ℓ columns and thus, $A \in \text{Avoid}(m, \{O_{k\ell}, J_{p\ell}\})$.

Note that this is a best case scenario with only one contribution per row, and so once we reach the boundary of $m = (k-1)\binom{2\ell-1}{\ell} + (p-1)\binom{2\ell-1}{\ell}$ the left hand side of the inequality in (2.1) will be equal to the right hand side. In the next row, because either $a_r \geq \ell$ or $b_r \geq \ell$, the left hand side will increase by at least one, contradicting

(2.1) and thereby proving that for $m > (k-1)\binom{2\ell-1}{\ell} + (p-1)\binom{2\ell-1}{\ell}$, the bound must drop to $2\ell - 2$ as already proved in Theorem 2.3.4 for m sufficiently large. ■

Similarly to the previous theorem, we can also show that if $k = p$, the column bound from Theorem 2.3.4 has a sharp boundary.

Theorem 2.3.8. *Let k, ℓ, q be given. Suppose $k = p$ and let $n = \ell + q$. Then if $m \leq (k-1)\binom{\ell+q}{\ell}$, there exists an $A \in \text{Avoid}(m, \{0_{k\ell}, J_{kq}\})$ where $\|A\| = n$. Furthermore, for $m > (k-1)\binom{\ell+q}{\ell}$, $\text{forb}(m, \{0_{k\ell}, J_{kp}\}) < \ell + q$.*

PROOF. Let $k = p$ and $n = \ell + q$. Let $A = [(k-1) \cdot K_{\ell+q}^\ell]^T$. By definition, A is an $(k-1)\binom{\ell+q}{\ell} \times (\ell + q)$ simple matrix. Now consider all ℓ -sized subsets of the columns of A . Each subset will contain all zeroes on exactly $k-1$ rows, thus avoiding $0_{k\ell}$. Similarly, each q -sized subset of columns will contain all zeroes on exactly $k-1$ rows of A , avoiding J_{kq} .

It remains to show that this boundary is sharp. We may apply (2.1) from Lemma 2.3.2. Because $k = p$ and $n = \ell + q$, where $\binom{\ell+q}{\ell} = \binom{\ell+q}{q}$, the right hand side of (2.1) can be written as $2(k-1)\binom{\ell+q}{\ell}$. On the left hand side of the inequality, we note that since $a_r + b_r = n = \ell + q$, we have $2 \leq \binom{a_r}{\ell} + \binom{b_r}{q}$, with equality when $a_r = \ell$ and $b_r = q$. Thus,

$$2m \leq \sum_{r=1}^m \left(\binom{a_r}{\ell} + \binom{b_r}{q} \right) \leq 2(k-1) \binom{\ell+q}{\ell}$$

Hence for $m > (k-1)\binom{\ell+q}{\ell}$, $\text{forb}(m, \{0_{k\ell}, J_{kp}\}) < \ell + q$. ■

A similar argument can be made for $n = \ell + q - 1$, with the same row bound.

Theorem 2.3.9. *Let k, ℓ, q be given. Suppose $k = p$ and let $n = \ell + q - 1$. Then if $m \leq (k-1)\binom{\ell+q}{\ell}$, there exists an $A \in \text{Avoid}(m, \{0_{k\ell}, J_{kq}\})$ where $\|A\| = n$. Furthermore, for $m > (k-1)\binom{\ell+q}{\ell}$, $\text{forb}(m, \{0_{k\ell}, J_{kp}\}) < \ell + q - 1$.*

PROOF. Let $k = p$ and $n = \ell + q - 1$. Let $A = [(k-1) \cdot K_{\ell+q-1}^\ell | (k-1) \cdot K_{\ell+q-1}^{\ell-1}]^T$. By definition, A is an $(k-1)\binom{\ell+q}{\ell} \times (\ell + q - 1)$ simple matrix. Now consider all ℓ -sized subsets of the columns of A . Each subset will contain all zeroes on exactly $k-1$ rows in the top half of the matrix, thus avoiding $0_{k\ell}$. Similarly, each q -sized subset of columns will contain all zeroes on exactly $k-1$ rows of the bottom half of A , avoiding J_{kq} .

It remains to show that this boundary is sharp. We may apply (2.1) from Lemma 2.3.2. Because $k = p$ and $n = \ell + q - 1$, after some manipulation of the binomials, the right hand side of (2.1) can be written as $(k - 1)\binom{\ell + q}{\ell}$. On the left hand side of the inequality, we note that since $a_r + b_r = n = \ell + q - 1$, we have $1 \leq \binom{a_r}{\ell} + \binom{b_r}{q}$, with equality when $a_r = \ell$ or $b_r = q$. Thus,

$$m \leq \sum_{r=1}^m \binom{a_r}{\ell} + \binom{b_r}{q} \leq (k - 1) \binom{\ell + q}{\ell}$$

Hence for $m > (k - 1)\binom{\ell + q}{\ell}$, $\text{forb}(m, \{0_{k\ell}, J_{kp}\}) < \ell + q - 1$. ■

2.3.3 Other bounds and constructions

We have now shown that for sufficiently large m , $\text{forb}(m, \{0_{k\ell}, J_{pq}\}) = \ell + q - 2$ and in certain cases, we know precisely how large m must be before this exact bound applies. However, it is not clear what $\text{forb}(m, \{0_{k\ell}, J_{pq}\})$ is for small m . The following results provide some additional insight into upper and lower bounds for $\text{forb}(m, \{0_{k\ell}, J_{pq}\})$.

The following proposition places an absolute upper bound on $\text{forb}(m, \{0_{k\ell}, J_{pq}\})$ for all values of m , provided that $k = p = 2$ and $\ell = q$.

Proposition 2.3.10. *Let $k = 2, p = 2$ and $\ell = q$. For $m \geq 3$, $\text{forb}(m, \{0_{2\ell}, J_{2\ell}\}) \leq 6(\ell - 1)$.*

PROOF. Suppose we have a matrix A which avoids $\{0_{2\ell}, J_{2\ell}\}$. On every subset of 3 rows, the columns of A will be one of the columns of K_3 , shown and labeled in (2.4).

$$\begin{array}{cccccccc} a & b & c & d & e & f & g & h \\ \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \end{array} \quad (2.4)$$

We now consider some subset S of three rows of A . The total number of columns in A can be expressed as a sum $a + b + c + d + e + f + g + h$, where each letter represents the total number of times that column appears on S . Note that we cannot have more than $\ell - 1$ copies of a and b in total, because if $a + b = \ell$ we will have $0_{2\ell}$. A similar argument applies to a, c ; a, d ; h, e ; h, f ; and h, g . So our question can be interpreted as a linear program where we want to maximize the sum $a + b + \dots + g + h$ under

the following constraints:

$$\begin{aligned} a + b &\leq \ell - 1, & h + e &\leq \ell - 1, \\ a + c &\leq \ell - 1, & h + f &\leq \ell - 1, \\ a + d &\leq \ell - 1, & h + g &\leq \ell - 1. \end{aligned}$$

If we add up all our constraints, we get that $3a+b+c+d+e+f+g+3h \leq 6(\ell-1)$ so the total number of columns $a+b+\dots+g+h$ is also bounded by $6(\ell-1)$. Furthermore, our sum will achieve this bound when $a = h = 0$ and $b = c = d = e = f = g = \ell - 1$.

■

While we demonstrated a construction that achieves our upper bound of $\ell + q - 2$ for m large, we are still searching for helpful constructions for m small. The following proposition describes one such candidate.

Proposition 2.3.11. *Let k, ℓ, p, q be given. Then $K_{k+p}^p \in \text{Avoid}(m, \{0_{k\ell}, J_{pq}\})$ for $\ell, q > 1$*

PROOF. Let $A = K_{k+p}^p$. Consider any k -sized subset of the rows of A . Because $A = K_{k+p}^p$, each k -sized subset will contain all zeroes in only one column of A . Similarly, each p -sized subset of rows will contain all ones in only one column of A . Thus, the rows of A avoid $\{0_{k\ell}, J_{pq}\}$ if $\ell, q > 1$. ■

This construction has $\binom{k+p}{p}$ columns on $k+p$ rows and thus shows that, for large k and p , there can be matrices with many columns in $\text{Avoid}(m, \{0_{k\ell}, J_{pq}\})$, even if the final constant bound as given in Theorem 2.3.4 is small.

2.4 Constant bound for another family

The previous bound for configurations of 0's and 1's is not the first refinement of the constant bound for families of forbidden configurations that have a constant bound. We present the following result of Balin Fleming, taken from his research notes with his permission.

Let the configurations $F_a, F_b,$ and F_c be defined as follows:

$$F_a = \left[\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & 1 \end{array} \middle| t \cdot \right], \quad F_b = \left[\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \end{array} \middle| t \cdot \right], \quad F_c = \left[\begin{array}{ccc} 0 & 1 & 1 \\ t \cdot & 0 & 1 \end{array} \right]$$

where t is some integer constant. The configurations F_a, F_b, F_c are contained in the matrices I_{3t}^c, I_{3t}, T_{3t} respectively, and are thus a family of constant bound as shown by Theorem 2.1.1. The value of the following result is a much smaller constant bound than that implied by Theorem 1.3.6.

Proposition 2.4.1. [14] *Let F_a, F_b, F_c be as given above, with $t \geq 2$. Then $\text{forb}(m, \{F_a, F_b, F_c\}) \leq 6t - 6$.*

PROOF. Let $A \in \text{Avoid}(m, \{F_a, F_b, F_c\})$. Either $\|A\| < 6t - 6$ (establishing the result) or $\|A\| > 6t - 5$. By pigeonhole principle we know that some row of A has at least $3t - 2$ zeroes or $3t - 2$ ones. Without loss of generality, we can assume there is some row of A with at least $3t - 2$ zeroes (if not, the following argument can proceed using the complement). We call this row i . We now consider the columns of A that have a zero in row i . We call this set of columns A' , and the next part of the proof will take place in A' , which, by definition, has at least $3t - 2$ columns.

Assume for a contradiction that no row in A' has $t \cdot [01]$. Assuming this, all rows in A' must have at least $2t - 1$ ones or $2t - 1$ zeroes. We choose rows j and k , with no $t \cdot [1]$, so each row has at least $2t - 1$ zeroes. With this number of zeroes, we have at least $t \cdot \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ on the row pair (j, k) . Since we have this many $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ columns on (j, k) , we cannot have an I_2 on the other columns of (j, k) as this would create the configuration F_b on these two rows. Hence, either rows j and k are the same or else they are comparable. In either case, each row j, k , has its ones confined to the same columns of A' , of which there are at most $t - 1$. This is true for all row pairs (j, k) where j and k contain no $t \cdot [1]$. Thus for these rows, all their ones are contained to the same set of $t - 1$ columns within A' . We now repeat this process for all pairs of rows (g, h) where g, h are a pair of rows with no $t \cdot [0]$. The same proof follows, that all zeroes in these rows must be confined to the same $t - 1$ (at most) columns in A' . Thus we have two sets of $t - 1$ columns, leaving behind a set of at least t columns that are identical, contradicting the simplicity of A . Therefore, there must be some row with $t \cdot [01]$ in A' . Call this row ℓ .

We now pull back to consider the full width of A , but now in the chosen row ℓ . As we have just shown, row ℓ has $[t \cdot 01]$. Assuming $\|A\| = 6t - 6$, then row ℓ has either $[(3t - 2) \cdot 0 | t \cdot 1]$ or $[(3t - 2) \cdot 1 | t \cdot 0]$. Assume the former. Following our previous argument with row i replaced by row ℓ we can show that there is some row p with $t \cdot [01]$ under the zeroes of row ℓ . Then there can be no zero in p under the ones of row

ℓ (else we will have F_b), so there must be t ones. But this gives a copy of F_c . Thus row ℓ cannot have $[(3t - 2) \cdot 0 | t \cdot 1]$. We obtain a similar contradiction if ℓ contains $[(3t - 2) \cdot 1 | t \cdot 0]$. Therefore $\|A\| < 6t - 6$. ■

Chapter 3

Families With a Linear Bound

In the previous chapter, we presented a characterization of finite forbidden families with a constant bound. The proof of this characterization indicated that there was a gap between forbidden families with constant bound and linear bound, that is, if a the bound for a forbidden family was not constant, it must be at least linear. Thus we now turn our attention to forbidden families with a linear bound. We begin with a set of single configurations with a quadratic bound and determine which families of these configurations yield a linear bound.

3.1 Families of minimal quadratic configurations

In our search for families with a linear bound, we first consider a list of minimal configurations with a quadratic bound, shown in Table 3.1. We use the word “minimal” for F with $\text{forb}(m, F)$ being $\Omega(m^2)$ to indicate that for any $F' \prec F$ and $F' \neq F$, then $\text{forb}(m, F')$ is $O(m)$. In a sense, these are the configurations “closest” to having a linear bound and thus seem like likely candidates for forming families with a linear bound. We hoped that the analysis of bounds would be more tractable with these minimal configurations and this expectation was correct.

We will focus our attention on families of size 2 from Table 3.1. By our characterization of constant families through Theorem 2.1.1, we know that some of these families will immediately yield a constant bound; for example any family containing both F_1 and F_2 will have a constant bound as detailed in Section 2.3. Furthermore, by Corollary 2.1.2 we know that any family without a constant bound must be $\Omega(m)$. Thus, our goal is to show that for some of these non-constant families the upper

	Configuration	Bound	Construction(s)	Reference
F_1	$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$	$\binom{m}{2} + \binom{m}{1} + \binom{m}{0}$	$I^c \times I^c$	[16]
F_2	$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$	$\binom{m}{2} + \binom{m}{1} + \binom{m}{0}$	$I \times I$	[16]
F_3	$\begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix}$	$\lfloor \frac{m^2}{4} \rfloor + m + 1$	$I \times I^c$	[3]
F_4	$\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$	$\binom{m}{2} + \binom{m}{1} + \binom{m}{0}$	$I^c \times I^c$	[20, 21, 23]
F_5	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$\binom{m}{2} + \binom{m}{1} + \binom{m}{0}$	$I^c \times I^c$ $I^c \times T$ $T \times T$	[19]
F_6	$\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$	$\binom{m}{2} + \binom{m}{1} + \binom{m}{0}$	$I \times I$ $I \times T$ $T \times T$	[19]
F_7	$\begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$	$\lfloor \frac{m^2}{4} \rfloor + m + 1$	$T \times T$	[1]
F_8	$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$	$\binom{m}{2} + \binom{m}{1} + \binom{m}{0}$	$I \times I$	[20, 21, 23]
F_9	$\begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}$	$\binom{m}{2} + 2m - 1$	$I \times T$ $I^c \times T$	[15]

Table 3.1: Minimal Quadratic Configurations

Pairs of configurations with a constant bound by Theorem 2.1.1.	$\{F_1, F_2\}, \{F_1, F_6\}, \{F_1, F_8\}, \{F_2, F_4\},$ $\{F_2, F_5\}, \{F_4, F_6\}, \{F_5, F_8\}$
Pairs of configurations with an $\Omega(m^2)$ construction.	$\{F_1, F_4\}, \{F_1, F_5\}, \{F_2, F_6\}, \{F_2, F_8\},$ $\{F_4, F_6\}, \{F_4, F_8\}, \{F_5, F_6\}, \{F_5, F_7\},$ $\{F_5, F_9\}, \{F_6, F_7\}, \{F_6, F_8\}, \{F_6, F_9\},$
Pairs of configurations known to have bound $\Omega(m)$.	$\{F_1, F_3\}, \{F_1, F_7\}, \{F_1, F_9\}, \{F_2, F_3\},$ $\{F_2, F_7\}, \{F_2, F_9\}, \{F_3, F_4\}, \{F_3, F_5\},$ $\{F_3, F_6\}, \{F_3, F_7\}, \{F_3, F_8\}, \{F_3, F_9\},$ $\{F_4, F_7\}, \{F_4, F_9\}, \{F_7, F_8\}, \{F_7, F_9\},$ $\{F_8, F_9\}$

Table 3.2: Pairs of Minimal Quadratic Configurations

bound is also linear. See row 1 of Table 3.2.

Furthermore, we note that some families of the configurations in Table 3.1 will continue to have a quadratic bound. Any family of configurations that shares a construction from column 3 will be $\Omega(m^2)$, and therefore not $\Theta(m)$. See row 2 of Table 3.2

Thus our method in examining families of the configurations in Table 3.1 was as follows: attempt to determine the upper bound for families which neither fall under our constant family classification, nor families where a quadratic product construction avoids all configurations in the family. Such families are listed in row 3 of Table 3.2.

3.1.1 Families proved using basic methods

The first three considered families all contain the configuration F_7 and have similar proofs.

Proposition 3.1.1. *We have $\text{forb}(m, \{F_4, F_7\}) = \text{forb}(m, \{F_7, F_8\}) \leq 2m$.*

PROOF. It suffices to show that $\text{forb}(m, \{F_4, F_7\}) \leq 2m$. Let $A \in \text{Avoid}(m, \{F_4, F_7\})$. In a standard decomposition of A using row r , C_r must avoid I_2 and $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$. To avoid I_2 , the columns of C_r must be pairwise comparable (recall from Section 1.2.4), making them the columns of a triangular matrix. To avoid $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$, the columns of C_r must have a minimum column sum of $m - 2$ (note that C_r has $m - 1$ rows). Because there are only two columns in a triangular matrix with column sum greater than or equal to $m - 2$, there are only two possible columns in C_r . Thus $\|C_r\| \leq 2$, and by standard

induction $\|A\| \leq 2m$. A construction of I^c will create a lower bound of m columns; we have not found a construction that shows the $2m$ upper bound to be tight. ■

Proposition 3.1.2. *We have $\text{forb}(m, \{F_1, F_7\}) = \text{forb}(m, \{F_2, F_7\}) \leq 2m$.*

PROOF. It suffices to show that $\text{forb}(m, \{F_2, F_7\}) \leq 2m$. Let $A \in \text{Avoid}(m, \{F_2, F_7\})$. In a standard decomposition of A using row r , C_r must avoid I_2 and $[1\ 1]$. To avoid I_2 , the columns of C_r must be pairwise comparable, but to avoid $[1\ 1]$ they must also be disjoint. The only possible choice for columns would be $\mathbf{0}$ and some non-zero column α , so $\|C_r\| \leq 2$ and thus by standard induction $\|A\| \leq 2m$. A construction of I^c will create a lower bound of m columns; we have not found a construction that shows the $2m$ upper bound to be tight. ■

Proposition 3.1.3. *We have $\text{forb}(m, \{F_3, F_7\}) \leq 3m$.*

PROOF. Let $A \in \text{Avoid}(m, \{F_3, F_7\})$. In a standard decomposition of A using row r , C_r must avoid I_2 and $[0\ 0\ 1\ 1]$. To avoid I_2 , the columns of C_r must be pairwise comparable, thus columns of a triangular matrix. However since we are also avoiding $[0\ 0\ 1\ 1]$ in C_r , this only allows for three columns in C_r . Thus $\|C_r\| \leq 3$ and by standard induction $\|A\| \leq 3m$. Any of I, I^c or T will create a lower bound of m columns; we have not found a construction showing the $3m$ upper bound to be tight. ■

The next proposition requires a graph theoretic argument to prove the bound.

Proposition 3.1.4. *We have $\text{forb}(m, \{F_3, F_4\}) = \text{forb}(m, \{F_3, F_8\}) \leq 3m + 1$.*

PROOF. It suffices to show that $\text{forb}(m, \{F_3, F_8\}) \leq 3m+1$. Let $A \in \text{Avoid}(m, \{F_3, F_8\})$. The presence of F_8 means that all columns in A have column sum at most two. There are $m+1$ possible columns of column sum 0 and 1. It remains to show that the number of columns of column sum 2 is bounded by $2m$.

Here we note that in our understanding of a $(0, 1)$ -matrix as a list of incidence vectors for a hypergraph (see Section 1.2.1), columns of column sum 2 correspond to edges on a graph of m vertices. Our second configuration F_3 is present when there are two vertices i, j , where the degree of i is at least 3, the degree of j is at least 3, there is an edge joining i, j , and there is an edge disjoint from i and j . Thus, edges from a vertex of large degree can only go to a vertex of degree less than 3, allowing

us to bound the number of edges on the graph (and thus the number of columns in A with column sum 2) by counting edges out of small degree vertices. Having selected i, j , it suffices to have more than $2(m - 1)$ edges to find an edge not going through i, j .

Thus a bound of $2m + m + 1 = 3m + 1$ suffices for $\text{forb}(m, \{F_3, F_8\})$. ■

Proposition 3.1.5. *We have $\text{forb}(m, \{F_3, F_5\})$ and $\text{forb}(m, \{F_3, F_6\})$ being $O(m)$.*

PROOF. It suffices to show $\text{forb}(m, \{F_3, F_5\})$ is $O(m)$. Let $A \in \text{Avoid}(m, \{F_3, F_5\})$. Then applying our decomposition seen in (1.1), we deduce that C_r has no configurations $[0011]$ or I_3 . Let $R(r)$ denote a minimal set of rows of C_r such that $C_r|_{R(r)}$ is simple and in particular $\|C_r\| = \|C_r|_{R(r)}\|$. We deduce that $C_r|_{R(r)}$ has no row of all 0's nor a row of all 1's nor a repeated row. Assume $C_r|_{R(r)}$ has at least 3 rows each of row sum 1. Then we deduce that $I_3 \prec C_r|_{R(r)}$. Thus we deduce that some row s of $C_r|_{R(r)}$ have only one 0 and at least two 1's. We then deduce that in A , we have at most one column containing $\begin{smallmatrix} r \\ s \end{smallmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, else A has F_3 . This works for every row $r \in [m]$. Consider a directed graph where for each row r we have a directed edge $r \rightarrow s$ in this case. We deduce that in the directed graph we have a directed cycle. Thus there are t rows a_1, a_2, \dots, a_t for which there is at most one column containing $\begin{smallmatrix} a_i \\ a_{i+1} \end{smallmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ for $i = 1, 2, \dots, t - 1$ and at most one column containing $\begin{smallmatrix} a_t \\ a_1 \end{smallmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. We now verify that apart from t columns, all columns of $A|_{\{a_1, a_2, \dots, a_t\}}$ are either all 0's or all 1's. Thus we may delete the $t - 1$ rows a_1, a_2, \dots, a_{t-1} and up to t columns and obtain a simple matrix. This yields the result by induction using $\text{forb}(m, \{F_3, F_5\}) \leq t + \text{forb}(m - t + 1, \{F_3, F_5\})$. ■

The proof of the bound for the following family uses a linear algebra reduction argument of Anstee and Fleming as seen in Section 1.4.3.

Proposition 3.1.6. *$\text{forb}(m, \{F_1, F_3\}) = \text{forb}(m, \{F_2, F_3\}) \leq 2m$*

PROOF. We take the case of $\{F_1, F_3\}$. Let $A \in \text{Avoid}(m, \{F_1, F_3\})$. In order to avoid F_1 , on each pair of rows in A must contain at most one $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ while to avoid F_3 , each pair of rows must have no $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ or no $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ or at most 1 $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$. In the case where all pairs of rows have no $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ or no $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$, then the result follows from Theorem 1.3.1. In the case where a pair of rows has only one $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and therefore, only one of $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$, we have that every pair of rows has two columns in short supply. By the result of Anstee and

Fleming (Theorem 1.4.1), we can find $O(m)$ columns to delete so that the resulting matrix has an absent column on each pair of rows. The resulting matrix has no K_2 and hence $O(m)$ columns by Theorem 1.3.1. Thus A has $O(m)$ columns. ■

3.1.2 Families proved using a structure argument

The results for several of our families taken from Table 3.1 use a result of Frankl, Füredi, and Pach [15]. The result is as follows: if a matrix A avoids F_9 , all columns of A with column sum i will have a particular structure. Let $A \in \text{Avoid}(m, F_9)$. Consider the columns of column sum i , labeled in (3.1) as T_i or S_i . Then the following structure must exist on these columns:

$$\begin{array}{c}
 T_i \\
 \left[\begin{array}{cccc}
 1 & 0 & \dots & 0 \\
 0 & 1 & \dots & 0 \\
 \vdots & \vdots & \ddots & \vdots \\
 0 & 0 & \dots & 1 \\
 \hline
 1 & 1 & \dots & 1 \\
 \vdots & & & \vdots \\
 1 & 1 & \dots & 1 \\
 \hline
 0 & 0 & \dots & 0 \\
 \vdots & & & \vdots \\
 0 & 0 & \dots & 0
 \end{array} \right]
 \end{array}
 \quad \text{or} \quad
 \begin{array}{c}
 S_i \\
 \left[\begin{array}{cccc}
 0 & 1 & \dots & 1 \\
 1 & 0 & \dots & 1 \\
 \vdots & \vdots & \ddots & \vdots \\
 1 & 1 & \dots & 0 \\
 \hline
 1 & 1 & \dots & 1 \\
 \vdots & & & \vdots \\
 1 & 1 & \dots & 1 \\
 \hline
 0 & 0 & \dots & 0 \\
 \vdots & & & \vdots \\
 0 & 0 & \dots & 0
 \end{array} \right]
 \end{array}
 \tag{3.1}$$

Note that for the columns of T_i , $|B_i| = i - 1$ and for S_i , $|B_i| = i - (|A_i| - 1)$. Also $\|T_i\| = |A_i|$ and $\|S_i\| = |A_i|$. We will use this structure and its properties in the proofs that follow.

Proposition 3.1.7. *We have $\text{forb}(m, \{F_5, F_6, F_9\}) \leq 2m$.*

PROOF. Let $A \in \text{Avoid}(m, \{F_5, F_6, F_9\})$. We sort the columns of A by column sum. Because A avoids F_9 , each set of columns with column sum i will have the structure shown in (3.1). However, because A must also avoid F_5 and F_6 , the maximum number of columns of any column sum i must be less than 3 (otherwise there will be an F_5 or F_6 on the rows A_i). Thus the upper bound bound on $\|A\|$ is $2m$. ■

Proposition 3.1.8. *We have $\text{forb}(m, \{F_1, F_9\}) = \text{forb}(m, \{F_2, F_9\}) = 2m + 1$.*

PROOF. It suffices to show $\text{forb}(m, \{F_1, F_9\}) = 2m + 1$. Let $A \in \text{Avoid}(m, \{F_1, F_9\})$. We sort the columns of A by column sum, and because A avoids F_9 , each set of columns with column sum i will have the structure shown in (3.1). Any matrix formed by two columns of sum at most $m - 3$ will either contain F_1 or F_9 or both. So consider the columns of sum $m - 2$, $m - 1$ and m . The maximum number of columns occurs with structure S_{m-2} , with $|A_{m-2}| = m - 1$, structure S_{m-1} with $|A_{m-1}| = m$ and $\mathbf{1}_m$, yielding $2m$ columns. If we use these columns then we can add precisely one column of smaller sum, namely the column with a single 1 in row m , without creating either F_1 or F_9 . Thus $\text{forb}(m, \{F_1, F_9\}) = 2m + 1$. ■

Proposition 3.1.9. *We have $\text{forb}(m, \{F_4, F_9\}) = \text{forb}(m, \{F_8, F_9\}) = 2m$.*

PROOF. It suffices to show that $\text{forb}(m, \{F_4, F_9\}) = 2m$. Let $A \in \text{Avoid}(m, \{F_4, F_9\})$. We sort the columns of A by column sum, and because A avoids F_9 , each set of columns with column sum i will have the structure shown in (3.1). However, we also know that A cannot contain F_4 and thus it must have column sum of $m - 2$ or greater on all columns. This means we can have the columns of sum $m - 2$, $m - 1$ in the two types of structures S_i and T_i and in fact, we can take S_{m-2} with $|A_{m-2}| = m - 1$, the structure S_{m-1} with $|A_{m-1}| = m$ and the column $\mathbf{1}_m$, yielding $2m$ columns total. ■

Proposition 3.1.10. *We have $\text{forb}(m, \{F_7, F_9\})$ is $O(m)$.*

PROOF. Let $A \in \text{Avoid}(m, \{F_7, F_9\})$. Let A_T denote the columns where columns of sum i form the structure T_i from (3.1) and let A_S denote the columns where columns of sum i form the structure S_i . Then $\|A\| \leq 2m + \|A_T\| + \|A_S\|$, where there at most 2 columns of column sum i not in A_T or A_S . It suffices to show that $\|A_T\| \leq 2m$; we will then know $\|A_S\| \leq 2m$ by taking $(0, 1)$ -complements, as F_7^c is F_7 and F_9^c is F_9 .

Note that for all i where columns of sum i are in A_T , $\|T_i\|$ is bounded by the number of rows containing the identity, or $|A_i|$. Thus, in order to show that $\text{forb}(m, \{F_7, F_9\})$ is $O(m)$, it suffices to show that $\sum |A_i|$ is $O(m)$.

If $A_i \cap A_j = \emptyset$ for all pairs i, j then $\sum_i |A_i| = m$ and we are done. Now suppose there are i, j column sums with $i > j$ such that $A_i \cap A_j \neq \emptyset$. Then $|A_i \cap A_j| < 2$. Suppose not. Then there are two rows in $A_i \cap A_j$ that contain $[I_2 | I_2]$ on the columns of sum i and j . Furthermore, because $i > j$, $|B_i| > |B_j|$, so there is some row $b \in B_i$, $b \notin B_j$. Therefore b must be all 1's in the columns of column sum i and mostly 0's in

the columns of column sum j . Combining row b with the rows in the intersection of $A_i \cap A_j$, we create F_7 , a contradiction. So $|A_i \cap A_j| = 1$.

We now consider B_i, B_j where $|B_i| > |B_j|$. If $B_j \setminus B_i \neq \emptyset$ then $B_i \setminus B_j \neq \emptyset$. We find a copy of F_7 on rows p, q, r where $p = A_i \cap A_j$, $q \in B_j \setminus B_i$ and $r \in B_i \setminus B_j$. So we may assume $B_j \subset B_i$.

Among the column sums of type T , choose three with $i > j > k$ and we may assume i is the largest column sum. Assume $A_i \cap A_j = \{a\}$, (a row index) and $A_i \cap A_k = \{b\}$. If $a \neq b$, then for some $c \in B_i \setminus B_j$, we find a copy of F_7 on the rows a, b, c . Recall that $B_k \subset B_j \subset B_i$ and so such a c can be chosen. Choosing two columns of sum i that have an I_2 on rows a and b , one column of sum j with a 1 in row a (and necessarily a 0 in row b) and one column of sum k with a 1 in row b , will generate an F_7 on rows a, b, c . So we may assume $a = b$.

With these properties in hand for A_T we now prove that $\|A_T\| \leq 2m$ by induction on m . In particular the sets A_t for t of type T , have the properties that $|A_i| \geq 3$ and $|A_i \cap A_j| \leq 1$. They can be ordered by the index i so that if we have three indices $i > j > k$ and $A_i \cap A_j = \{a\}$ and $A_i \cap A_k = \{b\}$, then $a = b$. By choosing the largest index i we reduce the family by eliminating A_i and the remaining sets are on the elements $[m] \setminus (A_i \setminus a)$ have the same properties so we can apply induction. Thus $\sum_t |A_t| = |A_i| + 2(m - |A_i| + 1) \leq 2m$. ■

Theorem 3.1.11. *We have that $\text{forb}(m, \{F_3, F_9\})$ is $O(m)$.*

PROOF. We establish that $\text{forb}(m, \{F_3, F_9\}) \leq 12m$. Let $A \in \text{Avoid}(m, \{F_3, F_9\})$. Let A_T denote the columns where columns of sum i form the structure T_i from (3.1) and let A_S denote the columns where columns of sum i form the structure S_i . Then $\|A\| \leq 2m + \|A_T\| + \|A_S\|$, where there at most 2 columns of column sum i not in A_T or A_S . We will show $\|A_T\| \leq 5m$ and then we deduce $\|A_S\| \leq 5m$ by taking (0,1)-complements (we note that F_3^c is F_3 and $F_9^c = F_9$ as configurations). This yields the bound.

We have that $\|A_T\| = \sum_i |A_i|$ where the sum is only over those i for which A_T has columns of sum i . We note that for $i < j$, we cannot have $|B_i \setminus B_j| \geq 2$. Else let $p, q \in B_i \setminus B_j$. Then since $|A_j| \geq 3$, there is always a column β of sum j with 0's in rows p, q . Now a column α of sum i has 1's in rows p, q . We deduce that $[\alpha \beta]$ has F_9 . Thus for $i < j$, we have $|B_i \setminus B_j| \leq 1$.

Now assume $\sum_i |A_i| \geq 5m$. Then there is a quintuple $i_1 < i_2 < i_3 < i_4 < i_5$

with some row $r \in A_{i_1} \cap A_{i_2} \cap A_{i_3} \cap A_{i_4} \cap A_{i_5}$. We will consider i_1, i_3, i_5 . We have $|B_{i_3}| \leq |B_{i_5}| - 2$ and so let $s, t \in B_{i_5} \setminus B_{i_3}$. Now $|B_{i_1}| \leq |B_{i_3}| - 2$ and yet $|B_{i_1} \setminus B_{i_3}| \leq 1$. Thus without loss of generality we may assume $s \notin B_{i_1}$. We now claim that we have F_3 in rows r, s . In the columns of sum i_5 with $r \in A_{i_5}$ and $s \in B_{i_5}$ we find

$$\begin{array}{r} r \\ s \end{array} \begin{bmatrix} 0 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

Now in both the columns of sum i_1 and also in columns of sum i_3 , we may find

$$\begin{array}{r} r \\ s \end{array} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix},$$

using the fact that $i_1 \notin B_{i_1}$ and $i_3 \notin B_{i_3}$. This yields a copy of F_3 and so we may deduce $\sum_i |A_i| \leq 5m$ which completes our proof. ■

3.2 Other linear results

Based on a result of Balogh, Keevash and Sudakov, we have the following theorem:

Theorem 3.2.1. [10] *Let $k \geq 2$ be given. Then $\text{forb}(m, \{I_k, T_k\})$ is $\Theta(m^{k-2})$*

From this theorem we can conclude the following.

Corollary 3.2.2. *We have that $\text{forb}(m, \{I_3, T_3\})$ is $O(m)$.*

Note that Conjecture 1.3.5 states that for single configurations, if $\text{forb}(m, F)$ is $\Theta(m^k)$, then there is a construction that avoids F consisting of a k -fold product of three matrices: I, I^c and T . Thus, if this conjecture were true for forbidden families, if a family of configurations was contained in all possible two-fold products of I, I^c and T , then its bound would have to be linear. The conjecture remains unproven (and likely requires a different statement for forbidden families - see Chapter 4), but its statement suggests which forbidden families would likely have a linear upper bound, a bound which must then be proved.

Chapter 4

Families With a Super-linear Bound

The first section of this chapter will contain a proof that refines a previous result, showing that a certain family of matrices defined as the product of 2×2 matrices has a bound of $\Theta(m^{3/2})$. In the second section of the chapter we will present another family of matrices, also obtained by taking the product of small 2-rowed matrices. We conjecture that this family has the same $\Theta(m^{3/2})$ bound and present the outline for a possible proof.

4.1 Unexpected bound for product configurations

4.1.1 Introduction

Conjecture 1.3.5 implies that for a single forbidden configuration, the power of m in the asymptotic bound is an integer. A logical question might be whether a similar (or even the same) conjecture applies to the asymptotic bound of forbidden families. However when considering a family of two configurations, Anstee and Sali found a bound of $\Theta(m^{3/2})$. The general idea is as follows. Let $ex(m, H)$ is the maximum number of edges in a (simple) graph G on m vertices that has no subgraph H . $A \in \text{Avoid}(m, \mathbf{1}_3)$ will be a matrix with up to $m+1$ columns of sum 0 or sum 1 plus columns of sum 2 which can be viewed as the vertex-edge incidence matrix of a graph.

Assume $p = |V(H)|$ and $q = |E(H)|$. Let $I(H)$ denote the $p \times q$ vertex-edge incidence matrix associated with H .

Remark 4.1.1. $\text{forb}(m, \{\mathbf{1}_3, I(H)\}) = m + 1 + \text{ex}(m, H)$.

We first encountered this with $I_2 \times I_2$ which is $I(C_4)$, the vertex-edge incidence matrix of the cycle of length 4.

Theorem 4.1.2. [1] $\text{forb}(m, \{\mathbf{1}_3, C_4\}) = m + 1 + \text{ex}(m, I(C_4))$ which is $\Theta(m^{3/2})$.

However we explored the result from the point of view of $I_2 \times I_2$ and various product constructions. This resulted in the statement and proof of Theorem 1.3.14. The matrices used in Theorem 1.3.14 are the products of a small identity matrix and a small triangular matrix as follows:

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad T_2 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix},$$

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

A preliminary version of [7] contained Theorem 1.3.14, which proved that $\text{forb}(m, \{I_2 \times I_2, T_2 \times T_2, I_2 \times T_2\})$ is $\Theta(m^{3/2})$. However, a closer examination of this family revealed that while there are $\Omega(m^2)$ constructions for the families $\{I_2 \times I_2, I_2 \times T_2\}$ and $\{T_2 \times T_2, I_2 \times T_2\}$, the best construction that could be achieved for the family $\{I_2 \times I_2, T_2 \times T_2\}$ was $\Omega(m^{3/2})$. This led to the conjecture of the theorem that follows, which we were able to prove by extending the result of Anstee, Raggi and Sali for $\mathcal{F} = \{I_2 \times I_2, T_2 \times T_2, I_2 \times T_2\}$.

4.1.2 Theorem and proof

Theorem 4.1.3. *We have $\text{forb}(m, \{I_2 \times I_2, T_2 \times T_2\})$ is $\Theta(m^{3/2})$.*

Before we proceed to the main body of the proof, we will make a few remarks and present three lemmas.

The proof will use a standard induction argument (recall from Section 1.4.1). Given our two forbidden configurations $I_2 \times I_2$ and $T_2 \times T_2$, when performing the

standard decomposition using any row r in A , then C_r cannot contain the following inductive children:

$$F_1 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix}, \quad F_2 = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 \end{bmatrix}.$$

Configuration F_1 in C_r will produce $T_2 \times T_2$ and configuration F_2 in C_r will produce $I_2 \times I_2$.

Lemma 4.1.4. *We have $\text{forb}(m, \{F_1, F_2\}) \leq 2m$.*

PROOF. Let $A' \in \text{Avoid}(m, \{F_1, F_2\})$ be a simple matrix. We proceed by standard induction. In order to avoid F_2 , the columns of C' must avoid its inductive child, I_2 . This implies that the columns of C' must be pairwise comparable (recall the definitions of Section 1.2.4), creating a tower matrix. By definition, C' is simple so all the columns must be unique. If $\|C'\| \geq 3$, there will be an F_1 contained in two of the columns. Thus $\|C'\| \leq 2$, and by induction A' has at most $2m$ columns. ■

It is worth noting that in order to achieve the bound of $\|C'\| = 2$ in the previous proof, it is necessary for the columns in C' to be the column of 0's and a column of sum 1. Of course, there may be only one of these columns present, but for our purposes there must be at least one, as a further restriction will preclude that C' is empty.

Definition 4.1.5. *Let A, r , be given with the decomposition (1.1). If there is a row r' in a subset of rows $S \subseteq [m]$ such that r' can be deleted and simplicity maintained in $C_r|_{S \setminus r'}$, then we call r' deletable.*

Definition 4.1.6. *Let A, r , be given with the decomposition (1.1). We define $R(r) \subseteq [m]$ to be a minimal set of rows with the property that $C_r|_{R(r)}$ is simple, with no deletable rows.*

We use the previous lemma to prove the following lemma about the size of $R(r)$.

Lemma 4.1.7. *Let A, r , be given with the decomposition (1.1). Given $R(r)$ as described in Definition 4.1.6, we have that $|R(r)| \geq \|C_r\|/2$.*

PROOF. Consider the matrix $C_r|_{R(r)}$. Note that $\|C_r|_{R(r)}\| = \|C_r\|$. We know that $C_r|_{R(r)}$ must avoid $\{F_1, F_2\}$ (because C_r avoids $\{F_1, F_2\}$) and so by Lemma 4.1.4, $\|C_r|_{R(r)}\| \leq 2|R(r)|$, which becomes $|R(r)| \geq \|C_r\|/2$, as desired. ■

Lemma 4.1.8. *Let A, r , be given with the decomposition (1.1). Given $R(r)$ as described in Definition 4.1.6, we can choose a set of rows $S(r) \subseteq R(r)$ such that $C_r|_{S(r)}$ contains $I_{|S(r)|}$ and $|S(r)| \geq |R(r)|/2$.*

PROOF. We choose $R(r)$ as described for the proof of Lemma 4.1.7. Let $s \in R(r)$ and perform the standard decomposition of $C_r|_{R(r)}$ using row s into E_s, G_s , and H_s ; see (4.1)

$$C_r|_{R(r)} = \begin{matrix} s \\ R(r) \setminus s \end{matrix} \begin{bmatrix} 0 \dots 0 & 1 \dots 1 \\ E_s & G_s & G_s & H_s \end{bmatrix}. \quad (4.1)$$

This scenario is analogous to the situation in Lemma 4.1.4, with $C_r|_{R(r)}$ as A' , E_s as B' , G_s as C' and H_s as D' . Therefore, as seen in the proof of Lemma 4.1.4, G_s is restricted to at most two columns, in fact, a column of sum 0, a column of sum 1 or both. Furthermore, in our case, with C_r restricted to the rows $R(r)$, G_s cannot be empty. Otherwise s could be deleted while preserving the simplicity of $C_r|_{R(r)}$, a contradiction to how we chose $R(r)$. Thus, in the matrix $C_r|_{R(r)}$, each row $s \in R(r)$ is associated with one or both of the following column types:

Case 1. There is a zero column in G_s , which yields a column of sum 1 in $C_r|_{R(r)}$ with its 1 in row s .

Case 2. There is a column of sum 1 in G_s , which yields a column of column sum 2 with 1's in rows s and some row t and a column of column sum 1 with a 1 in row t .

To choose our set $S(r) \subseteq R(r)$, with $I_{|S(r)|} \prec C_r|_{S(r)}$ and $|S(r)| \geq |R(r)|/2$, we will generate a directed graph on the rows of $R(r)$ and then use properties of the graph to choose an appropriate $S(r)$. We define D to be a directed graph, where there is an directed edge from vertex $s \rightarrow t$ if in the decomposition of $C_r|_{R(r)}$ based on row s , Case 2 occurs. We define the set T as the set of all rows $t \in R(r)$ such that $C_r|_{R(r)}$ has a column of column sum 1 with a 1 in row t . Let $U = R(r) \setminus T$. For each edge $u \in U$, there is an edge $u \rightarrow v$, where, by our choice of T , $v \in T$. Let V be a subset of T such that each vertex in V has an associated u -vertex in U , i.e. let $V = \{v \in T : \text{there is a } u \in U \text{ with } u \rightarrow v\}$. Finally, we define $S(r) = U \cup (T \setminus V)$. In words, we have chosen a subset of rows in $R(r)$, where each row is associated with

a column of sum 2 (U), but the row containing the second 1 (V) has been omitted. We have also included some extra rows associated with columns of column sum 1 that are distinct from the columns associated with this first set ($T \setminus V$). Thus each row in $S(r)$ is associated with a distinct column of column sum 1 and so $C_r|_{S(r)}$ contains an $I_{|S(r)|}$, potentially with some extra or repeated columns. We can confirm that $|S(r)| \geq |R(r)|/2$ by noting that in fact $S(r) = U \cup (T \setminus V) = R(r) \setminus V$, so $|S(r)| = |R(r)| - |V|$, and as $|V| \leq |U|$, this gives $|S(r)| \geq |R(r)|/2$. ■

With these lemmas proved, we can now proceed with proof of Theorem 4.1.3.

PROOF OF THEOREM 4.1.3. By a previous result of Anstee, Raggi and Sali, we know that $\text{forb}(m, \{I_2 \times I_2, I_2 \times T_2, T_2 \times T_2\})$ is $\Theta(m^{3/2})$. The construction that achieves this bound is as follows.

For each p a prime power, there is known to exist a projective plane of order p which consists of $p^2 + p + 1$ subsets each in $\binom{[p^2+p+1]}{p+1}$ such that for any two subsets L_1, L_2 , we have $|L_1 \cap L_2| = 1$ and for each pair of elements a, b there is a unique subset $L_{a,b}$ containing both. A major open problem is whether there exist projective planes of non-prime power order but we do not concern ourselves with that question. It is natural to think of elements as points and subsets as lines. The resulting $(p^2 + p + 1) \times (p^2 + p + 1)$ element-set incidence matrix has $(p + 1)(p^2 + p + 1)$ 1's and no 2×2 submatrix of four 1's. We may create a $2(p^2 + p + 1) \times (p + 1)(p^2 + p + 1)$ simple matrix so that for each pair of a element/point a and subset/line L with $a \in L$, we have a column with a 1 in row a and a 1 in row $p^2 + p + 1 + L$. We verify that the resulting matrix has no $I_2 \times I_2$ nor $T_2 \times T_2$ and that the number of columns $(p + 1)(p^2 + p + 1)$ is $\Theta((2(p^2 + p + 1))^{3/2})$. We can pad this matrix with rows of 0's to obtain a matrix of order m with $\Theta(m^{3/2})$ columns.

Because the construction for this bound avoids $\{I_2 \times I_2, I_2 \times T_2, T_2 \times T_2\}$, it also avoids the family of our theorem $\{I_2 \times I_2, T_2 \times T_2\}$. Thus $\text{forb}(m, \{I_2 \times I_2, T_2 \times T_2\})$ is $\Omega(m^{3/2})$. It remains to show that the bound is tight, i.e., that $\text{forb}(m, \{I_2 \times I_2, T_2 \times T_2\})$ is $O(m^{3/2})$.

If we perform the standard decomposition on A , we would be done by induction if we could show that $\|C_r\| \leq cm^{1/2}$ for some row r and some constant c . For this proof, we will use $c = 20$. We will assume $\|C_r\| \geq 20m^{1/2}$ for all r and arrive at a contradiction as follows: we will show that we can associate matrix C_r with a set of rows $S(r)$ where $|S(r)| \geq \|C_r\|/4$. Then we show that we can choose a set $Q \subseteq S(r)$

with $|Q(r)| \geq |S(r)|/3$ such that for every choice $r_p, r_q \in Q$ we have $|S(r_p) \cap S(r_q)| \leq 5$. Thus for $t = 5/3m^{1/2}$ choices $r_1, r_2, \dots, r_t \in S(r)$ we show that $|S(r_i) \cap S(r_j)| < 6$ and so obtain disjoint sets of rows

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

If we sum the size of these disjoint sets of rows, we get:

$$5m^{1/2} + (5m^{1/2} - 5) + (5m^{1/2} - 10) + \dots + (5m^{1/2} - 5(t-1)) > m$$

a contradiction given that we have m rows.

Let r be a row in A so $\|C_r\| > 20m^{1/2}$. We first choose a set of rows $R(r) \subseteq [m]$ as described in Definition 4.1.6. By Lemma 4.1.7 we know that $|R(r)| \geq \|C_r\|/2$. We continue by choosing a set of row $S(r) \subset R(r)$ as described in Lemma 4.1.8; again, we know that $|S(r)| \leq |R(r)|/2$, thus $|S(r)| \geq \|C_r\|/4$, and by our choice of $S(r)$, we know that $C_r|_{S(r)}$ contains an $I_{|S(r)|}$.

We now focus our attention on the matrix $A|_{S(r)}$. In particular, if we perform the standard decomposition on A using a row $r_j \in S(r)$, we wish to know what happens on the columns of $C_{r_j}|_{S(r) \setminus r_j}$.

Let $r_j \in S(r)$. Perform the standard decomposition using r_j on the original matrix A and then restrict to the rows in $S(r) \cup r$. In the decomposition, the columns of $C_{r_j}|_{S(r)}$ must correspond to columns of A which appear with a 1 and 0 in row r_j and are the same elsewhere, in particular, on the rows $r \cup (S(r) \setminus r_j)$. We thus have columns in C_{r_j} as follows:

$$\begin{array}{c} r_j \\ S(r) \setminus r_j \\ r \end{array} \begin{bmatrix} 0 & 1 \\ \alpha & \alpha \\ a & a \end{bmatrix}. \quad (4.2)$$

We now consider $A|_{S(r)}$ in this arrangement. We will extract two conclusions from the structure in (4.2), first from the case where $a = 0$ and second from the case where $a = 1$.

We begin with the case where $a = 0$. Note that since $a = 0$ is in row r , the columns of C_{r_j} with this feature correspond to columns in $[B_r C_r]$ in the original decomposition (1.1), and thus, by Lemma 4.1.10, these columns are laminar.

Returning to the decomposition based on r_j , as shown in (4.2); suppose there exist

two non-zero choices for α , say $\beta \neq \gamma$. We have the following situation for columns in C_{r_j} :

$$\begin{array}{c} r_j \\ S(r) \setminus r_j \{ \\ r \end{array} \begin{bmatrix} 0 & 0 & 1 & 1 \\ \beta & \gamma & \beta & \gamma \\ 0 & 0 & 0 & 0 \end{bmatrix}. \quad (4.3)$$

We know that on the rows $S(r)$, the columns from (4.3) must form a laminar family, as they have a zero in row r and are thus in $[B_r C_r]_{S(r)}$. Columns 3 and 4 have 1's in common on row r_j and we deduce without loss of generality that $\beta \leq \gamma$ and $\beta \neq \gamma$. Now, considering columns 2, 3, and the fact that $\mathbf{0} \neq \beta$, we violate the nested family property of laminar families, a contradiction. Thus there is only one non-zero choice for α in C_{r_j} when $a = 0$. Furthermore, if α is repeated, it must have column sum 1. If α has column sum greater than 1 and it is repeated, it creates a $T_2 \times T_2$. This happens because A is a simple matrix and so on some row r' the columns of C_{r_j} must differ, yielding the following, with $T_2 \times T_2$ on the first three rows plus row r' .

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

We prove a slightly different result when $a = 1$. This case will be simpler since the fact that $a = 1$ will automatically yield a row of ones on row r , naturally generating a row of $T_2 \times T_2$. We will show that the columns in C_{r_j} must be disjoint on the rows $S(r) \setminus r_j$.

Suppose there are two non-zero columns β, γ in $C_{r_j}|_{S(r) \setminus r_j}$.

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

Now suppose β and γ both have entry 1 on some row $r' \in S(r) \setminus r_j$. Also, like above, because A is a simple matrix, there is some row r'' ($r'' \neq r, r', r_j$) where the

columns of C_{r_j} differ. The situation is then as follows:

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

This creates a $T_2 \times T_2$, therefore all of our non-zero columns α in $C_{r_j}|_{S(r) \setminus r_j}$, with 1's in row r cannot have a 1 entry on the same row.

Thus we have shown that for any column in $C_{r_p}|_{S(r)}$, where $r_p \in S(r)$, there are two results. If the entry of a column in row r is 0, then those columns can be $\mathbf{0}$ or a single non-zero column α : if α is of size 1 it may be repeated, otherwise there is only one copy. If the entry of a column in row r is 1, the columns of $C_{r_p}|_{S(r)}$ must be disjoint. We now use these facts to consider what occurs in the rows of $C_{r_p}|_{S(r) \setminus r_p}$ (denoted X below).

On the columns of C_{r_p} with a 0 in row r , we have two cases for the rows of $S(r) \setminus r_j$. If α is of size 1 and repeated, there will be one “bad” row, say $r_s \in S(r) \setminus r_p$, that contains multiple ones in these columns. The remaining rows will contain all zeroes on these columns. In the other case, if the non-zero α is of size greater than 1 and therefore not repeated, all rows in $S(r) \setminus r_p$ will have at most one 1. Thus for all rows in $S(r) \setminus r_p$ with the exception of the previously described “bad” row, the remaining rows will have at most one 1 on the columns of C_{r_p} .

On the columns of C_{r_p} with a 1 in row r , the rows of $S(r) \setminus r_j$ will contain only one 1 each.

We can now show that once we choose $r_p \in S(r)$ for all but one $r_q \in S(r) \setminus r_p$, $|S(r_p) \cap S(r_q)| < 6$. We decompose the matrix A based on row $r_p \in S(r)$:

$$\begin{array}{l} r_p \\ S(r) \setminus r_p \end{array} \begin{bmatrix} C_{r_p} & C_{r_p} \\ \mathbf{0}^T & \mathbf{1}^T \\ X & X \end{bmatrix}.$$

We choose r_p and r_q so that neither is a “bad” row for the other. By Lemma 4.1.9 we know that we can choose at least a third of the rows from $S(r)$ such that each row r_q is not a “bad” row for all other rows chosen. We decompose A first using r_p and

then using r_q .

$$S(r_p) \cap S(r_q) \begin{matrix} & C_{r_p} & C_{r_p} \\ r_p & \begin{bmatrix} 1 & 0 \\ \beta^T & \beta^T \\ I & I \end{bmatrix} \\ r_q & \end{matrix}, \quad S(r_p) \cap S(r_q) \begin{matrix} & C_{r_q} & C_{r_q} \\ r_p & \begin{bmatrix} \gamma^T & \gamma^T \\ 1 & 0 \\ I & I \end{bmatrix} \\ r_q & \end{matrix}.$$

Note that by our choice of r_p and r_q , β^T and γ^T contain at most two 1's each and therefore must be mostly 0's. We consider the columns in C_{r_p} that have entry 0 in row r_q . By our assumption about the size of $S(r_p) \cap S(r_q)$, we know these columns must contain an I_2 , underneath 1's in row r_p and 0's in r_q . We now consider the same columns in C_{r_q} . We can find at least two columns with 0 in row r_p and 1 in r_q , containing an I_2 underneath. Furthermore, these two pairs of columns are distinct, as they have different entries on r_p and r_q . When put together they create an $I_2 \times I_2$. Therefore $|S(r_p) \cap S(r_q)| \leq 5$.

We finally generate the series of disjoint sets described at the beginning of this proof and arrive at our contradiction. ■

Lemma 4.1.9. *Given a row $p \in R(r)$, we say row $q \in R(r) \setminus p$ is a bad row for row p if there is more than one column of C_p with a 0 in row r and a 1 in row q . Assume for each row $r \in (s(r))$ there is at most one bad row $s \in S(r)$. Then there is a set of rows $Q \subseteq S(r)$ such that $|Q| \geq |S(r)|/3$ and for any pair of rows $r_p, r_q \in Q$, we have r_p and r_q are not bad rows for each other.*

PROOF. We will interpret this question as a graph theory problem. We create a graph with $|S(r)|$ vertices and assign each row of $S(r)$ to a vertex. We assign a directed edge to two vertices (r_i, r_j) if r_j is a "bad" row for r_i . Because for each edge/row r_i there can only be one "bad" row r_j , the out-degree of any vertex v_i is at most one. A set of rows that has our desired property is equivalent to choosing an independent set on our graph. We can choose an independent set by coloring the graph—each color will correspond to an independent set of vertices.

We will show that we can three-color our graph. We use induction on the number of vertices. Suppose there exists a $v' \in V$ such that v' has in-degree 0. Let G' be the induced subgraph of $G \setminus v'$. Directed graph G' maintains our property of maximum out-degree 1, so we color G' with 3 colors and then add v' back with an appropriate color. Suppose there is no vertex $v \in V$ with in-degree 0. In this case all vertices

have in-degree 1 and out-degree 1, and so the edges of G form the union of disjoint cycles, which we can color with 3 colors.

Since each color corresponds to an independent set, we can choose an independent set that is at least a third of the remaining rows. ■

Lemma 4.1.10. *Assume we have obtained $S(r)$ as described in Lemma 4.1.8. Then on all triples of rows in $S(r)$, the columns in $[B_r C_r]$ must form a laminar family.*

PROOF. Recall our description of the “what is missing” method in Section 1.4.2. Using a computer program written in C++ by Miguel Raggi (download at <http://www.math.ubc.ca/~anstee/FConfThesisVersion.tar.gz>) we are able to compute what is missing on any quadruple of rows in A in order to avoid $I_2 \times I_2$ and $T_2 \times T_2$. There are five possibilities, and on any quadruple of rows in A , one must apply.

$$\begin{aligned}
 Q_0 &= \begin{array}{c} \textit{no} \quad \textit{no} \quad \textit{no} \\ \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} \quad \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \quad \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \end{array}, \quad Q_1 = \begin{array}{c} \textit{no} \quad \textit{no} \quad \textit{no} \quad \textit{no} \quad \textit{no} \\ \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} \quad \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix} \quad \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \quad \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix} \quad \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} \end{array}, \\
 Q_2 &= \begin{array}{c} \textit{no} \quad \textit{no} \quad \textit{no} \quad \textit{no} \\ \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} \quad \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} \quad \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \quad \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix} \end{array}, \quad Q_3 = \begin{array}{c} \textit{no} \quad \textit{no} \quad \textit{no} \quad \textit{no} \\ \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} \quad \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \quad \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} \quad \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \end{array}, \\
 Q_4 &= \begin{array}{c} \textit{no} \quad \textit{no} \quad \textit{no} \quad \textit{no} \quad \textit{no} \quad \textit{no} \\ \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} \quad \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} \quad \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \quad \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} \quad \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \end{array}.
 \end{aligned}$$

When we perform the standard decomposition of A based on row r and consider a quadruple of rows r, i, j, k , we can determine what is missing in C_r on the triple of rows i, j, k . We would be able to delete rows from C_r and preserve simplicity if we find a copy of “ K_2 ” in what is missing. That is, if on the triple i, j, k , there is a pair

of rows i, j with all 4 columns of K_2 appearing as follows:

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

where $a, b, c, d \in \{0, 1\}$, then we could delete row k from C_r and preserve simplicity, a contradiction. The reason for this is that on the three rows i, j, k , the only columns possibly present would be

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

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

We note that Q_3 and Q_4 both have the property that if any row is deleted, there will be a K_2 contained on two of the remaining three rows and so we can ignore the cases represented by Q_3 and Q_4 . In the remaining cases to avoid leaving a K_2 after deleting a row we must delete a certain row(s) from a quadruple. We note that Q_2 has K_2 on rows 1 and 2 and so to avoid leaving a copy of K_2 row r must either be row 1 or row 2 of such quadruples. Similarly, r must be row 1 of Q_1 and could be any row of Q_0 . We use P_i to denote a triple arising from the quadruple Q_i in these ways.

$$P_0 = \begin{array}{c} \text{no} \quad \text{no} \quad \text{no} \\ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \quad \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \quad \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \quad \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad \text{or} \\ \text{no} \quad \text{no} \quad \text{no} \\ \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \quad \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \end{array}$$

$$P_1 = \begin{array}{ccccc} \text{no} & \text{no} & \text{no} & \text{no} & \text{no} \\ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \end{array} \quad P_2 = \begin{array}{ccc} \text{no} & \text{no} & \text{no} \\ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \end{array} \quad \text{or} \quad \begin{array}{ccc} \text{no} & \text{no} & \text{no} \\ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \end{array}.$$

These then, are the possibilities for what is missing in A , including $A|_{S(r)}$. Because we know that the columns of A contain an identity matrix on the rows $S(r)$ we know that the final three cases of P_0 and the second set of case of P_2 can not be what is missing on for $A|_{S(r)}$.

Among the remaining options for what is missing (the first case for P_0, P_1 and P_2), by looking at Q_0, Q_1 and Q_2 , we see that below the zeroes in row r , each triple of rows in $S(r)$ is missing two columns of sum 2. This means that on all triples of rows in $S(r)$, the columns in $[B_r C_r]$ must form a laminar family. ■

4.2 A similar bound for larger product configurations

Extending the previous result, we considered another family of configurations formed by the product of small matrices. As before, these small matrices were taken from the three matrices of Theorem 1.3.6 and Conjecture 1.3.5, I, I^c , and T . In this case, we used the following three configurations:

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

Note that $E_1 \prec I_3, E_2 \prec T_3$ and $E_3 \prec I_3^c$.

We then forbid all possible products of these sub-configurations, namely the family $\{E_1 \times E_1, E_1 \times E_2, E_1 \times E_3, E_2 \times E_2, E_2 \times E_3, E_3 \times E_3\}$. If only five of these products are forbidden, there is a product construction of size $\Omega(m^2)$ that avoids the five included matrices. However, when all six are forbidden, the best lower bound is $\Omega(m^{3/2})$ from a construction of Anstee, Raggi and Sali [7]. This leads us to the following conjecture:

Conjecture 4.2.1. *Given the family $\mathcal{F} = \{E_i \times E_j : i, j = 1, 2, 3\}$ we have $\text{forb}(m, \mathcal{F})$*

is $\Theta(m^{3/2})$.

The lower bound has already been given, thus the main goal is to establish the upper bound that $\text{forb}(m, \mathcal{F})$ is $O(m^{3/2})$. We here outline some steps of a potential proof that we were unable to complete.

The argument will proceed by induction. Let $A \in \text{Avoid}(m, \mathcal{F})$. Then using the decomposition shown in (1.1), C_r cannot contain the following inductive children of our family:

$$F_1 = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}, \quad F_2 = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix},$$

$$F_3 = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 \end{bmatrix}.$$

Similarly to the last proof, we have two lemmas relating to the inductive children and rows of A .

Lemma 4.2.2. *We have $\text{forb}(m, \{F_1, F_2, F_3\}) \leq 2m$.*

PROOF. Let $A' \in \text{Avoid}(m, \{F_1, F_2, F_3\})$ be a simple matrix. We proceed by standard induction. In order to avoid $\{F_1, F_2, F_3\}$, the columns of C' must avoid the sub-configurations E_1, E_2 and E_3 . But because $E_1 \prec I_3$, $E_2 \prec T_3$ and $E_3 \prec I_3^c$, by Lemma 1.3.20, $\|C'\|$ must be constant. A simple pigeonhole argument proves that $\|C'\| \leq 2$. ■

Using Definition 4.1.6 as in the previous proof, we define $R(r) \in [m]$ to be a minimal set of rows with the property that $C_r|_{R(r)}$ is simple, with no deletable rows. The result of Lemma 4.1.7 again applies to $|R(r)|$, so $|R(r)| \geq \|C_r\|/2$.

Similarly to the previous proof, we want to show that for all r , $\|C_r\|$ is $O(m^{1/2})$, which yields an upper bound of $O(m^{3/2})$ for the matrix A by induction. We will attempt to do so by contradiction: assume that for all r that $\|C_r\|$ is $\Omega(m^{1/2})$ and then choose a subset of the rows in such a way that we arrive at a contradiction regarding the number of rows.

We now need to show that there exists some constant c (say $c = 1/3$) and some $S(r) \subseteq R(r)$ such that $|S(r)| \geq c|R(r)|$ and $C_r|_{S(r)}$ contains either a $I_{|S(r)|}, T_{|S(r)|}$, or

$I_{|S(r)|}^c$. Such a statement would parallel the statement of Lemma 4.1.8. This is the primary missing element of the proof. If such a statement could be proved, the rest of the proof would proceed as follows.

Supposing we can choose this $S(r)$ for all rows $r \in [m]$, there must be at least a third of the rows in $[m]$ such that $S(r)$ contains $I_{|S(r)|}$, a third of rows such that $S(r)$ contains $T_{|S(r)|}$, or a third of rows in $S(r)$ contain $I_{|S(r)|}^c$. Thus without loss of generality, we can choose a set $Q(r) \subset [m]$ of size $\Theta(m^{1/2})$ such that for all $r \in Q(r)$ $C_r|_{S(r)}$ contains a $I_{|S(r)|}$.

We now consider the standard decomposition of A based on rows $r_p, r_q \in Q(r)$. Suppose $|S(r_p) \cap S(r_q)| \geq 5$. If we decompose A based on r_p (and then r_q), we have the following:

$$S(r_p) \cap S(r_q) \begin{array}{c} r_p \\ r_q \end{array} \left[\begin{array}{cc|cc} C_{r_p} & & C_{r_p} & \\ \hline 00\dots & \dots 00 & 11\dots & \dots 11 \\ 0\dots 0 & 1\dots 1 & 0\dots 0 & 1\dots 1 \\ I & I & I & I \end{array} \right].$$

Using our assumption about $|S(r_p) \cap S(r_q)|$, and the pigeonhole principle, we can assume that one of the two following configurations will occur in the above decomposition:

$$S(r_p) \cap S(r_q) \left\{ \begin{array}{c} r_p \\ r_q \end{array} \left[\begin{array}{cccccc} 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array} \right] \right. \text{ or } \left. \begin{array}{c} r_p \\ r_q \end{array} \left[\begin{array}{cccccc} 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array} \right] \right\}. \quad (4.4)$$

We now consider a decomposition again using r_p and r_q ; however, this time the decomposition begins with r_q .

$$S(r_p) \cap S(r_q) \begin{array}{c} r_p \\ r_q \end{array} \left[\begin{array}{cc|cc} C_{r_q} & & C_{r_q} & \\ \hline 0\dots 0 & 1\dots 1 & 0\dots 0 & 1\dots 1 \\ 00\dots & \dots 00 & 11\dots & \dots 11 \\ I & I & I & I \end{array} \right].$$

Note that the columns of C_{r_q} with a 1 in row r_q will be disjoint from the columns

of the first configuration in (4.4) and the columns of C_{r_q} will be disjoint from the columns of the second configuration in (4.4).

Now the columns of C_{r_q} with a 1 in row r_q will be forced to contain one of the following configurations:

$$S(r_p) \cap S(r_q) \left\{ \begin{array}{l} \begin{bmatrix} 0 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ or } \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right\}. \quad (4.5)$$

These configurations are both disjoint from the first configuration in (4.4) and thus, taken together, produce either an $E_1 \times E_1$ in the first case or $E_1 \times E_2$ in the second case. If we consider the cases of (4.5) with 0's in row r_q (one of which must occur), these will be disjoint from the 2nd case listed above, and again, taken together, produce a $E_1 \times E_2$ or an $E_1 \times E_3$. In each case the E_1 occurs in the rows of $S(r_p) \cap S(r_q)$ and the other matrix of the product occurs in $r_p \cup r_q$.

Thus $|S(r_p) \cap S(r_q)| \leq 4$.

Finally, we consider the following sequence of sets. We assume that $|S(r)| > Km^{1/2}$ for some suitably large K . For all $r_i \in S(r)$, we form disjoint sets $S'(r_i)$ as follows:

$$S'(r_i) = S(r_i) \setminus \sum_{j=1}^{i-1} S(r_j).$$

Note that $S'(r_i), S'(r_j)$ are disjoint for all $r_i, r_j \in Q(r)$. We then consider the size of these sets:

$$|S'(r_1)| + |S'(r_2)| + |S'(r_3)| + \dots + |S'(r_t)|,$$

where $t = m^{1/2}$. This sum size of these sets is thus:

$$Km^{1/2} + (Km^{1/2} - 4) + (Km^{1/2} - 8) + \dots$$

which is greater than m for K suitably large, a contradiction.

Chapter 5

A Case Study: Families of 2 x 2 Configurations

When first investigating families of forbidden configurations, we chose to take the set of all seven 2×2 configurations (described and indexed in Table 5.1) and find $\text{forb}(m, \mathcal{F})$ for all $2^7 - 1$ non-trivial families of these configurations, with the goal of using these families as case studies for other families with larger configurations.

	Configuration	Bound	Construction	Contained in
F_1	$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$	$\binom{m}{2} + \binom{m}{1} + \binom{m}{0}$	$I_m^c \times I_m^c$	I_4, T_4
F_2	$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$	$m + 1$	$[I_m^c \mathbf{1}]$	I_3, T_3
F_3	$\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$	$m + 2$	$[I_m \mathbf{0} \mathbf{1}]$ or $[I_m^c \mathbf{0} \mathbf{1}]$	T_3
F_4	$\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$	$m + 1$	$[I_m \mathbf{0}]$ or $[I_m^c \mathbf{1}]$	T_3
F_5	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$m + 1$	$[T_m \mathbf{0}]$	I_2, I_2^c
F_6	$\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$	$m + 1$	$[I_m \mathbf{0}]$	T_2, I_3^c
F_7	$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$	$\binom{m}{2} + \binom{m}{1} + \binom{m}{0}$	$I_m \times I_m$	T_3, I_4^c

Table 5.1: List of 2 x 2 Configurations

These families proved both a good starting point for exploration, and a helpful way to demonstrate the use of our results and techniques used to prove bounds.

5.1 Constant families

We can use many results previously shown in this paper in order to simplify our classification of families. The first is Theorem 2.1.1, which allows us to identify families with a constant bound. For example, the family $\{F_1, F_7\}$ has the property that $F_1 \prec I_4$, $F_7 \prec T_3$, and $F_7 \prec I_4^c$ and thus by Theorem 2.1.1 $\text{forb}(m, \{F_1, F_7\})$ is $O(1)$. There are many pairs of configurations that have this property and thus a constant bound. They are as follows:

$$\begin{aligned} &\{F_1, F_5\}, \{F_1, F_6\}, \{F_1, F_7\}, \{F_2, F_5\}, \{F_2, F_6\}, \\ &\{F_2, F_7\}, \{F_3, F_5\}, \{F_4, F_5\}, \{F_5, F_6\}, \{F_5, F_7\}. \end{aligned} \tag{5.1}$$

Now by Remark 1.3.19, we know that $\text{forb}(m, \mathcal{F})$ must be no larger than $\text{forb}(m, \mathcal{F}')$ for all $\mathcal{F}' \subseteq \mathcal{F}$. Thus any family that contains one of the families in (5.1) can have at most a constant bound.

We can determine the exact bound for families of 2×2 configurations that have a constant bound. We will prove bounds for ten minimal families of configurations taken from Table 5.1. Any family with a constant bound will contain a sub-family identical to one of our ten minimal families for which we have proved an exact upper bound. By Remark 1.3.19, the bound of the entire family will be the smallest bound of the subfamilies; thus the bound for a family with a constant bound will be given by one of our ten minimal families.

Proposition 5.1.1. *Let $m \geq 2$. We have $\text{forb}(m, \{F_1, F_5\}) = \text{forb}(m, \{F_3, F_5\}) = \text{forb}(m, \{F_5, F_7\}) = 3$.*

PROOF. Let $A \in \text{Avoid}(m, \mathcal{F})$, for one of the families above. Because our matrix A must avoid F_5 , our columns must be pairwise comparable, so that for any $\alpha, \beta \in A$, $\alpha \leq \beta$ or $\beta \leq \alpha$. Thus our matrix will be a selection of columns chosen from the triangular matrix $[T_m | \mathbf{0}]$. If we choose four or more columns from this matrix, on some subset of rows there will be a T_4 or $[T_3 | \mathbf{0}]$, both of which contain F_1, F_3 , and F_7 . Thus the bound must be less than 4. The following constructions achieve the bound

of 3:

$$\{F_1, F_5\} : \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \\ \vdots & \vdots & \vdots \end{bmatrix} \quad \{F_4, F_5\} : \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 0 \\ \vdots & \vdots & \vdots \\ 1 & 0 & 0 \end{bmatrix} \text{ or } \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ \vdots & \vdots & \vdots \\ 1 & 1 & 0 \end{bmatrix}.$$

The construction for $\{F_5, F_7\}$ is the complement of the construction for $\{F_1, F_5\}$, per Remark 1.3.16. ■

Proposition 5.1.2. *We have $\text{forb}(m, \{F_4, F_5\}) = \text{forb}(m, \{F_2, F_5\}) = \text{forb}(m, \{F_5, F_6\}) = 2$.*

PROOF. Let $A \in \text{Avoid}(m, \mathcal{F})$, for one of the families above. Similarly to the previous proof, in order to avoid F_5 , A will be taken from columns of the triangular matrix $[T_m|\mathbf{0}]$. But any three columns chosen from $[T_m|\mathbf{0}]$ will include the submatrices T_3 or $[T_2|\mathbf{0}]$, both of which contain F_2, F_4 , and F_6 . Thus the bound must be less than 3. The constructions that achieve the bound of for $\{F_4, F_5\}$ are as follows:

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ \vdots & \vdots \\ 1 & 1 \end{bmatrix}.$$

The construction that achieves the bound of 2 for $\{F_2, F_5\}, \{F_5, F_6\}$ is $[\mathbf{1}|\mathbf{0}]$. ■

Proposition 5.1.3. *Let $m \geq 7$. Then $\text{forb}(m, \{F_1, F_7\}) = 2$.*

PROOF. For proof of this bound, see Section 2.3 and Theorem 2.3.4. For any column α , a construction of $[\alpha|\alpha^c]$ achieves the bound. ■

Proposition 5.1.4. *We have $\text{forb}(m, \{F_2, F_6\}) = 2$.*

PROOF. Let $A \in \text{Avoid}(m, \{F_2, F_6\})$. Suppose our matrix A has 3 columns. Without loss of generality, we can suppose the first row has two 1's and one 0. Because A is simple, for some row below the columns containing these two 1's, these columns must contain a $[10]$. However, such a matrix contains F_6 . A similar argument works for the complement. Thus $\text{forb}(m, \{F_2, F_6\}) < 3$. For any column α , a construction of $[\alpha|\alpha^c]$ achieves the bound. ■

Proposition 5.1.5. *We have $\text{forb}(m, \{F_1, F_6\}) = \text{forb}(m, \{F_2, F_7\}) = 2$.*

PROOF. We will demonstrate the argument for $\text{forb}(m, \{F_1, F_6\}$; $\text{forb}(m, \{F_2, F_7\})$ follows by Remark 1.3.16. Let $A \in \text{Avoid}(m, \{F_1, F_6\})$. Suppose our matrix A has 3 columns. Suppose that the first row has two 1s and one 0. Then A cannot avoid F_6 by the same argument as the previous family. Now suppose the first row has one 1 and two 0s. We consider the two columns α and β containing the 0s. After the first row, $[\alpha|\beta]$ cannot contain $[00]$ on any row, otherwise A will contain F_1 . In addition, because A is simple, on some row $[\alpha|\beta]$ must contain $[01]$. Therefore $[\alpha|\beta]$ can never contain $[11]$ on a row (to avoid F_6). In the row containing $[10]$ in columns $[\alpha|\beta]$, the third column must contain a 0 in order to avoid F_6 . However, if this row is repeated more than once, we have F_1 . Thus an $A \in \text{Avoid}(m, \{F_1, F_6\})$ with 3 columns is not possible and $\text{forb}(m, \{F_1, F_6\}) < 3$. For any column α , a construction of $[\alpha|\alpha^c]$ achieves the bound. ■

Proposition 5.1.6. *For $m \geq 5$, We have $\text{forb}(m, \{F_1, F_4, F_6\}) = \text{forb}(m, \{F_1, F_4, F_7\}) = \text{forb}(m, \{F_2, F_4, F_6\}) = \text{forb}(m, \{F_2, F_4, F_7\}) = 1$.*

PROOF. Suppose there is a two-column construction A avoiding these families. There are four options for the rows of A : $[00]$, $[11]$, $[01]$ or $[10]$. If any one of these rows is repeated more than once, we will have a F_1, F_7, F_4 , or F_4 respectively. Thus for the family $\{F_1, F_4, F_6\}$, the only row that can be repeated is $[11]$. However, because this is a simple matrix, at some point we must have the row $[01]$ or $[10]$. This generates an F_6 , which is forbidden. A similar argument can be used for the other families listed above.

Any column α will achieve the bound. ■

5.2 Linear families

We know from Proposition 2.1.2 that if a family does not have a constant bound, its bound is at least linear. Thus all families of 2×2 configurations that do not have a constant bound must have at least a linear bound. What is the upper bound of these families?

In fact, most of the 2×2 configurations listed in Table 5.1 have a linear bound. Thus by Remark 1.3.19, we know that $\text{forb}(m, \mathcal{F})$ is $O(m)$ for any \mathcal{F} containing

F_2, F_3, \dots, F_6 . The only family with the potential to have a greater than linear bound would be the two configurations with a quadratic bound, namely $\{F_1, F_7\}$, which we have already shown to have a constant bound. Thus for all families of size greater than 1, $\text{forb}(m, \mathcal{F})$ has at most a linear upper bound. Therefore all families of 2×2 configurations that do not have a constant bound must have a linear bound, i.e. are $\Theta(m)$.

Carefully examining the list of pairs of configurations that give a constant bound, we find that the only families of 2×2 configurations that do not contain a family from our list of constant-bounded families (5.1) are the families $\{F_1, F_2, F_3, F_4\}$ or $\{F_3, F_4, F_6, F_7\}$ and all of their subfamilies. To show that $\text{forb}(m, \mathcal{F})$ is $\Theta(m)$ for these families, it remains to show that they admit a linear construction. We will do so by giving exact bounds for two scenarios.

First, we consider all possible pairs of 2×2 configurations that do not yield a constant bound. These pairs are contained in the family of either $\{F_1, F_2, F_3, F_4\}$ or $\{F_3, F_4, F_6, F_7\}$, or alternatively, are not contained in (5.1).

$$\{F_1, F_2\}, \{F_1, F_4\}, \{F_2, F_3\}, \{F_2, F_4\}, \{F_3, F_4\} \quad (5.2)$$

$$\{F_3, F_4\}, \{F_3, F_6\}, \{F_4, F_6\}, \{F_4, F_7\}, \{F_6, F_7\} \quad (5.3)$$

$$\{F_1, F_3\}, \{F_3, F_7\} \quad (5.4)$$

Remark 5.2.1. *The families in (5.2) and (5.3) have a bound of $m + 1$.*

The families listed in (5.2) and (5.3) have at least one configuration with a bound of $m + 1$. Thus the bound for each of the families listed in (5.2) and (5.3) is $m + 1$, a consequence of Remark 1.3.19. In all cases there is also a linear construction that achieves the bound. In the case of families in (5.2), the matrix $[I_m^c | \mathbf{1}]$ avoids each pair of configurations, as well as any union of them. Similarly, for families in (5.3), the matrix $[I_m | \mathbf{0}]$ avoids each family and any union of them.

Remark 5.2.2. *The families in (5.4) have a bound of $m + 2$.*

Each configuration individually has a bound of $m + 2$, enforcing a maximum upper bound of $m + 2$ (again by Remark 1.3.19) and the configurations $[I_m^c | \mathbf{1} | \mathbf{0}]$ and $[I_m | \mathbf{1} | \mathbf{0}]$ achieve this bound by avoiding $\{F_1, F_3\}$ and $\{F_3, F_7\}$ respectively.

5.3 Table of results

The following table (Table 5.2) summarizes the results preceding this section. These results should completely characterize $\text{forb}(m, \mathcal{F})$ where \mathcal{F} is a family of 2×2 configurations. Given such a family \mathcal{F} , the first line of the table which contains a subfamily contained in \mathcal{F} gives the bound and construction for the entire family \mathcal{F} .

\mathcal{F}	$\text{forb}(m, \mathcal{F})$	Construction	Proof
$\{F_1, F_4, F_6\}, \{F_1, F_4, F_7\},$ $\{F_2, F_4, F_6\}, \{F_2, F_4, F_7\}$	1	any column	Proposition 5.1.6
$\{F_1, F_7\}$	2	$[\alpha \alpha^c]$	Theorem 2.3.4
$\{F_2, F_6\}$	2	$[\alpha \alpha^c]$	Proposition 5.1.4
$\{F_1, F_6\}, \{F_2, F_7\}$	2	$[\alpha \alpha^c]$	Proposition 5.1.5
$\{F_2, F_5\}, \{F_4, F_5\}, \{F_5, F_6\}$	2	Proposition 5.1.2	Proposition 5.1.2
$\{F_1, F_5\}, \{F_3, F_5\}, \{F_5, F_7\}$	3	Proposition 5.1.1	Proposition 5.1.1
$\{F_1, F_2\}, \{F_1, F_4\}, \{F_2, F_3\},$ $\{F_2, F_4\}, \{F_3, F_4\}$	$m + 1$	$[I^c \mathbf{1}]$	Remark 5.2.1
$\{F_3, F_4\}, \{F_3, F_6\}, \{F_4, F_6\},$ $\{F_4, F_7\}, \{F_6, F_7\}$	$m + 1$	$[I \mathbf{0}]$	Remark 5.2.1
$\{F_1, F_3\}$	$m + 2$	$[I_m^c \mathbf{0} \mathbf{1}]$	Remark 5.2.2
$\{F_3, F_7\}$	$m + 2$	$[I_m \mathbf{0} \mathbf{1}]$	Remark 5.2.2

Table 5.2: Summary of Results

5.4 Critical subfamilies

Our results here suggest an interesting parallel to previous results from the study of single forbidden configurations. When forbidding a single configuration F , the configuration may have a subconfiguration $F' \prec F$ such that $\text{forb}(m, F') = \text{forb}(m, F)$. If this F' is minimal, that is, if for all $F'' \prec F'$, $\text{forb}(m, F'') < \text{forb}(m, F)$, F' is called a *critical substructure* of F [6]. Such a sub-configuration is called critical, because the larger configuration's bound somehow depends on the bound of this substructure and its presence in the larger configuration.

For example, when forbidding the complete object K_4 , there are seven critical substructures, whose bound is the same as that of K_4 .

Theorem 5.4.1. [18] *The critical substructures of K_4 are $\mathbf{0}_4, \mathbf{1}_4, I_4, I_4^c, K_4^2, [2 \cdot \mathbf{0}_3], [2 \cdot \mathbf{1}_3]$.*

This means that forbidding, say, $\mathbf{1}_4$, has the same “power” as forbidding all of K_4 , in that $\text{forb}(m, \mathbf{1}_4) = \text{forb}(m, K_4)$.

We can extend this idea of critical substructures to our forbidden families, where a subfamily $\mathcal{F}' \subset \mathcal{F}$ is called a *critical subfamily* if $\text{forb}(m, \mathcal{F}') = \text{forb}(m, \mathcal{F})$ and \mathcal{F}' is minimal, that is, for all $\mathcal{F}'' \subset \mathcal{F}'$, $\text{forb}(m, \mathcal{F}'') > \text{forb}(m, \mathcal{F})$. This definition can also be seen as a further refinement of Remark 1.3.19, which states this very idea: that the upper bound of a family is restricted by the bounds of all its subfamilies.

We can see this idea clearly at work in the bounds established in the previous sections of this chapter. For example the family $\{F_1, F_2, F_3, F_4\}$ has a bound of $m + 1$, with critical subfamilies $\{F_2\}$, and $\{F_4\}$, as they each also have a bound of $m + 1$. Our implicit classification of all families of 2×2 configurations with constant bound also uses this principal: if a family \mathcal{F} contains at least one of our ten proven families as a subfamily, then that subfamily serves as a critical subfamily, restricting the bound for \mathcal{F} .

One of the results described in Section 1.3 was that $\text{forb}(m, C_3) = \text{forb}(m, \{C_3, C_4, C_5, \dots\})$ and $\text{forb}(m, C_3) = \text{forb}(m, \{C_3, C_5, C_7, \dots\})$. Thus $\{C_3\}$ is a critical subfamily of $\{C_3, C_4, C_5, \dots\}$ and $\{C_3, C_5, C_7, \dots\}$.

It should be noted that while critical subfamilies can be a useful way to determine bounds for larger families, the bound for a forbidden family may not always arise from a critical subfamily. In the case of $\{F_1, F_7\}$, $\text{forb}(m, \{F_1, F_7\})$ is constant. However, the only subfamilies of $\{F_1, F_7\}$ are the single configurations $\{F_1\}$ and $\{F_7\}$, both of which have a quadratic bound. Thus the bound in this case arises from the interaction of the two configurations and Theorem 1.3.6, not the bound of a critical subfamily.

Chapter 6

Conclusion

6.1 Summary of results

The first type of result in this paper is a classification, seen in Chapter 2 where we classify all families with a constant bound. While the classification theorem is based on a result not original to this paper, its conjecture and proof is a strong statement both about families of forbidden configurations and the interaction of the three key matrices of Conjecture 1.3.5. First, the classification allows us to not only identify, almost at a glance, which forbidden families have a constant upper bound, but tells us therefore which families have at least a linear construction. Furthermore, the strength and completeness of the result supports the premise of Conjecture 1.3.5, that somehow the matrices I_m, I_m^c and T_m are integral to understanding forbidden configurations.

However, the bulk of this paper has been devoted to establishing and proving both exact and asymptotic bounds for specific families of forbidden configurations. It is by no means an exhaustive survey, and we did not reach any other grand classification schemes; however, many of our results present interesting aspects of the forbidden configuration problem when it is extended to families. For example, the theorem and proof of Section 4.1 both show that the asymptotic bound can be a fractional exponent of m , not merely an integer exponent of m . The results regarding forbidden blocks of zeroes and ones in Section 2.3 illustrate the failed monotonicity of the (constant) bound for a family of forbidden configurations.

However, almost as important as the final bounds discussed in the paper are the techniques used to prove them. Proofs in Section 2.3 used basic combinatorics, a

bucketing approach from computer science, and a variation of linear programming. The proofs of certain linear bounds used a previous result about the structure of matrices that avoid a certain configuration (F_9 in Table 3.1). The proof of the bound in Section 4.1 is a fresh spin on standard induction, and also includes the method of “what is missing.”

Finally, the results in this paper suggest various other ideas paralleling previous work in single forbidden configurations (for example, the idea of critical subfamilies discussed in Section 5.4). These are not results per se, but we have noted all such correspondences, or other patterns of note.

In conclusion, this paper presents both refinements and extensions of previous work, as well as original results. The results will provide either tools or a jumping off point for continued research in the field. Thus this paper could be used as a helpful starting reference for anyone seeking to continue working in the field of forbidden configurations, as it combines a summary of previous important results, presents an important classification, provides illustration of most known proof techniques, and suggests areas for exploration.

6.2 Future work

There are several possibilities for future work regarding families of forbidden configurations. The first would be an extension of Conjecture 1.3.5 from single forbidden configurations to families. As alluded to in Chapter 4, Conjecture 1.3.5 fails in the case of forbidden families. However, there might be an alternative statement for which it holds. One example might be that the conclusions of Conjecture 1.3.5 hold if the configurations in a forbidden family represent all the inductive children of a single forbidden configuration. There may be another claim to be made about the size or type of a forbidden family on the asymptotic power of its bound.

Another avenue of exploration would be more theorems to classify families of a certain bound, like Theorem 1.3.6 and Theorem 2.1.1. Similarly, in Section 5.4, we discussed families which “force” a certain asymptotic bound. Part of classifying families may be finding all of these critical families or types of critical families for which a certain bound occurs.

In the proof of Theorem 4.1.3, we used a new technique based on standard induction and proof by contradiction. This technique was originally used in a similar proof,

and its adaptability to the new result indicates that it may be a useful addition to the standard repertoire of methods used to solve problems in forbidden configurations. In addition, the analysis of “what is missing” was a useful tool in determining various bounds.

In our results, we have often used pre-established results from the study of single forbidden configurations. In return, the study of forbidden families might be of use when studying single configurations. In particular, the use of standard induction can change the problem from avoiding a single configuration in A to avoiding a family of inductive children in C_r . Knowledge of certain common families could then help prove upper bounds for these single configurations.

Bibliography

- [1] R.P. Anstee. A survey of forbidden configuration results. *The Electronic Journal of Combinatorics*, 20:DS20, 2013. (Dynamic Survey).
- [2] R.P Anstee and M. Farber. Characterizations of totally balanced matrices. *Journal of Algorithms*, 5(2):215 – 230, 1984.
- [3] R.P Anstee, R. Ferguson, and A. Sali. Small forbidden configurations II. *Electronic Journal of Combinatorics*, 8(1):R4, 2001.
- [4] R.P. Anstee and B. Fleming. Linear algebra methods for forbidden configurations. *Combinatorica*, 31:1–19, 2011.
- [5] R.P. Anstee and Z. Füredi. Forbidden submatrices. *Discrete Math*, 62:225–243, 1986.
- [6] R.P. Anstee and S.N. Karp. Forbidden configurations: exact bounds determined by critical substructures. *Electronic Journal of Combinatorics*, 17, 2012.
- [7] R.P. Anstee, Christina Koch, Miguel Raggi, and A. Sali. Forbidden configurations and product constructions. Submitted to *Graphs and Combinatorics*.
- [8] R.P Anstee and A. Sali. Small forbidden configurations IV: The 3 rowed case. *Combinatorica*, 25(5):503–518, 2005.
- [9] J. Balogh and B. Bollobás. Unavoidable traces of set systems. *Combinatorica*, 25(6):633–643, 2005.
- [10] J. Balogh, P. Keevash, and B. Sudakov. Disjoining representability of sets and their complements. *Journal of Combinatorial Theory*, 95:12–28, 2005.
- [11] Laura Dunwoody. Notes. Research notes, 2005.

- [12] P. Erdős and M. Simonovits. A limit theorem in graph theory. *Studia Scientiarum Mathematicarum Hungarica*, 1:51–57, 1966.
- [13] P. Erdős and A.H. Stone. On the structure of linear graphs. *Bulletin of the American Mathematical Society*, 52:1089 – 1091, 1946.
- [14] Balin Fleming. Some families of forbidden configurations. Research notes, 2003.
- [15] P. Frankl, Z. Füredi, and J. Pach. Bounding one-way differences. *Graphs and Combinatorics*, 3:341–347, 1987.
- [16] H.O.F. Gronau. An extremal set theory problem. *Studia Scientiarum Mathematicarum Hungarica*, 15:29–30, 1980.
- [17] W. Mantel. Problem 28 solution. *Wiskundige Opgavem*, 10:60–61, 1907.
- [18] Miguel Raggi. *Forbidden Configurations*. PhD thesis, University of British Columbia, 2011.
- [19] H.J. Ryser. A fundamental matrix equation for finite sets. *Proceedings of the American Mathematical Society*, 34:332–336, 1972.
- [20] N. Sauer. On the density of families of sets. *Journal of Combinatorial Theory, Series A*, 13(1):145 – 147, 1972.
- [21] S. Shelah. A combinatorial problem stability and order for models and theories in infinitary languages. *Pacific Journal of Mathematics*, 41(1):247 – 261, 1972.
- [22] Paul Turán. Eine extremalaufgabe aus der graphentheorie. *Matematikai és Fizikai Lapok*, 48:436 – 452, 1941.
- [23] V.N. Vapnik and A. Ya. Chervonenkis. On the uniform convergence of relative frequencies of events to their probabilities. *Theory of Probability and Applications*, 25:264–280, 1971.