

Forbidden Configurations

by

Miguel Raggi

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

Doctor of Philosophy

in

THE FACULTY OF GRADUATE STUDIES

(Mathematics)

The University Of British Columbia

(Vancouver)

August 2011

© Miguel Raggi, 2011

Abstract

In this work we explore the field of Forbidden Configurations, a problem in Extremal Set Theory. We consider a family of subsets of $\{1, 2, \dots, m\}$ as the corresponding $\{0, 1\}$ -incidence matrix. For $\{0, 1\}$ -matrices F, A , we write $F \prec A$ if A has a submatrix which is a row and column permutation of F . We say a $\{0, 1\}$ -matrix is *simple* if it has no repeated columns. Let $\|A\|$ denote the number of columns of A . A $\{0, 1\}$ -matrix F with row and column order stripped is a *configuration*. Given $m \in \mathbb{N}$ and a family of configurations \mathcal{F} , our main function of study is

$$\text{forb}(m, \mathcal{F}) := \max\{\|A\| : A \text{ simple and for all } F \in \mathcal{F} \text{ we have } F \not\prec A\}.$$

We give a general introduction to the main ideas and previous work done in the topic. We develop a new more computational approach that allows us to tackle larger problems. Then we present an array of new results, many of which were solved in part thanks to the new computational approach. We use both new ideas and new spins on old ideas to tackle the problems. The new results include finding exact bounds on small configurations that were previously unknown, and proving some previously conjectured asymptotic bounds for “boundary” configurations. We also develop a relationship between Forbidden Configurations and Patterns, which we use to prove some results.

Preface

Most of the results in this thesis were done jointly with both Dr. Richard Anstee and Dr. Attila Sali and we produced four (submitted) papers. There are three papers with Dr. Anstee and Dr. Sali: “Evidence for a Forbidden Configuration Conjecture; one more case solved,” [ARS10a], “Forbidden Configurations: Quadratic Bounds,” [ARS11] and “Forbidden Configurations and Product Constructions,” [ARS10b]. There is one with Dr. Anstee, “Genetic Algorithms applied to problems of Forbidden Configurations,” [AR11]. The preprints for these papers can be downloaded from <http://www.math.ubc.ca/~anstee/>. A more detailed description is in order.

Chapter 1 and Chapter 2 have an introductory character and the ideas contained in them were thought of before I was involved.

Chapter 3 is mostly my doing, under the supervision of Dr. Anstee. I wrote all the code and developed the algorithms, but of course, with good advice and encouragement from my supervisor.

From Chapter 4, Section 4.1 (from [AR11]) and Section 4.2 (from [ARS10b]) were done mostly by me as well, again with good advice from Dr. Anstee. In particular I produced a list of conjectures for all unknown exact bounds for 3×4 configurations (using a Genetic Algorithm as described in Chapter 3), and Dr. Anstee gave me advice on which ones to pursue, as he felt were more likely to produce results. He was right. Section 4.3 was done jointly with Dr. Anstee.

All results from Chapter 5 (from [ARS10a] and [ARS11]) were done jointly with Dr. Anstee and Dr. Sali.

The results on Chapter 6 and Chapter 7 (from [ARS10b]) were started off by Dr. Anstee and myself, but were completed by Dr. Anstee and Dr. Sali in Hungary (although using some of my computations).

Contents

- Abstract ii
- Preface iii
- Contents iv
- List of Tables vii
- Acknowledgements viii
- Dedication ix

- 1 Introduction 1**
 - 1.1 Summary 1
 - 1.2 Extremal Problems 3
 - 1.3 Basic Definitions 4
 - 1.3.1 Matrices of 0's and 1's 4
 - 1.3.2 Configurations 10
 - 1.3.3 Basic Properties 14
 - 1.4 The Conjecture 15
 - 1.5 Table of Notation 25
 - 1.6 History 29
 - 1.6.1 Extremal Combinatorics and Set Theory 29
 - 1.6.2 Forbidden Configurations 32

- 2 Basic Techniques 34**
 - 2.1 Standard Induction 34
 - 2.2 What Is Missing? 37
 - 2.3 Implications 39

3	Computer Program Developed	42
3.1	Representation of a Configuration	43
3.2	Subconfigurations	45
3.3	Determining $X(F)$	45
3.4	Boundary Cases to Classify Configurations	47
3.5	Finding What Is Missing and Forbidden	48
3.6	Guessing Forb	53
3.6.1	Brute Force Greedy Search	55
3.6.2	Hill Climbing	55
3.6.3	Genetic Algorithm	57
4	Exact Bounds	59
4.1	Two 3x4 Exact Bounds	59
4.1.1	Introduction	59
4.1.2	The Bound for W	60
4.1.3	The Bound for V	69
4.2	Exact Bound for Ten Products	79
4.3	Critical Substructures	82
5	Three Asymptotic Bounds	87
5.1	Introduction	87
5.2	Quadratic Bound for a 4-rowed Configuration	89
5.2.1	What Is Missing?	91
5.2.2	Case Analysis	93
5.2.3	Linear Bound for the Inductive Children	100
5.3	Quadratic Bound for a 5-rowed Configuration	103
5.3.1	Applying Standard Induction	103
5.3.2	What Is Missing?	105
5.3.3	Linear Bound for the Inductive Children	107
5.4	Classification of 6-rowed Quadratic Bounds	118
6	Patterns and Splits	123
6.1	Patterns and Splits in 2-Dimensions	123
6.2	Patterns and Splits in d -Dimensions	128
7	Products	132

7.1	Introduction	132
7.2	Submatrices of TxT	135
7.3	Submatrices of IxI	139
7.4	Submatrices of IxT	140
7.5	Submatrices of IxTxIc	142
7.6	Fractional Exponent Bound for a Family of Configurations	143
7.7	All Pairs of Columns	150
8	Conclusions	152
8.1	Open Problems	152
8.1.1	Monotonicity	152
8.1.2	A Common 4-rowed Subconfiguration	154
8.1.3	Other 3x4 Exact Bounds	154
8.1.4	Critical Substructures	156
8.2	Concluding Remarks	159
	Bibliography	161
	Index	166

List of Tables

Table 1.1	Classification: 1 row.	19
Table 1.2	Classification: 2 rows.	19
Table 1.3	Classification: 3 rows.	20
Table 1.4	Classification: 4 rows.	23
Table 1.5	Definitions	28
Table 8.1	Conjectured values of $\text{forb}(m, V_i)$	155
Table 8.2	Conjectured values for $\text{forb}(m, F_3)$ and $\text{forb}(m, K_4)$	157
Table 8.3	Conjectured values for $\text{forb}(m, 2 \cdot K_4^2)$	157

Acknowledgements

First, I would like to thank my supervisor, Richard Anstee, for his endless patience and the incredible amount of time he spent working with me. It was fun.

I would also like to thank my committee members, Stephanie Van Willigenburg, and Joel Friedman, as well the University Examiners and the External Examiner, all for taking the time to go through my rather long (but quite fun) thesis. Also deserving of a thankful mention is my other co-author, Attila Sali.

I would also like to thank the Department of Mathematics at UBC. In particular, I would like to thank all professors, specially those who I took (at least) one course from: Richard Anstee, Kai Behrend, Patrick Brosnan, Jingyi Chen, Joel Feldman, Joel Friedman, Richard Froese, Christoph Hauert, Kalle Karu, Richard Kenyon, Kevin Leyton-Brown (from Comp. Sci.), Greg Martin, Wayne Nagata, Ed Perkins, Malabika Pramanik, Laura Scull, Andrew Rechnitzer, Dale Rolfsen, Denis Sjerve, Lior Silberman, Jozsef Solymosi and Vinayak Vatsal.

The acknowledgements wouldn't be complete without thanking all the people who were my friends during my time in Vancouver. By no means a comprehensive list: Yuri, Dennis, Hardeep, Ramon, Itzia, Ignacio, Felipe, and all the others.

In my long life I have met many nice people who have done me good. I'll be eternally grateful to all of them. In my undergraduate thesis acknowledgements (or "agradecimientos") I gave my best shot to mention as many of them as I could. To mention all of them again would undoubtedly make me hit the much dreaded 173 page mark, quite unlucky indeed. Therefore, I would like to say the following:

```
\input{agradecimientos.tex}
```

Some individuals, though, were so instrumental in my (alleged) success that I feel they deserve the mention twice. My family: Malú, Gerardo, Tanja, Daniel, Grola, O'Kuri, Medo, Gole and, even though she doesn't live with us anymore, Kima. I'm sure she is running happily in a farm upstate somewhere.

I feel like I'm missing someone... someone important...

Dedication

A Teresa. *Aunque esta tesis represente el tiempo que no estuve contigo,
terminarla significa poder regresar a ti.*

Chapter 1

Introduction

1.1 Summary

We give a brief description of the contents of each chapter of the thesis. For a more in-depth description, each chapter contains some introductory remarks. For this summary we try to use as few concepts as possible, but in order to describe the contents of some chapters it is unavoidable to use some key concepts defined in Section 1.3. Forbidden Configurations is at its core an extremal set theory problem: We attempt to maximize the size of a family of subsets of $\{1, 2, \dots, m\}$ subject to some restriction, namely that the family doesn't contain a particular object which we call a *configuration*. The function $\text{forb}(m, F)$ represents the maximum size of a family of subsets of $\{1, 2, \dots, m\}$ with no configuration F . For proper definitions, see Section 1.3.

We start in Chapter 1 by giving a general introduction and motivation to the field of Forbidden Configurations (in Section 1.2) and some basic definitions (in Section 1.3) that are used throughout the whole thesis. In Section 1.4 we describe a conjecture of Anstee and Sali (Conjecture 1.4.1) which has been a driving force in studying Forbidden Configurations. There is a brief history of the topic in Section 1.6 including related extremal problems.

In Chapter 2, we explain three techniques that have been quite fruitful in the field of Forbidden Configurations, namely those of Standard Induction, What Is Missing and Implications. We have greatly increased the reach of Standard Induction (and to a lesser extent, Implications) with new ways to use them in solving problems. We have increased the usability of What Is Missing by developing a computer program to do it automatically.

In Chapter 3 we describe the computer program we use to push the boundaries by increasing the size of the problems we are able to deal with. We'll often refer to this section

throughout the rest of the thesis. We used the software in three fundamentally different ways. First and most importantly, we used it to obtain relevant information about configurations (or family of configurations), by finding What Is Missing as described in Section 2.2. Second, we used some specific-purpose code to perform case analysis when we found it hard or tedious to do so by hand, helping in various proofs. Lastly, we used some local search strategies (described in Section 3.6) to guess extremal matrices and structures. Section 3.6 could be removed entirely and all proofs would still be correct, but a reader might be left wondering “How did they come up with this answer?”. This would be a perfectly valid question for which the only reasonable answer was “the computer told us so.”

Chapter 4 provides some new exact values for $\text{forb}(m, F)$ (see Definition 1.3.24) for some configurations F for which the bound was not previously known. In Section 4.1 in particular we prove exact bounds for two of the smallest configurations for which the answer wasn’t known. These results can be found in [AR11]. In Section 4.2 we prove an exact bound for a family of 10 configurations, a result that can be found in [ARS10b]. Finally, in Section 4.3 we give all *critical substructures* (see Definition 4.3.1) of K_4 , as well as a conjecture for all critical substructures of K_k . In this chapter we make heavy use the program described in Section 3.6 to guess extremal configurations and then proceed to prove they are indeed extremal, mostly using Standard Induction and What Is Missing followed by case analysis.

In Chapter 5 we give proofs of some *boundary cases* (see Definition 1.4.3), which serve as further evidence for our motivating Conjecture 1.4.1. We give the asymptotic bound for a 4-rowed configuration $F_8(t)$ in Section 5.2 (appears in [ARS10a]), a 5-rowed configuration F_7 in Section 5.3 and a 6-rowed configuration $G_{6 \times 3}$ in Section 5.4. This last one in particular is proven to be the unique quadratic 6-rowed boundary case. The last two results appear in [ARS11].

For Chapter 6 we give an entirely different extremal combinatorics problem with a more geometric flavour. This chapter is independent of all previous chapters, but is used in Chapter 7 for some proofs, as some problems considered in Chapter 7 can be interpreted as problems of patterns.

We dedicate Chapter 7 to the study of product constructions. We define a sister function to ‘forb’ and prove various interesting results about it. Then we prove an unexpected bound (for ‘forb’) of a family of configurations, which on a first glance would seem to be a counterexample of Conjecture 1.4.1 (but isn’t). Most of the results of Chapter 6 and Chapter 7 can be found in [ARS10b].

Finally, Chapter 8 contains some open problems, as well as some ideas and thoughts on

Forbidden Configurations and what the future might yield.

The results of Section 4.1.3 and Section 5.3 and its application to Section 5.4 provide a good sample of the flavour of the new results given this thesis.

1.2 Extremal Problems

A general extremal problem in combinatorics has the following character: For a given property P , what is the maximum size of a family that satisfies P ? For example, if the objects of study are families of subsets of $\{1, 2, \dots, m\}$ and property P is that no two subsets have two elements in common, the question becomes: What is the maximum number of subsets of $\{1, 2, \dots, m\}$ such that no two subsets have two elements in common? For this particular example we will give the answer later in Example 1.3.14.

In the field of Forbidden Configurations, we consider a particular extremal problem that arises as a generalization to celebrated extremal results in combinatorics, namely those of Erdős and Stone [ES46] and Erdős and Simonovits [ES66]. They consider the following problem: Given $m \in \mathbb{N}$ and a fixed (small) graph F , find the maximum number of edges in a (simple) graph G with m vertices that avoids having a subgraph isomorphic to F . For example, Mantel's result answers the question "what is the maximum number of edges in a triangle-free graph on m vertices?" and can be viewed as a special case. We will discuss this further in Section 1.6.

There are a number of ways to generalize the above problem to hypergraphs, but in this thesis we consider the following generalization: Given $m \in \mathbb{N}$ and a hypergraph F , find the maximum number of edges in a simple hypergraph H on m vertices that avoids having a subhypergraph (or *trace*) isomorphic to F . We consider the notion of subhypergraph as follows: The vertices of a subhypergraph are a subset of the vertices of the original hypergraph. For the edges, we choose a subset of the edges from the original hypergraph and consider the multiset consisting of intersections of the chosen edges with the vertices of the subhypergraph (more on this later).

As stated before, there are alternate generalizations of this problem. For example, instead of considering the edges of the subhypergraph to be intersections of edges of the original hypergraph with the vertices of the subhypergraph, one might consider a subhypergraph to have only edges that are fully contained in the chosen subset of the vertices. This alternate definition makes sense especially when considering k -uniform hypergraphs, in the sense that all edges have size k (a graph is a 2-uniform hypergraph). This alternate view of subhyper-

graphs is consistent with the way subgraphs of graphs are usually defined. In contrast, the way we see subhypergraphs when restricted to 2-uniform hypergraphs would suggest that when considering a subset of the vertices, we might have to consider “edges” of size 1 or size 0, even though we started with all edges of size exactly 2, since the intersection of an edge of size 2 with the vertices of the subhypergraph might have size 1 or 0. Moreover, the subhypergraph need not be simple, even if the original hypergraph was.

When studying difficult extremal problems, such as Forbidden Configurations, an exact answer to the extremal question might not come easily. In such cases, it’s common practice to settle for *asymptotic* bounds, given in Big-O notation: If $f, g : \mathbb{N} \rightarrow \mathbb{N}$, we define the following notation:

- We say f is $O(g)$ if there exists a constant $c > 0$ and a number N for which for all $n \geq N$, we have $f(n) < c \cdot g(n)$.
- We say f is $\Omega(g)$ if there exists a constant $d > 0$ and a number N for which for all $n \geq N$, we have $f(n) > d \cdot g(n)$.
- We say f is $\Theta(g)$ if f is both $O(g)$ and $\Omega(g)$. That is, if there exists constants $c > 0$, $d > 0$ and N for which for all $n > N$, $d \cdot g(n) < f(n) < c \cdot g(n)$. In such case we say f has the same *asymptotic growth* as g .

1.3 Basic Definitions

1.3.1 Matrices of 0’s and 1’s

First some standard notation. Let X be a finite set. We denote by 2^X the **power set** of X : the set that consists of all subsets of X (note that there are $2^{|X|}$ of them). Denote all subsets of X of a given size t by $\binom{X}{t}$.

A **set system** \mathcal{A} on a set X is a family of subsets of X . In other words, $\mathcal{A} \subseteq 2^X$. We might view a set system on X as a **simple hypergraph** on the vertices X .

We find it convenient to encode hypergraphs or set systems in the language of matrices. For a given order of the elements of X , and an ordering of the elements of a family \mathcal{A} , we can encode a set system as a matrix in the following way: A single subset $\alpha \subseteq X$ can be thought of as the incidence vector of 0’s and 1’s with $|X|$ entries, where the entry i is 1 if and only if $i \in \alpha$. The family \mathcal{A} gives rise to a collection of $\{0, 1\}$ -vectors with $|X|$ components. Consider these $\{0, 1\}$ -vectors as column vectors and concatenate them together to form a

$\{0, 1\}$ -matrix. Since the elements of X don't play any role in our investigations, we might as well take $X = [m] := \{1, 2, \dots, m\}$ with m being the size of X .

For example, the matrix

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

represents the family $\mathcal{A} = \{\{1\}, \{2, 3\}, \{1, 3\}, \{2, 3, 4\}, \emptyset\}$. Notice that in this encoding no two columns are equal: if two columns were equal, they would represent the same subset of X .

Different orderings of the columns of A give rise to different ways of ordering the sets of the family \mathcal{A} . We usually think of the set $[m]$ as having a canonical order, but we can think of different row orders as giving rise to *equivalent* hypergraphs. We will define the notion of equivalence of $\{0, 1\}$ -matrices formally in the next section, but for now let us talk about matrices a little longer.

Definition 1.3.1 We say a $\{0, 1\}$ -matrix A is **simple** if no two columns of A are equal.

The size of the family \mathcal{A} (i.e. the number of edges in the hypergraph) is represented by the number of columns of matrix A . Since we will refer to this number often, we find it convenient to create a short notation for it.

Definition 1.3.2 For a $\{0, 1\}$ -matrix A , we define $\|A\|$ to be the number of columns of A .

The size of each *edge* in the hypergraph is represented by the number of 1's in the corresponding column.

Definition 1.3.3 For a $\{0, 1\}$ -matrix A we denote by $\sigma_0(A)$ the number of 0's in matrix A and by $\sigma_1(A)$ the number of 1's. In the case of a column α , we say α has **column sum** k (and write $\sigma_1(\alpha) = k$) if it has exactly k ones.

For m -rowed A , we have

$$\sigma_0(A) + \sigma_1(A) = m \cdot \|A\|.$$

Definition 1.3.4 Define $\mathbf{0}_m$ to be the m -rowed column with all 0's and $\mathbf{1}_m$ to be the m -rowed column of all 1's.

We now define some operations of $\{0, 1\}$ -matrices that will be useful for constructing new matrices from previously constructed ones.

Definition 1.3.5 Let A be a $\{0,1\}$ -matrix. We denote A^c the $\{0,1\}$ -**complement** of A . That is, the matrix that results from replacing every 0 in A by a 1, and every 1 by a 0.

For example,

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

Definition 1.3.6 Let A and B be $\{0,1\}$ -matrices with the same number of rows. Define the **concatenation** $[A|B]$ to be the configuration that results from taking all columns of A together with all columns of B . For $t \in \mathbb{N}$, we define the product

$$t \cdot A := \underbrace{[A \mid A \mid \cdots \mid A]}_{t \text{ times}}.$$

The operation \cdot has precedence over $|$, so that $[t \cdot A \mid B] = [(t \cdot A) \mid B]$.

Definition 1.3.7 Let A and B be $\{0,1\}$ -matrices. We construct the product $A \times B$ by taking each column of A , and putting it on top of each column of B . So if the columns of A are $\alpha_1, \dots, \alpha_a$ and the columns of B are β_1, \dots, β_b , then $A \times B$ is the matrix with ab columns:

$$A \times B = \begin{bmatrix} | & | & & | & | & | & & | & & & | & | & & | \\ \alpha_1 & \alpha_1 & \dots & \alpha_1 & \alpha_2 & \alpha_2 & \dots & \alpha_2 & \dots & \dots & \alpha_a & \alpha_a & \dots & \alpha_a \\ | & | & & | & | & | & & | & & & | & | & & | \\ | & | & & | & | & | & & | & & & | & | & & | \\ \beta_1 & \beta_2 & \dots & \beta_b & \beta_1 & \beta_2 & \dots & \beta_b & \dots & \dots & \beta_1 & \beta_2 & \dots & \beta_b \\ | & | & & | & | & | & & | & & & | & | & & | \end{bmatrix}.$$

Here is an example of a product:

$$A = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}, B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \implies A \times B = \begin{bmatrix} 0 & 0 & | & 1 & 1 & | & 1 & 1 \\ 0 & 0 & | & 0 & 0 & | & 1 & 1 \\ \hline 1 & 0 & | & 1 & 0 & | & 1 & 0 \\ 0 & 1 & | & 0 & 1 & | & 0 & 1 \end{bmatrix}.$$

Often we will consider single columns and rows of $\{0,1\}$ -matrices.

Definition 1.3.8 For a column α and a $\{0, 1\}$ -matrix A , we define the **multiplicity** of α in A , written as $\lambda(\alpha, A)$, as the number of columns of A which are equal to α .

For example,

$$\lambda \left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \right) = 2.$$

Definition 1.3.9 Let A and B be m -rowed $\{0, 1\}$ -matrices. We define **subtraction** $A - B$ to be the m -rowed $\{0, 1\}$ -matrix such that for each m -rowed column α , $A - B$ satisfies

$$\lambda(\alpha, A - B) = \max\{0, \lambda(\alpha, A) - \lambda(\alpha, B)\}.$$

The order in which the columns appear in $A - B$ isn't important, but just so that it is a well defined operation of matrices, we might choose the order of the columns of $A - B$ to be the same as for A . This operation corresponds to set difference of the corresponding set systems.

Here is an example of the difference of two matrices:

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

Of the columns of A , the column $(0, 1, 0)^T$ has multiplicity 1 in A and 0 in B , so it must have multiplicity 1 in $A - B$. The column $(1, 0, 0)^T$ has multiplicity 3 in A and 1 in B , so it must have multiplicity $3 - 1 = 2$ in $A - B$. Finally, the column $(1, 1, 1)$ has multiplicity 1 in A and 1 in B so it must have multiplicity 0 in $A - B$.

Definition 1.3.10 Let A be a $\{0, 1\}$ -matrix. Given a subset of the rows S , we define the **restriction** $A|_S$ to be a $\{0, 1\}$ -matrix formed from rows S of A .

For example,

$$A = \begin{bmatrix} 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 \end{bmatrix} \implies A|_{\{2,4\}} = \begin{bmatrix} 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 \end{bmatrix}.$$

The following notation is convenient.

Definition 1.3.11 Let α and β be $m \times 1$ columns. Then we say $\alpha \leq \beta$ if for every row r for which the r -th entry of α is a 1, the r -th entry of β is also a 1.

For example,

$$\begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \leq \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix} \quad \text{but} \quad \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \not\leq \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

Definition 1.3.12 Let α be an $m \times 1$ column and A be an m -rowed matrix. We use $\alpha \in A$ if α is a column of A .

The following three special matrices turn out to be remarkably useful in our investigations. We shall see why they are important in Conjecture 1.4.1

Definition 1.3.13 We define the following three matrices:

- The **identity matrix** I_m is be the $m \times m$ matrix with 1's in the diagonal and 0's everywhere else. It corresponds to the set system $\{\{1\}, \{2\}, \dots, \{m\}\}$. 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** I_m^c is the $\{0, 1\}$ -complement of I_m . That is, every entry is a 1, except for the diagonal entries, which are 0's. It corresponds to the set system $\{[m] \setminus \{1\}, [m] \setminus \{2\}, \dots, [m] \setminus \{m\}\}$. 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 **tower matrix** T_m is the $m \times (m + 1)$ matrix that corresponds to the set system $\{\emptyset, [1], [2], [3], \dots, [m]\}$. For example,

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

For these three matrices, when the number of rows is implicit we might omit the subindex m and write I, I^c or T .

Example 1.3.14 Returning to the problem of finding the maximum number of subsets of $[m]$ such that no two subsets have two elements in common, we can restate it as the problem of finding the most number of columns a simple matrix with m rows can have without having the following submatrix:

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

The answer to this particular example is straightforward: The maximum number of columns is

$$\binom{m}{0} + \binom{m}{1} + \binom{m}{2}.$$

Indeed, we can take one column with no 1's (the empty set), all columns with one 1 and all columns with two 1's.

For every pair of rows there is at most one column with two 1's in that pair. Thus the number of columns we can have with more than one 1 is at most $\binom{m}{2}$. For example, for $m = 3$, this is the matrix:

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

This concludes Example 1.3.14.

Example 1.3.15 Now consider the following question. What is the maximum number of subsets of $\{1, 2, \dots, m\}$ such that no three sets each have an element that the other two sets don't have? This is a very similar problem. We wish to find the most number of columns a

$\{0, 1\}$ -matrix can have, having neither of the following six submatrices:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{nor} \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad \text{nor} \quad \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{nor}$$

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \quad \text{nor} \quad \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad \text{nor} \quad \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$

In other words, we wish to find the maximum number of columns in a matrix that doesn't have any row and column permutation of I_3 as a submatrix. We will answer this question later (in Proposition 4.3.3). This demonstrates the need to get rid of row and column ordering. We do so in the next section.

1.3.2 Configurations

Definition 1.3.16 (informal) A **configuration** is a $\{0, 1\}$ -matrix with column and row order stripped.

More formally, let \mathcal{M}_m be the set of all m -rowed matrices. Define an equivalence relation: If $A, B \in \mathcal{M}_m$, we say $A \sim B$ iff A is a permutation of the rows and columns of B . Consider the set of configurations \mathcal{C}_m by taking the quotient:

$$\mathcal{C}_m := \mathcal{M}_m / \sim,$$

and if F is a $\{0, 1\}$ -matrix, define \tilde{F} to be the equivalence class (or **configuration**) to which F belongs.

We will usually abuse notation and not distinguish between a configuration and a $\{0, 1\}$ -matrix representative, always remembering that we can permute the order of the rows and columns without altering the configuration. For example, the six matrices given in Example 1.3.15 represent the same configuration.

Notice that if F, G are configurations and t is a number, the operations we defined for $\{0, 1\}$ -matrices $\|F\|$, $F \times G$, F^c , $t \cdot F$ are well defined in configurations, but concatenation $[F|G]$ and subtraction $F - G$ are not. We will therefore use the notation $[F|G]$ and $F - G$ carefully, when the order of rows is understood. The notion of being a **simple configuration**

is also well defined: a configuration for which any representative is a simple matrix. We have that $F \times G = G \times F$ as configurations, and if F and G are simple, $F \times G$ is also simple.

As with matrices, we might refer to the multiplicity of a column α (and use the notation $\lambda(\alpha, F)$). What we mean by this is that for a specific representative of a configuration F , there are k columns that are equal to α . Of course, if we choose a different representative for F (i.e. permute the rows and columns), we will have k copies of a row permutation of α .

Definition 1.3.17 Denote by K_k the unique $k \times 2^k$ simple “complete” configuration corresponding to the set system $2^{[k]}$. That is, the one representing the power set of $[k]$. The notation was chosen to mirror the common notation for the “complete” graph K_k . As an example, here is K_3 :

$$K_3 = \begin{bmatrix} 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{bmatrix}.$$

Denote by K_k^s the $k \times \binom{k}{s}$ configuration with k rows and all possible columns of column sum s . Here is K_3^2 :

$$K_3^2 = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}.$$

We also use notation such as $K_k^{\leq s}$ (resp. $K_k^{\geq s}$), meaning the configuration with k rows and all possible columns of column sum $\leq s$ (resp $\geq s$).

Observe that for $k, \ell \in \mathbb{N}$, we have $K_k \times K_\ell = K_{k+\ell}$.

Definition 1.3.18 For a configuration F and a $\{0, 1\}$ -matrix A (or a configuration A), we say that F is a **subconfiguration** of A , and write $F \prec A$ if there is a representative of F which is a submatrix of A . We say A **has no configuration** F (or doesn’t **contain** F as a **configuration** or **avoids** F) if F is not a subconfiguration of A .

The notion of subhypergraph in our matrix notation becomes the notion of a subconfiguration.

Observe that \prec defines a partial ordering in the set \mathcal{C}_m of configurations, since \prec is *reflexive* ($F \prec F$), *anti-symmetric* (if $F \prec G$ and $G \prec F$, then $F = G$ as configurations), and *transitive* ($F \prec G \prec H$ implies $F \prec H$).

Definition 1.3.19 Let

$$\text{Avoid}(m, F) := \{A \in \mathcal{C}_m : A \text{ is simple and } F \not\prec A\}.$$

In other words, the set of all m -rowed simple configurations with no F as a subconfiguration.

For example,

$$\text{Avoid}\left(2, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) = \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \right\}$$

Definition 1.3.20 Our main object of study is the following extremal function. We define

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

In other words, $\text{forb}(m, F)$ is the maximum number of columns a simple m -rowed $\{0, 1\}$ -matrix can have with no F as a subconfiguration.

For example, we may conclude that $\text{forb}\left(2, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) = 3$, by looking at all configurations in $\text{Avoid}\left(2, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right)$.

Remark 1.3.21 Observe that forb is an “increasing” function in the second variable, meaning

$$F \prec G \implies \text{forb}(m, F) \leq \text{forb}(m, G).$$

Indeed, if $F \prec G$, then every matrix containing G as a configuration also contains F . Thus, $\text{Avoid}(m, F) \subseteq \text{Avoid}(m, G)$, which means $\text{forb}(m, F) \leq \text{forb}(m, G)$. ■

Interestingly, it is an important open problem to see if $\text{forb}(m, F)$ is increasing in the first variable in the sense that $m < n$ implies $\text{forb}(m, F) \leq \text{forb}(n, F)$. We have strong evidence to believe this is the case, but it hasn’t been proven. More about this in Section 8.1.

Remark 1.3.22 Observe that $\text{forb}(m, F) = \text{forb}(m, F^c)$ by symmetry.

We may also consider forbidding a family of configurations \mathcal{F} :

Definition 1.3.23 Let \mathcal{F} be a family of configurations. Define

$$\text{Avoid}(m, \mathcal{F}) := \bigcap_{F \in \mathcal{F}} \text{Avoid}(m, F).$$

In other words, the set of all m -rowed simple matrices that avoid all configurations in \mathcal{F} .

Definition 1.3.24 We now define the analogous extremal function

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

Note that according to this definition, there exists a simple matrix of size $m \times \text{forb}(m, \mathcal{F})$ that doesn't contain any $F \in \mathcal{F}$ as a subconfiguration, but every matrix of size $m \times (\text{forb}(m, \mathcal{F}) + 1)$ contains at least one $F \in \mathcal{F}$. Also note that (\mathcal{F}, \prec) can be considered as a partially ordered set, and if $\mathcal{H} :=$ minimal elements of \mathcal{F} , then $\text{Avoid}(m, \mathcal{F}) = \text{Avoid}(m, \mathcal{H})$.

Definition 1.3.25 We define the extremal matrices

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

Problem 1.3.26 (THE MAIN PROBLEM) We are interested in determining $\text{forb}(m, \mathcal{F})$. We would like to find $\text{forb}(m, \mathcal{F})$ exactly for any family, but we'll settle for asymptotic bounds when it's too hard to find $\text{forb}(m, \mathcal{F})$ exactly. This problem was first introduced by Anstee in [Ans85] for $|\mathcal{F}| = 1$. We deal mostly with the case $|\mathcal{F}| = 1$.

We now give some definitions that will be helpful in our investigations.

Definition 1.3.27 We say a family \mathcal{A} is **laminar** if for every $X, Y \in \mathcal{A}$ we have that either $X \subseteq Y$, $Y \subseteq X$ or $X \cap Y = \emptyset$. Analogously, we say a $\{0, 1\}$ matrix A is **laminar** if for every two columns α, β of A we have either $\alpha \leq \beta$, or $\beta \leq \alpha$ or there are no rows in which both α and β have a 1. Equivalently, we say m -rowed A is laminar if $A \in \text{Avoid}(m, F)$ for

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

For example, $\begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix}$ is laminar but $\begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & \mathbf{0} & \mathbf{1} & 1 \\ 0 & \mathbf{1} & \mathbf{1} & 1 \\ 0 & \mathbf{1} & \mathbf{0} & 1 \end{bmatrix}$ isn't.

1.3.3 Basic Properties

While studying $\text{forb}(m, \mathcal{F})$ for some family of configurations \mathcal{F} , we usually give lower bounds and upper bounds on $\text{forb}(m, \mathcal{F})$. A lower bound usually consists of an example of a $\{0, 1\}$ -matrix with no $F \in \mathcal{F}$ as a configuration. An upper bound is usually accomplished by a theoretical argument.

Proposition 1.3.28 We have that

$$\text{forb}(m, \mathbf{0}_k) = \text{forb}(m, \mathbf{1}_k) = \binom{m}{k-1} + \binom{m}{k-2} + \dots + \binom{m}{0}$$

and

$$\text{forb}(m, 2 \cdot \mathbf{0}_k) = \text{forb}(m, 2 \cdot \mathbf{1}_k) = \binom{m}{k} + \binom{m}{k-1} + \dots + \binom{m}{0}.$$

Thus, if $\mathbf{0}_k \prec F$ or $\mathbf{1}_k \prec F$ for some configuration F , we have that $\text{forb}(m, F)$ is $\Omega(m^{k-1})$ and if $2 \cdot \mathbf{0}_k \prec F$ or $2 \cdot \mathbf{1}_k \prec F$, then $\text{forb}(m, F)$ is $\Omega(m^k)$.

Proof: The bound for $\text{forb}(m, \mathbf{1}_k)$ is easy. Consider an extremal matrix A such that $\mathbf{1}_k \not\prec A$. Note that in A , we cannot have columns with column sum k or more. This leaves only $\binom{m}{k-1} + \binom{m}{k-2} + \dots + \binom{m}{0}$ columns (the ones with column sum $k-1$ or less). The bound for $\text{forb}(m, 2 \cdot \mathbf{1}_k)$ is also easy. For the columns with column sum k or higher, for each k -subset of the rows there must be at most one column with 1's in that subset of the rows, so there are at most $\binom{m}{k}$ columns, plus the columns with column sum $< k$, which cannot contribute to create $2 \cdot \mathbf{1}_k$. The bounds for $\text{forb}(m, \mathbf{0}_k)$ and $\text{forb}(m, 2 \cdot \mathbf{0}_k)$ are analogous. ■

Here is an important result proved independently by V. Vapnik and A. Chervonenkis, N. Sauer and M. Perles and S. Shelah:

Theorem 1.3.29 [VC71][Sau72][She72] We have that

$$\text{forb}(m, K_k) = \text{forb}(m, \mathbf{1}_k) = \binom{m}{k-1} + \binom{m}{k-2} + \dots + \binom{m}{0},$$

and so $\text{forb}(m, K_k)$ is $\Theta(m^{k-1})$. Thus for any simple k -rowed configuration F , $\text{forb}(m, F)$ is $O(m^{k-1})$.

An important generalization of this result from Anstee and Füredi is:

Theorem 1.3.30 [AF86] For fixed k, t , we have that as $m \rightarrow \infty$,

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

This in particular says that if F has k rows, we have that $\text{forb}(m, F)$ is $O(m^k)$ [Für83], since F is contained in $t \cdot K_k$ for some t .

1.4 The Conjecture

Anstee and Sali conjectured that the “best” asymptotic constructions in terms of avoiding a single configuration F would be formed from products of I, I^c and T . There is ample evidence for this conjecture, but no proof or counterexample has been found yet. The research in forbidden configurations is often guided by this conjecture.

Conjecture 1.4.1 [AS05] Let F be a configuration. Let

$$P_r(a, b, c) := \underbrace{I_r \times \dots \times I_r}_a \times \underbrace{I_r^c \times \dots \times I_r^c}_b \times \underbrace{T_r \times \dots \times T_r}_c,$$

Define $X(F)$ to be the largest number such that there exist numbers $a, b, c \in \mathbb{N}$ with

$a + b + c = X(F)$ such that for all $r \in \mathbb{N}$,

$$F \not\prec P_r(a, b, c).$$

Then $\text{forb}(m, F)$ is $\Theta(m^{X(F)})$.

Observe that $X(F)$ is always an integer. Also note that $\|P_r(a, b, c)\| = r^{a+b} \cdot (r+1)^c$ which is $\Theta(r^{X(F)})$, so by taking $r = \lceil m/X(F) \rceil$ (and perhaps deleting some rows in case $X(F) \nmid m$), we have that $\|P_r(a, b, c)\|$ is $\Omega(m^{X(F)})$, so the fact that $\text{forb}(m, F)$ is $\Omega(m^{X(F)})$ is built into the conjecture. In order to prove the conjecture, all that would be required would be to prove that $\text{forb}(m, F)$ is $O(m^{X(F)})$ for every F . A disproof would be easier, as only a counterexample would be required.

A valid objection is that finding $X(F)$ given F is not a trivial task, but for relatively small configurations F we have a computer program that yields the answer very quickly. We can compute $X(F)$ for F having less than ~ 10 rows in just a few seconds. This task takes merely exponential time, not doubly exponential.

A simple (but surprising) corollary of the conjecture is that repeating columns more than twice in F has no effect on the asymptotic behavior of $\text{forb}(m, F)$. In other words, assuming the conjecture were true, the multiplicity of a column in a configuration would not affect the asymptotic bound, and it for asymptotic bounds, it would only matter if a column is not there (has multiplicity 0), appears once (has multiplicity 1), or appears “multiple times” (has multiplicity 2 or more).

Lemma 1.4.2 Let $F_t = [G|t \cdot H]$ with G and H simple $\{0, 1\}$ -matrices that have no columns in common. Then $X(F_2) = X(F_t)$ for all $t \geq 2$. In particular, if the conjecture were true, then $\text{forb}(m, F_t)$ and $\text{forb}(m, F_2)$ would have the same asymptotic behavior.

Proof: It suffices to show that given t, G, H, a, b and c , there exists an R such that for every $r \geq R$, we have

$$F_2 = [G|2 \cdot H] \prec P_r(a, b, c) \iff F_t = [G|t \cdot H] \prec P_r(a, b, c).$$

Since $F_2 \prec F_t$, we only need to prove that if $F_2 \prec P_r(a, b, c)$ for some r , then $F_t \prec P_R(a, b, c)$ for some R . Suppose then F_2 is contained in the product $P_r(a, b, c)$ for some r . The idea is to find a subconfiguration of $P_r(a, b, c)$ in which there are some columns with multiplicity 1, and for the columns with multiplicity 2 or more, the multiplicity depends on r . We need r large enough so that the multiplicity of any one column (with multiplicity of 2 or more) is larger than t . Let x be the number of rows of F_t . Notice the following three facts, which include definitions for E_I , E_{I^c} and E_T .

$$\begin{aligned} E_I(x, r) &:= [(r-x) \cdot \mathbf{0}_x \mid I_x] \prec I_r \\ E_{I^c}(x, r) &:= [(r-x) \cdot \mathbf{1}_x \mid I_x^c] \prec I_r^c \\ E_T(x, r) &:= \left\lfloor \frac{r}{x} \right\rfloor \cdot T_x \prec T_r. \end{aligned}$$

The first and second facts are easy to see; just take any subset of x rows from I_r or I_r^c . The third statement is true by taking the $\lfloor r/x \rfloor$ -th row of T_r , the $2\lfloor r/x \rfloor$ -th row of T_r , etcetera, up to the $x\lfloor r/x \rfloor$ -th row. For example, if $r = 5$ and $x = 2$, we may take the second and fourth row from T_5 :

$$T_5 = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \implies T_5|_{\{2,4\}} = \begin{bmatrix} 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} = E_T(2, 5)$$

Note that in the three configurations $E_I(x, r)$, $E_{I^c}(x, r)$ and $E_T(x, r)$, we have that there are some columns of multiplicity 1 and there are some columns for which their multiplicity can be made as large as we wish by making r large. Formally, let $E(x, r)$ be one of $E_I(x, r)$ or $E_{I^c}(x, r)$ or $E_T(x, r)$. We have that for every x -rowed column α there are three possibilities: either $\lambda(\alpha, E(x, r)) = 0$ for all r , or $\lambda(\alpha, E(x, r)) = 1$ for all r , or $\lim_{r \rightarrow \infty} \lambda(\alpha, E(x, r)) = \infty$.

If α is a column for which $\lim_{r \rightarrow \infty} \lambda(\alpha, E(x, r)) = \infty$, we may conclude that there is an R for which $\lambda(\alpha, E(x, r)) \geq t$ for every $r \geq R$.

Since F_2 is contained in $P_r(a, b, c)$ for some r , the columns in H will have multiplicity at least 2 in some subset of the rows of $P_r(a, b, c)$. By taking $P_R(a, b, c)$, we see that F_t is also a subconfiguration of $P_R(a, b, c)$. ■

Conjecture 1.4.1 has been a driving force behind the field of Forbidden Configurations, specifically when forbidding a single configuration (see [Ans]). We search for maximal and minimal forbidden configurations with a specific asymptotic bound, since we have that if both $\text{forb}(m, F)$ and $\text{forb}(m, G)$ are $\Theta(m^k)$, and we have a configuration H such that $F \prec H \prec G$, then $\text{forb}(m, H)$ is also $\Theta(m^k)$. For a given k and s , we search for minimal and maximal s -rowed configurations F for which $\text{forb}(m, F)$ is $\Theta(m^k)$. We hope to use this to classify the asymptotic bounds for all configurations.

Definition 1.4.3 We say a configuration $F_t = [G|t \cdot H]$ (for $t \geq 2$) with G and H simple with no columns in common is a **predicted boundary case** if for any column α not present in H , $X([F_t|\alpha]) > X(F_t)$. We say F_t is a **boundary case** if $\text{forb}(m, F_t)$ is $\Theta(m^k)$ but for any column α not present in H , $\text{forb}(m, [F_t|\alpha])$ is $\Omega(m^{k+1})$. Conjecture 1.4.1 is that predicted boundary cases are the same as boundary cases.

Boundary cases for s up to 4 were classified by Fleming ([Ans]). The conjecture has been proven for all $k \times \ell$ configurations F with $k = 1, 2, 3$ and many others in various papers. For $k = 2$ in [AGS97]. For $k = 3$ in [AGS97], [AFS01] and was completed in [AS05]. For $k = 4$, the case when F is simple was completed in [AF86]. For $k = 4$ and F non-simple, there were only three cases left to do. We completed one of them ([ARS10a]) and the proof is given in Section 5.2. For $\ell = 2$, the conjecture was verified in [AK06]. For $k = 5$, there are nine 5×6 predicted boundary configurations F for which $X(F) = 2$ ([ARS11]). We prove one of them in Section 5.3. For $k = 6$, we give a complete classification of the configurations F for which $\text{forb}(m, F)$ is quadratic ([ARS11]).

If we only consider simple configurations, then we may also consider a restricted version of a boundary case.

Definition 1.4.4 We say a simple configuration F is a **predicted s -boundary case** if for any column α not present in F , we have that $X([F|\alpha]) \geq X(F) + 1$. We say it is a **s -boundary case** if $\text{forb}(m, F)$ is $\Theta(m^k)$ but $\text{forb}(m, [F|\alpha])$ is $\Omega(m^{k+1})$ for any column α not present in F .

The classification of s -boundary cases for five rows was done by Ryan [Ans]. We now give here a classification of the minimal and maximal boundary cases. For $t \geq 2$,

	Minimal	Maximal
Constant	[0], [1]	[0 1]
Linear	[0 0], [1 1]	$t \cdot [0 1]$

Table 1.1: Classification: 1 row.

	Minimal	Maximal
Constant	$\begin{bmatrix} 1 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0 \end{bmatrix}$
Linear	$\begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$, $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} t \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} t \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$
Quadratic	$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, $\begin{bmatrix} 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}$	$t \cdot K_2$

Table 1.2: Classification: 2 rows.

	Minimal	Maximal
Linear	$\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$
Quadratic	$\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$ $\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix},$ $\begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$	$\begin{bmatrix} 0 & 0 & 1 & 1 & \begin{bmatrix} 1 & 0 & 1 & 0 \end{bmatrix} \\ 0 & 0 & 0 & 1 & t \cdot \begin{bmatrix} 0 & 1 & 1 & 1 \end{bmatrix} \\ 0 & 1 & 1 & 1 & \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix} \end{bmatrix},$ $\begin{bmatrix} 0 & 0 & 1 & 1 & \begin{bmatrix} 1 & 0 & 1 & 0 \end{bmatrix} \\ 0 & 0 & 1 & 1 & t \cdot \begin{bmatrix} 0 & 1 & 0 & 1 \end{bmatrix} \\ 0 & 1 & 0 & 1 & \begin{bmatrix} 0 & 0 & 1 & 1 \end{bmatrix} \end{bmatrix},$ $\begin{bmatrix} 0 & \begin{bmatrix} 1 & 0 & 0 & 1 & 1 \end{bmatrix} \\ 0 & t \cdot \begin{bmatrix} 0 & 1 & 0 & 1 & 0 \end{bmatrix} \\ 0 & \begin{bmatrix} 0 & 0 & 1 & 0 & 1 \end{bmatrix} \end{bmatrix},$ $\begin{bmatrix} 1 & \begin{bmatrix} 0 & 1 & 1 & 0 & 0 \end{bmatrix} \\ 1 & t \cdot \begin{bmatrix} 1 & 0 & 1 & 0 & 1 \end{bmatrix} \\ 1 & \begin{bmatrix} 1 & 1 & 0 & 1 & 0 \end{bmatrix} \end{bmatrix}$
Cubic	$2 \cdot \mathbf{0}_3, [2 \cdot K_3^1 K_3^2], [2 \cdot K_3^1 K_3^3], [K_3^0 2 \cdot K_3^2], [K_3^1 2 \cdot K_3^2], 2 \cdot \mathbf{1}_3$	$t \cdot K_3$

Table 1.3: Classification: 3 rows.

Before showing the table for 4 rows, let us define some configurations:

$$B_1 = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \end{bmatrix}, B_2 = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 \end{bmatrix}, B_3 = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 1 \end{bmatrix},$$

$$B_4 = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 \end{bmatrix}, B_5 = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 \end{bmatrix}, B_6 = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 \end{bmatrix}.$$

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

	Minimal	Maximal
Linear	$\begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$
Quadratic	$\begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix},$ $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$	$\left[\begin{array}{c} \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} t \cdot \begin{bmatrix} 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \\ \text{(Conjectured)} \end{array} \right]$ $\left[\begin{array}{c} \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} t \cdot \begin{bmatrix} 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix} \\ \text{(Conjectured)} \end{array} \right]$ $\left[\begin{array}{c} \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix} t \cdot \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 1 & 1 \\ 0 & 0 \end{bmatrix} \\ \text{Solved in Section 5.2} \end{array} \right]$

	Minimal	Maximal
Cubic	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix},$ $\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix},$ $\begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix},$ $\begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 1 & 1 \end{bmatrix}$ <p>All these are conjectured.</p>	$[K_4 t \cdot [K_4 - B_i]]$ for $i = 1, 2, \dots, 6,$ $[K_4^0 t \cdot D]$ $[K_4^0 t \cdot D]^c$
Quartic	All others (see [Ans] for a list)	$t \cdot K_4$

Table 1.4: Classification: 4 rows.

For five rows, the minimal and maximal matrices haven't been classified, except for the quadratic case, and these are the nine maximal five-rowed configurations which are conjectured to have a quadratic bound. Of these, only F_7 is known to have a quadratic bound, proved in Section 5.3.

$$F_3 = \begin{bmatrix} 1 & 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad F_4 = \begin{bmatrix} 1 & 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad F_5 = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 \end{bmatrix}$$

$$F_6 = \begin{bmatrix} 1 & 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$F_7 = \begin{bmatrix} 1 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

$$F_8 = \begin{bmatrix} 1 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

$$F_9 = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 \end{bmatrix}$$

$$F_{10} = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$F_{11} = \begin{bmatrix} 1 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}$$

1.5 Table of Notation

For the following definitions, A, B are m -rowed $\{0, 1\}$ matrices, F and G are configurations, \mathcal{F} is a family of configurations, α is a column, and S is a subset of rows. We only give informal definitions. For full definitions, see Section 1.3.

Name	Notation	Informal Defn.	Example
Simple Configuration	N/A	A configuration with no repeated columns.	$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \checkmark$ $\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix} \times$
Number of Columns	$\ F\ $	The number of columns of a representative.	$\left\ \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix} \right\ = 3$
Number of 0's/1's	$\sigma_0(\alpha), \sigma_1(\alpha)$ $\sigma_0(A), \sigma_1(A)$	The number of 0's/1's in a column or matrix.	$\sigma_0 \left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \right) = 3$ $\sigma_1 \left(\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right) = 2$
Column of 0's/1's	$\mathbf{0}_m, \mathbf{1}_m$	An m -rowed column of 0's/1's	$\mathbf{0}_3 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \mathbf{1}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$
Complement	F^c	Replace 1's for 0's and 0's for 1's.	$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}^c = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix}$

Name	Notation	Informal Defn.	Example
Concatenation	$[A B]$	The columns of A together with those of B .	$\left[\left[\begin{array}{c c} 1 & 0 \\ 0 & 1 \end{array} \right] \right] = \left[\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \end{array} \right]$
Product	$t \cdot F$	$\underbrace{[F \mid F \mid \dots \mid F]}_{t \text{ times}}$	$2 \cdot \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}$
Product	$F \times G$	The configuration that consists of all combinations of each column of F on top of each column of G	$A = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}, B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ \implies $A \times B = \left[\begin{array}{cc cc cc} 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \end{array} \right]$
Multiplicity	$\lambda(\alpha, A)$	Number of columns of A which are equal to α	$\lambda \left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix} \right) = 2$
Difference	$A - B$	Columns of A minus columns of B	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

Name	Notation	Informal Defn.	Example
Restriction	$A _S$	Rows S of A put together	$A = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$ \implies $A _{\{1,3\}} = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$
Identity	I_m	0's everywhere except for diagonal	$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$
Identity Complement	I_m^c	1's everywhere except for diagonal	$I_3^c = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$
Tower	T_m	1's on the upper-right corner and 0's on the lower left corner.	$T_3 = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$
Complete	K_m	All possible columns	$K_2 = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$
Complete of sum k	K_m^k	All columns of column sum k	$K_3^2 = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$

Name	Notation	Informal Defn.	Example
Sub-Configuration	$F \prec G$	There is a row and column permutation of F which is a submatrix of G	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \prec \begin{bmatrix} 0 & \mathbf{0} & \mathbf{1} \\ 0 & 1 & 1 \\ 1 & \mathbf{1} & \mathbf{0} \end{bmatrix}$ $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \not\prec \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}$
Avoid	$\text{Avoid}(m, \mathcal{F})$	The set of m -rowed simple matrices which avoid all $F \in \mathcal{F}$.	$\text{Avoid}\left(2, \begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} \right\}$
Forb	$\text{forb}(m, \mathcal{F})$	Maximum number of columns of a matrix in $\text{Avoid}(m, \mathcal{F})$.	$\text{forb}\left(m, \begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = 2$
Ext	$\text{ext}(m, \mathcal{F})$	Set of matrices in $\text{Avoid}(m, \mathcal{F})$ with exactly $\text{forb}(m, \mathcal{F})$ columns	$\text{ext}\left(2, \begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \left\{ \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} \right\}$
Laminar Configuration	N/A	$\begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \not\prec A$	$\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

Table 1.5: Definitions

1.6 History

1.6.1 Extremal Combinatorics and Set Theory

Extremal Combinatorics deals with problems of finding the maximum or minimum size of a combinatorial class. As such, the path taken by the study of Extremal Combinatorics is filled with alleyways, short roads with many stop signs and quite a few landmarks. The reason for the sinuous nature of this field is that coming up with extremal combinatorial problems is extremely easy, but coming up with answers to said problems often turns out to be very hard. Very similar problems can have vastly different solutions.

It is impossible to pinpoint exactly when the first time such questions were studied, but some (like [Juk01]) argue it was Euler who first studied such problems systematically.

The last few decades however have seen an explosion of problems of such ilk, one of the reasons being that the hungry monster of Computer Science has posed and keeps posing many such questions with practical applications.

Extremal Combinatorics (at least maximization problems) often involves a race to provide better and better upper bounds (which are generally accomplished by theoretical arguments), and better and better lower bounds (which are generally accomplished by constructions). The field of Forbidden Configurations is no different and we often try to “sandwich” $\text{forb}(m, \mathcal{F})$ by providing simple matrices with no configuration $F \in \mathcal{F}$ and then using a theoretical argument for the upper bound.

Many different problems have been studied by a diverse collection of mathematicians. Here we mention a few examples of problems and theorems.

Theorem 1.6.1 [Man07] Let $G = (V, E)$ be a graph on n vertices with the property that there is no triangle. Then $|E| \leq n^2/4$.

A generalization came from Turán.

Theorem 1.6.2 (Turán theorem)[Tur41] Let $G = (V, E)$ be a graph on n vertices such that G does not contain a subgraph isomorphic to K_k . Then,

$$|E| \leq \frac{k-2}{k-1} \cdot \frac{n^2}{2} = \left(1 - \frac{1}{k-1}\right) \cdot \frac{n^2}{2}.$$

And further generalization is due to Erdős and Stone, asking for the maximum number of edges that avoids a given graph H . This question was answered asymptotically by Erdős, Stone, and Simonovits.

Theorem 1.6.3 [ES46] [ES66] (Erdős-Stone-Simonovits) Given a number n and a graph H , let $\text{forb}(n, H)$ denote the maximum number of edges of a graph in n vertices with no H as a subgraph. Let $\chi(H)$ denote the chromatic number of H . Then for every $\epsilon > 0$ there exists N such that for all $n \geq N$ we have

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

The related problem of Zarankiewicz asks for the maximum number of edges from a given complete bipartite graph for which a smaller given complete bipartite graph is avoided. A bound was given by Kővári, Sós and Turán ([KST54]). A better bound was given by Füredi in [Für96]. We use this result in our investigations in Chapter 7.

There is a rich class of results involving intersecting families (i.e. every two sets in the family intersect). The following are some highlights.

Theorem 1.6.4 Let \mathcal{A} be an intersecting family on $[n]$. Then $|\mathcal{A}| \leq 2^{n-1}$.

Theorem 1.6.5 [EKR61] (Erdős, Ko, Rado) Let \mathcal{A} be an intersecting k -uniform family on $[n]$. Then $|\mathcal{A}| \leq \binom{n-1}{k-1}$.

A generalization of this theorem to t -intersecting families is due to Ahlswede and Kachatrian ([AK97]) who completely characterized the extremal families. A stability result by Anstee and Keevash ([AK06]) was obtained for certain values of the parameters. This was used to establish the asymptotics of $\text{forb}(m, F)$ for $k \times 2$ configurations F .

Here is a foundational antichain result.

Theorem 1.6.6 [Spe28] (Sperner’s Theorem) Let \mathcal{A} be a set system on $[n]$ for which there are no $A, B \in \mathcal{A}$ such that $A \subseteq B$. Then

$$|\mathcal{A}| \leq \binom{n}{\lceil n/2 \rceil}.$$

The following problem is one of many in the field of Ramsey Theory. Many generalizations of this problem also exist and there are many interesting results.

Problem 1.6.7 [Ram30] Given a number k , find the maximum number of vertices n for a graph G such that neither G nor G^c contain K_k .

By [Ram30], it is known that such number n exists for any k . The full Ramsey Theorem considers colouring all t -sets of $[n]$ using ℓ colours. The Theorem states that for every k there is n such that any such graph on n vertices has a full monochromatic k ‘clique’. We may use this, together with the idea of Section 2.2 to conclude that for m large enough there is a clique (of a fixed wanted size) of rows of the matrix A for which the same possibility of “What Is Missing” occurs.

We now give a very small sampling of extremal problems in many other areas of combinatorics. In graph theory,

Problem 1.6.8 Given a graph G , what is the minimum number of colors one needs in order to colour the vertices of G so that two vertices who share an edge do not have the same colour?

It is well-known that 4 colours is enough for any planar graph [AHK77]. For general graphs, the number of colours needed is called the *chromatic number*.

To additive combinatorics,

Problem 1.6.9 [TV06] Given an Abelian group G , what is the maximum size of a sum-free set $A \subseteq G$? (sum-free means there are no $a, b, c \in A$ with $a + b = c$).

In computational geometry,

Problem 1.6.10 [O'R87] (Art Gallery Problem) Given a polygon P , what is the minimum number of lights needed to fully illuminate the interior of the polygon?

As the reader might have realized by now, there are probably more problems in extremal combinatorics than there are mathematicians (find the minimum number of mathematicians to state all extremal combinatorics problems?) and so we could go on and on.

1.6.2 Forbidden Configurations

The field of Forbidden Configurations began when Vapnik and Chervonenkis, Sauer, Perles and Shelah independently proved Theorem 1.3.29, each for different purposes. The theorem was proved first by Vapnik and Chervonenkis in 1968 in [VC68]. Sauer and Bollobas attribute the original problem to Erdős. This theorem and the related concepts of VC-dimension and shattered sets have applications in fields as diverse as Machine Learning and Pattern Recognition ([BEHW89], [Vap00], [WD81]), Probability ([Ste78]), Combinatorial Geometry ([Ver05], [Mat02]), Extremal Set Theory ([MZ07], [BKS05]), etc.

The general question: “For a $\{0, 1\}$ -matrix F , what is the maximum number of columns a simple $\{0, 1\}$ -matrix A can have so that A doesn't contain a row and column permutation of F as a submatrix” was first posed by Anstee in [Ans85]. Since then, many people have studied this question.

The first paper studying this problem in a systematic way is [AGS97]. Both asymptotic and exact bounds were considered for small forbidden configurations (small in the number of rows). A large number of exact bounds have been proven in [AFS01], [AK07], [ABS11], [AK10]. A particular 4×2 configuration is given in [ABS11] for which finding an exact bound is highly unlikely. A clearer explanation is in [AK10]. In a private communication, Dukes has obtained tight bounds on the leading coefficient of the associated quadratic bound.

Conjecture 1.4.1 arose from the various investigations of asymptotic bounds. Product constructions were introduced in [AGS97]. The importance of three building blocks I , I^c and T were also noted in results by Balogh and Bollobás [BB05]. The reader might note that I , I^c and T arise repeatedly in proofs, such as in the results of Section 5.3, and product constructions arise even in exact bounds, such as in the results of Section 4.1.3. Conjecture 1.4.1 was first stated in [AS05], but Anstee already believed that forbidden configuration bounds would be $\Theta(m^k)$ for some integer k (that depended on F) long before that. Since then, much of the progress in asymptotic forbidden configurations has been guided by this conjecture.

Chapter 2

Basic Techniques

We have a number of powerful techniques that work in many cases and that we will use them repeatedly in this thesis. A careful look at the particular properties of each family of configurations we wish to study is typically required.

We will describe some of these techniques in generality here and we will refer to them when used to solve particular examples.

2.1 Standard Induction

Let \mathcal{F} be a family of configurations and let $A \in \text{Avoid}(m, \mathcal{F})$. We wish to do induction on m , but if we delete a row r from A , we might run into trouble: the resulting matrix after the deletion might not be simple.

Let $C_r = C_r(A)$ be the matrix that consists of the $m - 1$ -rowed columns that have multiplicity 2 in the matrix that results from deleting row r from A . If we permute the rows and columns of A (i.e. we take a representative of the configuration A) so r becomes the first row, the columns of $A \setminus \{\text{row } r\}$ can be divided into four blocks:

$$A = {}^r \begin{bmatrix} 0 & \cdots & 0 & 1 & \cdots & 1 \\ B_r & C_r & C_r & D_r \end{bmatrix}. \quad (2.1.1)$$

where $B_r = B_r(A)$ are the columns that appear with a 0 on row r , but don't appear with a 1, $D_r = D_r(A)$ are the columns that appear with a 1 but not a 0, and C_r are the columns that appear with both. We call this the **standard decomposition** of A .

Note that $[B_r|C_r|D_r]$ is simple and since $F \not\prec A$, we have that $F \not\prec [B_r|C_r|D_r]$. So $[B_r|C_r|D_r] \in \text{Avoid}(m - 1, \mathcal{F})$. This means any upper bound on $\|C_r\|$ (as a function of m),

automatically yields an inductive upper bound on $\|A\|$:

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

We now study the structure of C_r . Note that $[0 \ 1] \times C_r$ is in A .

Definition 2.1.1 We say a configuration H is an **inductive child** of a configuration F if

$$F \prec \left[\begin{array}{ccc|ccc} 0 & \dots & 0 & 1 & \dots & 1 \\ \hline & & H & & & H \end{array} \right].$$

Note that if H is an inductive child of F and $F \not\prec A$, then $H \not\prec C_r$. Therefore if we define $\mathcal{H}_{\mathcal{F}}$ to be the minimal elements of the family of inductive children of \mathcal{F} under order \prec , we know $C_r \in \text{Avoid}(m-1, \mathcal{H}_{\mathcal{F}})$.

Proposition 2.1.2 Let \mathcal{F} be a family of configurations. Then

$$\text{forb}(m, \mathcal{F}) \leq \sum_{i=1}^{m-1} \text{forb}(i, \mathcal{H}_{\mathcal{F}}).$$

Proof: Let $A \in \text{ext}(m, \mathcal{F})$. Then we know that for any row r ,

$$\begin{aligned} \text{forb}(m, \mathcal{F}) &= \|A\| \\ &= \|C_r\| + \|[B_r|C_r|D_r]\| \\ &\leq \|C_r\| + \text{forb}(m-1, \mathcal{F}) \\ &\leq \text{forb}(m-1, \mathcal{H}_{\mathcal{F}}) + \text{forb}(m-1, \mathcal{F}) \\ &\leq \sum_{i=1}^{m-1} \text{forb}(i, \mathcal{H}_{\mathcal{F}}) \end{aligned}$$

This concludes the proof. ■

Since this proposition gives an upper bound on $\text{forb}(m, \mathcal{F})$ with respect to $\text{forb}(m, \mathcal{H}_{\mathcal{F}})$, and $\mathcal{H}_{\mathcal{F}}$ consists of smaller configurations (but perhaps more of them), we may repeat this process and find a bound instead on $\text{forb}(m, \mathcal{H}_{\mathcal{H}_{\mathcal{F}}})$, and so on.

Given \mathcal{F} , constructing $\mathcal{H}_{\mathcal{F}}$ by hand is easy. We have a computer program that does this for us, but it almost never gets used since it's so easy to do it by hand. The inductive children of a configuration must have either the same number of rows, or one less than the original. Then we must look at columns which get “repeated”. Perhaps an example would help to clarify this.

Example 2.1.3 The inductive children of

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

are

$$\begin{bmatrix} 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \quad \text{and} \quad \begin{bmatrix} 0 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}.$$

These are obtained by deleting a row of F and only taking one for each column that appears with both a 0 and a 1 in the deleted row of F , as well as taking the repeated columns of F only once. ■

Sometimes this bound is not enough because $\text{forb}(m, \mathcal{H}_{\mathcal{F}})$ is too large. What we actually need in order to proceed as above by induction is that $\|C_r(A)\|$ is small for some choice of r , not necessarily that $\text{forb}(m, \mathcal{H}_{\mathcal{F}})$ is small. Often it is the case that we must search for the row r with the least amount of repetition, one for which $\|C_r(A)\|$ is as small as possible. If we can prove that for every $A \in \text{Avoid}(m, \mathcal{F})$ there must always be a row r with $\|C_r(A)\|$ being “small enough”, we may appeal to induction and proceed as above.

There is another method we might use when $\|C_r(A)\|$ is too large. We may delete a limited number of columns (without deleting any row) before proceeding to do induction. For example, if $A \in \text{Avoid}(m, \mathcal{F})$, we might select some columns U from A , so that

$$A = [A' \mid U]$$

which satisfy the property that $\|U\|$ is small, and $C_r(A')$ is also small (and therefore we might do induction on A').

In this thesis we also use a new method in order to prove the results of Section 5.3 and Section 7.6 that involves considering only “essential” rows (in some sense).

Definition 2.1.4 For a matrix A and a row r , let $L(r)$ be a minimal subset of the rows of $C_r(A)$ such that $C_r|_{L(r)}$ is a simple configuration. This involves some choice. For each row r we fix this choice.

This concept turns out to be quite useful in our investigations, and it arose by considering deleting “non-essential” rows of $C_r(A)$ and looking at the structure of the remaining rows.

In the next section we will give an alternate way to study the structure of a matrix $A \in \text{Avoid}(m, \mathcal{F})$.

2.2 What Is Missing?

Let \mathcal{F} be a (finite) family of forbidden configurations and let t be the maximum multiplicity of any column in any $F \in \mathcal{F}$. In other words,

$$t := \max\{\lambda(\alpha, F) : \alpha \text{ is a column, and } F \in \mathcal{F}\}.$$

Suppose we have a $\{0, 1\}$ -matrix $A \in \text{Avoid}(m, \mathcal{F})$. Given $s \in \mathbb{N}$, consider all s -tuples of rows from A , and for each s -tuple of rows S , consider the matrix $A|_S$ formed from rows S of A as in Definition 1.3.10.

Without any restriction, $A|_S$ could have all 2^s possible columns, each with “high” multiplicity. But with the restriction that $F \not\leq A$ for each $F \in \mathcal{F}$, in particular $F \not\leq A|_S$, so some columns have to be missing: we can’t have more than t of all 2^s columns, or else we would definitely have all $F \in \mathcal{F}$ as subconfigurations. That is, some columns have multiplicity 0 and some must have multiplicity less than t .

Perhaps an example might clarify this idea. Suppose \mathcal{F} consists of a single configuration

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

so in this case, $t = 2$. Let

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

For $S = \{2, 3, 4\}$ we see that

$$A|_S = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 \end{bmatrix}.$$

In $A|_S$, the columns $(1, 0, 0)^T$ and $(0, 0, 1)^T$ have multiplicity 0, the columns $(0, 0, 0)^T$, $(1, 1, 1)^T$, $(0, 1, 1)^T$ and $(1, 1, 0)^T$ have multiplicity 1, and the columns $(0, 1, 0)^T$ and $(1, 0, 1)^T$ have multiplicity 2. Now imagine what would happen if all columns had multiplicity 2 in $A|_S$ instead. Then F would be a subconfiguration of $A|_S$, which would be a contradiction. So *some* columns must have multiplicity less than 2. This motivates the following definition.

Definition 2.2.1 Given a matrix A , a number $s \in \mathbb{N}$ and an s -tuple S of the rows of A , we say an s -rowed column α is **absent** if $\lambda(\alpha, A|_S) = 0$. We say it is in **short supply** if $\lambda(\alpha, A|_S) < t$. We say α is in **long supply** if $\lambda(\alpha, A|_S) \geq t$.

In the example, we can conclude that for any $A \in \text{Avoid}(m, F)$, in each triple of rows (a, b, c) of A there is an ordering (x, y, z) of (a, b, c) , for which the columns marked by “no” must be absent, the columns marked with < 2 must be in short supply and the rest may potentially be in long supply.

$$\begin{array}{ccc} \text{no} & \text{no} & \text{no} \\ x & \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \\ y & & & \\ z & & & \end{array} \quad \text{or} \quad \begin{array}{ccc} < 2 & < 2 & < 2 \\ x & \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \\ y & & & \\ z & & & \end{array} \quad \text{or} \quad \begin{array}{ccc} \text{no} & \text{no} & < 2 & < 2 \\ x & \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \\ y & & & \\ z & & & \end{array}.$$

Of course if there are no columns of sum 1 in $A|_S$ (the first case), then $F \not\prec A|_S$. The same is true if all columns of sum 2 are in short supply (the second case). The third case might be a little harder to see, but if we take a look at the columns potentially in long supply, we see why:

$$\begin{array}{ccc} \text{absent} & \text{short supply} & \text{long supply} \\ \text{no} & \text{no} & < 2 & < 2 \\ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} & , & \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} & \implies & \begin{array}{cccc} \text{l.s.} & \text{l.s.} & \text{l.s.} & \text{l.s.} \\ \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \end{array} \end{array}$$

Clearly F is not a subconfiguration: no matter how many times we repeat each column marked in long supply, if we are only allowed to have the columns marked in short supply only once. Notice that in our example, since A does not have F as a subconfiguration, in $A|_S$ we see that the third case is the one that occurs, with row order (x, y, z) being $(2, 3, 1)$, so $A|_S$ satisfies

$$\begin{array}{cccc} \text{no} & \text{no} & < 2 & < 2 \\ \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \end{array} .$$

Some effort is required to determine What Is Missing when we avoid all $F \in \mathcal{F}$ for a given \mathcal{F} . In the case above, it is straightforward to check the list of 3 cases is complete.

In this particular example $s = 3$ and the number of rows of F was also 3, but s and the number of rows of F do not need to be equal. When s is less than the number of rows of F , every column is in long supply. We can choose s to be anything, but s being the number of rows of F is often used. Larger s provides more information, but there is a trade off: it's often harder to analyze such information.

2.3 Implications

Sometimes while studying the set of possibilities for What Is Missing from an 3-tuple of rows $\{x, y, z\}$, we find that one of the possibilities has this restriction:

$$\begin{array}{ccc} & < t & < t \\ x & \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \\ y & & \\ z & & \end{array}$$

Notice that this means in particular that in rows x, y there are at most $2t - 2$ columns that have a 0 in row x and a 1 in row y :

$$\begin{array}{ccc} & < t & < t \\ x & \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \\ y & & \\ z & & \end{array} \implies \begin{array}{ccc} & & \leq 2t - 2 \\ x & & \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ y & & \end{array}$$

This observation motivates the following definition.

Definition 2.3.1 Given a matrix A and a function $f : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$, we say for a row x and a row y that we have the **implication** $x \rightarrow y$, if the following is satisfied on the pair of rows x, y :

$$\begin{array}{c} x \\ y \end{array} \leq \begin{array}{c} f(t) \\ \left[\begin{array}{c} 0 \\ 1 \end{array} \right] \end{array} \quad (2.3.1)$$

Notice that $x \rightarrow y$ means that if in some column of A there is a 0 in row x , then there is usually a 0 in row y , except perhaps for $f(t)$ columns.

Definition 2.3.2 We say a column **violates an implication** $x \rightarrow y$ if it has a 0 in row x and a 1 in row y . Thus, $x \rightarrow y$ means there are at most $f(t)$ columns that violate the implication.

We call implications that never get violated **pure** implications, and implications that get violated at least once **impure** implications.

Proposition 2.3.3 Let A be a $\{0, 1\}$ -matrix and consider the directed graph G , where the vertices are the rows, and the arrows are the implications. Suppose we had an implication $x \rightarrow y$ and we also had a path of implications

$$x = x_0 \rightarrow x_1 \rightarrow x_2 \rightarrow \dots \rightarrow x_n = y.$$

Then we may conclude that if a column of A violates $x \rightarrow y$, it must also violate an implication of the form $x_{i-1} \rightarrow x_i$ for some i with $1 \leq i \leq n$.

Proof: Indeed, if a column α has a 0 in row x and a 1 in row y , consider the smallest index i for which α has a 1 in row i . Then the column must have a 0 in row $i - 1$ and therefore violates $x_{i-1} \rightarrow x_i$. ■

Given A , using this technique we might be able in some cases to select just some small set \mathcal{I} of implications, so that if a column of A violates *any* implication, then it must necessarily violate an implication in our set \mathcal{I} . The number of columns violating any implication will

then be at most $f(t) \cdot |\mathcal{I}|$. The power of implications comes from the fact that if $|\mathcal{I}|$ and $f(t)$ are small enough, even if the total number of implications is large, we can delete every column that violates any implication, thus making the implication pure. The result is that columns previously marked as being “in short supply” are now completely absent.

This is a powerful technique, because many times we are able to find an upper bound on $\text{forb}(m, [G|H])$ with G, H simple with no columns in common, and we wish to find an upper bound on $\text{forb}(m, [G|t \cdot H])$. Considering a matrix A that has no $[G|t \cdot H]$, we are able to delete at most $f(t) \cdot |\mathcal{I}|$ columns from A , and then this new matrix won't have $[G|H]$. So under these assumptions, we are able to conclude that

$$\|A\| \leq f(t) \cdot |\mathcal{I}| + \text{forb}(m, [G|H]).$$

Chapter 3

Computer Program Developed

In this chapter we will describe a C++ program we used extensively in the results that follow. We give a description of the algorithms and data structures. Our code computes the following tasks:

Sub-Configuration: Answers the question of whether or not a configuration F is a sub-configuration of a configuration A .

Conjecture: Determines the asymptotic bound predicted by Conjecture 1.4.1 for a given configuration F . In other words, it finds $X(F)$.

Classification: Once we know how to find $X(F)$, given a certain number of rows, this program finds all boundary cases.

What Is Missing? Given s and a family of configurations \mathcal{F} , this program finds the list of What Is Missing in each s -tuple of rows if we forbid \mathcal{F} .

Finding forb: Given \mathcal{F} and a small value of m (specifically, $m \leq 5$), this program computes $\text{forb}(m, \mathcal{F})$.

Guessing forb: For $m = 6, 7, 8, 9$, this program uses local search (in particular, a Genetic Algorithm among others) to guess the extremal matrices $\text{ext}(m, \mathcal{F})$.

In the subsequent sections we'll deal with all these problems and describe our own approach to these tasks. All these have been implemented in C++ (and some of them in sage as well) and are available for download from <http://www.math.ubc.ca/~ansteel/>

For some of these tasks, complexity is unfortunately doubly exponential in the number of rows (for What Is Missing, Finding forb, Classification), while for others (Sub-Configuration,

Conjecture) it is merely exponential in the number of rows. In any case, we only deal with these problems for small configurations, and even for such cases it is often the case we aren't able to use the information given by the computer to prove results, so better algorithms would not necessarily result in new results in Forbidden Configurations.

3.1 Representation of a Configuration

We are interested in an efficient representation for configurations, in order to perform the tasks described above. In the progress of our investigations, we have had various versions of the program. We will record here how the program works based on the posted version, and recorded for posterity in the above link. Various things were done in a different manner, but we have mainly improved usability and speed with each subsequent version. Each version of the program was tested against many different configurations for correctness, and some tasks were implemented in sage (<http://www.sagemath.org/>) as well, in order to check the answers when they were too large to check by hand.

For the “What Is Missing” calculation it is easy to check that each possibility, if satisfied, indeed implies that the given configurations are avoided, but there is no easy way to check that the list of possibilities is complete.

First of all, we explain how the program stores configurations. Most of what we want the program to do involves performing a huge number of *configuration comparison operations*, which is testing whether or not a configuration F is a subconfiguration of a configuration A . As a first approach it would seem as if, for this task, we would be required to test each row and column permutation of F against each submatrix of A . This is of course a very slow way to do this. A simple trick to speed up the computations is to keep the columns of a configuration always in some canonical order. Then, to test whether or not a configuration F is contained in another A , we just need to permute rows of F and take subsets S of rows of A and place the columns of $A|_S$ in canonical order.

Most of the tasks we described above involve checking whether or not a given (fixed) configuration F is a subconfiguration of a vast number of configurations A . In particular, any pre-processing we do on F can be considered as almost free. For example, finding all row permutations of F and storing them would need to be done once for each configuration F , and not at all for configurations A .

After many attempts, it seems that the best (fastest) way to store a configuration that would make many of the other tasks reasonably fast is this: Maintain an array of integers

where the *indices* of the array, written in binary, are the *columns* of the configuration, and the actual numbers of the array represent the number of times a column appears. That is to say, a configuration F in m rows is represented by an array (C++ vector) \mathbb{F} of size 2^m . For a number α , consider the binary representation of α and consider it as a column with m rows. If necessary, put enough 0's at the beginning of the binary representation in order to have the required m bits. The number $\mathbb{F}[\alpha]$ (the α -th number of the array) represents the number of times that column α appears in configuration F . In the implementation, we use an array of *unsigned characters* instead of integers, since we never need a configuration with the same column repeated more than 255 times. An unsigned character consists of 1 byte (8 bits).

For example, the array $\mathbb{F} = [1, 0, 0, 2, 0, 1, 0, 1]$ represents the following configuration (notice it has 3 rows, since the array has size $8 = 2^3$):

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

To see this, remember we start from 0. There is a one in position $0 = 000_b$, meaning the column $(0, 0, 0)^T$ gets repeated one time. A two in position $3 = 011_b$, a one in position $5 = 101_b$ and a one in position $7 = 111_b$. The columns of this matrix are the representations of these numbers in binary form.

An observant reader might complain that this has the disadvantage that it requires storing 2^m bytes, and if F doesn't have many columns, most of those will be 0's. But it's a minor disadvantage, because even at 10 rows we would only need 1024 bytes, and we usually have configurations for which the number of rows is 5 or less (32 bytes). Perhaps this would become more of an issue with configurations with a high number of rows, but for those configurations, most of our tasks would require too much time to be of any practical use.

We came to this representation after an implementation which represented columns as an array of bits (C++ bitset) and storing them into an ordered tree-like structure (C++ STL multiset). This might be a more natural implementation, but profiling the code made clear that the program was spending most of its time counting how many columns of a certain type appeared in a configuration, and was also spending a considerable amount of time navigating the tree. Explicitly storing the number of times each column appears, and making that number instantly accessible by storing it in an array (for random access) gives a very noticeable speedup and allows us to consider larger problems. By representing columns

as numbers, we can do a lot of preprocessing and compute large tables in which we have almost instant access time.

For example, consider the following problem, which has to be done many times for our tasks: Given a column α and a subset S (represented by an integer as written in binary), what column is α_S ? This is relatively slow to compute, but we can fill out a table by preprocessing to speed up any further access to it. Since we do this a few million times, the investment is sound.

The other advantage is that it becomes immediately clear how to compare two configurations with the same number of rows to see if one is a column-permutation submatrix of the other; check if for any column (index) the integer at position c of the first array is bigger than that of the second array. To check if F is a subconfiguration of A , we would need to find all permutations of the rows of F (which we need to do just once per configuration).

3.2 Subconfigurations

Suppose F and A are configurations and we want to decide if $F \prec A$. Then for every s -tuple of rows S (where s is the number of rows of F), we can extract from A the configuration $A|_S$ easily with our pre-stored table of columns and subsets. Once we've done this for each column of A and found $A|_S$, then for each permutation of rows of F , we check if every column α in the array corresponding to configuration F appears less than or equal to the corresponding number for column α in the array of $A|_S$. We can check every subset S like this. If at any point this is so, we can return true.

There are a few speedups. Sometimes it's immediately obvious a configuration can't be contained in another. For example, if there are more 1's in F than in A , or if F has more columns or rows than A , then $F \not\prec A$.

3.3 Determining $X(F)$

Given a configuration F , we wish to find $X(F)$. In other words, we wish to find the conjectured asymptotic bound for $\text{forb}(m, F)$.

We may make a simplification using Lemma 1.4.2 and assume the multiplicity of any column of F is at most 2.

First, suppose we wanted to test whether or not there exists r such that configuration F

is contained in the product

$$P_r(a, b, c) = \underbrace{I_r \times \dots \times I_r}_{a \text{ times}} \times \underbrace{I_r^c \times \dots \times I_r^c}_{b \text{ times}} \times \underbrace{T_r \times \dots \times T_r}_{c \text{ times}}.$$

Building this object with $r = R$ as calculated in Lemma 1.4.2 would be prohibitively slow. Instead, we build a set \mathcal{X} of subconfigurations from $P_R(a, b, c)$ such that if $F \prec P_R(a, b, c)$, then $F \prec X$ for some $X \in \mathcal{X}$.

Notice that if $F \prec P_R(a, b, c)$, then the rows of F get “partitioned” into $a + b + c$ parts (a part can be empty), where each part belongs to a factor of the product $P_R(a, b, c)$. Because of Lemma 1.4.2, we can assume each column appears at most twice. Given $s \in \mathbb{N}$, consider the following matrices:

$$A_I(s) = \begin{bmatrix} 0_s & 0_s & I_s \end{bmatrix}, \quad A_{I^c}(s) = \begin{bmatrix} 1_s & 1_s & I_s^c \end{bmatrix}, \quad A_T(s) = \begin{bmatrix} T_s & T_s \end{bmatrix}.$$

We see that an s -rowed configuration F with each column repeated at most twice is contained in I_m for some large m , if and only if $F \prec A_I(s)$. We can then consider all partitions of rows of F and see if each part is contained in the corresponding A_I , A_{I^c} or A_T .

For example, to test whether

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

is contained in $I \times T \times T$, we would partition the rows of F in three parts. In this case, F has five rows, so consider, for example, the following partition of 5: $(2, 2, 1)$, for example. Consider the following representatives:

$$A_I(2) = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad A_T(2) = \begin{bmatrix} 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}, \quad A_T(1) = \begin{bmatrix} 0 & 0 & 1 & 1 \end{bmatrix},$$

and build a 5-rowed matrix $A := A_I(2) \times A_T(2) \times A_T(1)$. If $F \prec A$, then $F \prec I \times T \times T$. If we do this for every possible partition of 5, we get the desired result.

To find $X(F)$ we can build a tree of possibilities of products and return the largest for which F not a subconfiguration of a product of this form, observing that if F is a

subconfiguration of a product $P_r(a, b, c)$, then it will be a subconfiguration in any product $P_r(a', b', c')$ with $a' \geq a$, $b' \geq b$ and $c' \geq c$.

3.4 Boundary Cases to Classify Configurations

We wish to find all boundary cases (Definition 1.4.3) for a given number of rows s and a number k using the computer. We use the program described in the previous section to find $X(F)$ for many different F .

The method we use is very straightforward: start adding columns, one by one. Build a tree of configurations, where the children of a configuration F are the ones that consist of F plus a column. Find $X(F)$ for each configuration in the tree. Store all those configurations F for which $X(F) = k$, and then find only the maximal and minimal configurations in the ordering \prec . If at some point we add a column and the bound jumps to $k + 1$ or higher, discard and go to the next configuration. Columns will be repeated at most twice because of Lemma 1.4.2.

Unfortunately this method is very slow because the same configuration is searched multiple times, since each configuration may have many representatives. To get rid of repetition we might check for equivalence of configurations against everything we have stored so far. But checking if two configurations are equivalent is usually slow, so doing it every time is also very slow.

What seems to work best is to do check for equivalence, but only up to a point. For example, we can consider all pairs of columns, test those and only take one representative of each equivalence class. Then from each pair, start building the tree as described above. After that, it will be relatively unlikely that two configurations we search are equivalent, so the amount of repetition will be relatively low. Of course, much repetition will still occur, but much less than in the original tree.

Notice that it isn't as critical that classifying configurations be a fast calculation. We need to do it once for every s and k , but no more. Once we know the maximal and minimal quadratics for five rows, we never need to calculate them again. Other calculations, such as finding $X(F)$ or What Is Missing, have to be performed for each configuration (or family) we wish to study, so decreasing the running time is more of a priority in those cases since faster programs allow us to study bigger configurations.

3.5 Finding What Is Missing and Forbidden

We describe the part of the program whose input is a family of configurations \mathcal{F} and a number s , and its output is the list of possibilities for columns absent or in short supply. This computation has been the most useful for the thesis.

To find a list of possibilities of What Is Missing in s rows after forbidding $F \in \mathcal{F}$, we present the problem in a complementary way and instead find what can be present.

To find this we may start with $S_t = t \cdot K_s$ (where t is the maximum multiplicity of a column in any configuration of \mathcal{F}) and remove columns from it doing breadth first search, to find all the maximal configurations A for which $F \not\prec A$. Then the $\{0, 1\}$ -complements of such A 's are the list of What Is Missing.

This is very slow, since the search space could have $2^{2^{ts}}$ configurations. But we can speed up the calculations. For example, by carefully choosing which columns we start our search with (i.e. removing unnecessary columns from S_t). The main speedup for this program comes from choosing a different starting point by considering each column sum separately in the following manner. For a sequence $(t_0, t_1, \dots, t_s) \in \mathbb{N}^s$, consider

$$S := [t_0 \cdot K_s^0 \mid t_1 \cdot K_s^1 \mid \dots \mid t_s \cdot K_s^s].$$

Instead of making $t_i = t$, we can make t_i the biggest number a column of column sum i appears in any configuration \mathcal{F} . Of course, since configurations $F \in \mathcal{F}$ might have less than s rows, each column might count for many different column sums, in all the ways it could be filled. In other words, if F has k rows, we consider the s -rowed configuration $F \times T_{s-k}$ and then count how many times each column appears. Then for each column sum take the maximum, and then take the maximum over all configurations $F \in \mathcal{F}$.

Perhaps an example would be useful. Suppose $s = 4$ and \mathcal{F} consist of the following two configurations:

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

We wish to find a starting point S for the search. We are searching then for t_0, t_1, t_2, t_3 and t_4 . As instructed above, since F_1 only has 3 rows, it acts as $F_1 \times T_1 = F_1 \times [0 \ 1]$ for the purpose of considering column sums.

We can deduce that $t_0 = 0$, since there are no columns of column sum 0. Since $F_1 \times T_1$

with a 0 on the bottom contains columns of column-sum 1, but no repeated columns of column sum 1, we can set $t_1 = 1$. F_2 has a column of sum two with multiplicity two, so $t_2 = 2$. And because of F_1 , we have both $t_3 = 1$ and $t_4 = 1$. Then instead of starting with $2 \cdot K_4$, which has 32 columns, we may start with the following matrix:

$$S = [K_4^1 \mid 2 \cdot K_4^2 \mid K_4^3 \mid K_4^4]$$

$$= \begin{bmatrix} 100011110011000011101 \\ 010011001100110011011 \\ 001000111100001110111 \\ 000100000011111101111 \end{bmatrix},$$

which has only 22 columns. The search space has size 2^{22} ; 1024 times smaller than 2^{32} .

This alone provides a massive reduction for most families \mathcal{F} . Using Lemma 1.4.2, if the multiplicity of a column is 3 or more, we can usually assume it is 2 for our purposes. We only distinguish between columns absent, those in short supply and those in long supply.

Once we have the list of subconfigurations A of S such that $F \not\prec A$, we take only the maximals with respect to order \prec . What Is Missing can be obtained by considering $S - A$.

For the example above, the computer gives us the list of What Is Missing in about one second.

$$P_0 = \begin{array}{cccccccc} \text{no} & \text{no} \\ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \end{array}$$

$$P_1 = \begin{array}{cccccccc} \text{no} & \text{no} & < 2 & \text{no} & < 2 & \text{no} & \text{no} & \text{no} & \text{no} \\ \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \end{array}$$

$$P_2 = \begin{array}{cccccccc} \text{no} & \text{no} \\ \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \end{array}$$

$$P_{13} = \begin{array}{cccccccccccc} \text{no} & \text{no} & \text{no} & \text{no} & \text{no} & \text{no} & < 2 & \text{no} & < 2 & \text{no} & \text{no} & \text{no} \\ \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \end{array}$$

$$P_{14} = \begin{array}{cccccccc} \text{no} & \text{no} \\ \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \end{array}$$

In practice checking configurations with $s \leq 4$ is almost instantaneous, $s = 5$ takes, depending on the configuration, anywhere from a few minutes to a couple of hours, and with $s = 6$ it's typically hopeless, although it can be done if for example the configuration only has columns of column sum 3 and is simple, since $2^{\binom{6}{3}}$ is still a reasonable number.

Finally, note that finding $\text{forb}(m, \mathcal{F})$ is a subproblem of this, when $s = m$. Start with the complete matrix K_m and remove columns one by one until we stop having $F \in \mathcal{F}$. Perform breadth-first on the tree of column deletions in order to search for the configuration with the most number of columns for which no $F \in \mathcal{F}$ is a subconfiguration.

Algorithm 3.5.1 Here is some pseudo-code to find What Is Missing (WIM).

Input: A family of configurations \mathcal{F} , a number s

Output: A list of possibilities for What Is Missing

for $i = 0$ to s **do**

$t_i := \max\{\lambda(\alpha, F \times T_{s-k}) : \alpha \text{ column of column sum } i \text{ and } F \in \mathcal{F} \text{ with } k \text{ rows}\}$

end for

$S := [t_0 \cdot K_s^0 \mid t_1 \cdot K_s^1 \mid \dots \mid t_s \cdot K_s^s]$

$List := \emptyset$

for all s -rowed $K \prec S$ **do**

if $F \not\prec K$ **then**

Add K to $List$

end if

```
end for  
List = maximals(List)  
ListWIM := {S - K : K ∈ List}  
return ListWIM
```

3.6 Guessing Forb

Consider \mathcal{F} a family of (small) configurations and m a (small) fixed integer. Suppose we wish to find (or rather, guess) $\text{forb}(m, \mathcal{F})$ using the computer. This approach has helped considerably with the proofs of 3 exact bounds (described in Section 4.1 and Section 4.3). For these results, once we had an extremal matrix that avoids \mathcal{F} it was relatively straightforward to construct a proof of this fact. We defer the details of the proofs to the subsequent sections. For now, we will describe the methods used to provide us with what seems to be a good and dependable guess of $\text{forb}(m, \mathcal{F})$.

The idea is to consider all the 2^m columns in some order and add them one by one into a matrix A , making sure at each step we don't create any $F \in \mathcal{F}$ as a configuration. The order in which to add columns is what will determine the size of $\|A\|$ at the end of the process. To do this, enumerate all the columns and for each permutation of $[2^m]$ find the number of columns that would be added while avoiding all $F \in \mathcal{F}$ if we were to add them in that order. We call the columns that would get added in this procedure *good* columns and the columns that get discarded *bad* columns, with respect to the given permutation. We wish then to find a permutation that maximizes the number of *good* columns. So the search space has size $(2^m)!$, which is too big to search exhaustively, but our experiments suggest local search can often find the correct answer when m isn't too large. We'll describe three methods.

Perhaps this explanation could benefit by an example. Suppose we wanted to forbid

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

and suppose $m = 3$. Consider all eight 3-rowed columns in the following order:

$$\begin{array}{cccccccc}
 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\
 \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}
 \end{array}$$

Then consider the permutation $(4, 1, 3, 6, 5, 8, 2, 7)$. We start adding columns in the order given by this permutation. If adding a certain column would rise to a copy of F , then don't add it and continue to the next one. This is how it would work for this permutation:

1. Add column 4.

$$A = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

2. Add column 1:

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

3. Can't add column 3.

4. Add column 6:

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

5. Can't add column 5.

6. Add column 8:

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

7. Can't add column 2.

8. Can't add column 7.

We end up with 4 columns, with the *good* columns being $\{4, 1, 6, 8\}$ and the *bad* columns being $\{3, 5, 2, 7\}$.

3.6.1 Brute Force Greedy Search

The brute force method does the following:

- Choose a permutation at random.
- Separate into *good* and *bad* columns.
- Count the number of *good* columns.
- Repeat, while keeping track of the “best” so far (i.e. the one with the most number of *good* columns)

The strength of this method is that we may do this thousands of times in a relatively short time. This method is good for simple configurations in which there are many different ways to achieve the bound, but in our experience, it fails to find the optimal bound in many cases.

An easy bound on the probability p of finding the best configuration in a single try is the following. Let $f = \text{forb}(m, \mathcal{F})$. Then

$$p \geq \frac{(2^m - f + 1)}{\binom{2^m}{f}}.$$

Indeed, if $A \in \text{ext}(m, \mathcal{F})$ and g_i (for $i \in \{1, 2, \dots, f\}$) is any permutation of the columns of A and b_i (for $i \in \{1, 2, \dots, (2^m - f)\}$) is a permutation of the columns that are not in A , we see that for any $j \in \{0, 1, \dots, (2^m - f)\}$ any permutation of the form:

$$(g_1 g_2 \dots g_{f-1} b_1, b_2, \dots, b_j, g_f, b_{j+1}, \dots, b_{2^m - f})$$

gives rise to a matrix in $\text{ext}(m, \mathcal{F})$, since by separating good columns from bad columns, as described in the previous section, we see that at least we are allowed to add g_1, g_2, \dots, g_{f-1} to the good columns without interruption, and we must be allowed to add to the good columns at least one of b_1, b_2, \dots, b_i or g_f .

The above bound might not be very accurate, as many other permutations might give rise to extremal matrices, but at least it grounds the problem.

3.6.2 Hill Climbing

We describe a method that improves over brute force. Start with some permutation and separate the 2^m columns into *good* and *bad* columns. Then for each column b in *bad*, form a

new permutation by putting b at the beginning, but leaving all others in place. This ensures that the chosen column b is selected. From all the possibilities for choice of b , choose the best one, that is, the one that gives the most columns in *good*. If there are ties, pick one at random. Repeat this process until there is no improvement. In our example, we would have to consider all the following permutations:

1. (3, 4, 1, 6, 5, 8, 2, 7)
2. (5, 4, 1, 3, 6, 8, 2, 7)
3. (2, 4, 1, 3, 6, 5, 8, 7)
4. (7, 4, 1, 3, 6, 5, 8, 2)

In this case all of them give size 4, but in general some might be better than others. Local search as described above gets stuck very quickly, but it does perform better than brute force.

Algorithm 3.6.1 Here is some pseudo-code to perform hill climbing. This code gives up if it gets stuck. It's easy to modify it to be able to "walk" while being stuck, and if there is no improvement after, say, 20 iterations, stop.

Input: A family of configurations \mathcal{F} , a number m , a permutation of 2^m , σ .

Output: A configurations in $\text{Avoid}(m, \mathcal{F})$ that is thought to be extremal.

stuck := false

Find $good(\sigma)$ and $bad(\sigma)$

while not *stuck* **do**

for all $b \in bad(\sigma)$ **do**

$\sigma_b := (b, \sigma \setminus b)$

 Separate σ_b into $good(\sigma_b)$ and $bad(\sigma_b)$.

end for

$\phi := \operatorname{argmax}\{|good(\sigma_b)| : b \in bad(\sigma)\}$

if $|good(\phi)| > |good(\sigma)|$ **then**

$\sigma \leftarrow \phi$

else

stuck = true

end if

```
end while
return  $\sigma$ 
```

3.6.3 Genetic Algorithm

An even better way to search the space is using a Genetic Algorithm. The idea is to mimic evolution. We maintain a population of ‘individuals’ and assign ‘fitness’ to them in some way (in our case, the number of ‘good’ columns). Then pick two at random, but giving higher probability to those with higher fitness. Call them father and mother. Then combine them in some way to produce offspring, hopefully with better results than either father or mother. Do this for many generations. There are many ways to combine permutations of course. The one we found that works well is as follows:

1. For both father and mother, separate the columns into *good* and *bad*.
2. Take a random number of the *good* part of mother, and consider all columns from the start up to the chosen random number.
3. Permute the entries of father so that the numbers in the chosen part of mother are in the same order as those in mother.

This is better shown with an example. Suppose we had a father and a mother like this:

$$\text{father} = (2, 3, 6, 1 \mid 7, 4, 8, 5) \text{ and } \text{mother} = (5, 3, 8, 1, 4 \mid 7, 2, 6)$$

with the numbers shown before the vertical line being the *good* part and the numbers shown after the vertical line being the *bad* part. Pick a random entry in the *good* part of mother. For example, pick *1* and look at the numbers that appear (in mother) to the left of the picked number. In this case, (5, 3, 8, 1). In father, select these numbers.

$$\text{father} = (2, \mathbf{3}, 6, \mathbf{1} \mid 7, 4, \mathbf{8}, \mathbf{5})$$

Make child by shuffling the selected entries in father so that they match the order of mother.

$$\text{child} = (2, \mathbf{5}, 6, \mathbf{3}, 7, 4, \mathbf{8}, \mathbf{1}).$$

And finally, separate the columns of *child* into *good* and *bad*. This approach has given very good results in our experience.

Algorithm 3.6.2 Here is some pseudo-code to perform genetic search as described in this section. After some empirical experimentation, we find that using a population size of 40 and 200 generations are good numbers.

Input: A family of configurations \mathcal{F} , a number m

Output: A list of configurations in $\text{Avoid}(m, \mathcal{F})$ that are thought to be extremal.

Define *Population* to be a list 40 random permutations of 2^m .

for $i = 0$ to 200 **do**

for all $\sigma \in \text{Population}$ **do**

 Separate σ into $\text{good}(\sigma)$ and $\text{bad}(\sigma)$ columns.

end for

 Define *NewGeneration* := 10 best from *Population*.

while $|\text{NewGeneration}| < 40$ **do**

 Pick *father* and *mother* from *Population* {Pick random permutations σ according to how big $\text{good}(\sigma)$ is}

 Mix *father* and *mother* to produce *child*

 Add *child* to *NewGeneration*

end while

Population \leftarrow *NewGeneration*

end for

Chapter 4

Exact Bounds

4.1 Two 3x4 Exact Bounds

A list of 3×4 configurations for which the exact bounds are not known are listed in [AK10]. We solve two of them, and give conjectured bounds for the rest in Section 8.1.3.

4.1.1 Introduction

We study the following configurations V, W (where $\text{forb}(m, V)$ and $\text{forb}(m, W)$ weren't previously known), and compare to the common submatrix X (where $\text{forb}(m, X)$ was found in [ABS11]).

$$X = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 0 & 0 \end{bmatrix}, \quad V = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}, \quad W = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}.$$

We used our Genetic Algorithm to seek extremal matrices $A \in \text{ext}(m, V)$ and $A \in \text{ext}(m, W)$ for small m . From these examples we guess the structure of extremal matrices in general and then we are subsequently able to prove these matrices are indeed extremal. Guessing such structures would have been challenging without the help of the Genetic Algorithm.

Theorem 4.1.1 [ABS11] Let $m \geq 3$. Then $\text{forb}(m, X) = \binom{m}{2} + \binom{m}{1} + \binom{m}{0} + 1$.

Note that $X \prec V$ and $X \prec W$. We initially thought (before using the Genetic Algorithm) that $\text{forb}(m, V) = \text{forb}(m, W) = \text{forb}(m, X)$. Related results are in [ABS11], [AK06].

The structure of the apparently extremal matrices generated by the Genetic Algorithm provided the strategy to tackle the bounds for $\text{forb}(m, V)$ and $\text{forb}(m, W)$. Interestingly, the Genetic Algorithm was used again in each example in the inductive proof to guess $\text{forb}(m, \mathcal{H}_V)$ and $\text{forb}(m, \mathcal{H}_W)$.

4.1.2 The Bound for W

Recall that

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

We can use the computer to compute $\text{forb}(m, W)$ for $m \leq 5$ as described in Section 3.5. We then use the Genetic Algorithm to compute guesses for $\text{forb}(m, W)$ for $m = 6, 7, 8$ and also guess the structure of the extremal matrices. We then proceed to prove these guesses.

Theorem 4.1.2 Let $m \geq 2$. Then $\text{forb}(m, W) = \binom{m}{2} + \binom{m}{1} + \binom{m}{0} + m - 2$.

In order to prove this we proceed using the Standard Induction of Section 2.1: Let $A \in \text{Avoid}(m, W)$. If we could find some row r for which the number of repeated columns $\|C_r(A)\| \leq m + 1$, then we would be done:

$$\begin{aligned} \|A\| &\leq m + 1 + \binom{m-1}{2} + \binom{m-1}{1} + \binom{m-1}{0} + (m-1) - 2 \\ &= \binom{m}{2} + \binom{m}{1} + \binom{m}{0} + m - 2. \end{aligned}$$

So we may assume that $\|C_r(A)\| \geq m + 2$ for every r . The minimal inductive children of W are:

$$G_1 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad G_2 = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}, \quad G_3 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \end{bmatrix}.$$

Given that $C_r(A) \in \text{Avoid}(m-1, \{G_1, G_2, G_3\})$, under these assumptions the following Lemma establishes that $\|C_r(A)\| = (m-1) + 3 = m+2$.

Lemma 4.1.3 Let $m \geq 4$. Then $\text{forb}(m, \{G_1, G_2, G_3\}) = m+3$.

Proof: For the lower bound $\text{forb}(m, \{G_1, G_2, G_3\}) \geq m+3$ an example (which was again found using our Local Search strategies) suffices. Consider the matrix

$$A = [\mathbf{0}_m \mid K_m^1 \mid \mathbf{1}_m \mid \alpha],$$

where α is any other column. Clearly $A \in \text{Avoid}(m, \{G_1, G_2, G_3\})$ and $\|A\| = m+3$. To prove the upper bound, we proceed by induction on m . Let $A \in \text{Avoid}(m, \{G_1, G_2, G_3\})$. Then, if we forbid $\{G_1, G_2, G_3\}$, below are the 16 possible cases of What Is Missing on each quadruple of rows, found using the program described in Section 3.5. The following list is complete because the search is exhaustive.

no	≤ 1	> 1	> 1												
$\begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$

no	≤ 1	> 1	> 1												
$\begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$

no	≤ 1	> 1	> 1												
$\begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix}$

$$\begin{array}{cccccccccccccccc}
\text{no} & \leq 1 & \leq 1 & > 1 & > 1 \\
\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}
\end{array}$$

We check this list to see that there are at most seven columns present on four rows (at least nine of sixteen are absent) and so $\text{forb}(4, \{G_1, G_2, G_3\}) = 7$. Now consider $m \geq 5$. It's easy to check by considering every quadruple of rows that there is a row and a column we can delete from A and keep the remaining $(m - 1)$ -rowed matrix A' simple. Then by induction, $\|A'\| \leq (m - 1) + 3 = m + 2$ and then $\|A\| \leq \|A'\| + 1 \leq m + 3$. To find such a row and column, look at the columns marked ≤ 1 and > 1 , and see that there is a row we can delete such that the only repeat (if there is one) has one of the columns marked ≤ 1 . We used the help of the free software sage (<http://www.sagemath.org/>), but it could also be done by hand. For example, for the first possibility,

$$\begin{array}{cccccccc}
& \leq 1 & > 1 & > 1 \\
i & \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \\
j & & & & & & & \\
k & & & & & & & \\
\ell & & & & & & &
\end{array} ,$$

deleting the row of A corresponding to row ℓ and the column of A corresponding to the fifth column above $(0, 0, 0, 1)^T$ (if it exists), keeps the remaining matrix simple. In all cases above one can delete the final row (together with at most one column in short supply). Thus $\|C\| \leq m + 3$. ■

We need a more detailed lemma about an $m \times (m + 3)$ matrix in $\text{Avoid}(m, \{G_1, G_2, G_3\})$, one that was predicted using our Genetic Algorithm.

Lemma 4.1.4 Let $A \in \text{Avoid}(m, \{G_1, G_2, G_3\})$ with $m \geq 3$ and $\|A\| = m + 3$. Then $K_m^1 \prec A$. Moreover, the remaining three columns are $\mathbf{0}_m$ and two additional columns α, β with $\alpha < \beta$ (meaning that on each row for which α has a 1, β also has a 1).

Proof: We proceed by induction on m . We checked all cases with $m = 3, 4$ using a computer.

Assume $A \in \text{Avoid}(m, \{G_1, G_2, G_3\})$ with $\|A\| = m + 3$ and $m \geq 5$. From our proof of Lemma 4.1.3, there is a row and a column from A we can delete to obtain a $(m - 1) \times (m + 2)$ simple matrix A' . We may assume $K_{m-1}^1 \prec A'$ by induction. Assume we deleted the last row from A to obtain A' and that the deleted column was the last column of A . If we restrict ourselves to the first $m - 1$ rows, the deleted column has to be a repeat of one of the columns of A' , else we would have an $(m - 1) \times (m + 3)$ simple matrix contradicting Lemma 4.1.3. At this point the proof is finished save for a case analysis on each of the possible columns to repeat.

Aside from the column of zeros, there are two columns α, β which aren't in K_{m-1}^1 . Given $\alpha < \beta$, we call α the *small* column and β the *big* column. We consider 3 different types of rows:

- Row type 1: Both α, β has 0 in the row.
- Row type 2: Column β has a 1 and α has a 0 in the row.
- Row type 3: Both α, β has 1 in the row.

There may not be any rows of type 1, but there has to be at least one row of type 2 (in order to differentiate between α and β) and at least two rows of type 3 (in order to differentiate between α and a column of column-sum 1. Consider the generic rows below. We've included the appropriate parts of the copy of K_{m-1}^1 and the column of 0's. The entries marked $c_1, c_2, \dots, c_8, r_1, \dots, r_4$ are the entries of the deleted row and column.

							α	β		
0	1	0	0	0	0	0	0	r_1	type 1	
0	0	1	0	0	0	0	1	r_2	type 2	
0	0	0	1	0	1	1	1	r_3	type 3	
0	0	0	0	1	1	1	1	r_4	type 3	
c_1	c_2	c_3	c_4	c_5	c_6	c_7	c_8			

Of course there might be many rows of each of the types 1,2,3, but there is no loss of generality if we focus on these rows. There have to be at least two rows of type 3 so it is possible to have two rows which correspond to the entries r_3, r_4 . We have to be careful because row r_1 might not exist.

There are some cases for which column gets repeated. In each case we attempt to find either G_1, G_2 or G_3 . The fact that A has no G_1 means A is laminar, as in Definition 1.3.27.

The following case analysis is probably much easier as an exercise for the reader than it is to either write or read about.

Case 1: $\mathbf{0}_{m-1}$ is the repeated column. Then we have

$$\begin{array}{ccccccc|c} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ \hline 0 & c_2 & c_3 & c_4 & c_5 & c_6 & c_7 & 1 \end{array} .$$

So either $K_m^1 \prec A$ or some of c_2, c_3, c_4, c_5 are 1. If $c_2 = 1$ then $c_6 = 0$ and $c_7 = 0$ in order to have a laminar matrix. But then we have G_2 in the last and next-to-last rows. So we may assume $c_2 = 0$. If $c_3 = 1$ then $c_5 = 0$, $c_6 = 0$ and $c_7 = 1$ by the laminar property and we have then

$$\begin{array}{ccccccc|c} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ \hline 0 & 0 & 1 & c_4 & 0 & 0 & 1 & 1 \end{array} .$$

But then we have G_2 in the last two rows. We may assume then that $c_3 = 0$. If $c_4 = 1$ then both c_6 and c_7 have to be 1, and so we get

$$\begin{array}{ccccccc|c} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ \hline 0 & 0 & 0 & 1 & c_5 & 1 & 1 & 1 \end{array} ,$$

which has G_3 in the last two rows. This completes Case 1.

Case 2: The repeated column has column sum 1. Then there are three sub-cases, depending on the position of the 1 in the new column. Let r be the row on which, other than the last row, the new column has a 1.

Subcase 2a: r is of type 1. We have

$$\begin{array}{ccccccc|c}
 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\
 \hline
 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\
 \hline
 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\
 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\
 \hline
 c_1 & 0 & c_3 & c_4 & c_5 & c_6 & c_7 & 1
 \end{array} ,$$

which contains G_2 in the first two rows.

Subcase 2b: r is of type 2. We have

$$\begin{array}{ccccccc|c}
 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
 \hline
 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 \\
 \hline
 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\
 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\
 \hline
 c_1 & c_2 & 0 & c_4 & c_5 & c_6 & c_7 & 1
 \end{array} ,$$

which contains G_2 in the second and third rows.

Subcase 2c: r is of type 3. We have

$$\begin{array}{ccccccc|c}
 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
 \hline
 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\
 \hline
 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \\
 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\
 \hline
 c_1 & 0 & c_3 & c_4 & c_5 & c_6 & c_7 & 1
 \end{array} ,$$

which contains G_3 in the third and fourth row.

Case 3: The repeated column is the *small* column α . Then we have this:

$$\begin{array}{ccccccc|c}
 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
 \hline
 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\
 \hline
 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \\
 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\
 \hline
 c_1 & c_2 & c_3 & c_4 & c_5 & 0 & c_7 & 1
 \end{array} .$$

So c_7 has to be 1 in order to have a laminar matrix. If either c_4 or c_5 were 0, then we

would have G_3 in the last row together with one of the next-to-last rows. So both have to be 1. But this contradicts the fact that we have a laminar matrix.

Case 4: The repeated column is the *big* column β : Then we have this:

$$\begin{array}{cccccc|c} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ \hline 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ \hline c_1 & c_2 & c_3 & c_4 & c_5 & c_6 & 0 & 1 \end{array} .$$

This yields G_3 in the second and third rows. This completes all cases.

We've concluded that $K_m^1 \prec A$. We deduce that, apart from the columns of K_m^1 and $\mathbf{0}_m$, A has two other columns α and β . Since $G_2 \not\prec A$ we deduce that either $\alpha < \beta$ or $\beta < \alpha$. ■

Proof of Theorem 4.1.2: We use induction on m . The result is true for $m = 2, 3$ so we may assume $m \geq 4$. Let $A \in \text{Avoid}(m, W)$. Apply the decomposition (2.1.1). If $\|C_r(A)\| \leq m + 1$, then we can apply the Standard Induction Section 2.1 to establish the bound for $\|A\|$. So assume $\|C_r(A)\| \geq m + 2$ for all choices of r . By Lemma 4.1.3, we deduce that $\|C_r(A)\| = m + 2$ for every row r , and by Lemma 4.1.4, we have that $K_{m-1}^1 \prec C_r(A)$ also for every row r . Thus $K_m^2 \prec A$, since all columns of column sum 1 in $C_r(A)$ appear with a 1 in row r (and this happens for every row r). We also have $[K_m^1 \mid K_m^0] \prec A$. Now in every triple of rows of K_m^2 we have the matrix G_1 once in every ordering of the triple. Given that $W = 2 \cdot G_1$, the columns of A of column sum at least 3 have no configuration G_1 . We appeal to Lemma 4.1.5 below to deduce that the number of columns in A of column sum at least 3 is at most $m - 2$. Then

$$\|A\| \leq \binom{m}{2} + \binom{m}{1} + \binom{m}{0} + m - 2,$$

which yields the desired bound. ■

A quick counting argument reveals the following general result about laminar families which we use in the proof of Theorem 4.1.2.

Lemma 4.1.5 Let $m \geq 3$ and let \mathcal{Z} be a laminar family of subsets of $[m] = \{1, 2, \dots, m\}$ with the property that for all $Z \in \mathcal{Z}$ we have $|Z| \geq 3$. Then $|\mathcal{Z}| \leq m - 2$ and furthermore, this bound is tight (i.e. there exists a family \mathcal{Z} for which $|\mathcal{Z}| = m - 2$). Thus, if $A \in \text{Avoid}(m, G_1)$ satisfies that all column sums are at least 3, then $\|A\| \leq m - 2$.

Proof: Let $f(x)$ denote the size of the biggest laminar family of $[x]$ with no sets of size less than or equal to 2. Recall that a laminar family is equivalent to a configuration which avoids G_1 . Assume $A \in \text{Avoid}(m, G_1)$ with the property that all column sums are at least 3. We wish to show that $f(m) = m - 2$.

The family $\{[3], [4], \dots, [m]\}$ has size $m - 2$, which proves $f(m) \geq m - 2$. We wish to prove $f(m) \leq m - 2$.

Let \mathcal{Z} be such that $|\mathcal{Z}| = f(m)$. We may assume $[m] \in \mathcal{Z}$: if $[m] \notin \mathcal{Z}$, observe that $\mathcal{Z} \cup \{[m]\}$ is also a laminar family of bigger size. Suppose then that the next biggest set Z in \mathcal{Z} has size k . We partition $[m]$ into two disjoint sets: Z and $[m] \setminus Z$. Every set $Y \in \mathcal{Z}$ satisfies either $Y \subseteq Z$ or $Y \subseteq [m] \setminus Z$ or $Y = [m]$. This gives the recurrence

$$f(m) \leq 1 + f(k) + f(m - k).$$

If $k \neq m - 1$, then by induction $f(k) = k - 2$ and $f(m - k) = m - k - 2$, and so $f(m) \leq 1 + k - 2 + m - k - 2 = m - 3$, a contradiction. When $k = m - 1$ we have $f(m) \leq m - 2$. Moreover if $f(m) = m - 2$ we observe that \mathcal{Z} is “equivalent” to $\{[3], [4], \dots, [m]\}$. ■

4.1.3 The Bound for V

Recall that

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

Using the computer, we can prove by exhaustively looking at all the possibilities that $\text{forb}(3, V) = 8$, $\text{forb}(4, V) = 13$, and $\text{forb}(5, V) = 18$. Using the Genetic Algorithm of Section 3.6, we obtained large matrices with no subconfiguration V which suggested to us that $\text{forb}(6, V) = 25$, $\text{forb}(7, V) = 32$, $\text{forb}(8, V) = 40$. For $m \geq 6$, this suggests $\text{forb}(m, V) = \binom{m}{2} + \binom{m}{1} + \binom{m}{0} + 3$, two more than our first guess that $\text{forb}(m, V) = \text{forb}(m, X)$.

Theorem 4.1.6 Let $m \geq 6$. Then $\text{forb}(m, V) = \binom{m}{2} + \binom{m}{1} + \binom{m}{0} + 3$.

The Genetic Algorithm also predicted the following structure of matrices in $\text{ext}(m, V)$. Recall the product \times defined in Definition 1.3.7. Consider a partition of the m rows into two disjoint sets, say $U = \{1, 2, \dots, u\}$ and $L = \{u + 1, u + 2, \dots, m\}$. Suppose $|U| = u$ and $|L| = \ell$ with $u + \ell = m$. Let A have the following structure:

$$A = \begin{array}{c} U \\ L \end{array} \left[\begin{array}{c|c|c|c|c|c|c|c|c|c} \mathbf{0}_u & \mathbf{0}_u & \mathbf{0}_u & \mathbf{0}_u & K_u^1 & K_u^1 & K_u^2 & \mathbf{1}_u & \mathbf{1}_u \\ \times & \times \\ \mathbf{0}_\ell & \mathbf{1}_\ell & K_\ell^{\ell-1} & K_\ell^{\ell-2} & \mathbf{1}_\ell & K_\ell^{\ell-1} & \mathbf{1}_\ell & \mathbf{0}_\ell & \mathbf{1}_\ell \end{array} \right]. \quad (4.1.1)$$

We easily check that for $A \in \text{Avoid}(m, V)$ and $3 \leq u, \ell \leq m - 3$, $\|A\| = \binom{m}{2} + \binom{m}{1} + \binom{m}{0} + 3$. We will prove that $A \in \text{ext}(m, V)$ and hence establish Theorem 4.1.6.

To prove this, consider $A \in \text{Avoid}(m, V)$ and apply the standard decomposition (2.1.1). The minimal inductive children of V are:

$$H_1 = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad H_2 = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}, \quad H_3 = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix}$$

Thus, $C_r(A) \in \text{Avoid}(m - 1, \{H_1, H_2, H_3\})$. We used the computer again to conjecture a structure on a matrix in $\text{Avoid}(m, \{H_1, H_2, H_3\})$.

Lemma 4.1.7 Let $m \geq 4$ and $A \in \text{Avoid}(m, \{H_1, H_2, H_3\})$. We have $\|A\| \leq m + 2$.

Proof: Using the program described in Section 3.5, we find that one of the following must hold for each quadruple of rows of A :

$$\begin{array}{cccccccccccccccccc} \text{no} & \leq 1 & \leq 1 & \leq 1 & \leq 1 & > 1 & > 1 \\ \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \end{array}$$

$$\begin{array}{cccccccccccccccc}
\text{no} & \leq 1 & \leq 1 & \leq 1 & > 1 \\
\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix}
\end{array}$$

This verifies in particular that $\text{forb}(4, \{H_1, H_2, H_3\}) = 6$ (at least 10 columns are absent in each of the twenty cases). We proceed as we did in Lemma 4.1.3 and verify that each of these possibilities yields one row (the final row in each case) and at most one column that we can delete from A , maintaining simplicity. This means that $\|A\| \leq m + 2$ by induction.

■

Lemma 4.1.8 Assume $m \geq 4$. Let $A \in \text{Avoid}(m, \{H_1, H_2, H_3\})$ with $\|A\| = m + 2$. Then there exists a partition $U \subseteq [m]$ and $L = [m] \setminus U$ with $|U| = u \geq 1$ and $|L| = \ell = m - u \geq 1$ so that if we permute rows,

$$A = \begin{array}{c} U \\ L \end{array} \left[\begin{array}{c|c|c|c} \mathbf{0}_u & K_u^1 & \mathbf{0}_u & \mathbf{0}_u \\ \times & \times & \times & \times \\ \mathbf{1}_\ell & \mathbf{1}_\ell & K_\ell^{\ell-1} & \mathbf{0}_\ell \end{array} \right] \text{ or } A = \begin{array}{c} U \\ L \end{array} \left[\begin{array}{c|c|c|c} \mathbf{0}_u & K_u^1 & \mathbf{0}_u & \mathbf{1}_u \\ \times & \times & \times & \times \\ \mathbf{1}_\ell & \mathbf{1}_\ell & K_\ell^{\ell-1} & \mathbf{1}_\ell \end{array} \right] \quad (4.1.2)$$

Note that in the former case we must have $\ell \geq 2$ and in the latter $u \geq 2$.

Proof: We use induction on m . The computer is able to show that the result is true for $m = 3$. Assume $m \geq 4$. Let $A \in \text{Avoid}(m, \{H_1, H_2, H_3\})$. By our argument in Lemma 4.1.7, there is a row and a single column we can delete, leaving the remainder simple. Let A' be the resulting simple matrix. We may assume by induction that there exists disjoint sets U', L' such that $|U'| = a \geq 1$, $|L'| = b \geq 1$ where $a + b = m - 1$ so that after permuting rows and columns,

$$A' = \begin{array}{c} U' \\ L' \end{array} \left[\begin{array}{c|c|c|c} \mathbf{0}_a & K_a^1 & \mathbf{0}_a & \mathbf{0}_a \\ \times & \times & \times & \times \\ \mathbf{1}_b & \mathbf{1}_b & K_b^{b-1} & \mathbf{0}_b \end{array} \right], \text{ or } A' = \begin{array}{c} U' \\ L' \end{array} \left[\begin{array}{c|c|c|c} \mathbf{0}_a & K_a^1 & \mathbf{0}_a & \mathbf{1}_a \\ \times & \times & \times & \times \\ \mathbf{1}_b & \mathbf{1}_b & K_b^{b-1} & \mathbf{1}_b \end{array} \right].$$

We will assume $b \geq 2$. Assume the last column ($m + 2$) and last row (m) of A is deleted to obtain A' . After deleting row m , the last column of A must be one of the columns of A'

given that $\|A'\| = \text{forb}(m, \{H_1, H_2, H_3\})$. In order to avoid H_1, H_2 and H_3 in A , we can show that we have the desired structure (4.1.1) in A (with either a or b one larger than before).

We can make a few general comments about row m . If we have both a 0 and a 1 in row m under the columns containing K_a^1 , then using the two columns containing the 0 and 1 and two rows of the U' together with row m , we obtain a copy of H_1 , a contradiction. Similarly we cannot have both a 0 and a 1 in row m under the columns containing K_b^{b-1} .

It is also true that in row m we cannot have a 0 under K_a^1 and a 1 under K_b^{b-1} else we find a copy of H_1 in two columns containing 0 and 1 in row m and in a row of U' , a row of L' and row m .

We will consider the two cases that A' has either the column $\mathbf{0}_{m-1}$ or $\mathbf{1}_{m-1}$ together. Note that if $a = 1$, then A' has both $\mathbf{0}_{m-1}$ and $\mathbf{1}_{m-1}$.

As in the proof of Theorem 4.1.2, we must do some case analysis for which column gets repeated. Let γ be the last (repeated) column of A .

Case 1: $\gamma = \mathbf{0}_a \times \mathbf{1}_b$. We deduce that column $\mathbf{0}_{m-1}$ of A' (if present) appears with a 0 in row m else we have the subconfiguration H_1 in two rows of L' and row m . Similarly, if column $\mathbf{1}_{m-1}$ is present and $a \geq 2$ then $\mathbf{1}_{m-1}$ appears with a 1 in row m else we have a copy of H_1 in two rows of U' and row m .

Now if we have 1's in row m under both K_a^1 and K_b^{b-1} , then we have the desired structure with $U = U'$ and $L = L' \cup \{m\}$. Similarly, if we have 0's in row m under both K_a^1 and under K_b^{b-1} then we have the desired structure with $U = U' \cup \{m\}$ and $L = L'$.

Thus we may assume we have 1's under K_a^1 and 0's under K_b^{b-1} . Recall that either A' has $\mathbf{0}_{m-1}$, or A' has $\mathbf{1}_{m-1}$ but in that case $a \geq 2$. We have two subcases.

Subcase 1a: A' has $\mathbf{0}_{m-1}$ (so A has $\mathbf{0}_m$). If $a = 1$ we note that A' has $\mathbf{1}_{m-1}$ (and therefore we are in Subcase 1b). Else we find a copy of H_3 in A in a row of L' and row m in the column $\mathbf{0}_m$ of A , a column from $\mathbf{0}_a \times K_b^{b-1} \times \mathbf{0}_1$ with a 0 in the chosen row of L' , the column with $\mathbf{0}_a \times \mathbf{1}_b \times \mathbf{0}_1$ and a column from $K_a^1 \times \mathbf{1}_{b+1}$.

Subcase 1b: A' has $\mathbf{1}_{m-1}$ (so A has $\mathbf{1}_m$) and $a \geq 2$. We have $a \geq 2$ and find a copy of H_3 in A in a row of U' and row m in the column $\mathbf{1}_m$ of A , a column from $K_a^1 \times \mathbf{1}_{b+1}$ with a 1 in the chosen row of U' , the column with $\mathbf{0}_a \times \mathbf{1}_b \times \mathbf{0}_1$ and a column from $\mathbf{0}_a \times K_b^{b-1} \times \mathbf{0}_1$.

This completes Case 1.

Case 2: $\gamma = \mathbf{0}_a \times \beta_b$, where $\beta_b \in K_b^{b-1}$. Given that $\mathbf{0}_a \times \beta_b$ is repeated in A on rows $[m-1]$, we deduce that it appears both with a 1 and with a 0 on row m . Assume any other column $\mathbf{0}_a \times \beta'_b$ of A' (with $\beta'_b \in K_b^{b-1}$) has a 0 in row m . Then we find a copy of H_1 in row m and the two rows of L' where β_b and β'_b differ and in the column of A containing $\mathbf{0}_a \times \beta'_b$ and one of the two columns of A containing $\mathbf{0}_a \times \beta_b$ which differs from the first column in row m .

Case 3: $\gamma = \alpha_a \times \mathbf{1}_b$, where $\alpha_a \in K_a^1$. For $a \geq 2$ we may follow the argument of the previous case and find a copy of H_1 . Given $a = 1$, we have that A' contains $\mathbf{0}_{m-1}$ as well as $\alpha_a \times \mathbf{1}_b = \alpha_1 \times \mathbf{1}_{m-2} = \mathbf{1}_{m-1}$. We deduce that A has $\mathbf{1}_m$ and $\mathbf{1}_{m-1} \times \mathbf{0}_1$. Thus we can find a copy of H_3 in a row of U' together with a row of L' , in the two columns with $\mathbf{1}_{m-1}$ in A' , the column with $\mathbf{0}_{m-1}$ in A' and a column of $\mathbf{0}_1 \times K_b^{b-1}$ selected in order to have a 0 on the chosen row of L' .

Case 4: $\gamma = \mathbf{0}_a \times \mathbf{0}_b$. We find H_3 in two rows of L' since $b \geq 2$.

Case 5: $\gamma = \mathbf{1}_a \times \mathbf{1}_b$. If $a \geq 2$, we can find H_3 in two rows of U' . In case $a = 1$, we know $\mathbf{0}_{m-1}$ as well as $\mathbf{1}_{m-1}$ are in A' . We find H_3 using a row of U' and a row of L' where we choose a column of $\mathbf{0}_a \times K_b^{b-1}$ that has a 0 in the chosen row of L' , plus $\mathbf{0}_{m-1}$ of A' , $\mathbf{1}_{m-1} \times 0$ and $\mathbf{1}_{m-1} \times 1$.

This completes the proof of Lemma 4.1.8. ■

Proof of Theorem 4.1.6: We use induction on m for $m \geq 6$. We established by computer that $\text{forb}(5, V) = 18$ (which is smaller than the bound of Theorem 4.1.6). Noting that $\text{forb}(5, \{H_1, H_2, H_3\}) = 7$, we deduce using Section 2.1 that

$$\text{forb}(6, V) \leq \text{forb}(5, V) + \text{forb}(5, \{H_1, H_2, H_3\}) = 18 + 7 = 25$$

and so $\text{forb}(6, V) = 25$, because of construction (4.1.1). Thus, we may assume $m \geq 7$.

By induction, assume $\text{forb}(m-1, V) = \binom{m-1}{2} + \binom{m-1}{1} + \binom{m-1}{0} + 3$. Let $A \in \text{Avoid}(m, V)$ with $\|A\| = \text{forb}(m, V)$. Apply the standard decomposition (2.1.1) to A for some row r . If

$\|C_r(A)\| \leq m$, we obtain

$$\begin{aligned} \|A\| &\leq \|C_r(A)\| + \text{forb}(m-1, V) \\ &\leq m + \binom{m-1}{2} + \binom{m-1}{1} + \binom{m-1}{0} + 3 \\ &= \binom{m}{2} + \binom{m}{1} + \binom{m}{0} + 3 \end{aligned}$$

Thus, we may assume that for every r we have $\|C_r(A)\| \geq m+1$. Using Lemma 4.1.7 (with m replaced by $m-1$), we may assume $\|C_r(A)\| = m+1$ for each r . Then using Lemma 4.1.8 we can assume $C_r(A)$ has the structure of (4.1.2) for every r so that every $C_r(A)$ partitions $[m] \setminus r$ rows into sets U_r, L_r with $|U_r|, |L_r| \geq 2$. Note also that the only difference between the two possible structures is a column of 0's or a column of 1's neither of which is used in the case analysis below. Furthermore, we will prove that there is a partition of the rows $[m]$ of A into U, L where $U_r = U \setminus r$ and $L_r = L \setminus r$.

Take two rows, say s and t . Consider $C_s(A)$ and $C_t(A)$ as determined by (2.1.1) by setting $r = s$ and $r = t$. Applying Lemma 4.1.8 when considering $C_s(A)$ and $C_t(A)$ we obtain the partitions U_s, L_s, U_t, L_t (Upper and Lower) of rows. Remember that $C_s(A)$ and $C_t(A)$ both appear twice in A with 0's and 1's in rows s and t respectively. We now define partitions $U'_s = U_s \setminus t, L'_s = L_s \setminus t, U'_t = U_t \setminus s, L'_t = L_t \setminus s$ so that $U'_s \cup L'_s = [m] \setminus \{s, t\} = U'_t \cup L'_t$. We assumed $m \geq 7$ and so $|[m] \setminus \{s, t\}| \geq 5$. Hence we may assume that at least one of U'_s and L'_s has size at least 3. Without loss of generality, assume $|U'_s| \geq 3$.

Let

$$X = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 0 & 0 \end{bmatrix}, \quad Y = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 1 \end{bmatrix}.$$

Consider the following three cases:

$|U'_s \cap L'_t| \geq 3$: We can find V in rows $U'_s \cap L'_t$ in A (since A contains two copies of K_3^1 in each triple of rows of U'_s and two copies of K_3^2 in each triple of rows of L'_t).

$|U'_s \cap L'_t| = 2$: Then $|U'_s \cap U'_t| \geq 1$, and so we can find V in $A|_{U'_s}$ by taking two rows of $U'_s \cap L'_t$ together with any row in the intersection $U'_s \cap U'_t$. We find Y as a submatrix in any row order (A contains two copies of K_3^1 in each triple of rows of U'_s) and we also have X as a submatrix whose first two rows are from $U'_s \cap L'_t$ and the last one from $U'_s \cap U'_t$.

This yields V .

$|U'_s \cap L'_t| = 1$: We have $|U'_s \cap U'_t| \geq 2$, and so we can find V in $A|_{U'_s}$ by taking the row of $U'_s \cap L'_t$ together with two rows in the intersection $U'_s \cap U'_t$. We find Y as a submatrix in any row order (A contains two copies of K_3^1 in each triple of rows of U'_s) and we also have X as a submatrix whose first row is $U'_s \cap L'_t$ and last two rows are from $U'_s \cap U'_t$. This yields V .

This means that $U'_s \subseteq U'_t$, but then $|U'_t| \geq 3$ and so analogously $U'_t \subseteq U'_s$. So $U'_s = U'_t$, and then $L'_s = L'_t$. The same conclusion will hold if $|L'_s| \geq 3$. Thus for all $s, t \in [m]$, $U'_s = U'_t$, and then $L'_s = L'_t$.

Using $m \geq 4$, we now may deduce that there is a partition U, L of $[m]$ so that for any r , $U_r = U \setminus r$ and $L_r = L \setminus r$. This proves that the partition for each C_r is really a global partition. Let $|U| = u$ and $|L| = \ell$. We may argue $u, \ell \geq 2$ since for example if $|U'_r| \geq 1$ and $U'_r = \{s\}$ then $U'_s \cup \{s\} \subseteq U$ and we have $|U| \geq 2$.

Note that for every row r , we have that $[0 \ 1] \times C_r(A) \prec A$. We deduce A contains the following columns:

$$B = \begin{array}{c} U \\ L \end{array} \left[\begin{array}{c|c|c|c|c|c|c|c} \mathbf{0}_u & \mathbf{0}_u & \mathbf{0}_u & \mathbf{0}_u & K_u^1 & K_u^1 & K_u^2 & \mathbf{1}_u \\ \times & \times \\ \mathbf{0}_\ell & \mathbf{1}_\ell & K_\ell^{\ell-1} & K_\ell^{\ell-2} & \mathbf{1}_\ell & K_\ell^{\ell-1} & \mathbf{1}_\ell & \mathbf{1}_\ell \end{array} \right]. \quad (4.1.3)$$

We have included the column of 0's and the column of 1's since such columns can be added to any matrix without creating V . What other columns might we add to this? For $u \geq 3$, matrix B contains

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

in any triple of rows or U in any row order. So $(A - B)|_U$ must not contain the configuration $(1, 1, 0)^T$, else A has subconfiguration V . Similarly for $\ell \geq 4$, $(A - B)|_L$ must not contain the configuration $(1, 0, 0)^T$.

Thus, for $u, \ell \geq 3$, all columns of $(A - B)$ are in $[\mathbf{0}_u \mid K_u^1 \mid \mathbf{1}_u] \times [\mathbf{0}_\ell \mid K_\ell^{\ell-1} \mid \mathbf{1}_\ell]$. The only column not already in B is $\mathbf{1}_u \times \mathbf{0}_\ell$ which is a column of the hypothesized structure (4.1.1). Thus, without loss of generality, we need only consider the case $u = 2$, $\ell \geq 5$. Let $U = \{a, b\}$ and consider any two rows $c, d \in L$. We know B contains $K_u^1 \times [\mathbf{1}_\ell \mid K_\ell^{\ell-1}]$ and

$[\mathbf{0}_u \mid K_u^1] \times K_\ell^{\ell-1}$. So B has:

$$\begin{array}{c|cccc} a & 0 & 0 & 1 & 0 & 0 & 1 \\ b & 1 & 1 & 0 & & & \\ \hline c & 1 & 1 & 0 & 0 & 0 & 1 \\ d & & & & 1 & 1 & 0 \end{array} .$$

Note we can interchange a with b and c with d . To avoid V we must not have columns in $(A - B)$ with

$$\begin{array}{c|cc} a & 1 & 1 \\ b & 0 & \\ \hline c & 0 & 1 \\ d & & 0 \end{array} .$$

Thus, the only possible columns of $A - B$ are

$$\begin{array}{c|ccccc} a & 0 & 1 & 0 & 1 & 1 \\ b & 0 & 0 & 1 & 1 & 1 \\ \hline L & \alpha & \mathbf{1}_\ell & \mathbf{1}_\ell & \mathbf{0}_\ell & \mathbf{1}_\ell \end{array} .$$

where α is any column. Recall that since $\ell \geq 3$, any such α must avoid configuration $(1, 0, 0)^T$. All these columns are already in B , except for $\mathbf{1}_2 \times \mathbf{0}_\ell$, which together with B completes the hypothesized structure (4.1.1). The desired bound follows. \blacksquare

Interestingly, the structure of (4.1.1) falls short of the bound in the case $u = 2$.

4.2 Exact Bound for Ten Products

This section is dedicated to proving an exact bound. The following four matrices are all 2×2 simple matrices (up to row and column permutations). Let

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

We note

$$\text{forb}(m, \{I_2, T_2, U_2, V_2\}) = 1.$$

Define $\{I_2, T_2, U_2, V_2\} \times \{I_2, T_2, U_2, V_2\} = \{X \times Y : X, Y \in \{I_2, T_2, U_2, V_2\}\}$ as the 10

possible products of these matrices (note some of the 16 are equivalent).

Theorem 4.2.1 We have $\text{forb}(m, \{I_2, T_2, U_2, V_2\} \times \{I_2, T_2, U_2, V_2\}) = m + 3$.

Proof: Use the notation $\mathcal{F} = \{I_2, T_2, U_2, V_2\} \times \{I_2, T_2, U_2, V_2\}$. We establish the lower bound by construction. Let $\alpha = \mathbf{1}_{m-1} \times \mathbf{0}_1$. The $m \times (m+3)$ matrix A consisting of $[\mathbf{0}_m \mid I \mid \alpha \mid \mathbf{1}_m]$ avoids all configurations in \mathcal{F} , hence $\text{forb}(m, \mathcal{F}) \geq m + 3$.

We use induction on m for the upper bound. We verified $\text{forb}(4, \mathcal{F}) = 7$ using the computer program described in Section 3.5. To prove the bound for $m \geq 5$, we will proceed by induction on m . For an m -rowed matrix A that doesn't contain any configuration in \mathcal{F} it suffices by induction to show there exists a row r for which $\|C_r\| \leq 1$, using the standard decomposition as in (2.1.1). If this were so, we could delete row r and perhaps one column (one instance of the column forming C_r) from A , keeping the remaining matrix simple. This would yield $\text{forb}(m, \mathcal{F}) \leq 1 + \text{forb}(m-1, \mathcal{F}) = 1 + (m-1) + 3 = m + 3$ as desired.

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

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

But α and β must differ in some row. Without loss of generality, assume they differ on row 2, and suppose $\alpha_2 = 0$ and $\beta_2 = 1$. We will prove that α and β must be complements of each other. Suppose otherwise and suppose they had something in common, say in row 3. The first four rows of A would look like this:

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

for some values of a, b, c (we are using the fact that the matrix has at least 4 rows). Then in

rows 1 and 3 and rows 2 and 4 we get that this matrix contains

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

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

Now $C_2(A)$ must have two repeated columns, say γ and δ . As argued above, $\delta = \bar{\gamma}$. Here is part of the matrix A :

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

We have abused notation and refer to α and γ on $m - 2$ rows as α and γ again. This abuse will continue throughout this proof.

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

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

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

$$(d = c) : \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ a & a & a & \bar{a} & \bar{a} & a & \bar{a} & \bar{a} \\ b & \bar{b} & b & b & \bar{b} & \bar{b} & \bar{b} & b \\ d & d & d & \bar{d} & \bar{d} & d & \bar{d} & \bar{d} \end{bmatrix} \quad \text{or} \quad (d = \bar{c}) : \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ a & a & a & \bar{a} & \bar{a} & a & \bar{a} & \bar{a} \\ b & \bar{b} & b & b & \bar{b} & \bar{b} & \bar{b} & b \\ d & \bar{d} & d & d & \bar{d} & \bar{d} & \bar{d} & d \end{bmatrix} .$$

$$\alpha \quad \gamma \quad \alpha \quad \bar{\gamma} \quad \bar{\alpha} \quad \gamma \quad \bar{\alpha} \quad \bar{\gamma} \qquad \alpha \quad \gamma \quad \alpha \quad \bar{\gamma} \quad \bar{\alpha} \quad \gamma \quad \bar{\alpha} \quad \bar{\gamma}$$

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

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

This is a contradiction to $\|C_r\| \geq 2$ and hence for $m \geq 5$, there must be some row r for which $\|C_r\| \leq 1$. The bound is achieved by induction. \blacksquare

4.3 Critical Substructures

In view of Remark 1.3.21, we know that $G \prec F$ implies $\text{forb}(m, G) \leq \text{forb}(m, F)$. So given F , it is a natural question to try to find minimal configurations $G \prec F$ such that $\text{forb}(m, F) = \text{forb}(m, G)$. In this section we give a conjecture for all such subconfigurations of K_k and prove our conjecture for $k = 4$. Remember that by Theorem 1.3.29 we have that $\text{forb}(m, K_k) = \binom{m}{k-1} + \binom{m}{k-2} + \dots + \binom{m}{0}$.

Definition 4.3.1 Let G be a configuration. We say $G \prec F$ is a **critical substructure** if $\text{forb}(m, G) = \text{forb}(m, F)$.

We give a conjecture that has been verified for $k = 2, 3$ and we verify it for $k = 4$ in Proposition 4.3.3, Proposition 4.3.7 and Proposition 4.3.9.

Conjecture 4.3.2 Let k be given. Then the minimal critical substructures of K_k are the $k + 3$ configurations $2 \cdot \mathbf{1}_{k-1}$, $2 \cdot \mathbf{0}_{k-1}$ and K_k^ℓ for $0 \leq \ell \leq k$.

Previous results established this conjecture for $k = 3$. The fact that all these configurations are critical substructures of K_k isn't hard to prove. The difficulty then lies in proving every critical substructure contains one of these as a configuration.

We show that Conjecture 4.3.2 is equivalent to Conjecture 4.3.8 (for any k) and then prove Conjecture 4.3.8 for $k = 4$.

The following proposition was proven by Füredi and Quinn [FQ83], and Gronau [Gro80].

Proposition 4.3.3 [FQ83][Gro80] For all ℓ, k with $0 \leq \ell \leq k$ we have

$$\text{forb}(m, K_k) = \text{forb}(m, 2 \cdot \mathbf{1}_{k-1}) = \text{forb}(m, 2 \cdot \mathbf{0}_{k-1}) = \text{forb}(m, K_k^\ell).$$

Notice in particular that this proposition answers the question posed in Example 1.3.15. We have $I_3 = K_3^1$ and so $\text{forb}(m, I_3) = \text{forb}(m, K_3) = \binom{m}{2} + \binom{m}{1} + \binom{m}{0}$.

We now consider a *shifting* argument first used by Alon for this context in [Alo83]. Let A be a simple $\{0, 1\}$ -matrix. We construct a **shifted matrix** \bar{A} by performing the following operation repeatedly. Consider an entry of the matrix that has a 1 and make it into a 0 if the resulting matrix is still simple. Keep doing this until there is no entry with a 1 which can be made 0 with the matrix remaining simple. If a column α appears in \bar{A} , then every column with 0's in positions where α is 0 (and perhaps others) also appears in \bar{A} . In other words, the associated family for the shifted matrix \bar{A} is a downset (if a set is in the family, every subset of that set is in the family as well).

Lemma 4.3.4 [Alo83] Let A be a simple $\{0, 1\}$ -matrix and let \bar{A} be a shifted version of A . For any set of rows S , we have that the number of different columns of $A|_S$ is at least the number of different columns of $\bar{A}|_S$.

Lemma 4.3.5 Let $A \in \text{ext}(m, K_k)$. Then for any set $S \in \binom{[m]}{k}$, we have that there is a $k \times 1$ column α such that $(K_k - \alpha) \prec A|_S$.

Proof: Let $A \in \text{ext}(m, K_k)$. Then for any set $S \in \binom{[m]}{k}$, $A|_S$ has at most $2^k - 1$ different columns since it does not have K_k . Let \bar{A} be a shifted version of A and so $\bar{A}|_S$ has at most $2^k - 1$ different columns. Now the columns of \bar{A} form a downset and so cannot have a column of sum k or larger. Now $|\bar{A}|$ is $\text{forb}(m, K_k)$ which is equal to the number of columns of sum $k - 1$ or smaller and so \bar{A} consists of all columns of sum at most $k - 1$. Thus $\bar{A}|_S$ has exactly $2^k - 1$ different columns (it must have at least $2^k - 1$ but cannot have 2^k since it would contain K_k). Then $A|_S$ has exactly $2^k - 1$ different columns, establishing our result. ■

We establish the k -rowed critical substructures of K_k using the following Lemma.

Lemma 4.3.6 Let B be an $k \times (k + 1)$ matrix consisting of one column of each column sum i for $0 \leq i \leq k$. Let $F = K_k - B$. Assume $m \geq k$. Then $\text{forb}(m, F) < \text{forb}(m, K_k)$.

Proof: Let $A \in \text{ext}(m, F)$. Assume $\text{forb}(m, F) = \text{forb}(m, K_k)$. Then also $A \in \text{ext}(m, K_k)$. But then by Lemma 4.3.5 we have that there is a $k \times 1$ column α so that A has $K_k - \alpha$. Now F is a configuration in $K_k - \alpha$ (for any choice of α), contradicting our assumption. Thus $\text{forb}(m, F) < \text{forb}(m, K_k)$. ■

Proposition 4.3.7 Let F be a minimal k -rowed critical substructure of K_k . Then $F = K_k^\ell$ for some ℓ .

Proof: Consider any k -rowed critical substructure F of K_k . Then if there exists an ℓ such that $K_k^\ell \prec F$, since we know $\text{forb}(m, K_k^\ell) = \text{forb}(m, K_k)$, then F cannot be minimal unless $F = K_k^\ell$. Consider the case where $K_k^\ell \not\prec F$ for all ℓ . In this case F is contained in $K_k - B$ for

some B a collection of columns with one column of each column sum. Using Lemma 4.3.6, we can conclude

$$\text{forb}(m, F) \leq \text{forb}(m, K_k - B) < \text{forb}(m, F).$$

Thus, F is not a critical substructure. \blacksquare

For $(k-1)$ -rowed critical substructures we have more work to do. To prove this conjecture, it would suffice to show that the only $(k-1)$ -rowed minimal critical substructures are $2 \cdot \mathbf{0}_{k-1}$ and $2 \cdot \mathbf{1}_{k-1}$. Given the previous Lemmas, we conclude Conjecture 4.3.2 is equivalent to the following:

Conjecture 4.3.8 Let

$$F_{k-1} = [\mathbf{0}_{k-1} \mid 2 \cdot K_{k-1}^1 \mid 2 \cdot K_{k-1}^2 \mid \cdots \mid 2 \cdot K_{k-1}^{k-2} \mid \mathbf{1}_{k-1}].$$

Then there exists M so that for $m \geq M$, $\text{forb}(m, F_{k-1}) < \text{forb}(m, K_k)$.

Let $A \in \text{ext}(m, K_k)$. A consequence of Lemma 4.3.5 is that for every $S \in \binom{[m]}{k}$ there is some choice of α for which $K_k - \alpha \prec A|_S$. Assume that α has p 1's and $(k-p)$ 0's. Then for some $T \subseteq S$ with $|T| = k-1$, we have that for some choice of β , $K_{k-1} - \beta \prec A|_T$, where β has either p 1's and $k-p-1$ 0's or has $p-1$ 1's and $k-p$ 0's. Let C be a $(k-1)$ -rowed matrix with at most one column of each column sum and having column of 0's, column of 1's and for each $1 \leq i \leq k-3$, we have either a column of sum i or a column of sum $i+1$. Then $A|_T$ has $2 \cdot K_{k-1} - C$.

This handles $k=3$ for which we can take C to consist of a column of 0's and a column of 1's. We can handle the case $k=4$ as well.

Proposition 4.3.9 Let $m \geq 6$. Let $F_3 = [\mathbf{0}_3 \mid 2 \cdot K_3^1 \mid 2 \cdot K_3^2 \mid \mathbf{1}_3]$ as in Conjecture 4.3.8. Then $\text{forb}(m, F_3) < \text{forb}(m, K_4)$.

Proof: Let $A \in \text{ext}(m, F_3)$ and assume $\text{ext}(m, F_3) \subseteq \text{ext}(m, K_4)$. Consider any column α of $[K_m^0 \mid K_m^1 \mid K_m^{m-1} \mid K_m^m]$. By extremality of A , $[A \mid \alpha]$ has K_4 as a subconfiguration, say on rows S . Then for any row $j \in S$ with $R_j := S \setminus \{j\}$, we have that $A|_{R_j}$ contains

$2 \cdot K_3 - \alpha|_{R_j}$. Selecting $j \in S$ so that $\alpha|_{R_j}$ is either $\mathbf{1}_3$ or $\mathbf{0}_3$ we have then that A contains F_3 , a contradiction. So we deduce that $[K_m^0 \mid K_m^1 \mid K_m^{m-1} \mid K_m^m] \prec A$ already and hence A contains K_3 in every set of 3 rows.

Let B consist of the remaining columns of A (not in $[K_m^0 \mid K_m^1 \mid K_m^{m-1} \mid K_m^m]$). We deduce that B has no $[K_3^1 \mid K_3^2]$. Then

$$\|B\| \leq \text{forb}(m, [K_3^1 \mid K_3^2]) = \binom{m}{2} + \binom{m}{1} + \binom{m}{0}.$$

Thus $\|A\| \leq 2m + 2 + \binom{m}{2} + \binom{m}{1} + \binom{m}{0}$. For $m \geq 6$, we have $2m + 2 < \binom{m}{3}$, a contradiction.

■

In Section 8.1 we give some further ideas on how to deal with $k \geq 5$.

Chapter 5

Three Asymptotic Bounds

5.1 Introduction

When we are not able to obtain an exact bound on $\text{forb}(m, F)$ for some F , we might settle for asymptotic bounds. The research in this topic is often guided by Conjecture 1.4.1. In this chapter we prove three main results, each of which completes some part which would be required in order to establish Conjecture 1.4.1.

In Section 5.2 we prove the quadratic asymptotic bound predicted by Conjecture 1.4.1 for the configuration $F_8(t)$, one of the three maximal 4-rowed configurations for which the conjecture predicts a quadratic bound. Here are these three configurations (called $F_6(t)$, $F_7(t)$, $F_8(t)$ for historical reasons).

$$F_6(t) = \left[\begin{array}{cccc} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{array} t \cdot \left[\begin{array}{cccc} 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{array} \right] \right]$$

$$F_7(t) = \left[\begin{array}{cccc} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array} t \cdot \left[\begin{array}{cccc} 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \end{array} \right] \right]$$

$$F_8(t) := \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix} t \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \\ 0 & 0 \end{bmatrix}.$$

As of this writing, quadratic bounds for $F_6(t)$ and $F_7(t)$ have not been found. Note that both $F_6(t)$ and $F_7(t)$ contain the following:

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

It is not known (but The Conjecture predicts it) whether or not the above matrix has a quadratic bound for $t \geq 3$, but for $t = 2$ and $t = 1$, the quadratic bound was proven by Anstee in [Ans90]. The methods described in the following chapters have so far failed to produce results for these other configurations. There are some cases in their respective lists of What Is Missing which we don't know how to deal with.

Then in Section 5.3 we prove the bound for F_7 , one of the nine maximal 5-rowed configurations below which are predicted to be quadratic by Conjecture 1.4.1.

$$F_3 = \begin{bmatrix} 1 & 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad F_4 = \begin{bmatrix} 1 & 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad F_5 = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 \end{bmatrix}$$

$$F_6 = \begin{bmatrix} 1 & 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad F_7 = \begin{bmatrix} 1 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \quad F_8 = \begin{bmatrix} 1 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

$$F_9 = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 \end{bmatrix} \quad F_{10} = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad F_{11} = \begin{bmatrix} 1 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}$$

It is not known whether or not any of the other eight boundary configurations will satisfy the conjecture.

Finally, in Section 5.4 we use the result about F_7 to classify all 6-rowed quadratic configurations by proving the unique boundary case:

$$G_{6 \times 3} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

5.2 Quadratic Bound for a 4-rowed Configuration

Let $F_8(t)$ be the following $4 \times (4 + 2t)$ configuration (as in the Introduction)

$$F_8(t) := \left[\begin{array}{cccc} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{array} \quad t \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \\ 0 & 0 \end{bmatrix} \right].$$

Theorem 5.2.1 Let $t \geq 1$ be given. Then $\text{forb}(m, F_8(t))$ is $\Theta(m^2)$. More precisely, $\text{forb}(m, F_8(t)) \leq 12tm^2$. Furthermore, $F_8(t)$ is a boundary quadratic case.

We have several ingredients to the proof. The first is our Standard Induction given in Section 2.1: We find the inductive children $\mathcal{H}_{F_8(t)}$. By Standard Induction, it suffices then to prove that $\text{forb}(m, \mathcal{H}_{F_8(t)})$ is $\Theta(m)$. We then find What Is Missing as in Section 2.2 in a

triple of rows if $\mathcal{H}_{F_8(t)}$ is avoided. Lastly, we use Implications, as described in Section 2.3, to delete a linear number of columns from a matrix $A \in \text{Avoid}(m, \mathcal{H}_{F_8(t)})$ so that $C_r(A)$ becomes constant.

First, let us calculate $\mathcal{H}_{F_8(t')}$ for some number t' . Let $A \in \text{Avoid}(m, F_8(t'))$ and consider $C_r(A)$. The fact that A contains no configuration $F_8(t')$ means $C_r(A)$ doesn't have any of the following three configurations, for $t = \lfloor \frac{t'+1}{2} \rfloor + 1$:

$$H_1(t) = \begin{bmatrix} 1 & 0 & \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ 0 & 1 & t \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ 0 & 0 & \begin{pmatrix} 1 & 1 \end{pmatrix} \end{bmatrix}, \quad H_2(t) = \begin{bmatrix} 1 & 0 & \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ 0 & 1 & t \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ 1 & 1 & \begin{pmatrix} 0 & 0 \end{pmatrix} \end{bmatrix},$$

$$H_3(t) = \begin{bmatrix} 0 & 1 & 0 & 1 & \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \\ 1 & 1 & 0 & 0 & t \cdot \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \\ 1 & 1 & 0 & 0 & \begin{pmatrix} 0 & 0 \end{pmatrix} \end{bmatrix}.$$

So we may conclude that $\mathcal{H}_{F_8(t)} = \{H_1(t), H_2(t), H_3(t)\}$. We now focus our attention to matrices $A \in \text{Avoid}(m, \{H_1(t), H_2(t), H_3(t)\})$. We make the following bold claim:

Lemma 5.2.2 Let t be given. Then $\text{forb}(m, \{H_1(t), H_2(t), H_3(t)\}) \leq 12tm$.

Note that Theorem 5.2.1 is not equivalent to the previous lemma is false, and this is indeed the case for $F_6(t)$ and $F_7(t)$, for which we know the bound for the inductive children is quadratic and not linear. The proof of Lemma 5.2.2 appears at the end of this subsection. But by using this lemma we are ready to prove Theorem 5.2.1.

Proof of Theorem 5.2.1 : Let $A \in \text{Avoid}(m, F_8(t))$. We simply use induction as in Proposition 2.1.2 (replacing t by $\frac{t+1}{2} + 1$) using Lemma 5.2.2 to deduce $\|A\|$ is linear and hence $\text{forb}(m, F_8(t))$ is $O(m^2)$. The fact that $F_8(t)$ is a boundary case follows from the constructions in Conjecture 1.4.1. If $\alpha \in F_8(0)$, then α has column sum 1 or 3, but then $[F_8(t) | \alpha] \not\prec I \times I \times I$. Consider then the possibilities for α not in $F_8(1)$. If α consists of all 1's or three 1's or two 1's except on the first two rows then each pair of rows of $[F_8(1) | \alpha]$ has a column of two 1's and so $[F_8(1) | \alpha]$ is not in $I_{m/3} \times I_{m/3} \times I_{m/3}$. This also handles the complementary case, i.e. where α is all 0's or one 1 or two 1's on the last two rows using $I_{m/3}^c \times I_{m/3}^c \times I_{m/3}^c$. There are two remaining cases: α having two 1's in the first and fourth rows or in the second and fourth rows. Then $[F_8(1) | \alpha]$ has the 2×2 matrix I_2 on each pair

of rows and so $[F_8(1) | \alpha]$ is not in $T_{m/3} \times T_{m/3} \times T_{m/3}$. Thus $\text{forb}(m, [F_8(1) | \alpha])$ is $\Omega(m^3)$.

■

To prove Lemma 5.2.2, we will need some additional Lemmas and properties.

5.2.1 What Is Missing?

The following are all the possibilities of columns that are either missing or in short supply on three rows if we forbid $H_1(t), H_2(t), H_3(t)$. This was computed using the C++ program referred to in Chapter 3. Checking that if a triple of rows satisfies each P_i , then configurations $H_1(t), H_2(t), H_3(t)$ are avoided is quite easy, but the computer is used to avoid the enormous amount of work that would be required to establish that the list is complete.

Lemma 5.2.3 Let $A \in \text{Avoid}(m, \{H_1(t), H_2(t), H_3(t)\})$. For each triple of rows (x, y, z) satisfies (at least) one of the 21 cases P_1, P_2, \dots, P_{21} described below for some ordering of (x, y, z) .

The cases are:

$$P_1 = \begin{array}{cc} \text{no} & \text{no} \\ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \end{array}, \quad P_2 = \begin{array}{cccc} < t & < t & \text{no} & \text{no} \\ \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \end{array}$$

$$P_3 = \begin{array}{cccc} < t & < t & \text{no} & \text{no} \\ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \end{array}, \quad P_4 = \begin{array}{cccc} < t & < t & < t & \text{no} \\ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \end{array}$$

$$P_5 = \begin{array}{cccc} < t & < t & < t & \text{no} \\ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \end{array}, \quad P_6 = \begin{array}{ccc} \text{no} & < t & \text{no} \\ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \end{array}$$

$$\begin{aligned}
P_{19} &= \begin{array}{cccc} \text{no} & < t & \text{no} & < t \\ \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \end{array}, & P_{20} &= \begin{array}{cccc} \text{no} & < t & \text{no} & < t \\ \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \end{array} \\
P_{21} &= \begin{array}{cc} \text{no} & \text{no} \\ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \end{array}.
\end{aligned}$$

Proof: By exhaustive computer search as described in Section 3.5. ■

5.2.2 Case Analysis

We now proceed to do some case analysis of the 21 cases mentioned in Lemma 5.2.3. The first trick to discard some cases involves attempting to do induction one more time. For some of the 21 cases, if they happened to occur in A we may apply the Standard Induction again.

Definition 5.2.4 We say that a row r of A is **non-essential** if $\|C_r(A)\| \leq 4t$.

The motivation for this definition is that if we could find a non-essential row of A , we could use Standard Induction as in Section 2.1 for the case where we are forbidding $\{H_1(t), H_2(t), H_3(t)\}$ to prove Lemma 5.2.2 (in fact $\|C_r(A)\| \leq 12t$ would suffice).

Lemma 5.2.5 If $A \in \text{Avoid}(m, \{H_1(t), H_2(t), H_3(t)\})$ has a triple of rows (i, j, k) that satisfies one of $P_2, P_4, P_8, P_9, P_{10}, P_{12}, P_{17}, P_{18}$ and P_{19} , then there is a non-essential row r of A .

Proof: If we analyze the columns that could be in long supply for each of the cases, we see in each case that one of the rows of A isn't necessary to distinguish between columns in long supply.

Perhaps an example would be useful. The other cases are similar. Suppose A is missing P_2 in rows i, j, k , in that order. So A satisfies the following from rows i, j, k :

$$P_2 = \begin{array}{c} i \\ j \\ k \end{array} \begin{array}{cccc} < t & < t & \text{no} & \text{no} \\ \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \end{array}$$

Let us analyze the columns that could be in long supply (denote l.s.). These columns are:

$$\begin{array}{c} i \\ j \\ k \end{array} \begin{array}{cccc} \text{l.s} & \text{l.s} & \text{l.s} & \text{l.s} \\ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \end{array} .$$

We can see that $\|C_k(A)\| \leq 2t$, and therefore row k is non-essential. The same happens for the other cases: in each of them, there is a row r , like row k , for which $\|C_r(A)\| \leq 4t$.

■

If we assume there are no non-essential rows then we may restrict our attention to matrices which for all triples of rows satisfy one of the cases $P_1, P_3, P_5, P_6, P_7, P_{11}, P_{13}, P_{14}, P_{15}, P_{16}, P_{20}$ or P_{21} .

We will now use the technique of Implications described in Section 2.3 to delete a linear number of columns from A (without deleting any row) in order to obtain a matrix A' for which $\|C_r(A')\|$ is bounded by a constant (for any row r).

First we give a lemma which is purely a property of directed graphs.

Lemma 5.2.6 Let G be a directed graph on m vertices. Then we can colour the edges of G using three colours (blue, red and green) in such a way that for vertices r, a, b we have that G satisfies the following properties:

- (R) There are at most $2m$ red edges.
- (B) If $r \rightarrow a$ and $r \rightarrow b$ are blue, then neither $a \rightarrow b$ nor $b \rightarrow a$ (of any colour).
- (G) If $a \rightarrow b$ is green, there is a blue-red path from a to b .

Proof: The idea for this colouring came from an idea first introduced by Anstee and Sali in [AS05], although the actual colouring is different. We provide an algorithmic proof.

1. Divide G into strongly connected components X_1, X_2, \dots, X_k ordered in a way consistent with the order given by the acyclic ordering (so that if $i < j$, there might be a path between a vertex of X_i and a vertex of X_j , but there is no path back).
2. Pick a strongly connected component X_i . It is a well known property of directed graphs that there is a strongly connected subgraph Y_i of X_i that uses all the vertices of X_i and the number of edges is at most $2|X_i|$. For every edge of X_i , see whether it is in Y_i or not. If it is, colour it with red, and if it isn't, colour it with green.
3. Colour every remaining edge with blue.

Notice that currently the only property that may not be satisfied is **(B)**. We will change some of the blue ones to green (leaving the red ones intact) until we get the desired property, but never breaking **(R)** nor **(G)**.

Notice also that red edges always stay in the same strongly connected component, while blue edges always go to a higher level. To make this statement precise, define a level function $\lambda : V(G) \rightarrow \mathbb{N}$ as

$$\lambda(v) := i \quad \text{if} \quad v \in X_i.$$

We have the property that if $v \rightarrow u$ is a red edge, then $\lambda(v) = \lambda(u)$, and if $v \rightarrow u$ is a blue edge, then $\lambda(v) < \lambda(u)$. This property will be preserved during all steps of the colouring algorithm. In particular, it is true when applying steps 1 through 3.

4. Go through all the strongly connected components and for each component go through each vertex. Suppose we are at vertex v . Look at the set of blue edges coming out of v . Say their endpoints are v_1, v_2, \dots, v_d .
5. If there is an edge $v_i \rightarrow v_j$ (of any colour) and $v \rightarrow v_i$ and $v \rightarrow v_j$ are (still) blue, paint $v \rightarrow v_j$ with green.

It is easy to check that step 5 preserves property **(G)**. We only need to prove that there is a blue-red path from v to v_j . We know there is a blue-red path P_{v_i, v_j} from v_i to v_j by **(G)**, and since $v \rightarrow v_i$ is blue, we can consider the path $v \rightarrow v_i$ plus P_{v_i, v_j} . This can be done as long as P_{v_i, v_j} doesn't contain $v \rightarrow v_j$, and indeed it can't, because red edges always stay in the same connected component, while blue edges always go to a higher level, and since

$v \rightarrow v_i$ is blue and the path P_{v_i, v_j} is blue-red, $\lambda(v) < \lambda(v_i) \leq \lambda(v_j)$. This implies that P_{v_i, v_j} cannot contain vertex v . ■

Recall our definition of implications of Section 2.3.

Lemma 5.2.7 Let $A \in \text{Avoid}(m, \{H_1(t), H_2(t), H_3(t)\})$ have no non-essential rows. Form the directed graph G on m vertices whose edges are the implications in A and colour G using Lemma 5.2.6. Then we can delete at most $4tm$ columns from A so that all red implications are pure.

Proof: There are at most $2m$ red edges and there are at most $2t$ columns that violate any given implication, hence there are at most $4tm$ columns that violate red implications. We can delete at most $4tm$ columns to make the red implications pure. ■

If a column violates a green implication, it must also violate an implication in a blue-red path, so it must violate either a blue or a red implication. So if we manage to purify the blue implications, no column could violate a green implication either.

We will devote some time to proving there are at most $4tm$ columns that violate at least one blue implication.

Lemma 5.2.8 Let $A \in \text{Avoid}(m, \{H_1(t), H_2(t), H_3(t)\})$ with no non-essential rows (i.e. for every row r , we have $\|C_r(A)\| > 4t$). Colour the associated implications graph as in Lemma 5.2.6. In any triple of rows r, r_i, r_j where $r \rightarrow r_i$ and $r \rightarrow r_j$ are blue and impure, then either P_{16} or P_{21} is satisfied.

Proof: By Lemma 5.2.5, we know that in any triple of rows one of $P_1, P_3, P_5, P_6, P_7, P_{11}, P_{13}, P_{14}, P_{15}, P_{16}, P_{20}, P_{21}$ has to be satisfied on the triple r, r_i, r_j .

second category, \mathcal{C}_1 consists of all columns with all the entries in rows r_1, r_2, \dots, r_ℓ being 1. And the last, \mathcal{C} , consists of columns that start with some number of zeros and end with ones, like this:

$$\begin{array}{c} r \\ r_1 \\ \vdots \\ r_i \\ r_{i+1} \\ \vdots \\ r_\ell \end{array} \begin{array}{c} \left[0 \right] \\ \left[0 \right] \\ \left[\vdots \right] \\ \left[0 \right] \\ \left[1 \right] \\ \left[\vdots \right] \\ \left[1 \right] \end{array} .$$

We deal with these three categories separately.

Columns in \mathcal{C}_0 : These columns already don't violate any implication $r \rightarrow r_i$.

Columns in \mathcal{C} : Rows r_1 and r_ℓ , in addition to satisfying $\text{row}(r_1) \subseteq \text{row}(r_\ell)$, must also satisfy

$$\begin{array}{c} r \\ r_1 \\ r_\ell \end{array} \begin{array}{c} < t \\ \left[0 \right] \\ \left[0 \right] \\ \left[1 \right] \end{array} ,$$

which means the number of columns with 0 in row r , 0 in row r_1 and 1 in row r_ℓ is at most $t - 1$. This means $|\mathcal{C}| < t$.

Columns in \mathcal{C}_1 : We may use the fact that each triple r, r_i, r_j satisfies

$$\begin{array}{c} r \\ r_i \\ r_j \end{array} \begin{array}{c} < 2t \\ \left[0 \right] \\ \left[1 \right] \\ \left[1 \right] \end{array} ,$$

so $\mathcal{C}_1 < 2t$ as well.

In conclusion, for each row r , we can delete $3t$ columns, and then every blue implication $r \rightarrow r_i$ is pure. We repeat this for every row r , and by deleting at most $3tm$ columns of A , every blue implication is pure. ■

After these column deletions every blue implication is pure and we may now assume every implication is pure: No column violates either blue or red implications, which means no column violates any green implications. We've managed to delete only a linear number of columns of A without deleting any row, and now no implication gets violated.

5.2.3 Linear Bound for the Inductive Children

We are now ready to prove Lemma 5.2.2.

Proof of Lemma 5.2.2 : We will show that $\text{forb}(m, \{H_1(t), H_2(t), H_3(t)\}) \leq 12tm$. Let $A \in \text{Avoid}(m, \{H_1(t), H_2(t), H_3(t)\})$. By Proposition 2.1.2 as described above, we may assume A has no non-essential rows. For every triple of rows of A , we choose P_i if P_i is satisfied in that triple (other P_j might be satisfied as well, but pick one for every triple). This yields a map $p : S_3(A) \rightarrow \{P_i : i \in \{1, 3, 5, 6, 7, 11, 13, 14, 15, 16, 20, 21\}\}$ (by Lemma 5.2.5) where $S_3(A)$ is the set of triples of rows of A .

Appealing to Lemma 5.2.7 and Lemma 5.2.9 we can delete at most $8tm$ columns and conclude that all implications associated to one of the P_i in the image of $S_3(A)$ are pure. We now do induction again with a new hypothesis.

We wish to show that for $A \in \text{Avoid}(m, \{H_1(t), H_2(t), H_3(t)\})$ satisfying that each triple of rows has a chosen satisfied condition P_i (for some $i \in \{1, 3, 5, 6, 7, 11, 13, 14, 15, 16, 20, 21\}$) and any implications arising from these chosen conditions is pure, then $\|A\| \leq 4tm$. We note that any submatrix of A satisfies the same hypotheses. Thus it suffices to appeal to our induction argument Proposition 2.1.2 and show that A has a non-essential row. We will in fact show that row 1 is non-essential.

Let $A \in \text{Avoid}(m, \{H_1(t), H_2(t), H_3(t)\})$ satisfying that each triple of rows has a chosen satisfied condition P_i (for some $i \in \{1, 3, 5, 6, 7, 11, 13, 14, 15, 16, 20, 21\}$) and any implications arising from these chosen conditions is pure. Consider a triple of rows $1, r, s$ from A . In this triple, one of the 12 cases will have been chosen to be satisfied, so for the pair r, s in $C_1(A)$, two rows corresponding to two rows of one of the 12 P_i 's must be satisfied, since $C_1(A)$ consists of the repeated columns.

Consider a $(0,1)$ -row with n columns as the incidence vector of a subset of $[n] = \{1, 2, \dots, n\}$. We now use $\text{row}(r)$ to denote row r in $C_1(A)$, since we've restricted our attention to this matrix. We show by the case analysis below that if two columns are absent on a pair of rows r, s , then either $\text{row}(r) = \emptyset$ or $\text{row}(r) = [n]$, (or $\text{row}(s) = \emptyset$ or $\text{row}(s) = [n]$), or $\text{row}(r) = \text{row}(s)$ or $\text{row}(r) = \text{row}(s)^c$.

If two columns are absent, then only two columns can be present. Thus,

1. If the two columns absent are

$$\begin{array}{cc} & \begin{array}{c} no \\ no \end{array} \\ r & \begin{bmatrix} 1 \\ 0 \end{bmatrix} \end{array} \quad \text{or} \quad \begin{array}{cc} & \begin{array}{c} no \\ no \end{array} \\ r & \begin{bmatrix} 0 \\ 1 \end{bmatrix} \end{array}$$

then $\text{row}(r) = \emptyset$ (row r is 0's) in the former, or $\text{row}(r) = [n]$ (row r is 1's) in the latter.

2. If the columns missing are

$$\begin{array}{cc} & \begin{array}{c} no \\ no \end{array} \\ r & \begin{bmatrix} 0 \\ 1 \end{bmatrix} \end{array}$$

then $\text{row}(r) = \text{row}(s)$.

3. If the columns missing are

$$\begin{array}{cc} & \begin{array}{c} no \\ no \end{array} \\ r & \begin{bmatrix} 0 \\ 0 \end{bmatrix} \end{array}$$

then $\text{row}(r) = \text{row}(s)^c$.

We may check each case $P_1, P_3, P_5, P_6, P_7, P_{11}, P_{13}, P_{14}, P_{15}, P_{16}, P_{20}, P_{21}$ to find that two columns are absent for each pair of rows of $C_1(A)$, except for P_6 and P_{14} when row 1 of A corresponds to row 3 of P_6 or row 2 of P_{14} ,

In the case when $1, r, s$ form a P_6 or a P_{14} , and row 1 of A corresponds to row 3 of P_6 or row 2 of P_{14} , we have (for some order of r and s)

$$\begin{array}{cc} & \begin{array}{c} no \\ < t \end{array} \\ r & \begin{bmatrix} 0 \\ 1 \end{bmatrix} \end{array}$$

which means $\text{row}(s) \subseteq \text{row}(r)$, and the difference $\text{row}(r) \setminus \text{row}(s)$ is at most t .

Construct the following coloured semi-directed graph:

- The vertices are the rows r of $C_1(A)$ with $\text{row}(r) \neq \emptyset$ and $\text{row}(r) \neq [n]$.
- Place a purple edge between two rows r, s if $\text{row}(r) = \text{row}(s)$.
- Place a yellow edge between two rows r, s if $\text{row}(r) = \text{row}(s)^c$.

- Place a directed edge $r \rightarrow s$ if $\text{row}(r) \subseteq \text{row}(s)$.

If some rows are equal, we will treat them as being just one row. So we can take the quotient of the graph over the purple edges and work in the new graph.

If two yellow edges share a vertex, the non-shared vertices must have a purple edge between them, because the complement of the complement is itself. Since we did the quotient over purple edges, we can assume no two yellow edges share a vertex.

So we are left with only directed and yellow edges. We will prove there are no yellow edges. We proceed by contradiction. Suppose we have a yellow edge between rows r_1 and r_2 so that $\text{row}(r_1) = \text{row}(r_2)^c$. If there is no other row then the matrix has at most 2 columns and we are done. Assume r is another row, different from r_1 or r_2 . Consider the edge between r and r_1 and between r and r_2 . Let us analyze the four possibilities. Clearly it can't be yellow or purple.

- If $r \rightarrow r_1$ and $r \rightarrow r_2$, then $\text{row}(r) \subseteq \text{row}(r_1)$ and $\text{row}(r) \subseteq \text{row}(r_2)$ contradicts $\text{row}(r_1) = \text{row}(r_2)^c$ if $\text{row}(r) \neq \emptyset$. So we conclude $\text{row}(r) = \emptyset$, contradicting our construction.
- If $r \rightarrow r_1$ and $r_2 \rightarrow r$, then $\text{row}(r_2) \subseteq \text{row}(r_1)$, a contradiction.
- If $r_1 \rightarrow r$ and $r \rightarrow r_2$, then $\text{row}(r_1) \subseteq \text{row}(r_2)$, a contradiction.
- If $r_1 \rightarrow r$ and $r_2 \rightarrow r$, then we have that $\text{row}(r)$ contains both a set and its complement. This means $\text{row}(r) = [n]$ contradicting our construction.

Every pair of rows r, s has a directed edge and therefore we have a tournament. We note that the graph has no directed cycles (since a directed edge means containment of rows) and hence it is a transitive tournament. This in particular yields a path that goes through all the vertices (a total ordering of the rows).

We have more information: A directed edge only occurs in cases P_6 and P_{14} , when row 1 of A corresponds to row 3 of P_6 or row 2 of P_{14} . In these two cases, when a row r contains row s , we also have that we get that rows r and s differ in at most t columns. And since the first row in the path (the one with the least number of ones) and the last (the one with the most number of ones) have to differ in at most t places, there must be at most $t + 2$ columns in $C_1(A)$, and so row 1 is non-essential, which proves the lemma. ■

5.3 Quadratic Bound for a 5-rowed Configuration

Previous work of Ryan, reported in [Ans], computed nine 5-rowed simple matrices F which by Conjecture 1.4.1 should be boundary cases and for which $\text{forb}(m, F)$ should be $\Theta(m^2)$. One of them, named F_7 in [Ans], is

$$F_7 = \begin{bmatrix} 1 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

Note that $F_7 = F_7^c$.

Theorem 5.3.1 We have that $\text{forb}(m, F_7)$ is $\Theta(m^2)$. Moreover, for any 5×1 $\{0, 1\}$ -column α , $\text{forb}(m, [F_7 \mid \alpha])$ is $\Omega(m^3)$.

The proof uses Standard Induction (Section 2.1) and a linear bound for three smaller matrices in Lemma 5.3.2 which in a novel way uses Standard Induction. We give the proof of Theorem 5.3.1 from Lemma 5.3.2 in Section 2.1 and the proof of Lemma 5.3.2 in Section 5.3.3.

5.3.1 Applying Standard Induction

Let $A \in \text{Avoid}(m, F_7)$ and apply the standard decomposition of (2.1.1) for $r = 1$.

Our goal is to show $\|A\|$ is quadratic by showing that $\|C_1(A)\|$ is linear. We find the inductive children of F_7 (basically, delete each row of F_7 in turn) and note that $C_1(A)$ cannot contain any of the configurations H_1, H_2, H_3, H_4, H_5 :

$$H_1 = \begin{bmatrix} 1 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}, \quad H_2 = \begin{bmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}, \quad H_3 = \begin{bmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

$$H_4 = \begin{bmatrix} 0 & 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad H_5 = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \end{bmatrix}.$$

We observe that $H_3^c = H_3$, $H_4 = H_1^c$, $H_2^c = H_5$. Also $H_3 \prec H_1$ (columns 2,3,5,6) and so $H_3 \prec H_4$ and so we may ignore H_1, H_4 .

Standard Induction (Section 2.1) gives the bound $\text{forb}(m, F_7) < m \cdot \text{forb}(m, \{H_2, H_3, H_5\})$. The next lemma states that $\text{forb}(m, \{H_2, H_3, H_5\})$ has a linear bound, which means $\text{forb}(m, F_7)$ has a quadratic bound.

Lemma 5.3.2 We have that $\text{forb}(m, \{H_2, H_3, H_5\})$ is $O(m)$.

We will prove Lemma 5.3.2 in Section 5.3.3. We can now prove that $\text{forb}(m, F_7)$ is quadratic.

Proof of Theorem 5.3.1: The fact that $\text{forb}(m, F_7)$ is $\Omega(m^2)$ comes directly out of the conjecture, as $F_7 \not\prec I \times I$. We show $\text{forb}(m, F_7)$ is $O(m^2)$ using induction on m . Consider $A \in \text{Avoid}(m, F_7)$ with $\|A\| = \text{forb}(m, F_7)$. Then using Proposition 2.1.2, we have

$$\text{forb}(m, F_7) = \|A\| \leq \sum_{i=1}^{m-1} \text{forb}(i, \{H_2, H_3, H_5\}).$$

Given that there is a constant c so that $\text{forb}(i, \{H_2, H_3, H_5\}) \leq ci$ by Lemma 5.3.2, we deduce the bound $\text{forb}(m, F_7) < c \cdot m(m-1)/2$ which is $\Theta(m^2)$.

Now consider any 5×1 column α . We immediately deduce that $\text{forb}(m, [F_7 | \alpha])$ is $\Omega(m^3)$ for α having zero, one, four or five 1's, or if α is a column in F_7 (considered as a matrix). It is a computational exercise to show that every other α results in $\text{forb}(m, [F_7 | \alpha])$ being $\Omega(m^3)$ and the computer program described in Section 3.4 indeed gave us this result. We give a proof here for completeness.

We need only consider α having two 1's since $F_7^c = F_7$. If α has 0's on rows 2,3 then $[F_7 | \alpha] \not\prec I^c \times I^c \times I^c$ (each pair of rows from the four rows 1,2,3,4 of $[F | \alpha]$ has $(0,0)^T$) or two 0's on rows 1,4 then $[F_7 | \alpha] \not\prec I^c \times I^c \times I^c$ (each pair of rows from the four rows 1,3,4,5 has $(0,0)^T$). This only leaves four columns, three of which are already in F_7 . Only

$\alpha = (0, 0, 1, 1, 0)^T$ is not already in F_7 , and in such case $[F_7 | \alpha] \not\prec T \times T \times T$ since every pair of rows from the four rows 1,2,3,4 has the 2×2 configuration I_2 . ■

5.3.2 What Is Missing?

Applying the technique of What Is Missing described in Section 2.2 to $\mathcal{F} = \{H_2, H_3, H_5\}$, we get the following lemma.

Lemma 5.3.3 Let $A \in \text{Avoid}(m, \{H_2, H_3, H_5\})$. Then there are 13 possibilities Q_0, Q_1, \dots, Q_{12} for What Is Missing on each 4-set of rows.

$$Q_0 = \begin{array}{c} \text{no} \quad \text{no} \quad \text{no} \quad \text{no} \quad \text{no} \quad \text{no} \\ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} \end{array} \quad Q_1 = \begin{array}{c} \text{no} \quad \text{no} \quad \text{no} \quad \text{no} \quad \text{no} \quad \text{no} \\ \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix} \end{array}$$

$$Q_2 = \begin{array}{c} \text{no} \quad \text{no} \quad \text{no} \quad \text{no} \quad \text{no} \quad \text{no} \\ \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix} \end{array} \quad Q_3 = \begin{array}{c} \text{no} \quad \text{no} \quad \text{no} \quad \text{no} \quad \text{no} \quad \text{no} \\ \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \end{array}$$

$$Q_4 = \begin{array}{c} \text{no} \quad \text{no} \quad \text{no} \quad \text{no} \\ \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} \end{array} \quad Q_5 = \begin{array}{c} \text{no} \quad \text{no} \\ \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} \end{array}$$

$$Q_6 = \begin{array}{c} \text{no} \quad \text{no} \quad \text{no} \quad \text{no} \quad \text{no} \quad \text{no} \\ \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix} \end{array}$$

$$Q_7 = \begin{array}{c} \text{no} \quad \text{no} \quad \text{no} \quad \text{no} \quad \text{no} \quad \text{no} \\ \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix} \end{array}$$

$$Q_8 = \begin{array}{c} \text{no} \quad \text{no} \quad \text{no} \quad \text{no} \quad \text{no} \quad \text{no} \\ \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \end{array}$$

$$Q_9 = \begin{array}{c} \text{no} \quad \text{no} \quad \text{no} \quad \text{no} \quad \text{no} \quad \text{no} \\ \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix} \end{array}$$

$$Q_{10} = \begin{array}{c} \text{no} \quad \text{no} \quad \text{no} \quad \text{no} \quad \text{no} \quad \text{no} \\ \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \end{array}$$

$$Q_{11} = \begin{array}{c} \text{no} \quad \text{no} \quad \text{no} \quad \text{no} \quad \text{no} \quad \text{no} \\ \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \end{array}$$

$$Q_{12} = \begin{array}{c} \text{no} \quad \text{no} \\ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} \end{array}$$

Proof of Lemma 5.3.3: An exhaustive computer search by the program described in Section 3.5 yields the result. ■

5.3.3 Linear Bound for the Inductive Children

The rest of the section is a proof of Lemma 5.3.2. Let $A \in \text{Avoid}(m, \{H_2, H_3, H_5\})$. We will use special features of H_2, H_3, H_5 to obtain a linear bound on $\|A\|$. The forbidden configuration H_3 is used most often in this proof. We will show $\|A\| \leq 7m$ by induction on m . We analyze the 13 cases of Lemma 5.3.3 one by one and have special arguments for the three troublesome cases Q_2, Q_3, Q_{11} .

Lemma 5.3.4 Let $A \in \text{Avoid}(m, \{H_2, H_3, H_5\})$. Consider the standard decomposition (2.1.1) of A based on row r . Let $L(r) \neq \emptyset$ be a minimal set of rows such that $C_r|_{L(r)}$ is simple. Then each triple of rows $\{i, j, k\}$ in $L(r)$ yield a quadruple of rows $\{r, i, j, k\}$ on which one of the cases Q_2, Q_3, Q_{11} occurs, with row r being the first row of each of the cases Q_2, Q_3, Q_{11} as given in Lemma 5.3.3.

Proof: For each Q_i we record pairs of rows containing “a copy of K_2 ”: namely in the columns marked absent we find

$$\begin{array}{cccc}
 & \text{no} & \text{no} & \text{no} & \text{no} \\
 r & \begin{bmatrix} a \\ e \\ 0 \\ 0 \end{bmatrix} & , & \begin{bmatrix} b \\ f \\ 1 \\ 0 \end{bmatrix} & , & \begin{bmatrix} c \\ g \\ 0 \\ 1 \end{bmatrix} & , & \begin{bmatrix} d \\ h \\ 1 \\ 1 \end{bmatrix} . \\
 i & & & & & & & \\
 j & & & & & & & \\
 k & & & & & & &
 \end{array}$$

Suppose A had these columns missing on the quadruple of rows r, i, j, k and that rows i, j, k belong to $L(r)$. Then the simple matrix C_r from (2.1.1) has the four 3×1 columns $(e, 0, 0)^T, (f, 1, 0)^T, (g, 0, 1)^T$ and $(h, 1, 1)^T$ missing on the triple of rows $\{i, j, k\}$. We deduce that row i cannot belong to $L(r)$, a contradiction.

By analyzing the cases, we find that $Q_0, Q_1, Q_5, Q_6, Q_7, Q_8, Q_{10}, Q_{12}$ have 3 rows each pair of which have a “ K_2 ” and Q_4, Q_9 have two disjoint pairs of rows each with a “ K_2 ”. Thus in any of these cases, What Is Missing on a triple of rows in C_r will contain a copy of “ K_2 ” and so we can delete a row from C_r without disturbing simplicity of the remainder of C_r . In cases Q_2, Q_3, Q_{11} , if we choose row r to be any row but the first row in each of the cases then there is a “ K_2 ” on the remaining triple. ■

We would like to show that for all $A \in \text{Avoid}(m, \{H_2, H_3, H_5\})$ we can choose row r so

that $\|C_r\| \leq 7$ as in (2.1.1). Then by Proposition 2.1.2 and induction, $\|A\| \leq 7m$. We will assume the contrary, namely that there is $A \in \text{Avoid}(m, \{H_2, H_3, H_5\})$ such that for every row r , $\|C_r\| \geq 8$.

In each of the troublesome cases Q_2, Q_3, Q_{11} , we end up with the following sets of columns missing on a triple of rows in C_r (arising from What Is Missing in A on a quadruple of rows involving r) and we name the cases correspondingly P_2, P_3, P_{11} . Note the implications arising from P_3 .

$$P_2 : \begin{array}{ccc} \text{no} & \text{no} & \text{no} \\ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \end{array} \quad (5.3.1)$$

$$P_3 : \begin{array}{ccccccc} & \text{no} \\ i & \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} & \text{yielding} & i & \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \\ j & & & & & & j & & & \\ k & & & & & & k & & & \end{array} \quad (5.3.2)$$

$$P_{11} : \begin{array}{ccc} \text{no} & \text{no} & \text{no} \\ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \end{array} \quad (5.3.3)$$

Lemma 5.3.5 Let $A \in \text{Avoid}(m, \{H_2, H_3, H_5\})$. Consider the standard decomposition (2.1.1) of A based on row r . Let $L(r) \neq \emptyset$ be a minimal set of rows such that $C_r|_{L(r)}$ is simple. Then each triple of rows $\{i, j, k\}$ in $L(r)$ is in one of the cases P_2, P_3 or P_{11} . Moreover, if any triple in $L(r)$ is in case P_2 , then all triples of rows of $L(r)$ are in case P_2 . Similarly if any triple in $L(r)$ is in case P_3 (respectively P_{11}), then all triples of rows are in case P_3 (resp. P_{11}).

Proof: By Lemma 5.3.4, every triple of rows of $L(r)$ satisfies one of P_2, P_3 or P_{11} . A triple of rows $\{a, b, c\}$ in case P_3 can't overlap with a triple of rows in case P_2 (respectively P_{11}) on two rows $\{a, b\}$ since on the two rows $\{a, b\}$ What Is Missing (by (5.3.2)) will extend to one

new column missing on the triple from P_2 (resp. P_{11}) yielding a “ K_2 ”. This would allow us to delete a further row from $C_r|_{L(r)}$ while preserving simplicity, a contradiction to the fact that $L(r)$ is minimal with $C_r|_{L(r)}$ simple. Thus, if any triple of rows of $L(r)$ is in case P_3 , then all triples of rows of $L(r)$ are in case P_3 . Assume all triples of rows are in case P_2 or P_{11} .

We can't have a triple of rows in case P_2 overlap with a triple of rows in case P_{11} on two rows as shown below. On the quadruple of rows we have marked ‘OK’ over the columns which can occur on the quadruple of rows. At most 6 columns can be present in $C_r|_{L(r)}$ and we note that we can delete the second or third row from $C_r|_{L(r)}$ and not affect simplicity of $C_r|_{L(r)}$, a contradiction. Hence such an overlap cannot occur.

no	no	no	no	no	no	OK	OK	OK	OK	OK	OK
1	1	0	*	*	*	0	1	0	0	1	1
1	0	1	1	0	0	0	0	1	0	1	1
0	1	1	0	1	0	0	0	0	1	1	1
*	*	*	0	0	1	0	0	1	1	0	1

Given that each triple of the remaining rows of C_r rows must be in case P_2 or P_{11} , we must have all triples satisfy only one of the two. ■

Lemma 5.3.6 Assume all triples in $L(r)$ are in case P_3 . Then the rows of C_r can be ordered so that each triple of rows $a < b < c$ corresponds to $a = i$, $b = j$, and $c = k$ in P_3 .

Proof: In this case there is an ordering of the rows $L(r)$ so that all triples are consistent with the ordering given. We had noted in (5.3.2) that having P_3 on rows i, j, k in that order correspond to three columns, each on two rows, being absent. If we cannot find a consistent ordering of the rows of $L(r)$, then on some pair of rows we will be missing two columns and this implies that one of the two rows can be deleted while preserving simplicity of $C_r|_{L(r)}$. This contradiction proves the result. ■

In view of Lemma 5.3.5, we will say $L(r)$ is *type i* if each triple of rows in $L(r)$ is in case P_i for $i = 2, 3$ or 11 . Recall we assumed $\|C_r\| \geq 8$. We obtain $M(r)$ from $L(r)$ as follows

where the type of $M(r)$ is the type of $L(r)$.

$$M(r) = \begin{cases} L(r) & \text{if } L(r) \text{ is type 2 or 11} \\ L(r) \setminus \{\text{first and last row in ordering}\} & \text{if } L(r) \text{ is type 3} \end{cases} \quad (5.3.4)$$

Note $C_r|_{M(r)}$ need not be simple.

Lemma 5.3.7 Let $A \in \text{Avoid}(m, \{H_2, H_3, H_5\})$ with (2.1.1) applied for row r and $M(r)$ from (5.3.4).

- (i) If $M(r)$ is type 2, then $C_r|_{M(r)}$ must consist of $[\mathbf{0}_{|M(r)|} I_{|M(r)|}]$ and possibly column $\mathbf{1}_{|M(r)|}$ and no other column. Thus $\|C_r\| - 2 \leq |M(r)| \leq \|C_r\| - 1$. In addition, columns of $A|_{M(r)}$ are from $[\mathbf{0}_{|M(r)|} I_{|M(r)|} \mathbf{1}_{|M(r)|}]$.
- (ii) If $M(r)$ is type 11, then $C_r|_{M(r)}$ must consist of $[I_{|M(r)|}^c \mathbf{1}_{|M(r)|}]$ and possibly column $\mathbf{0}_{|M(r)|}$ and no other column. Thus $\|C_r\| - 2 \leq |M(r)| \leq \|C_r\| - 1$. In addition columns of $A|_{M(r)}$ are from $[\mathbf{0}_{|M(r)|} I_{|M(r)|}^c \mathbf{1}_{|M(r)|}]$.
- (iii) If $M(r)$ is type 3, then $C_r|_{M(r)}$ must consist of $[\mathbf{0}_{|M(r)|} T_{|M(r)|} \mathbf{1}_{|M(r)|}]$. Thus $|M(r)| = \|C_r\| - 3$. In addition, columns of $A|_{M(r)}$ are from $T_{|M(r)|}$.

Proof: For $M(r)$ being type 2, we observe that columns of $C_r|_{M(r)}$ must belong to

$$[\mathbf{0}_{|M(r)|} I_{|M(r)|} \mathbf{1}_{|M(r)|}].$$

By minimality of $L(r)$ (which is $M(r)$), we cannot delete any rows from $C_r|_{M(r)}$ and preserve simplicity. Thus all columns of $[\mathbf{0}_{|M(r)|} I_{|M(r)|}]$ must be present. A quick count reveals

$$\|C_r\| - 2 \leq |M(r)| \leq \|C_r\| - 1.$$

Similarly for $M(r)$ being type 11, $C_r|_{M(r)}$ must consist of $[I_{|M(r)|}^c \mathbf{1}_{|M(r)|}]$, and possibly column $\mathbf{0}_{|M(r)|}$ and no other column.

For $M(r)$ being type 3 then, with the row ordering of Lemma 5.3.6, $C_r|_{L(r)}$ must consist of $T_{|L(r)|}$. Hence $C_r|_{M(r)}$ must consist of

$$[\mathbf{0}_{|M(r)|} T_{|M(r)|} \mathbf{1}_{|M(r)|}],$$

and so $|M(r)| = \|C_r\| - 3$.

The restricted columns on $C_r|_{M(r)}$ extend to restricted columns on $A|_{M(r)}$ as follows. If $M(r)$ is type 2 then for any $H \subseteq M(r)$ with $|H| = 3$, the 6 forbidden columns on rows $r \cup H$ of Q_2 yield the restrictions P_2 of 3 forbidden columns on rows H of A . Thus the columns of $A|_{M(r)}$ are all contained in $[\mathbf{0}_{|M(r)|} I_{|M(r)|} \mathbf{1}_{|M(r)|}]$. In a similar way, if $M(r)$ is type 11 then the columns of $A|_{M(r)}$ are all contained in $[\mathbf{0}_{|M(r)|} I_{|M(r)|}^c \mathbf{1}_{|M(r)|}]$.

If $L(r)$ is type 3 we noted $C_r|_{L(r)}$ is $T|_{L(r)}$. Indeed, by Lemma 5.3.6, Q_3 has each triple $i, j, k \in L(r)$ ordered consistent with the ordering of the rows of $L(r)$ yielding T . We deduce the following columns are absent in A on rows $i < j < k$:

$$\begin{array}{cc} i \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} & i \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \\ j \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} & j \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \\ k \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} & k \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \end{array}$$

The following two columns are also forbidden on the 4 rows r, i, j, k of A by Q_3 :

$$\alpha = \begin{array}{c} r \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \\ i \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \\ j \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \\ k \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \end{array} \quad \beta = \begin{array}{c} r \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \\ i \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \\ j \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \\ k \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \end{array}$$

Thus, using α , under the 0's in row r in $[B_r C_r]|_{L(r)}$ we may only have the columns of $T|_{L(r)}$ plus one additional column consisting of all 0's except a 1 in the last row of $L(r)$. Similarly using β , under the 1's in row r in $[C_r D_r]|_{L(r)}$ we may only have the columns of $T|_{L(r)}$ plus one additional column consisting of all 1's except a 0 in the first row of $L(r)$. Thus if $M(r)$ is $L(r)$ with the first and last row deleted then $C_r|_{M(r)} = [\mathbf{0} T \mathbf{1}]$ and the columns of $A|_{M(r)}$ are contained in $[T_{M(r)}]$. ■

Proof of Lemma 5.3.2: Let $A \in \text{Avoid}(m, \{H_2, H_3, H_5\})$. Use the decomposition of A given in (2.1.1). Our procedure is as follows. We use Lemma 5.3.5 to deduce the possible cases we need to consider. Under the assumption that $\|C_r\| \geq 8$ for all rows r , we will establish by induction an infinite sequence r_1, r_2, r_3, \dots and associated sets of rows $N(r_1), N(r_2), N(r_3), \dots$ with $|N(r_i)| \geq 4$ for each i . The sets $N(r)$ differ very little from $L(r)$ and $M(r)$. We are able to show that the sets $N(r_1) \setminus r_2, N(r_2) \setminus r_3, \dots, N(r_i) \setminus r_{i+1}$ are all disjoint (see the beginning of Case 1a) and yet $|N(r_j) \setminus r_{j+1}| \geq 3$. This yields a contradiction (there are only m rows!) and so we may conclude that for some r , $\|C_r\| \leq 7$. Hence by our induction we deduce that

$$\|A\| \leq 7m.$$

Assume for all rows r that $\|C_r\| \geq 8$ and hence find the sets $M(r)$ with $|M(r)| \geq 5$ (checking the three cases of Lemma 5.3.7). Let r_1 be some row of A . We form $M(r_1)$. Note that if $M(r_1)$ was type 3 then we have deleted the first and last rows (in the ordering) from the originally determined $L(r_1)$. We determine the sets $N(r_i)$ from $M(r_i)$ as follows

$$N(r) = \begin{cases} M(r) & \text{if } M(r) \text{ is type 2 or 11} \\ M(r) \setminus \text{first row in ordering} & \text{if } M(r) \text{ is type 3} \end{cases} \quad (5.3.5)$$

Our general step commences with $N(r_i)$. We select a row $r_{i+1} \in N(r_i)$, making sure that when $N(r_i)$ is of type 3, we select the first row in the ordering of Lemma 5.3.6.

We obtain $M(r_{i+1})$ applying Lemma 5.3.4, Lemma 5.3.5, Lemma 5.3.6 and Lemma 5.3.7. Given our assumption that $\|C_r\| \geq 8$ we have $|M(r_{i+1})| \geq 5$. Now by (5.3.5) we deduce $|N(r_{i+1})| \geq 4$ in all cases. We hope identifying $L(r), M(r), N(r)$ makes the proof clearer.

To show the desired properties of the sets $N(r_i)$, we set up an inductive hypothesis concerning the structure of A . In what follows let Z denote a matrix of 0's (or perhaps a matrix of no columns) and J denote a matrix of 1's (or perhaps a matrix of no columns). The critical inductive structure is the following. The middle columns correspond to the columns of C_{r_i} as shown in (5.3.6). For each p with $p < i$ with $N(r_p)$ of type 2 or 3 we have the structure shown in rows $N(r_p) \setminus r_{p+1}$. For each q with $q < i$ with $N(r_q)$ of 11 we have the structure shown in rows $N(r_q) \setminus r_{q+1}$. We have three cases depending on the type of $N(r_i)$.

When $N(r_i)$ is type 2: We have $S = [\mathbf{0} I]$ or $[\mathbf{0} I \mathbf{1}]$ and the columns of U_i and V_i are in $[\mathbf{0} I \mathbf{1}]$.

When $N(r_i)$ is type 11: We have $S = [I^c \mathbf{1}]$ or $[\mathbf{0} I^c \mathbf{1}]$ and the columns of U_i, V_i are in $[\mathbf{0} I^c \mathbf{1}]$.

When $N(r_i)$ is type 3: We have $S = [\mathbf{0} T \mathbf{1}]$ and the columns of U_i, V_i are in T .

$$A = \begin{array}{c} r_i \rightarrow \\ \vdots \\ N(r_p) \setminus r_{p+1} \{ \\ \vdots \\ N(r_q) \setminus r_{q+1} \{ \\ \vdots \\ N(r_i) \{ \\ \vdots \end{array} \left[\begin{array}{cc|c|c|cc} 0 \cdots 0 & 0 \cdots 0 & 0 \cdots 0 & 1 \cdots 1 & 1 \cdots 1 & 1 \cdots 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ Z & W_p^0 & Z & Z & W_p^1 & Z \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ J & W_q^0 & J & J & W_q^1 & J \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ U_i & ZJ & S & S & ZJ & V_i \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{array} \right] \quad (5.3.6)$$

We proceed to verify that we have the same inductive structure for r_{i+1} . There will be cases to explore. It is helpful to display representatives of H_2, H_3, H_5 that we will use in our arguments. For $M(r_{i+1})$ type 2 or 11 we will use

$$H_2 = \begin{array}{c} r_{i+1} \\ s \\ i \\ j \end{array} \left[\begin{array}{c|ccc} 0 & 0 & 0 & 1 & 1 \\ \hline 1 & 0 & 0 & 0 & 0 \\ \hline 1 & 1 & 0 & 1 & 0 \\ \hline 0 & 0 & 1 & 0 & 1 \end{array} \right], \quad H_3 = \begin{array}{c} r_{i+1} \\ s \\ i \\ j \end{array} \left[\begin{array}{c|ccc} 0 & 1 & 1 & 1 \\ \hline 0 & 0 & 0 & 1 \\ \hline 1 & 1 & 0 & 1 \\ \hline 0 & 0 & 1 & 0 \end{array} \right] \quad (5.3.7)$$

$$H_3 = \begin{array}{c} r_{i+1} \\ t \\ i \\ j \end{array} \left[\begin{array}{c|ccc} 0 & 0 & 0 & 1 \\ \hline 0 & 1 & 1 & 1 \\ \hline 1 & 0 & 1 & 1 \\ \hline 0 & 1 & 0 & 0 \end{array} \right], \quad H_5 = \begin{array}{c} r_{i+1} \\ t \\ i \\ j \end{array} \left[\begin{array}{c|ccc} 0 & 0 & 1 & 1 & 1 \\ \hline 1 & 1 & 1 & 1 & 0 \\ \hline 0 & 1 & 0 & 1 & 1 \\ \hline 1 & 0 & 1 & 0 & 0 \end{array} \right] \quad (5.3.8)$$

For $M(r_{i+1})$ type 3 we will use

$$H_3 = \begin{array}{c} r_{i+1} \\ s \\ i \\ j \end{array} \left[\begin{array}{c|ccc} 0 & 1 & 1 & 1 \\ \hline 1 & 0 & 0 & 0 \\ \hline 1 & 0 & 1 & 1 \\ \hline 0 & 0 & 0 & 1 \end{array} \right], \quad H_3 = \begin{array}{c} r_{i+1} \\ s \\ i \\ j \end{array} \left[\begin{array}{c|ccc} 0 & 1 & 1 & 1 \\ \hline 0 & 0 & 0 & 1 \\ \hline 1 & 0 & 1 & 1 \\ \hline 1 & 0 & 0 & 0 \end{array} \right] \quad (5.3.9)$$

$$H_3 = \begin{matrix} & r_{i+1} & \left[\begin{array}{c|ccc} 0 & 0 & 0 & 1 \\ \hline 0 & 1 & 1 & 1 \\ \hline 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \end{array} \right] \\ t & & \\ i & & \\ j & & \end{matrix}, \quad H_3 = \begin{matrix} & r_{i+1} & \left[\begin{array}{ccc|c} 0 & 0 & 0 & 1 \\ \hline 1 & 1 & 1 & 0 \\ \hline 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \end{array} \right] \\ t & & \\ i & & \\ j & & \end{matrix} \quad (5.3.10)$$

Case 1: $N(r_i)$ is type 2.

Begin with inductive structure of (5.3.6). Given $N(r_i)$ of type 2 we have $S = [\mathbf{0} I]$ or $[\mathbf{0} I \mathbf{1}]$. Choose a row $r_{i+1} \in N(r_i)$. Now consider the decomposition (2.1.1) applied to A using row $r = r_{i+1}$. Apply Lemma 5.3.4, Lemma 5.3.5, Lemma 5.3.6 and Lemma 5.3.7 to obtain $M(r_{i+1})$.

Case 1a: $M(r_{i+1})$ is type 2.

The columns of $C_{r_{i+1}}$ must appear once with a 0 in row r_{i+1} and once with a 1 in row r_{i+1} . By Lemma 5.3.7 we know that columns of $A|_{N(r_i)}$ are contained in $[\mathbf{0} I \mathbf{1}]$. The only columns of $A|_{N(r_i)}$ which differ only in row r_{i+1} would be the column of 0's and the column of all 0's except a 1 in row r_{i+1} . Thus the repeated columns of $C_{r_{i+1}}$, when restricted to rows $N(r_i) \setminus r_{i+1}$, must be all 0's. By examining (5.3.6), the only columns of A which on rows $N(r_i)$ that have a single 1 (on row r_{i+1}) respectively on the rows $N(r_i)$ are the columns which are Z in rows $N(r_p) \setminus r_{p+1}$ for those $p < i$ with $N(r_p)$ being type 2 or 3 and J in rows $N(r_q) \setminus r_{q+1}$ for those $q < i$ with $N(r_q)$ being type 11.

We need to show that $N(r_{i+1})$ is disjoint from $N(r_j) \setminus r_{j+1}$ for all $j < i + 1$. All columns in W^0 or W^1 of (5.3.6) are either all 0's or all 1's on the rows of $N(r_i)$ and so won't give rise to columns of $C_{r_{i+1}}$. We deduce that the columns of $C_{r_{i+1}}$ are all 0's in rows $N(r_p) \setminus r_{p+1}$ for those $p < i$ with $N(r_p)$ being type 2 or 3 and all 1's in rows $N(r_q) \setminus r_{q+1}$ for those $q < i$ with $N(r_q)$ being type 11. Recalling that we form $L(r_{i+1})$ by deleting rows of $C_{r_{i+1}}$ while preserving simplicity, we deduce that $L(r_{i+1})$ (and hence $M(r_{i+1})$ and $N(r_{i+1})$) is disjoint from $N(r_j) \setminus r_{j+1}$ for all $j < i + 1$.

This gives us the structure of $C_{r_{i+1}}$ given below in (5.3.11) where the two copies of $C_{r_{i+1}}$ occupy the central columns. To complete (5.3.11) we define W^0 and W^1 (likely different from those in (5.3.6) in the paragraph above). We choose from the columns of $B_{r_{i+1}}$ and $D_{r_{i+1}}$, all columns which for some $\ell < i$, where $N(r_\ell)$ is type 2 or 3 (and hence rows $N(r_\ell)$ is Z in $C_{r_{i+1}}$), have a 1 in some row of $N(r_\ell)$ or for some $\ell < i$, with $N(r_\ell)$ is type 11 (and hence rows $N(r_\ell)$ is J in $C_{r_{i+1}}$), have a 0 in some row of $N(r_\ell)$. We identify such columns in $B_{r_{i+1}}$ as W^0 and such columns in $D_{r_{i+1}}$ as W^1 . Moreover let W_t^0 (respectively W_t^1) denote the submatrix of W^0 (respectively W^1) in rows $N(r_t) \setminus r_{t+1}$ for $t = 1, \dots, i$ or in rows $M(r_t)$

for $t = i + 1$.

All remaining columns of $B_{r_{i+1}}$ and $D_{r_{i+1}}$ are all 0's on rows of each $N(r_\ell)$ where $N(r_\ell)$ is type 2 or 3 and all 1's on rows of each $N(r_\ell)$ where $N(r_\ell)$ is type 11 for $\ell < i$.

$$\begin{array}{l}
r_{i+1} \rightarrow \\
\vdots \\
N(r_p) \setminus r_{p+1} \{ \\
\vdots \\
N(r_q) \setminus r_{q+1} \{ \\
\vdots \\
N(r_i) \setminus r_{i+1} \{ \\
M(r_{i+1}) \{ \\
\vdots
\end{array}
\left|
\begin{array}{cccccc}
0 \cdots 0 & 0 \cdots 0 & 0 \cdots 0 & 1 \cdots 1 & 1 \cdots 1 & 1 \cdots 1 \\
Z & W_p^0 & Z & Z & W_p^1 & Z \\
J & W_q^0 & J & J & W_q^1 & J \\
Z & W_i^0 & Z & Z & W_i^1 & Z \\
U_{i+1} & W_{i+1}^0 & \mathbf{0} I \mathbf{1} & \mathbf{0} I \mathbf{1} & W_{i+1}^1 & V_{i+1} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots
\end{array}
\right.
\quad (5.3.11)$$

By Lemma 5.3.7 we know that columns of $A|_{N(r_i)}$ are contained in $[\mathbf{0} I \mathbf{1}]$ and so we deduce that columns of U_{i+1}, V_{i+1} are in $[\mathbf{0} I \mathbf{1}]$. Our remaining goal is to show that $W_{i+1}^0 = ZJ$ and $W_{i+1}^1 = ZJ$ to complete the induction. We will use the four forbidden matrices of (5.3.7), (5.3.8) which have been ordered and labelled to assist the reader in seeing the occurrence of the forbidden objects H_2, H_3, H_5 . Assume for some column α of W^0 that α has a 1 in row $s \in N(r_p) \setminus r_{p+1}$ where $N(r_p)$ is type 2 or 3. We will give this first case in greater detail. All columns of $C_{r_{i+1}}$ have 0's in the rows of $N(r_p)$ and in particular in row s . Given that $M(r_{i+1})$ is type 2 or 11 we deduce $C_{r_{i+1}}|_{M(r_{i+1})}$ contains either I or I^c . Thus each pair of rows $i, j \in M(r_{i+1})$ will contain $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ in each copy of $C_{r_{i+1}}$. We find the following entries in A in the rows r, s, i, j where the left column comes from α and the remaining columns are from the two copies of $C_{r_{i+1}}$:

$$\begin{array}{l}
r_{i+1} \\
s \\
i \\
j
\end{array}
\left[
\begin{array}{c|ccc}
0 & 0 & 0 & 1 & 1 \\
\hline
1 & 0 & 0 & 0 & 0 \\
\hline
a & 1 & 0 & 1 & 0 \\
b & 0 & 1 & 0 & 1
\end{array}
\right].$$

If $\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ or $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ then we have a representative of H_2 as noted in the left matrix of (5.3.7). Thus the column α which contains a 1 in some row s of W_p^0 must either be all 0's or all 1's on the rows $M(r_{i+1})$. Assume for some column β of W^0 that β has a 0 in row $t \in N(r_q) \setminus r_{q+1}$ where $N(r_q)$ is type 11. Using the left matrix of (5.3.8) we may argue as above that column

β must either be all 0's or all 1's on the rows $M(r_{i+1})$. Given our choice of W^0 , this is enough to show that W_{i+1}^0 is ZJ .

Assume for some column α of W^1 that α has a 1 in row $s \in N(r_p) \setminus r_{p+1}$ where $N(r_p)$ is type 2 or 3 and hence we find 0's in $C_{r_{i+1}}$ in row s . Hence by the right matrix in (5.3.7) we cannot have the matrix $\begin{smallmatrix} i \\ j \end{smallmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ in α for any choices $i, j \in M(r_{i+1})$. As above, the column α is either all 1's or all 0's on the rows of $M(r_{i+1})$. Similarly, using the right matrix of (5.3.8), we can show that for any column β of W^1 that has a 0 in row $t \in N(r_q) \setminus r_{q+1}$ where $N(r_q)$ is type 11 that β cannot have the matrix $\begin{smallmatrix} i \\ j \end{smallmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ in α for any choices $i, j \in M(r_{i+1})$. Hence β is either all 0's or all 1's on the rows of $M(r_{i+1})$. Thus $W_{i+1}^1 = ZJ$ as desired. Setting $N(r_{i+1}) = M(r_{i+1})$ results in the same structure of (5.3.6) with r_i replaced by r_{i+1} and $S = [\mathbf{0} I]$ or $[\mathbf{0} I \mathbf{1}]$.

Case 1b: $M(r_{i+1})$ is type 11.

We can use the argument of Case 1a if $M(r_{i+1})$ is type 11 since any two rows of I^c contain I_2 allowing us to use the matrices of (5.3.7),(5.3.8) as above. We would obtain (5.3.6) with r_i replaced by r_{i+1} , $N(r_{i+1}) = M(r_{i+1})$ and $S = [I^c \mathbf{1}]$ or $[\mathbf{0} I^c \mathbf{1}]$.

Case 1c: $M(r_{i+1})$ is type 3.

We follow the argument at the beginning of Case 1a) to obtain most of the structure of (5.3.12). Given that we form $L(r_{i+1})$ by deleting rows of $C_{r_{i+1}}$ while preserving simplicity, we deduce that $L(r_{i+1})$ (and hence $M(r_{i+1})$) is disjoint from $N(r_j) \setminus r_{j+1}$ for all $j < i + 1$. We will use (5.3.9) and (5.3.10) and, arising from the left matrix of (5.3.10), we discover a row of $M(r_{i+1})$ that must be deleted.

$$\begin{array}{l}
r_{i+1} \rightarrow \\
\vdots \\
N(r_p) \setminus r_{p+1} \{ \\
\vdots \\
N(r_q) \setminus r_{q+1} \{ \\
\vdots \\
N(r_i) \setminus r_{i+1} \{ \\
M(r_{i+1}) \{
\end{array}
\left|
\begin{array}{cccccc}
0 \cdots 0 & 0 \cdots 0 & 0 \cdots 0 & 1 \cdots 1 & 1 \cdots 1 & 1 \cdots 1 \\
Z & W_p^0 & Z & Z & W_p^1 & Z \\
J & W_q^0 & J & J & W_q^1 & J \\
Z & W_i^0 & Z & Z & W_i^1 & Z \\
U_{i+1} & W_{i+1}^0 & \mathbf{0} T \mathbf{1} & \mathbf{0} T \mathbf{1} & W_{i+1}^1 & V_{i+1}
\end{array}
\right.
\tag{5.3.12}$$

Do not be concerned that $C_{r_{i+1}}$ as shown is not simple, as we have deleted two rows from $L(r_{i+1})$ to obtain $M(r_{i+1})$ which are not displayed here. As before, we note that by Lemma 5.3.7, the columns of $U_{i+1}, V_{i+1}, W_{i+1}^0, W_{i+1}^1$ are contained in T . Our goal to complete the induction is to show $W_{i+1}^0 = ZJ$ and $W_{i+1}^1 = ZJ$. We use the four forbidden matrices of

(5.3.9),(5.3.10).

Given that $C_{r_{i+1}}|_{M(r_{i+1})} = [\mathbf{0} T \mathbf{1}]$, each pair of rows $i, j \in M(r_{i+1})$ with $i < j$ in the special row ordering of $M(r_{i+1})$ will contain $\begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$ in each copy of $C_{r_{i+1}}$.

If we have a column α of W^1 with a 1 in a row $s \in N(r_j) \setminus r_{j+1}$ where $N(r_j)$ is type 2 or 3 and hence we find 0's in columns of $C_{r_{i+1}}$ in row s . Hence by the right matrix in (5.3.9), α cannot have the submatrix $\begin{matrix} i \\ j \end{matrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ for each pair of rows $i, j \in M(r_{i+1})$ with $i < j$. Given that $\alpha|_{M(r_{i+1})}$ is a column in T , we deduce that column α is either all 1's or all 0's on the rows of $M(r_{i+1})$. If we have a column β of W^1 with a 0 in a row $t \in N(r_j) \setminus r_{j+1}$ where $N(r_j)$ is type 11, we find 1's in row t of $C_{r_{i+1}}$. Hence by the right matrix in (5.3.10), β cannot have the submatrix $\begin{matrix} i \\ j \end{matrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ for each pair of rows $i, j \in M(r_{i+1})$ with $i < j$. As above, the column β is either all 1's or all 0's on the rows of $M(r_{i+1})$. This considers all columns of W^1 and so $W_{i+1}^1 = ZJ$.

If we have a column α of W^0 with a 1 in a row $s \in M(r_p) \setminus r_{p+1}$ where $M(r_p)$ is type 2 or 3, we find 0's in row s of $C_{r_{i+1}}$. Hence by the left matrix in (5.3.9), α cannot have the submatrix $\begin{matrix} i \\ j \end{matrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ for each pair of rows $i, j \in M(r_{i+1})$ with $i < j$ and so the column α is either all 1's or all 0's on the rows of $M(r_{i+1})$. If we have a column β of W^0 with a 0 in row $t \in N(r_q)$ where $N(r_q)$ is type 11 then we follow a different argument that we explain more carefully. For $i, j \in M(r_{i+1})$ with $i < j$, we find the entries as given below in the rows r_{i+1}, t, i, j in the given column β (the column on the left) and selected columns of $C_{r_{i+1}}$ (on the right).

$$\begin{array}{c} r_{i+1} \\ t \\ i \\ j \end{array} \left[\begin{array}{c|ccc} 0 & 0 & 0 & 1 \\ \hline 0 & 1 & 1 & 1 \\ \hline a & 0 & 1 & 1 \\ \hline b & 0 & 0 & 0 \end{array} \right]$$

If $\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ then this yields H_3 in A as noted in the left matrix in (5.3.10). Now $\beta|_{M(r_{i+1})}$ is a column in T and yet cannot have the submatrix $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Thus β on the rows of $M(r_{i+1})$ is either all 0's or possibly the column of all 0's except a single 1 in the first row of $M(r_{i+1})$. It is for this case that we need to delete the first row of $M(r_{i+1})$ to obtain $N(r_{i+1})$ (as in (5.3.5)) so that on the rows $N(r_{i+1})$, the matrix $W_{i+1}^0 = ZJ$. We now have obtained (5.3.6) with r_i replaced by r_{i+1} , and $S = T$.

Case 2: $N(r_i)$ is type 11.

We use the same argument as Case 1. When $N(r_i)$ is type 11 we would have to replace I by I^c in S in (5.3.6) and then proceed to $M(r_{i+1})$ of type 2 or 11 (essentially Case 1a or 1b) or $M(r_{i+1})$ of type 3 (essentially Case 1c).

Case 3: $N(r_i)$ is type 3.

Begin with inductive structure of (5.3.6) where $N(r_i)$ is type 3 and $S = [\mathbf{00T11}]$. Now choose the *first* row $r_{i+1} \in N(r_i)$ using the ordering on $N(r_i)$. Now consider Standard Induction applied to A using row r_{i+1} . We deduce that in rows $N(r_i) \setminus r_{i+1}$, the repeated columns in $C_{r_{i+1}}$ are Z since for a column to be repeated it extension to row r_{i+1} with both a 0 and a 1 must be present in $C_{r_{i+1}}$. Given that the repeated columns under the 1's in row r_{i+1} must correspond to columns of a single 1 on rows $N(r_i) \setminus r_{i+1}$ and by (5.3.6) that means we can deduce the structure of the other rows of the columns in $C_{r_{i+1}}$. Note that in what follows I have rearranged the columns of $B_{r_{i+1}}$ and $D_{r_{i+1}}$ so that we have put in the columns of the W_j 's all columns which either have a 0 in a row of $N(r_j)$ where $N(r_j)$ is type 2 or 3 (and hence is Z in $C_{r_{i+1}}$), and all columns which have a 1 in a row of $N(r_j)$ where $N(r_j)$ is type 11 (and hence is J in $C_{r_{i+1}}$). This yields (5.3.11) or (5.3.12) .

If $M(r_{i+1})$ is type 2 we follow the same argument as in Case 1a) to deduce that W_{i+1}^0 and W_{i+1}^1 have only constant columns. Similarly the case $M(r_{i+1})$ is type 11 can use the argument of Case 1b) by switching I with I^c . In either case we set $N(r_{i+1}) = M(r_{i+1})$. If $M(r_{i+1})$ is type 3, we follow the same argument as in Case 1c) and again may have to delete the first row of $M(r_{i+1})$ to obtain $N(r_{i+1})$ and yields (5.3.6) with r_i replaces by r_{i+1} . This concludes the induction and so have proven that we can find rows r_1, r_2, r_3, \dots and disjoint sets $|N(r_i) \setminus r_{i+1}| \geq 3$ yielding a contradiction. As noted this proves the result. ■

We still have eight 5×6 simple F for which the conjecture predicts they are boundary cases with $\text{forb}(m, F)$ being $O(m^2)$. Given the complicated case analysis of this proof, it seems challenging to prove such bounds. One positive observation is that Lemma 5.3.5 may not be necessary. We were only interested in having a large set $L(r)$, say $|L(r)| \geq 8$, for which each triple is in a given case. We could appeal to Ramsey Theory [Ram30] and given a finite number of cases, we can identify a large (!) constant c so that if $\|C_r\| \geq c$ then there are say 8 rows such that every triple is in the same case and in the same row ordering. This would avoid appealing to the particular structures of cases P_2, P_3, P_{11} but is not advantageous for our proof.

5.4 Classification of 6-rowed Quadratic Bounds

This section is devoted to proving a quadratic bound for the 6×3 -rowed configuration $G_{6 \times 3}$. We prove $\text{forb}(m, G_{6 \times 3})$ is $\Theta(m^2)$, and furthermore, we prove $G_{6 \times 3}$ is the only 6-rowed

boundary quadratic case, therefore classifying all quadratic 6-rowed cases.

The following theorem classifies all 6-rowed configurations F for which $\text{forb}(m, F)$ is $\Theta(m^2)$ by giving the unique boundary case.

Theorem 5.4.1 Let F be any 6-rowed configuration. Then $\text{forb}(m, F)$ is $\Theta(m^2)$ if and only if F is a configuration in

$$G_{6 \times 3} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Furthermore, if $F \neq G_{6 \times 3}$, then $\text{forb}(m, F)$ is $\Omega(m^3)$.

Given the classification and Remark 1.3.22, we are not surprised that $G_{6 \times 3}^c = G_{6 \times 3}$.

Anstee and Keevash [AK06] established the asymptotic bounds for all $k \times 2$ configurations, and in particular concluded that

$$\text{forb}\left(m, \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}\right) \text{ and } \text{forb}\left(m, \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}\right) \text{ are both } \Theta(m^2).$$

The proof of the second of these begins to use the full power of the proof in [AK06] and so it is interesting that Theorem 5.4.1 provides a generalization for both of them using an inductive proof (admittedly, a rather complicated one for Theorem 5.3.1) that is quite different than that in [AK06].

In order to prove Theorem 5.4.1, we will use three results. First, Lemma 5.4.2 is the “only if” part of the theorem. The second, Lemma 5.4.3, generalizes Lemma 3.2 in [AK06]. Lastly, we will use the main result of the previous chapter, Theorem 5.3.1.

Lemma 5.4.2 Let F be a 6-rowed configuration such that $F \not\prec G_{6 \times 3}$. Then $\text{forb}(m, F)$ must be $\Omega(m^3)$.

Proof: We may assume all of F 's columns have column sum 3, otherwise, if F had a column of column sum 4 or more, then $F \not\prec I \times I \times I$, and if F had a column sum of 2 or less, then $F \not\prec I^c \times I^c \times I^c$.

Without loss of generality, let the first column of F be $(1, 1, 1, 0, 0, 0)^T$. With these assumptions, there are only a few cases left to check, and an exhaustive computer search revealed the lemma to be true. But we present here a more constructive proof, if for no other reason than to check the computer code.

Note that the following 2-columned matrices have at least a cubic bound:

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \not\prec I \times I \times I, \quad \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \not\prec I \times I \times T.$$

This means that to in order to have F for which $\text{forb}(m, F)$ is not $\Omega(m^3)$, we must put together columns of sum 3 such that for each pair of columns, the number of rows where both columns have 1's is either one or two. Here are all the possibilities for (the first) two columns having 1's in (the first) two rows in common:

$$\begin{array}{ccccc} \left[\begin{array}{c|c|c} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{array} \right] & \left[\begin{array}{c|c|c} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array} \right] & \left[\begin{array}{c|c|c} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{array} \right] & \left[\begin{array}{c|c|c} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{array} \right] & \left[\begin{array}{c|c|c} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{array} \right] . \\ \not\prec I^c \times I^c \times I^c & \not\prec I \times I \times I & = G_{6 \times 3} & \not\prec I^c \times I^c \times I^c & \not\prec I \times I \times I \end{array}$$

The only other possibility is that each pair of columns has a 1 in only one row in common.

$$\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \not\sim I \times I \times T.$$

Thus, the only four-columned matrices F for which $\text{forb}(m, F)$ could be $O(m^2)$ have to contain $G_{6 \times 3}$ in every three-columned subset. The only possibility is then

$$\begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \not\sim I \times I \times T,$$

which means $\text{forb}(m, F)$ is $\Omega(m^3)$ for the above matrix F . This concludes the lemma. ■

The following lemma generalizes Lemma 3.2 in [AK06].

Lemma 5.4.3 Let

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

Then we can conclude that

$$\text{forb}(m, F) \leq \text{forb} \left(m, \begin{bmatrix} 1 & \cdots & 1 \\ F' \end{bmatrix} \right) + \text{forb} \left(m, \begin{bmatrix} 0 & \cdots & 0 \\ F' \end{bmatrix} \right).$$

Proof: Let $A \in \text{Avoid}(m, F)$ with $\|A\| = \text{forb}(m, F)$. Then permute the columns of A (take

another representative in the equivalence class) and write it as

$$A = \begin{bmatrix} 0 & \cdots & 0 & 1 & \cdots & 1 \\ & & A' & & & A'' \end{bmatrix}.$$

Note that A' and A'' are simple. Since A' cannot have $\begin{bmatrix} 1 & \cdots & 1 \\ & F' & \end{bmatrix}$ as a subconfiguration,

and A'' cannot have $\begin{bmatrix} 0 & \cdots & 0 \\ & F' & \end{bmatrix}$ as a subconfiguration, the bound follows. \blacksquare

From the previous lemma, we note that $G_{6 \times 3}$ has a row of 0's and a row of 1's, and therefore the quadratic bound for $\text{forb}(m, G_{6 \times 3})$ would follow from quadratic bounds for $\text{forb}(m, G)$ and $\text{forb}(m, G')$, with G and G' obtained by removing the row of 1's and the row of 0's from $G_{6 \times 3}$ respectively:

$$G = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad G' = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

We will prove more, as both are contained in the boundary case F_7 . Observe that $G' = G^c$ as configurations. We are now ready to prove Theorem 5.4.1.

Proof of Theorem 5.4.1: To prove $\text{forb}(m, G_{6 \times 3})$ is $O(m^2)$ we use Lemma 5.4.3 and so we only need to prove $\text{forb}(m, G)$ and $\text{forb}(m, G')$ are both $O(m^2)$. We check that $G \prec F_7$ and $G' \prec F_7$. Now Theorem 5.3.1 shows that $\text{forb}(m, F_7)$ is $O(m^2)$. Applying Remark 1.3.21 yields the bound for $G_{6 \times 3}$. Lemma 5.4.2 verifies that every configuration F not contained in $G_{6 \times 3}$ has $\text{forb}(m, F)$ being $\Omega(m^3)$. \blacksquare

Chapter 6

Patterns and Splits

6.1 Patterns and Splits in 2-Dimensions

We move away from the world of Forbidden Configurations to consider another extremal problem which will have applications to Forbidden Configurations (in Chapter 7), but no knowledge of Forbidden Configurations is required in order to understand this chapter.

Instead of considering the maximum number of columns an m -rowed $\{0, 1\}$ -matrix can have, we are going to consider the maximum number of 1's an $m \times n$ $\{0, 1\}$ -matrix can have subject to some property. These problems are about geometric patterns in a grid and are close relatives of Zarankiewicz' problem [KST54],[Für96] and the investigations of *patterns* in [FH92],[MT04],[Tar05],[KM07].

Consider the grid $[m] \times [n]$ in the Euclidean 2-dimensional space. Let $\mathbb{A} \subseteq [m] \times [n]$. Any such subset can be represented by an $m \times n$ $\{0, 1\}$ -matrix A , where A has a 1 in position (x, y) iff $(x, y) \in \mathbb{A}$. In order to give the following definitions, we find it convenient to talk about the subset \mathbb{A} instead of the matrix A , but in our applications and proofs it becomes more convenient to talk about the matrix A . Since A and \mathbb{A} are equivalent, we will switch back and forth. Sometimes it will be more convenient to use the language of grids and other times to use the language of matrices.

Recall that $\sigma_1(A)$ denotes the number of 1's of A . Observe that $\sigma_1(A) = |\mathbb{A}|$. We consider horizontal and vertical lines on integer values (i.e. $y = \ell$ or $x = \ell$ for some integer k). For a horizontal line $y = \ell$, we partition the points of the grid into two parts. The first part consists of those points which lie in the bottom region of the line (points (x, y) such that $y < \ell$) and the second part consists of points which lie above the line or in the line itself (points (x, y) with $y \geq \ell$).

Let p and q be numbers. If we consider $p - 1$ horizontal lines and $q - 1$ vertical lines in the plane, we naturally divide the plane into $p \cdot q$ rectangular regions. The horizontal lines can be represented by $p - 1$ numbers $h_1 < h_2 < \dots < h_{p-1}$ and the vertical lines by $q - 1$ numbers $v_1 < v_2 < \dots < v_{q-1}$. Set $h_0 = v_0 = 1$ and set $h_p = n, v_q = m$. Then the regions are of the form $R(i, j) := [h_i, h_{i+1}) \times [v_j, v_{j+1})$.

Definition 6.1.1 Let p and q be numbers and let $\mathbb{A} \subseteq [m] \times [n]$. We say \mathbb{A} has a (p, q) -**split** if there exist p vertical lines and q horizontal lines such that for every one of the pq regions $R(i, j)$ we have $R(i, j) \cap \mathbb{A} \neq \emptyset$.

Below is an example of a 3, 3-split in A , where a 1 from each block is indicated

$$\begin{bmatrix} & & 1 & & \\ & 1 & & & 1 \\ \hline 1 & & 1 & & 1 \\ & & & & \\ \hline & 1 & & & \\ & & & 1 & 1 \end{bmatrix}$$

A 3,3 split

Definition 6.1.2 Let $\text{NoSplit}(m, n; p, q)$ denote the maximum number of 1's in a $m \times n$ $\{0, 1\}$ -matrix that does *not* have a p, q -split.

Here is a definition that is related to splits, but differs. Note that row and column order is important.

Definition 6.1.3 Let F be a $\{0, 1\}$ -matrix. We say a $\{0, 1\}$ matrix A has F as a **pattern** if there is a submatrix G of A with $F \leq G$. That is, if the entry (i, j) is 1 in F then entry (i, j) is also 1.

The following theorem was proven by Marcus and Tardos, but using different language.

Theorem 6.1.4 Marcus and Klazar [KM07], Marcus and Tardos [MT04], Balogh, Bollobás and Morris [BBM06]. Let k be given. Then there exists a constant c_k such that $\text{NoSplit}(n, n; k, k) \leq c_k n$.

The result in [MT04] involving forbidden permutation patterns implies the above result by choosing the permutation appropriately. Moreover the proof of [MT04] directly extends to the above result. The papers [KM07], [BBM06] note this as well as deriving higher dimension generalizations. While the constants involved in [MT04] are not optimal (in fact, they are very far from optimal), we can produce best possible constants for small values:

Theorem 6.1.5 Let m, n be given with $m, n \geq 2$. Then $\text{NoSplit}(m, n; 2, q) = m + (q - 1) \cdot n - (q - 1)$ and $\text{NoSplit}(m, n; 3, q) = 2m + (q - 1)n - 2(q - 1)$.

Proof: For any m, n, p, q , an $m \times n$ matrix A with $(p - 1)m + (q - 1)n - (p - 1)(q - 1)$ 1's can be constructed with 1's in the first $p - 1$ rows and the first $q - 1$ columns. Then A has no p, q split. We show an example of this for $m = 7, n = 6, p = 3$ and $q = 2$.

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

Of course the matrix shown above has no 3, 2-split. If we take any two horizontal lines and one vertical line, the bottom-right region will have no 1's. This yields a lower bound for $\text{NoSplit}(p, q; m, n)$.

For $p = 2$, consider a matrix A' by first deleting the bottommost 1 from each column. Thus the bottom row has all 0's. At most m 1's were deleted. Then delete the $q - 1$ rightmost 1's from each non-zero row. At most we delete $(n - 1)(q - 1)$ 1's. Then we have

$$\sigma_1(A) \leq \sigma_1(A') + m + (n - 1)(q - 1)$$

Now, if $\sigma_1(A') > 0$, pick a 1 from A' . It must have $q - 1$ 1's to its right, and each of those must have a 1 below. This produces a 2, q split, a contradiction.

For $p = 3$ we proceed in a similar way. Form A' by deleting from A the bottommost 1 from each column. Then delete the topmost 1 from each column. Note that the bottom row as well as the top row cannot have 1's. Then delete the $q - 1$ rightmost 1's in each non-zero

row. So

$$\sigma_1(A) \leq \sigma_1(A') + 2m + (n - 2)(q - 1)$$

Now, if $\sigma_1(A') > 0$, pick a 1 from A' . It must have $(q - 1)$ 1's to its right, and each of those must have a 1 below and a 1 above. This produces a $3, q$ split, a contradiction. ■

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

We do not have good bounds for $\text{NoSplit}(m, n; p, q)$ with $p, q \geq 4$. We conjecture that $\text{NoSplit}(m, n; 4, 4) = 3m + 3n + \min(m, n) - 13$. We can prove that $\text{NoSplit}(m, n; 4, 4) \geq 3m + 3n + \min(m, n) - 13$ by giving an example with no $(4, 4)$ -split. Here is the example for $m = n = 8$, but it can be easily generalized.

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

The construction above can be generalized for bigger p, q . We believe this to be the best possible construction, but have so far been unable to prove it.

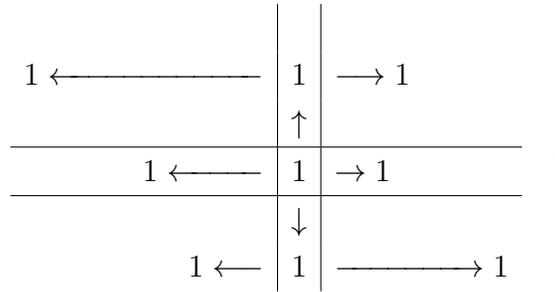
Conjecture 6.1.6 Let n, p be given. Then for $p \geq 4$,

$$\text{NoSplit}(n, n; p, p) = (3p - 5)n + O(p^2)$$

The lower bound is given by a construction like the one given above for $p = 4, q = 4$.

We use the following lemma about a particular 3, 3 split in Chapter 7.

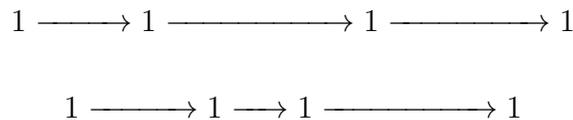
Lemma 6.1.7 Let A be a $\{0, 1\}$ -matrix with $\sigma_1(A) > 4m - 4$. Then A has a special 3,3 split for which the 1's in rows are in the same line. In other words, A has nine 1's that satisfy the following structure,



Proof: The proof is the same as Theorem 6.1.5. Form a matrix A' from A by deleting from A in turn the leftmost and rightmost 1's in each row and then delete topmost and bottommost 1's in each column. We have that $\sigma_1(A) \leq \sigma_1(A') + 4m - 4$. Hence $\sigma_1(A') > 0$. Select a 1 in A' . There is a 1 above and a 1 below. Then for each of these three 1's, there is a 1 to the right and a 1 to the left. This yields the desired structure. ■

We also study a particular 2, 4 split.

Lemma 6.1.8 There exists a $\{0, 1\}$ -matrix A with $\sigma_1(A)$ being $\Omega(m\alpha(m))$, where $\alpha(m)$ is the inverse Ackermann function, and which doesn't have a 2, 4 split on which the topmost 1's lie in the same row and the bottommost 1's lie in the same row as well. In other words, there are no eight 1's with the following structure:



The following observation was suggested to us by Tardos. Observe that if A has a special

split like the one described above, then it must also have the pattern

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

The following proposition of Füredi and Hajnal provides a construction.

Proposition 6.1.9 [FH92] There exists a $\{0, 1\}$ -matrix A with no pattern P for which $\sigma_1(A)$ is $\Omega(m\alpha(m))$, where $\alpha(m)$ denotes the inverse Ackermann function.

Since the Ackermann function grows very quickly the inverse grows very slowly, but it still is bigger than a constant. This means $m\alpha(m)$ is more than linear.

6.2 Patterns and Splits in d-Dimensions

The papers [KM07],[BBM06] consider Theorem 6.1.4 generalized to d -dimensional arrays.

Unfortunately, the notation for talking about these splits becomes admittedly cumbersome as we consider many dimensions, but the ideas remain simple. We use the following notation. Given integers n_1, n_2, \dots, n_d we can consider the positions $\prod_{i=1}^d [n_i]$ in an $n_1 \times n_2 \times \dots \times n_d$ $\{0, 1\}$ -array A . Our main interest is in the case $n_1 = n_2 = \dots = n_d$. Let $p_1, p_2, \dots, p_d \geq 2$ be given. Assume we have d sets of indices $I(j) = \{r_1(j), r_2(j), \dots, r_{p_j-1}(j)\}$ for coordinate j , for $j = 1, 2, \dots, d$. We can form d sets $R_1(j), R_2(j), \dots, R_{p_j}(j)$ with $\cup_{i=1}^{p_j} R_i(j) = [n_j]$ as follows: $R_1(j) = \{1, 2, \dots, r_1(j)\}$, $R_2(j) = \{r_1(j) + 1, r_1(j) + 2, \dots, r_2(j)\}, \dots, R_{p_j}(j) = \{r_{p_j-1}(j) + 1, r_{p_j-1}(j) + 2, \dots, n_j\}$.

We generalize to many dimensions the notion of splits and NoSplit.

Definition 6.2.1 We say A has a p_1, p_2, \dots, p_d **split** if we can choose the sets as above and for each $j \in [d]$ and for each possible choice $t \in [p_j]$ with $R(j) = R_t(j)$, the $\prod_{i=1}^d p_j$ block $A|_{R(1) \times R(2) \times \dots \times R(d)}$ contains at least one 1.

Perhaps an example would be useful to understand this rather cumbersome notation.

Example 6.2.2 The following matrix

$$\begin{array}{c}
 R_1(1) \quad R_2(1) \quad R_3(1) \\
 R_1(2) \left[\begin{array}{ccc|ccc}
 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\
 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
 \hline
 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
 R_2(2) \left[\begin{array}{ccc|ccc}
 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
 \end{array} \right]
 \end{array}$$

in $[8] \times [8]$ has a 3,2 split, where $p_1 = 3$, $p_2 = 2$, and $I(1) = \{3, 6\}$ and $I(2) = \{3\}$, and so $R_1(1) = \{1, 2, 3\}$, $R_2(1) = \{4, 5, 6\}$, $R_3(1) = \{7, 8\}$, $R_1(2) = \{1, 2, 3\}$, $R_2(2) = \{4, 5, 6, 7, 8\}$,

Definition 6.2.3 Let $\text{NoSplit}(n_1, n_2, \dots, n_d; p_1, p_2, \dots, p_d)$ be the maximum number of 1's in $n_1 \times n_2 \times \dots \times n_d$ $\{0, 1\}$ -array that has no p_1, p_2, \dots, p_d split.

The following describes the asymptotic behavior of NoSplit .

Theorem 6.2.4 Klazar and Marcus [KM07], Balogh, Bollabás and Morris [BBM06].

Let k, d be given. Then there exists a constant $c_{k,d}$ so that $\text{NoSplit}(\overbrace{n, \dots, n}^d; \overbrace{k, \dots, k}^d) \leq c_{k,d} n^{d-1}$.

We may also extend the argument in Theorem 6.1.5 to $\overbrace{3, 3, \dots, 3}^{d-1}, q$ splits of d -dimensional arrays. It is surprising that we get exact results here yet the results of [KM07] and [BBM06] do not have a reasonable bound for $\text{NoSplit}(m, n; 4, 4)$.

Theorem 6.2.5 Let B be the $\prod_{i=1}^d [n_i]$ $\{0, 1\}$ -array with 1's in entries whose j coordinate is 1 or 2 for some $j = 1, 2, \dots, d-1$ or whose d th coordinate is 1, 2, \dots or $q-1$. The matrix B has no $\overbrace{3, 3, \dots, 3}^{d-1}, q$ split. Let A be any $\prod_{i=1}^d [n_i]$ $\{0, 1\}$ -array with no $\overbrace{3, 3, \dots, 3}^{d-1}, q$ split. Then $\sigma_1(A) \leq \sigma_1(B)$, and hence for $q = 3$, $\text{NoSplit}(n_1, n_2, \dots, n_d; \overbrace{3, 3, \dots, 3}^{d-1}, q) = \sigma_1(B)$.

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

Now if A' has a 1 in position (y_1, y_2, \dots, y_d) then we note that A has 1's in positions $(y_1, y_2, \dots, y_d(j))$ for $j \in [q]$ where $y_d = y_d(1)$ and $y_d(1) < y_d(2) < \dots < y_d(q)$. It is now straightforward to show that choosing indices $I(1) = \{y_1 - 1, y_1\}$, $I(2) = \{y_2 - 1, y_2\}$, ..., $I(d-1) = \{y_{d-1} - 1, y_d\}$ and $I(d) = \{y_d(1), y_d(2), \dots, y_d(q-1)\}$ yields a $\overbrace{3, 3, \dots, 3}^{d-1}, q$ split.

We can show that $\sigma_1(B) \leq \sigma_1(A) - \sigma_1(A')$, hence if $\sigma_1(A) > \sigma_1(B)$, then A would have the desired split. ■

The following notation is helpful.

Definition 6.2.6 Thinking of the positions in B as elements of $[n]^d$, we let the coordinates of B be x_1, x_2, \dots, x_d and for a position $\mathbf{y} \in [n]^d$ we define $x_i(\mathbf{y})$ to be the value of coordinate x_i in \mathbf{y} . Let $\mathbf{proj}_i(B)$ denote the $[n]^{d-1}$ $(d-1)$ -dimensional $\{0, 1\}$ -array C obtained from B by projecting in the direction e_i (where e_i is the d -dimensional $\{0, 1\}$ -vector with a single 1 in coordinate x_i). For each $\mathbf{y} \in [n]^d$ we form $\mathbf{y}_{\bar{i}}$ in $[n]^{d-1}$ by deleting the i th coordinate of \mathbf{y} and placing a 1 in a position $\mathbf{y}_{\bar{i}}$ of C if and only if there is at least one 1 in a position $\mathbf{z} \in [n]^d$ of B with $\mathbf{z}_{\bar{i}} = \mathbf{y}_{\bar{i}}$.

We are able to use the proof technique of Theorem 6.1.5 to obtain the following result about arrangements of 1's.

Lemma 6.2.7 Let C be an $(n/3) \times (n/3) \times (n/3)$ 3-dimensional $\{0, 1\}$ -array with more than $6(n/3)^2$ 1's. Then there are twenty-seven 1's as follows. There are three values a, b, c for x_1 coordinate and each plane $x_1 = a$, $x_1 = b$ and $x_1 = c$ of C contains nine points. The nine points in each plane form a special 3,3 split as in Lemma 6.1.7 with the central 1's having the same x_2, x_3 coordinates in each of the three planes.

Proof: Form a matrix C' from C by deleting from C in turn the top and bottom 1 in each line in direction x_3 , the top and bottom 1 in each line in the direction x_2 and the top two

1's in each line in the direction x_1 . We obtain

$$\sigma_1(C) \leq \sigma_1(C') + 6(n/3)^2.$$

If C' has a 1 in position \mathbf{y}_1 , we can find twenty-seven 1's yielding a special 3,3,3 split as follows. There are 2 1's of C in positions $\mathbf{y}_2, \mathbf{y}_3$ with $x_1(\mathbf{y}_1) < x_1(\mathbf{y}_2) < x_1(\mathbf{y}_3)$, $x_2(\mathbf{y}_1) = x_2(\mathbf{y}_2) = x_2(\mathbf{y}_3)$ and $x_3(\mathbf{y}_1) = x_3(\mathbf{y}_2) = x_3(\mathbf{y}_3)$. Then there are 1's of C in positions $\mathbf{x}_j, \mathbf{z}_j$ for $j = 1, 2, 3$, where $x_2(\mathbf{x}_j) < x_2(\mathbf{y}_j) < x_2(\mathbf{z}_j)$ and $x_1(\mathbf{x}_j) = x_1(\mathbf{y}_j) = x_1(\mathbf{z}_j)$, $x_3(\mathbf{x}_j) = x_3(\mathbf{y}_j) = x_3(\mathbf{z}_j)$. Now we obtain positions $\mathbf{v}_j, \mathbf{v}_j''$ for $j = 1, 2, 3$ and $\mathbf{v} = \mathbf{x}, \mathbf{y}, \mathbf{z}$ with $x_3(\mathbf{v}'_j) < x_3(\mathbf{v}_j) < x_3(\mathbf{v}''_j)$ and $x_1(\mathbf{v}'_j) = x_1(\mathbf{v}_j) = x_1(\mathbf{v}''_j)$, $x_2(\mathbf{v}'_j) = x_2(\mathbf{v}_j) = x_2(\mathbf{v}''_j)$. In particular there are three planes $x_1 = a$, $x_1 = b$, $x_1 = c$ each with nine 1's and each plane has a special 3,3 split as in Lemma 6.1.7 with the central 1's of each plane (namely $\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3$) having the same x_2, x_3 coordinates and the horizontal direction corresponding to the x_2 direction. ■

Chapter 7

Products

In this chapter we will study another related extremal problem related to Conjecture 1.4.1. Instead of considering how many columns out of all possible columns we can have avoiding a configuration F , we now turn our attention to the problem of finding how many columns we can have out of only a particular, restricted set of columns, while avoiding F . For this we will make heavy use of splits and patterns of Chapter 6. At the end of this section we conclude with forbidden configuration bounds for certain families arising as products.

7.1 Introduction

The following are the maximal 2-rowed simple submatrices of the matrices I, T, I^c . Let

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

Note that $E_2 = T_2$. Let $T'_m = T_m - \mathbf{0}_m$. That is, T'_m is the tower matrix T_m except for the column of 0's.

We now give an analogous definition to forb.

Definition 7.1.1 Let \mathcal{F} be a family of configurations, and let P be an m -rowed $\{0, 1\}$ -matrix. We define

$$\text{MaxChoiceCols}(\mathcal{F}, P) = \max\{\|A\| : A \prec P \text{ and } A \in \text{Avoid}(m, \mathcal{F})\}.$$

Notice that according to this definition, $\text{forb}(m, F) = \text{MaxChoiceCols}(F, K_m)$.

Theorem 7.1.2 $\text{MaxChoiceCols}(E_1 \times E_1, I_{m/2} \times I_{m/2})$ is $\Theta(m^{3/2})$.

Theorem 7.1.3 $\text{MaxChoiceCols}(E_1 \times E_2, I_{m/2} \times T'_{m/2}) \leq 2m$.

Theorem 7.1.4 $\text{MaxChoiceCols}(E_2 \times E_2, T'_{m/2} \times T'_{m/2}) \leq 2m$.

The bound of Theorem 7.1.2 is perhaps unexpected in view of Conjecture 1.4.1 but it is not a counterexample.

The remaining three cases ($E_1 \times E_3$ in $I_{m/2} \times I_{m/2}^c$, $E_2 \times E_3$ in $T'_{m/2} \times I_{m/2}^c$ and $E_3 \times E_3$ in $I_{m/2}^c \times I_{m/2}^c$) essentially follow by taking appropriate $\{0, 1\}$ -complements. The proof of Theorem 7.1.4 is in Section 7.2, the proof of Theorem 7.1.2 is in Section 7.3 and the proof of Theorem 7.1.3 is in Section 7.4. Related results such as

Theorem 7.1.5 $\text{MaxChoiceCols}(E_1 \times E_2 \times E_3, I_{m/3} \times T'_{m/3} \times I_{m/3}^c)$ is $\Theta(m^2)$.

are proved in Section 7.5.

A central idea to studying these products is to encode columns of a p -fold product $A_1 \times A_2 \times \cdots \times A_p$ as entries in a p dimensional $\{0, 1\}$ -array B whose i th coordinate is indexed by the columns of A_i . Then the problem of finding $\text{MaxChoiceCols}(F, A_1 \times A_2 \times \cdots \times A_p)$ can be transformed to the problem of finding the maximum number of 1's in a $\{0, 1\}$ -array that avoids a certain pattern of 1's.

In Chapter 6 we gave results about patterns that we will use in this chapter. The following basic result is proven in Section 7.2.

Proposition 7.1.6 Let p, q, r, u, v, w be given positive integers. Define $x^+ = \max\{0, x\}$.

The configuration given by the product

$$F(u, v, w) = \overbrace{E_1 \times \cdots \times E_1}^u \times \overbrace{E_2 \times \cdots \times E_2}^v \times \overbrace{E_3 \times \cdots \times E_3}^w$$

is contained in the product

$$P_m(p, q, r) = \overbrace{I_m \times \cdots \times I_m}^p \times \overbrace{T'_m \times \cdots \times T'_m}^q \times \overbrace{I_m^c \times \cdots \times I_m^c}^r$$

for some m if and only if

$$2((u - p)^+ + (v - q)^+ + (w - r)^+) \leq (p - u)^+ + (q - v)^+ + (r - w)^+.$$

For example with $u = 2, q = 3$ and the rest being 0, Proposition 7.1.6 yields that $E_1 \times E_1 \not\subseteq T' \times T' \times T'$ and hence $\text{forb}(m, E_1 \times E_1) = \Omega(m^3)$.

Proof: We note that any row from E_i contains $[0 \ 1]$ and also note that $[0 \ 1] \times [0 \ 1] = K_2$. Since none of our product terms I, T', I^c contain K_2 then two rows of F chosen from two different products of the $u+v+w$ 2-rowed products, will necessarily contain K_2 . This implies that if F is contained in the $(p + q + r)$ -fold product

$$\overbrace{I \times \cdots \times I}^p \times \overbrace{T' \times \cdots \times T'}^q \times \overbrace{I^c \times \cdots \times I^c}^r,$$

then each product I, T', I^c has at most 2 rows of F and if it has two rows then they come from the same 2-rowed product term E_i of F . Of the three matrices I, T', I^c , we note that we can find E_1 only in I , E_2 only in T' and E_3 only in I^c . ■

We now consider forbidden families of configurations. It was noted in [AF86] that $\text{forb}(m, \{E_1, E_2, E_3\}) = 2$. Balogh and Bollobás [BB05] have the much more general result that for a given k , there is a constant c_k such that $\text{forb}(m, \{I_k, T'_k, I_k^c\}) = c_k$.

Let $\{E_1, E_2, E_3\} \times \{E_1, E_2, E_3\}$ denote the 6 possible 2-fold products whose terms are chosen from $\{E_1, E_2, E_3\}$. We would like to compute $\text{forb}(m, \{E_1, E_2, E_3\} \times \{E_1, E_2, E_3\})$ but in the interest of a more tractable proof we consider I_2 as a replacement for both

E_1 and E_3 ($I_2^c = I_2$) and T_2' as replacement for E_2 . It is likely (but unknown) that $\text{forb}(m, \{E_1, E_2, E_3\} \times \{E_1, E_2, E_3\})$ is $O(m^{3/2})$. One might ask the relationship of Theorem 7.1.7 to Conjecture 1.4.1. The Conjecture (which applies only for a single forbidden configuration) says that only product constructions are needed for best possible asymptotic bounds, but this case $\{I_2 \times I_2, I_2 \times T_2', T_2' \times T_2'\}$ are simultaneously missing from all 1-fold products but not simultaneously missing from any 2-fold product of I, I^c and T . In particular $I \times I$ avoids $I_2 \times T_2'$ and $T_2' \times T_2'$ but does not avoid $I_2 \times I_2$ (Proposition 7.1.6). Surprisingly there is an $O(m^{3/2})$ construction contained in $I \times I$ and yet avoiding $I_2 \times I_2$ (Theorem 7.1.2) and of course also avoiding $I_2 \times T_2'$, and $T_2' \times T_2'$. The other 2-fold products $I_2 \times T_2'$ (Theorem 7.1.3) and $T_2' \times T_2'$ (Theorem 7.1.4) behave as the conjecture might suggest. We note $\text{forb}(m, \{I_2, T_2'\}) = 2$ and the bounds of Theorem 7.1.2, Theorem 7.1.3 and Theorem 7.1.4 apply.

In Section 7.6 we prove a rather surprising result in which the exponent turns out to be fractional, when in most other cases the exponents were always integral:

Theorem 7.1.7 $\text{forb}(m, \{I_2 \times I_2, I_2 \times T_2', T_2' \times T_2'\})$ is $\Theta(m^{3/2})$.

Let

$$I_2 \times I_2 = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}, \quad I_2 \times T_2' = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix}, \quad T_2' \times T_2' = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix}. \quad (7.1.1)$$

We make use of our Standard Induction (Section 2.1) in a way similar to the way it was used in Section 5.3.3.

7.2 Submatrices of TxT

In this section we'll show how to exploit the results about splits in the context of forbidden configurations.

Proof of Theorem 7.1.4 that $\text{MaxChoiceCols}(E_2 \times E_2, T_{m/2}' \times T_{m/2}') \leq 2m$: Let $F = E_2 \times E_2$. Recall the i th column of T_k' is the column with 1's in rows $1, 2, \dots, i$ and 0's in the

remaining rows. Let A be an m -rowed submatrix of $T'_{m/2} \times T'_{m/2}$. We can create an $m/2 \times m/2$ $\{0, 1\}$ -matrix B from A by placing a 1 in position r, c if A has the column obtained from the r th column of $T'_{m/2}$ placed on top of the c th column of $T'_{m/2}$ namely the column with 1's only in rows $1, 2, \dots, r$ and $m + 1, m + 2, \dots, m + c$. We note that $\|A\|$ is $\sigma_1(B)$.

We claim that $F \prec A$ if and only if B has a 3, 3 split. The only way for a submatrix of $T'_{m/2} \times T'_{m/2}$ to be a row and column permutation of F is to lie in rows $r_1, r_2, m/2 + c_1, m/2 + c_2$ for some choices $2 \leq r_1 < r_2 \leq m/2$ and $2 \leq c_1 < c_2 \leq m/2$ (using the argument of Proposition 7.1.6 and noting that first row of $T'_{m/2}$ is 1's). We have that any two rows of $T'_{m/2}$ (not including the first) have a copy of E_2 . We note that the t th column of $T'_{m/2}$ on rows r_1, r_2 (with $r_1 < r_2$) has

$$r_1 \begin{bmatrix} 0 \\ 0 \end{bmatrix} \text{ for } 1 \leq t < r_1, \quad r_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ for } r_1 \leq t < r_2, \quad \text{and } r_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ for } r_2 \leq t.$$

Assume A has a copy of F in the 4 rows $r_1, r_2, m/2 + c_1, m/2 + c_2$. We discover that the nine columns of F would correspond to nine 1's, one 1 in each of the nine blocks in the 3, 3 split of B given by $I(1) = \{r_1 - 1, r_2 - 1\}$ and $I(2) = \{c_1 - 1, c_2 - 1\}$ (notation from Chapter 6). Similarly a 3, 3 split of B yields a copy of F in A . We now appeal to the bound in Lemma 6.1.5. ■

An immediate generalization is the following.

Lemma 7.2.1 We have $\text{MaxChoiceCols}(\overbrace{E_2 \times E_2 \times \dots \times E_2}^d, \overbrace{T'_{m/d} \times T'_{m/d} \times \dots \times T'_{m/d}}^d)$ is $\Theta(m^{d-1})$.

Proof: We use the d -dimensional generalization of splits Theorem 6.2.4 where the d -fold product $E_2 \times E_2 \times \dots \times E_2$ will correspond to a $\overbrace{3, 3, \dots, 3}^d$ split. We have an exact bound from Theorem 6.2.5 if needed. ■

A further generalization considers the matrix T_k .

Lemma 7.2.2 We have that

$$\text{MaxChoiceCols}(\overbrace{T_k \times T_k \times \cdots \times T_k}^d, \overbrace{T'_{m/d} \times T'_{m/d} \times \cdots \times T'_{m/d}}^d)$$

is equal to

$$\text{NoSplit}(m/d, m/d, \dots, m/d; k+1, k+1, \dots, k+1),$$

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

Proof: We again use the d -dimensional generalization of splits Theorem 6.2.4 where the d -fold product $T_k \times T_k \times \cdots \times T_k$ will correspond to a $\overbrace{k+1, k+1, \dots, k+1}^d$ split. \blacksquare

A rather interesting version of Theorem 7.1.4 and Lemma 7.2.1 that uses the idea of ‘peeling’ from Theorem 6.1.5 is the following.

$$\text{Lemma 7.2.3 } f(E_2 \times E_2, T'_{m/p} \times T'_{m/p} \times \cdots \times T'_{m/p}) \leq 2 \cdot 4^{p-2}m.$$

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

We first handle the case $p = 3$. By Theorem 7.1.4, we have that for $i = 1, 2, 3$, $\sigma_1(\text{proj}_i(B)) < 2m$. In fact if $\sigma_1(\text{proj}_i(B)) > 2m$ then we have a $3, 3$ split in $\text{proj}_i(B)$ and that yields F in A where no rows of F come from the i th term $T_{m/3}$ of the product and 2 rows of F come from another $T'_{m/3}$ and the other 2 rows of F come from remaining part $T'_{m/3}$.

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

$$\sigma_1(B) \leq \sigma_1(B') + \sigma_1(\text{proj}_3(B)) + \sigma_1(\text{proj}_2(B)) + 2\sigma_1(\text{proj}_1(B)) \leq \sigma_1(B') + 8m.$$

Let \mathbf{y}_1 be a 1 of B' . Then, by our construction, there are 2 further 1's of B in positions $\mathbf{y}_2, \mathbf{y}_3$ with $x_1(\mathbf{y}_1) < x_1(\mathbf{y}_2) < x_1(\mathbf{y}_3)$, $x_2(\mathbf{y}_1) = x_2(\mathbf{y}_2) = x_2(\mathbf{y}_3)$ and $x_3(\mathbf{y}_1) = x_3(\mathbf{y}_2) = x_3(\mathbf{y}_3)$. For each \mathbf{y}_j we will have two 1's in positions $\mathbf{y}'_j, \mathbf{y}''_j$ of B where \mathbf{y}'_j agrees with \mathbf{y}_j except in coordinate x_2 where $x_2(\mathbf{y}'_j) < x_2(\mathbf{y}_j)$ and \mathbf{y}''_j agrees with \mathbf{y}_j except in coordinate x_3 where $x_3(\mathbf{y}_j) < x_3(\mathbf{y}''_j)$. Then these nine 1's in B correspond to a copy of F in A as follows.

We choose two coordinates a, b from x_1 so that when we consider the columns of A corresponding to \mathbf{y}_1 (and $\mathbf{y}'_1, \mathbf{y}''_1$ respectively), \mathbf{y}_2 (and $\mathbf{y}'_2, \mathbf{y}''_2$ resp.), \mathbf{y}_3 (and $\mathbf{y}'_3, \mathbf{y}''_3$ resp.) we have

$$\begin{array}{c} a \\ b \end{array} \begin{array}{ccc} \mathbf{y}_1 & \mathbf{y}_2 & \mathbf{y}_3 \\ \begin{bmatrix} 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \end{bmatrix} \end{array}.$$

For the next part note that column t of $T'_{m/3}$ has a 0 in row r if and only if $t < r$. Noting that $x_2(\mathbf{y}'_j) < x_2(\mathbf{y}_j) = x_2(\mathbf{y}_j)$ and that $x_3(\mathbf{y}'_j) = x_3(\mathbf{y}_j) < x_3(\mathbf{y}''_j)$, we choose a value c for x_2 and a value d for x_3 and so that in A , the columns corresponding to the 1's $\mathbf{y}_j, \mathbf{y}'_j, \mathbf{y}''_j$ have

$$\begin{array}{ccc} m/3 + c & \mathbf{y}_j & \mathbf{y}'_j & \mathbf{y}''_j \\ 2(m/3) + d & \begin{bmatrix} 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \end{bmatrix} \end{array}.$$

This yields a copy of F in A in rows a, b, c, d , a contradiction. We deduce that $\sigma_1(B') = 0$ and hence $\sigma_1(B) \leq 8m$, concluding the proof for $p = 3$.

For $p \geq 4$, we proceed in a similar fashion. By induction on p , $\sigma_1(\text{proj}_i(B))$ is at most $2 \cdot 4^{p-3}m$. We form a matrix B' from B by deleting from B in turn the top 1 in each line in the direction x_4 and then deleting the bottom 1 in each line in the direction x_3 and then deleting the top 1 in each line in the direction x_2 finally deleting the bottom 1 in each line in the direction x_1 . We have

$$\sigma_1(B) \leq \sigma_1(B') + \sum_{i=1}^4 \sigma_1(\text{proj}_i(B)) \leq \sigma_1(B') + 4 \cdot 2 \cdot 4^{p-3}m.$$

Let \mathbf{y}_1 be an 1 of B' . Then, by our construction, there are 2 further 1's of B in positions $\mathbf{y}_2, \mathbf{y}_3$ with $x_1(\mathbf{y}_2) < x_1(\mathbf{y}_1)$ and $x_i(\mathbf{y}_2) = x_i(\mathbf{y}_1)$ for $i \neq 1$, and $x_2(\mathbf{y}_1) < x_2(\mathbf{y}_3)$ and $x_i(\mathbf{y}_1) = x_i(\mathbf{y}_3)$ for $i \neq 2$. For each \mathbf{y}_j we will have two 1's in positions $\mathbf{y}'_j, \mathbf{y}''_j$ of B where \mathbf{y}'_j agrees with \mathbf{y}_j except in coordinate x_3 where $x_3(\mathbf{y}'_j) < x_3(\mathbf{y}_j)$ and \mathbf{y}''_j agrees with \mathbf{y}_j except in coordinate x_4 where $x_4(\mathbf{y}_j) < x_4(\mathbf{y}''_j)$. Then these nine 1's in B correspond to a copy of F . In particular

we can choose coordinates values a, b so that in A the columns contain

$$\begin{array}{c} a \\ b + (m/4) \end{array} \begin{array}{c} \mathbf{y}_2 \\ \mathbf{y}_1 \end{array} \begin{array}{c} \mathbf{y}_1 \\ \mathbf{y}_3 \end{array} \begin{array}{c} \mathbf{y}_3 \\ \mathbf{y}_2 \end{array} \begin{array}{c} \left[\begin{array}{c} 0 \\ 0 \end{array} \right] \\ \left[\begin{array}{c} 1 \\ 0 \end{array} \right] \\ \left[\begin{array}{c} 1 \\ 1 \end{array} \right] \end{array}.$$

As above we can choose values c, d and obtain the copy of F from the nine columns of A given by the nine 1's of B , a contradiction.

We deduce that $\sigma_1(B') = 0$ and hence $\sigma_1(B) \leq 2 \cdot 4^{p-2}m$ ■

Some growth in the bound with respect to p is to be expected since $\text{forb}(m, E_2 \times E_2)$ is $\Theta(m^3)$. Note that the proof technique suggests generalizations of Theorem 6.1.4 and Theorem 6.2.4 where we forbid a larger class of arrangements of 1's in a $(m/d) \times (m/d) \times \dots \times (m/d)$ d -dimensional $\{0, 1\}$ -array and are able to conclude that the array has only a linear number of 1's. Our geometric argument above has no obvious generalization which would allow us to generalize Theorem 7.2.2. A first step would be a simple geometric proof of the linear bound for $\text{NoSplit}(m/2, m/2; 4, 4)$ but Lemma 6.1.8 suggests other difficulties.

7.3 Submatrices of IxI

Proof of Theorem 7.1.2 that $\text{MaxChoiceCols}(E_1 \times E_1, I_{m/2} \times I_{m/2})$ is $\Theta(m^{3/2})$. Let $F = E_1 \times E_1$. Let A be a submatrix of $I_{m/2} \times I_{m/2}$ that has no configuration $E_1 \times E_1$. We consider A as an $(m/2) \times (m/2)$ $\{0, 1\}$ -matrix B whose rows are indexed by the columns of $I_{m/2}$ and whose columns are indexed by the columns of $I_{m/2}$. Then $\|A\| = \sigma_1(B)$. Now 4 rows of A contain F if and only if 2 rows of A chosen from the first $m/2$ contain the first two rows of F (and so correspond to one copy of E_1) and 2 rows of A chosen from the second $m/2$ correspond to the second two rows of F (and the other copy of E_1). Now the nine columns of A containing F corresponds to B having nine 1's as follows: a 2×2 submatrix of 1's and at least one more 1 in each row of the 2×2 submatrix and at least one more 1 in each column of the submatrix and at least one more 1 in neither of the two chosen rows or two chosen columns.

To see this, consider the 2 chosen rows of A from the first $m/2$ to be i, j and the two other chosen rows are $k + m/2, \ell + m/2$. Now $I_{m/2}$ restricted to rows i, j has $\begin{array}{c} i \\ j \end{array} \left[\begin{array}{c} 1 \\ 0 \end{array} \right]$ in column

i , has $\begin{matrix} i \\ j \end{matrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ in column j and $\begin{matrix} i \\ j \end{matrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ in all columns not equal to i, j . Thus we get a copy of $I_2 \times I_2$ in rows $i, j, k + m/2, \ell + m/2$ of A from 4 1's in B in rows i, j and columns k, ℓ . Similar observations yield the other 5 columns of F from the 5 1's of B as described.

We initially process B by deleting any row or column with at most two 1's (and hence up to $2m$ 1's) repeating the deletion process if necessary so that the resulting matrix \overline{B} has row and column sums at least 3. We note that $\sigma_1(B) \leq \sigma_1(\overline{B}) + 2m$. We now appeal to Kóvari, Sós and Turán [KST54] for a solution of Zarankiewicz' problem and deduce that if the number of 1's in \overline{B} is $\Omega(m^{3/2})$, then \overline{B} has 2×2 block of 1's and then B has the configuration of nine 1's yielding $E_1 \times E_1$ in A . Moreover from [KST54] we can point out that a construction using projective planes establishes $\text{MaxChoiceCols}(E_1 \times E_1, I_{m/2} \times I_{m/2})$ is $\Omega(m^{3/2})$. ■

Problem 7.3.1 Determine $\text{MaxChoiceCols}(E_1 \times E_1, I_{m/3} \times I_{m/3} \times I_{m/3})$.

We note that $\text{MaxChoiceCols}(E_1 \times E_1, I_{m/3} \times I_{m/3} \times I_{m/3})$ is $\Omega(m^{3/2})$ by Theorem 7.1.2 and is also $O(m^{7/4})$.

Problem 7.3.2 Determine $\text{MaxChoiceCols}(E_1 \times E_1 \times E_1, I_{m/3} \times I_{m/3} \times I_{m/3})$.

The core of this problem would be determining the maximum number of 1's in a 3-dimensional $(m/3) \times (m/3) \times (m/3)$ $\{0, 1\}$ -matrix which has no $2 \times 2 \times 2$ submatrix of eight 1's. Erdős [Erd64] has obtained a bound $O(m^{11/4})$ for this but no matching construction. Note the sharp contrast with results such as Theorem 7.1.4, Lemma 7.2.3, Lemma 7.2.1.

7.4 Submatrices of $I \times T$

Proof of Theorem 7.1.3 that $\text{MaxChoiceCols}(E_1 \times E_2, I_{m/2} \times T'_{m/2}) \leq 2m$. Let $F = E_1 \times E_2$. Let A be an m -rowed submatrix of $I_{m/2} \times T'_{m/2}$ with no F . We consider A as an $(m/2) \times (m/2)$ $\{0, 1\}$ -matrix B whose rows are indexed by the columns of $I_{m/2}$ and whose columns are indexed by the columns of $T'_{m/2}$ in the usual order. We note that $\|A\|$ is $\sigma_1(B)$.

Now A contains F if and only if there are three rows and in each row there are three 1's such that we can divide the nine 1's as a sort of "1,3 split", where the three leftmost 1's in each row are entirely to the left of the middle three 1's in each row which are entirely to the left of the rightmost three 1's in each row. By Lemma 6.1.7 we have the desired result.

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

Lemma 7.4.1 $\text{MaxChoiceCols}(E_1 \times E_2, I_{m/3} \times T'_{m/3} \times T'_{m/3})$ is $\Theta(m^2)$.

Proof: Let $F = E_1 \times E_2$. Let A be an m -rowed submatrix $I_{m/3} \times T'_{m/3} \times T'_{m/3}$ with no F . We consider A as an $(m/3) \times (m/3) \times (m/3)$ 3-dimensional $\{0, 1\}$ -array B with coordinate x_1 indexed by the columns of $I_{m/3}$ and with coordinates x_2, x_3 indexed by columns of $T'_{m/3}$ in the usual order. We note that $\|A\|$ is $\sigma_1(B)$.

By Theorem 7.1.3 we have $|\text{proj}_2(B)|$ and $|\text{proj}_3(B)|$ being $O(m)$, and of course $|\text{proj}_1(B)|$ is $O(m^2)$. Form a matrix B' from B by deleting from B in turn the top 1 in each line in direction x_3 , the bottom 1 in each line in the direction x_2 and the top two 1's in each line in the direction x_1 . We have that the number of deleted 1's is at most $O(m^2)$.

Now assume B' has a 1 in position \mathbf{y}_1 . Then, by our construction, there are 2 further 1's of B in positions $\mathbf{y}_2, \mathbf{y}_3$ with $x_1(\mathbf{y}_1) < x_1(\mathbf{y}_2) < x_1(\mathbf{y}_3)$, $x_2(\mathbf{y}_1) = x_2(\mathbf{y}_2) = x_2(\mathbf{y}_3)$ and $x_3(\mathbf{y}_1) = x_3(\mathbf{y}_2) = x_3(\mathbf{y}_3)$. For each \mathbf{y}_j we will have two 1's in positions $\mathbf{y}'_j, \mathbf{y}''_j$ of B where \mathbf{y}'_j agrees with \mathbf{y}_j except in coordinate x_2 where $x_2(\mathbf{y}'_j) < x_2(\mathbf{y}_j)$ and \mathbf{y}''_j agrees with \mathbf{y}_j except in coordinate x_3 where $x_3(\mathbf{y}_j) < x_3(\mathbf{y}''_j)$. Then these nine 1's in B correspond to a copy of $F = E_1 \times E_2$. We choose r_1, r_2 so that in the columns of $I_{m/3}$ indexed by $x_1(\mathbf{y}_1), x_1(\mathbf{y}_2), x_1(\mathbf{y}_3)$ we find E_1 :
$$\begin{matrix} r_1 \\ r_2 \end{matrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$
. We choose two additional rows b, c following the discussion for $p = 3$ in Lemma 7.2.3 when looking for E_2 . Thus $\sigma_1(B') = 0$ and then $\sigma_1(B)$ is $O(m^2)$. ■

Lemma 7.4.2 $\text{MaxChoiceCols}(E_1 \times E_2 \times E_2, I_{m/3} \times T'_{m/3} \times T'_{m/3})$ is $\Theta(m^2)$.

Proof: Let A be an m -rowed submatrix $I_{m/3} \times T'_{m/3} \times T'_{m/3}$ with no $E_1 \times E_2 \times E_2$. As above, we translate A into the 3-dimensional array B with $\|A\| = \sigma_1(B)$.

Now by Lemma 6.2.7, if $\sigma_1(B) > 6(m/3)^2$ there will be twenty-seven 1's in B as described and this will yield a copy of $E_1 \times E_2 \times E_2$ in A . Thus $\sigma_1(B) \leq 6(m/3)^2$. ■

Using an analogous argument one obtains

Lemma 7.4.3 $\text{MaxChoiceCols}(E_1 \times \overbrace{E_2 \times \cdots \times E_2}^{p-1}, I_{m/p} \times \overbrace{T'_{m/p} \times \cdots \times T'_{m/p}}^{p-1})$ is $\Theta(m^{p-1})$.

■

We note some difficulty may arise when considering T_k instead of $E_2 = T_2$.

Theorem 7.4.4 $\text{MaxChoiceCols}(T_3 \times I_2, T \times I)$ is $\Omega(m\alpha(m))$, where $\alpha(m)$ denotes the inverse Ackermann function.

Proof: Form our usual array B and observe that B can't have the structure described in Lemma 6.1.8. ■

Initially we had thought $\text{MaxChoiceCols}(T_3 \times I_2, T \times I)$ would be linear, but it isn't.

7.5 Submatrices of $I \times T \times I^c$

Proof of Theorem 7.1.5 that $f(E_1 \times E_2 \times E_3, I_{m/3} \times T'_{m/3} \times I^c_{m/3})$ is $\Theta(m^2)$. Let $F = E_1 \times E_2 \times E_3$. Let A be an m -rowed submatrix of $I_{m/3} \times T'_{m/3} \times I^c_{m/3}$ with no F . We consider A as an $(m/3) \times (m/3) \times (m/3)$ 3-dimensional $\{0, 1\}$ -array B as follows. Let x_1, x_2, x_3 be the three coordinate directions in B . The entries in coordinate direction x_1 are indexed by the columns of $I_{m/3}$, the entries in coordinate direction x_2 are indexed by the columns of $T'_{m/3}$ in that order and the entries in coordinate direction x_3 are indexed by the columns of $I^c_{m/3}$. As before $\|A\|$ is $\sigma_1(B)$.

By Lemma 6.2.7, we know that if $\sigma_1(B) > 6(m/3)^2$, then there is the configuration of twenty-seven points as described. Then these twenty-seven 1's in B correspond to a special 3, 3, 3 split that yield copy of F in A following our usual arguments. This contradiction implies that $\sigma_1(B) \leq 6(m/3)^2$. ■

7.6 Fractional Exponent Bound for a Family of Configurations

Proof of Theorem 7.1.7 that $\text{forb}(m, \{I_2 \times I_2, I_2 \times T'_2, T'_2 \times T'_2\})$ is $\Theta(m^{3/2})$.

By Theorem 7.1.2 there exists a matrix $A \prec I_{m/2} \times I_{m/2}$ with $\|A\|$ being $\Theta(m^{3/2})$ and no $I_2 \times I_2$. Because of Proposition 7.1.6, neither $I_2 \times T'_2 \prec A$ nor $T'_2 \times T'_2 \prec A$.

Let $A \in \text{Avoid}(m, \{I_2 \times I_2, I_2 \times T'_2, T'_2 \times T'_2\})$. We begin using our usual standard decomposition (2.1.1) on row r

$$A = \begin{matrix} r \rightarrow \\ \left[\begin{array}{cccc} 00 \cdots 0 & 11 \cdots 1 \\ B_r & C_r & C_r & D_r \end{array} \right] \end{matrix}$$

We would be done by induction if we could show that $\|C_r\| \leq 36m^{1/2}$ for some r , so we may assume $\|C_r\| \geq 36m^{1/2}$ for all r . Our proof will show that we can associate matrix C_r with a certain set of rows $M(r)$ (to be defined later), where $|M(r)| \geq \|C_r\|/4$. Consider a given choice r and set $M(r)$ of rows with $|M(r)| \geq 9m^{1/2}$. Then for $t = 9m^{1/2}$ choices $r_1, r_2, \dots, r_t \in M(r)$ we will show that

$$|M(r_i) \cap M(r_j)| \leq 9 \tag{7.6.1}$$

and so we obtain disjoint sets

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

which yields that $M(r_1) \cup M(r_2) \cup M(r_3) \cdots$ is of size

$$9m^{1/2} + (9m^{1/2} - 9) + (9m^{1/2} - 18) + \dots > m,$$

a contradiction given that we only have m rows.

We can find the inductive children of $\{I_2 \times I_2, I_2 \times T'_2, T'_2 \times T'_2\}$ easily:

$$F_4 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad F_5 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}. \tag{7.6.2}$$

This follows since if C_r has F_4 , then A has $T'_2 \times T'_2$, and if C_r has F_5 , then A has $T'_2 \times I_2$, both forbidden configurations. Because of F_5 we know C_r is a laminar matrix (as in

$$Q_8 = \begin{matrix} & \begin{matrix} no \\ \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \end{matrix} & \begin{matrix} no \\ \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} \end{matrix} & \begin{matrix} no \\ \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} \end{matrix} & \begin{matrix} no \\ \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix} \end{matrix} & \begin{matrix} no \\ \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} \end{matrix} & , & Q_9 = \begin{matrix} no \\ \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \end{matrix} & \begin{matrix} no \\ \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} \end{matrix} & \begin{matrix} no \\ \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix} \end{matrix} & \begin{matrix} no \\ \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix} \end{matrix} & \begin{matrix} no \\ \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} \end{matrix} \end{matrix}$$

$$Q_{10} = \begin{matrix} & \begin{matrix} no \\ \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \end{matrix} & \begin{matrix} no \\ \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} \end{matrix} & \begin{matrix} no \\ \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix} \end{matrix} & \begin{matrix} no \\ \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix} \end{matrix} & \begin{matrix} no \\ \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \end{matrix} \end{matrix}$$

As in the proof given in Lemma 5.3.2, choose $L(r)$ to be a minimal subset of rows of C_r such that $C_r|_{L(r)}$ is a simple matrix. Note that $C_r|_{L(r)}$ also has the property that the columns of sum at least 2 are disjoint as sets.

Now consider the family of sets \mathcal{C} corresponding to the columns of C_r . We can have the empty set as well as sets of size 1 and some disjoint family of sets of size at least 2. Let $V \in \mathcal{C}$ be a set of size at least 2 and let U_1, U_2, \dots, U_t be the sets of \mathcal{C} of size 1 contained in V . Now, if $t < |V|$ and $|V| \geq 3$, we could delete a row of $C_r|_{L(r)}$ and preserve simplicity, a contradiction. As an example, consider deleting the last row of G below. For $|V| = 2$, we can have $t = 1 < |V|$ and still have no row to delete, as shown for example in H below.

$$G = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad H = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

Our choice of $L(r)$ ensures that no row of $C_r|_{L(r)}$ can be deleted (while preserving simplicity) and hence $C_r|_{L(r)}$ contains the column of 0's ($\mathbf{0}_{|L(r)|}$) and at least half the columns of $I_{|L(r)|}$, as well as possibly up to $|L(r)|/2$ disjoint columns of sum 2 or more. Thus we may estimate $|L(r)| \geq \|C_r\|/2$.

Thus the number of sets of size 1 in \mathcal{C} is at least $|L(r)|/2$. We choose $M(r)$ to correspond to the sets of size 1, namely $C_r|_{M(r)}$ contains $I_{|M(r)|}$ as well as $\mathbf{0}_{|M(r)|}$. Note that $C_r|_{M(r)}$ need not be a simple matrix and may in fact have up to 2 copies of each column from $I_{|M(r)|}$ (imagine deleting the second row of H). We obtain $|M(r)| \geq \|C_r\|/4 \geq 9m^{1/2}$.

In what follows we analyze closely each of the 11 cases above to deduce the structure of $C_r|_{L(r)}$ in order to prove (7.6.1).

It's possible to determine What Is Missing in $C_r|_{L(r)}$ on the triple of rows $i, j, k \in L(r)$ by considering What Is Missing on the quadruple of rows r, i, j, k , as we did for F_7 in Section 5.3. We obtain a contradiction if we find a copy of “ K_2 ” in What Is Missing, namely if on the triple i, j, k there is a pair of rows i, j with all 4 columns of K_2 appearing as follows:

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

where $a, b, c, d \in \{0, 1\}$. Perhaps other columns are missing on rows i, j, k . Note that we could delete row k from $C_r|_{L(r)}$ and preserve simplicity, contradicting our choice of $L(r)$. The reason for this is that on the three rows i, j, k , the columns present would possibly 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 $C_r|_{L(r)}$ has no repeated columns.

We note that Q_0, Q_1, Q_2 each have 3 rows, each pair of which has a K_2 in What Is Missing (rows 1, 2, 3 for Q_0 , rows 2, 3, 4 for Q_1 , rows 2, 3, 4 for Q_2) and Q_5, Q_6 have two disjoint pairs of rows, each of which has K_2 (rows 1, 2 and rows 3, 4 for Q_5 and rows 1, 3 and rows 2, 4 for Q_6) so deleting any row of the quadruple will leave a K_2 leaving a row to delete from C_r and so by our choice of $L(r)$ we cannot have the cases represented by Q_0, Q_1, Q_2, Q_5, Q_6 in quadruples of rows consisting of r plus a triple of rows from $L(r)$.

For the remaining cases, to avoid leaving a K_2 after deleting a row we find that we must have deleted particular rows from each quadruple. Note that Q_3 and Q_8 have K_2 on rows 3, 4 and so if present, to avoid leaving a copy of K_2 row r must either correspond to row 3 or to row 4. The cases Q_4, Q_7, Q_9, Q_{10} all have K_2 on rows 2, 4 and rows 3, 4 and so row r must be row 4 of such a quadruple. We have used P_i to denote the condition on a triple of rows arising from the quadruple Q_i in these ways.

$$P_3 : \begin{array}{cccc} \text{no} & \text{no} & \text{no} & \text{no} \\ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} & \text{or} & \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \end{array}$$

$$P_4 \text{ or } P_9 : \begin{array}{ccc} \text{no} & \text{no} & \text{no} \\ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \end{array}$$

$$P_8 : \begin{array}{cccc} \text{no} & \text{no} & \text{no} & \text{no} \\ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} & \text{or} & \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \end{array}$$

$$P_7 \text{ or } P_{10} : \begin{array}{ccc} \text{no} & \text{no} & \text{no} \\ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \end{array}$$

Note that the presence of $I_{|M(r)|}$ means that on each triple of rows in $M(r)$ we have all three columns of sum 1 present and so cannot have P_3 (arising from Q_3 with row r being either the third or fourth row) or the second of the two cases for P_8 (arising from Q_8 with row r being the third row) on triples of rows in $M(r)$ since each forbids a column of sum 1. This careful detail is used in what follows.

We now consider what is possible in the full width of A on the rows of $M(r)$. By considering the corresponding Q_i 's and the row of Q_i that corresponds to r , we find that under the 0's in row r , any triple of rows $[B_r C_r]_{|M(r)|}$ must have two columns of sum 2 absent and hence $[B_r C_r]_{|M(r)|}$ cannot have the configuration F_5 . Thus the columns form a laminar family on rows $M(r)$, namely, considered as subsets of $M(r)$, any two sets are either disjoint or one is contained in the other. One has to consider Q_4, Q_8, Q_9, Q_{10} and then in each case the row corresponding to row r must be the fourth row of each such Q_i (else a K_2 remains on the triple). Similarly under the 1's in row r , in $[C_r D_r]_{|M(r)|}$ we cannot have the configuration

F_5 and so the columns also form a laminar family on $M(r)$.

Had we used row r as the third or fourth row of Q_3 then $[C_r D_r]|_{M(r)}$ might have F_5 and had we used row r as the third row of Q_8 then $[B_r C_r]|_{M(r)}$ might have the configuration F_5 . Fortunately these cases are not possible.

Now consider $r_j \in M(r)$ and standard decomposition based on row r_j . The columns of C_{r_j} , must correspond to columns of A which appear with a 1 and with a 0 in row r_j and are the same elsewhere and hence this is also true when restricted to the rows $M(r) \cup r$. We have pairs of columns as follows:

$$M(r) \setminus r_j \left\{ \begin{array}{l} r_j \begin{bmatrix} 0 & 1 \\ \alpha & \alpha \end{bmatrix} \\ r \begin{bmatrix} a & a \end{bmatrix} \end{array} \right.$$

We can show that there are at most three possible choices for pairs of columns and hence at most three choices for α . In fact consider $[B_r C_r]$ and $\beta \neq \gamma$ be two nonzero choices for α .

$$M(r) \setminus r_j \left\{ \begin{array}{l} r_j \begin{bmatrix} 0 & 1 & 0 & 1 \\ \beta & \beta & \gamma & \gamma \end{bmatrix} \\ r \begin{bmatrix} 0 & 0 & 0 & 0 \end{bmatrix} \end{array} \right.$$

The columns of A with 0's in row r , when restricted to the rows $M(r)$, must form a laminar family. Columns 2 and 4 have 1's in common on row r_j and we deduce, without loss of generality by considering columns 2 and 4, that $\beta \leq \gamma$. Now considering columns 2,3 and using $\mathbf{0} \neq \beta \leq \gamma$, we violate the laminar property, a contradiction to $\beta \neq \gamma$. Thus there is at most one nonzero choice for α in $[B_r C_r]$ and at most one nonzero choice for α in $[C_r D_r]$.

Given that there are at most three choices for columns of C_{r_j} restricted to the rows $M(r) \setminus r_j$, we can deduce that $|L(r_j) \cap M(r)| \leq 2$ (we would be able to delete all but two rows of $M(r) \setminus r_j$ from $C_{r_j}|_{L(r_j)}$ without affecting simplicity). So let $M'(r_j) = M(r_j) \setminus M(r)$ and note that $|M(r_j)| - 2 \leq |M'(r_j)| \leq |M(r_j)|$.

From our previous observations we have that the columns of A indexed by the rows of $M'(r_j)$ consist of the column of 0's, the columns of the identity matrix, and (possibly) columns of sum at least 2 which are disjoint when considered as sets. We use the following

notation.

$$C_{r_j} = \begin{matrix} M(r) \setminus r_j \{ \\ M'(r_j) \{ \\ \\ \end{matrix} \begin{bmatrix} X \\ Y \\ \vdots \end{bmatrix}, \quad A = \begin{matrix} r_j \rightarrow \\ M(r) \setminus r_j \{ \\ M'(r_j) \{ \\ \\ \end{matrix} \begin{bmatrix} 0 \cdots 0 & 0 \cdots 0 & 1 \cdots 1 & 1 \cdots 1 \\ & X & X & \\ & Y & Y & \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}.$$

By our previous observations for $C_r|_{L(r)}$ and $C_r|_{M(r)}$ we know that columns of Y are either columns of 0's, columns of sum 1 and possibly columns of sum 2 (disjoint as sets). By reducing to $M'(r_j)$ and deleting the possible row of overlap with $M(r)$, we may assume that Y contains $I_{|M'(r_j)|}$. We can show that two different columns of Y of sum 1 (i.e. with 1's in different rows) on rows $M'(r_j)$ cannot lie under two identical nonzero columns (of X) on rows $M(r) \setminus r_j$, else we have $I_2 \times T'_2$ in A as follows. Let i, k correspond to the rows of $M'(r_j)$ which contain the 1's of the two selected columns of Y and let $a \in M(r) \setminus r_j$ be a row containing 1's in the repeated nonzero column of X . This gives $I_2 \times T'_2$:

$$\begin{matrix} r_j \\ a \\ i \\ k \end{matrix} \begin{bmatrix} 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}$$

Recall that there are at most 3 different columns in X and Y has $I_{|M'(r_j)|}$. Let $M''(r_j) \subseteq M'(r_j)$ denote the rows corresponding to the 1's in the columns of the identity $I_{|M'(r_j)|}$ which lie under columns of 0's in rows of $M(r) \setminus r_j$. We have that $|M'(r_j)| - 2 \leq |M''(r_j)| \leq |M'(r_j)|$ else we create $I_2 \times T'_2$ as above since there are at most two different nonzero columns of X . Thus $|M(r_j)| - 4 \leq |M''(r_j)| \leq |M(r_j)|$.

Now consider two rows $r_p, r_q \in M(r)$. Let $i, j \in M''(r_p) \cap M''(r_q)$. From the standard decomposition using row r_p , we can find two columns with 1's in row r_p , 0's in row r_q (and also 0's in all other rows of $M(r)$) and I_2 on rows i, j . Similarly we can find two columns with 1's in row r_q , 0's in row r_p (and all other rows of $M(r)$) and I_2 on rows i, j . Then we can find a copy of $I_2 \times I_2$:

$$\begin{matrix} r_p \\ r_q \\ i \\ j \end{matrix} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}$$

We deduce that $|M''(r_p) \cap M''(r_q)| \leq 1$, which in turn establishes (7.6.1) that $|M(r_i) \cap$

$M(r_j)| \leq 9$, completing our proof. ■

7.7 All Pairs of Columns

We now establish an asymptotic bound that is a generalization of the exact bound of the bound proved in Section 4.2.

Consider all pairs of possible $\ell \times 1$ columns.

$$\mathcal{F}_\ell := \{[\alpha \beta] : \alpha, \beta \in \{0, 1\}^\ell, \alpha \neq \beta\}.$$

For example, \mathcal{F}_2 is the set of matrices considered in Section 4.2, where we proved $\text{forb}(m, \mathcal{F}_2 \times \mathcal{F}_2) = m + 3$.

$$\mathcal{F}_2 = \left\{ \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \right\}$$

and

$$\mathcal{F}_3 = \left\{ \begin{bmatrix} 11 \\ 11 \\ 10 \end{bmatrix}, \begin{bmatrix} 11 \\ 10 \\ 10 \end{bmatrix}, \begin{bmatrix} 10 \\ 10 \\ 10 \end{bmatrix}, \begin{bmatrix} 11 \\ 10 \\ 01 \end{bmatrix}, \begin{bmatrix} 11 \\ 10 \\ 00 \end{bmatrix}, \begin{bmatrix} 10 \\ 10 \\ 01 \end{bmatrix}, \begin{bmatrix} 10 \\ 10 \\ 00 \end{bmatrix}, \begin{bmatrix} 10 \\ 01 \\ 00 \end{bmatrix}, \begin{bmatrix} 10 \\ 00 \\ 00 \end{bmatrix} \right\}.$$

Theorem 4.2.1 has an asymptotic version.

Observe that $\text{forb}(m, \mathcal{F}_\ell) = 1$ for $m \geq \ell$ (and 2^m for $m < \ell$).

Theorem 7.7.1 Let \mathcal{T} be a family of forbidden configurations with $\text{forb}(m, \mathcal{T})$ being $O(m^k)$. Then for any ℓ , we have $\text{forb}(m, \mathcal{F}_\ell \times \mathcal{T})$ is $O(m^{k+1})$.

Proof: We consider a matrix A with no configuration in $\mathcal{T} \times \mathcal{F}_\ell$. We do an analogous operation to the standard decomposition (2.1.1) on ℓ rows. For any two columns α, β with $\alpha \neq \beta$ we let $A_{\alpha, \beta}$ denote the $(m - \ell)$ -rowed matrix formed as the set of columns on $A|_{\{\ell+1, \ell+2, \dots, m\}}$ which appear under both α and β . For example, for $\ell = 2$, we would consider

the following representative

$$A = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ A_{\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}} & A_{\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}} & A_{\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}} & A_{\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}} \end{bmatrix}$$

where A_α are all simple matrices for $\alpha \in \{\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}, \begin{smallmatrix} 1 \\ 0 \end{smallmatrix}, \begin{smallmatrix} 0 \\ 1 \end{smallmatrix}, \begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\}$. In such matrix, $A_{\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}, \begin{smallmatrix} 1 \\ 0 \end{smallmatrix}}$ would be matrix formed by the columns that are in both $A_{\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}}$ and $A_{\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}}$.

We deduce that $A_{\alpha, \beta}$ is a simple matrix with no configuration in \mathcal{T} since if $F \in \mathcal{T}$ is in $A|_{\alpha, \beta}$ then A has $[\alpha \beta] \times F \in \mathcal{F}_\ell \times \mathcal{T}$. Hence at $A|_{\alpha, \beta}$ has at most $\text{forb}(m - \ell, \mathcal{T})$ columns.

There are at most a constant number of choices for pairs α and β and so after deleting at most $O(m^k)$ columns from A to produce A' we can ensure $A|_{\{\ell+1, \ell+2, \dots, m\}}$ is simple and hence, by induction, has at most $\text{forb}(m - \ell, \mathcal{T} \times \mathcal{F}_\ell)$ columns. This yields the result. \blacksquare

The $O(m)$ bound of Theorem 4.2.1 would follow from taking $\mathcal{T} = \mathcal{F}_2$, although to get the exact bound we use the proof given in that section.

Chapter 8

Conclusions

8.1 Open Problems

8.1.1 Monotonicity

It is quite reasonable to conjecture that the function $\text{forb}(m, F)$ is increasing in the first variable m , as we know it is in the second variable F (Remark 1.3.21).

Problem 8.1.1 Let F be a k -rowed $(0,1)$ -matrix and let $m \geq 1$. Is it true that $\text{forb}(m, F) \leq \text{forb}(m + 1, F)$?

We can answer positively for single forbidden configurations with certain properties. We know this is false for families of forbidden configurations.

Proposition 8.1.2 Let F be a configuration that doesn't have either a row of 0's, a row of 1's or a repeated row. Then for all m , $\text{forb}(m, F) \leq \text{forb}(m + 1, F)$.

Proof: Let $A \in \text{ext}(m, F)$. Then consider the $(m + 1)$ -rowed matrix

$$A' = \begin{cases} A \times [0] & \text{if } F \text{ doesn't have a row of 0's,} \\ A \times [1] & \text{if } F \text{ doesn't have a row of 1's,} \\ \begin{bmatrix} A \\ r \end{bmatrix} & \text{if } F \text{ doesn't have a repeated row,} \end{cases}$$

where r is any row of A . Then clearly $A' \in \text{Avoid}(m+1, F)$ and $\|A'\| = \|A\|$, and $F \not\prec A'$.
 ■

Note that for families of configuration this is not always the case. Here is an amusing case.

Theorem 8.1.3

$$\text{forb}(m, \left\{ \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \right\}) = \begin{cases} 2 & \text{for } m = 1 \text{ or } m \geq 7 \\ 6 & \text{for } m = 3, 4 \\ 4 & \text{for } m = 2, 5, 6 \end{cases}$$

Proof: The result is easy for $m = 1, 2$. First consider pairs of columns. Let F_{abcd} denote the $(a+b+c+d) \times 2$ matrix of a rows $[00]$, b rows $[10]$, c rows $[01]$ and d rows $[11]$. Consider two columns of A and suppose they are F_{abcd} . We have $a, d \leq 1$. We deduce that $b, c \leq 2$ else with $b \geq 3$ (F_{0300}), by the pigeonhole principle a third column will create either $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ or $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ on those 3 rows. For $m \geq 7$ we deduce then that there are at most 2 columns. For $m = 6$ we can use this to deduce that every pair of columns must form F_{1221} and then the 6×4 configuration $(K_4^2)^T$ establishes the bound. For $m = 5$ we can use this to deduce that every pair of columns must form F_{1211} or F_{1220} or F_{0221} . We can create 4 columns by deleting one row from $(K_4^2)^T$.

Let $A \in \text{ext}(m, \left\{ \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \right\})$. Using the technique of What Is Missing, we have that on any pair of rows we have

$$\begin{matrix} \leq 1 & \leq 1 \\ \begin{bmatrix} 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \end{bmatrix} \end{matrix}$$

For any three rows we find that any choice of columns for those three rows will violate one of the restrictions but there are at most 6 restrictions to violate on the three rows so there are at most 6 columns. For $m = 3$ we may use the construction $[K_3^2 \mid K_3^1]$ and for $m = 4$ we may use the configuration K_4^2 .
 ■

One could potentially circumvent this example by only seeking monotonicity for m large enough.

Conjecture 8.1.4 Let \mathcal{F} be a family of forbidden configurations. Then there exists M for which for all $m \geq M$, $\text{forb}(m, \mathcal{F}) \leq \text{forb}(m + 1, \mathcal{F})$.

8.1.2 A Common 4-rowed Subconfiguration

As mentioned in the introduction to Chapter 5, there are two 4-rowed configurations for which we do not know the asymptotic bound. Both configurations have the following configuration in common:

$$F(t) = t \cdot \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}$$

This is the smallest configuration F for which an asymptotic bound on $\text{forb}(m, F)$ is not known. Interestingly, for $t = 1$ and $t = 2$ a quadratic bound has been proven.

Problem 8.1.5 Prove that $\text{forb}(m, F(t))$ is $\Theta(m^2)$ (or find a counterexample).

8.1.3 Other 3x4 Exact Bounds

In Section 4.1, we found two exact bounds on 3×4 configurations. There are only 10 others for which the answer isn't known. We can apply the local search techniques described in Section 3.6 to find conjectured extremal matrices.

The following table shows the conjectured answers. The entries of the table represent $\text{forb}(m, V_i)$. The “?” symbol represents that the computer has found an example (i.e. a lower bound) but we haven't proven the upper bound.

$m \rightarrow$	3	4	5	6	7	8
$V_1 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}$	8	16	22	32?	43?	55?
$V_2 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}$	8	12	18	26?	36?	45?
$V_3 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$	8	12	18	26?	36?	45?
$V_4 = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}$	8	12	18	26?	36?	45?
$V_5 = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$	8	12	18	26?	36?	44?
$V_6 = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$	8	16	22	29?	37?	45?
$V_7 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$	8	16	22	29?	37?	46?
$V_8 = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$	8	16	22	29?	37?	46?
$V_9 = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$	8	16	22	29?	37?	46?
$V_{10} = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$	8	12	17	23?	30?	38?

Table 8.1: Conjectured values of $\text{forb}(m, V_i)$

The last numbers on this table (for 8 rows) take a few hours to find. The next number, for 9 rows, might take a few weeks of computer time.

One could conjecture that $\text{forb}(m, V_{10}) = \binom{m}{2} + \binom{m}{1} + \binom{m}{0} + 1$.

8.1.4 Critical Substructures

We give some ideas extending those of Section 4.3.

Problem 8.1.6 Prove Conjecture 4.3.2 or the equivalent Conjecture 4.3.8 or give a counterexample.

What follows are ideas and attempts at proving Conjecture 4.3.2 for $k \geq 5$. We prove some of the properties required for general k and we prove some for $k = 5$.

Consider applying the idea for the proof of Proposition 4.3.9 for general k to try to show that $\text{forb}(m, F_{k-1}) < \text{forb}(m, K_k)$. If we could show this is true for any m , then it will be true for all larger m by induction. One might try for $m = k$ as a base case, but unfortunately we find that $\text{forb}(k, 2 \cdot K_k - \mathbf{1}_{k-1} - \mathbf{0}_{k-1}) = \text{forb}(k, K_k) = 2^k - 1$, namely if we delete any column of K_k which has column sum not 0 or k , then the resulting matrix has no $2 \cdot K_k - \mathbf{1}_{k-1} - \mathbf{0}_{k-1}$. One might try next to obtain the inequality for $m = k + 1$.

Proceed by contradiction and assume $\text{forb}(k + 1, F_{k-1}) = \text{forb}(k + 1, K_k)$. Let $A \in \text{ext}(k + 1, F_{k-1})$. As argued before, for any $(k + 1) \times 1$ column α which is not a column in A , $[A|\alpha]$ has a copy of K_k say on rows S . Similarly for a column β not in A , we have that $[A|\beta]$ has a copy of K_k say on rows T .

Now if $S = T$, then $A|_S$ has the submatrix $K_k - \alpha|_S$ and also the submatrix $K_k - \beta|_S$. If $\alpha_S \neq \beta|_S$, then we deduce A has K_k , a contradiction. There are at most 2 columns on $k + 1$ rows agreeing on a k -set of rows S . Now if $S \neq T$ then $|S \cap T| = k - 1$. A similar argument yields a contradiction if $\alpha|_{S \cap T} \neq \beta|_{S \cap T}$.

Table 8.2 shows some computer-found lower bounds (using the algorithms described in Chapter 3). Remember these are conjectures for $m \geq 6$.

We can show $\text{forb}(k + 2, F_k) = \text{forb}(k + 2, K_{k+1})$ for $k \geq 4$ as follows. The $k + 3$ columns to delete from K_{k+2} are

$$B = \begin{bmatrix} \mathbf{1}_3 & I_3^c & \mathbf{1}_3 \\ \times & \times & \times \\ \mathbf{0}_{k-1} & \mathbf{0}_{k-1} & I_{k-1} \end{bmatrix}$$

One can show that this matrix B has the property that for every k -set of rows S , there is a column repeated three times in $B|_S$, while K_{k+2} has four copies of each column on a k -set

m	$\text{forb}(m, F_3)$	$\text{forb}(m, K_4)$
4	15	15
5	25	26
6	40(?)	43
7	59(?)	64

m	$\text{forb}(m, F_4)$	$\text{forb}(m, K_5)$
5	31	31
6	57	57
7	98(?)	99

Table 8.2: Conjectured values for $\text{forb}(m, F_3)$ and $\text{forb}(m, K_4)$

m	$\text{forb}(m, 2 \cdot K_4^2)$	$\text{forb}(m, K_5)$
4	16	16
5	30	31
6	53(?)	57
7	86(?)	99

Table 8.3: Conjectured values for $\text{forb}(m, 2 \cdot K_4^2)$

of rows.

We should also expect that $\text{forb}(m, 2 \cdot K_4^2)$ is much less than $\text{forb}(m, K_5)$. Several proof techniques will work but the easiest is to consider $\text{forb}(5, 2 \cdot K_4^2)$ and show that $\text{forb}(5, 2 \cdot K_4^2) \leq 2^5 - 2$, i.e you must delete at least 2 columns in order to avoid $2 \cdot K_4^2$. This is relatively easy. A column of sum 2 can only yield $\mathbf{1}_2 \times \mathbf{0}_2$ on three 4-sets of rows but there are five 4-sets of rows. One way to achieve this is to delete from K_5 , a column of sum 2 and a column of sum 3 that overlap in two rows. Table 8.3 shows some more computer experiments.

Lemma 8.1.7 We can establish that

$$\text{forb}(m, 2 \cdot K_4^2) \leq \text{forb}(m, K_5) - m + 4.$$

Proof: We use induction as follows. We first note $\text{forb}(m, [2 \cdot K_3^1 \mid 2 \cdot K_3^2]) < \text{forb}(m, K_4)$. Now applying the standard decomposition (2.1.1) repeatedly we have $\text{forb}(m, K_4^2) \leq \text{forb}(m - 1, K_4^2) + \text{forb}(m - 1, [2 \cdot K_3^1 \mid 2 \cdot K_3^2])$ and then $\text{forb}(m, K_4^2) \leq \text{forb}(m - 2, K_4^2) + \text{forb}(m - 2, [2 \cdot K_3^1 \mid 2 \cdot K_3^2]) + \text{forb}(m - 1, [2 \cdot K_3^1 \mid 2 \cdot K_3^2])$ and then repeat this $m - 4$ times using that $[2 \cdot K_3^1 \mid 2 \cdot K_3^2]$ is in F_3 and $\text{forb}(5, F_3) = \text{forb}(5, K_4) - 1$. ■

The following idea is probably of no help in establishing these inequalities but it should be noted.

Proposition 8.1.8 If $\text{forb}(m, F) < \text{forb}(m, K_k)$ for $m \geq m_0$, then $\text{forb}(m, [0 \ 1] \times F) < \text{forb}(m, K_{k+1}) - m + m_0$.

Proof: Simply use induction $\text{forb}(m, [0 \ 1] \times F) \leq \text{forb}(m - 1, [0 \ 1] \times F) + \text{forb}(m - 1, F)$.

■

Let $\text{forb}(m, F_4) = \text{forb}(m, K_5)$ and let $A \in \text{ext}(m, F_4) \subseteq \text{ext}(m, K_5)$. Then on each 5-set of rows exactly one column is missing as described above. So each 5-set contains $K_5 - \alpha$ for some choice of 5×1 α . Now if α has column sum 0,1,4,5 then we find that $K_5 - \alpha$ has F_4 . So we deduce that α has column sum 2,3. Thus we deduce that A has no $K_5^2 K_5^3$ on the grounds that one column is missing. Immediately we deduce that A has all columns of sum 0,1, $m-1$, m (adding these columns can't create a K_5 . But then $A - [K_m^0 \mid K_m^1 \mid K_m^{m-1} \mid K_m^m]$ has no $[K_4^1 \mid 2 \cdot K_4^2 \mid K_4^3]$. Now applying our inductive decomposition on $[K_4^1 \mid 2 \cdot K_4^2 \mid K_4^3]$ yields F_3 . We would be done if we could show $\text{forb}(m, F_3) \leq \text{forb}(m, K_4) - 3$.

Here's an idea that might not help at all, but gives rise to an interesting problem. Let F be a k -rowed configuration for which $\text{forb}(m, F)$ is known. Let $G \prec F$ be an ℓ -rowed subconfiguration. Suppose we wish to find out if G is a critical substructure of F . That is, suppose we wished to find out if

$$\text{forb}(m, G) = \text{forb}(m, F).$$

Proceeding by contradiction, assume this was true and consider a matrix $A \in \text{ext}(m, G) \subseteq \text{ext}(m, F)$. If we add a column α to A , by hypothesis we should get that $[A \mid \alpha]$ contains F as a configuration. When does this mean A contains G ? In particular, this means A contains $F - \phi$ for some column ϕ of F with $\phi = \alpha|_S$ for some subset of the rows S . Note then that $G \not\prec F - \phi$. Thus, any copy of G in $[A \mid \alpha]$ must be missing at most 1 column of F .

Consider the following reinterpretation of the problem.

Lemma 8.1.9 If for every column ϕ of F we have that $G \prec F - \phi$, then

$$\text{forb}(m, G) < \text{forb}(m, F).$$

Proof: If A is in $\text{ext}(m, G) \subseteq \text{ext}(m, F)$, then by adding any column not already present in A to A we must obtain the configuration F since $A \in \text{ext}(m, F)$, which means that for some column ϕ of F , we have that $F - \phi$ is present in A , which means $G \not\subseteq F - \phi$, a contradiction. ■

8.2 Concluding Remarks

We have seen many results concerning a single Forbidden Configuration and just a peek at results concerning multiple Forbidden Configurations. An observant reader might have noticed that the tools we use to study a single forbidden configuration are mostly the same tools we use to study a family of forbidden configurations, so, the observant reader might ask, why then have we neglected our families and opted for singles? A perhaps unsatisfying answer could be that single configurations are hard enough.

Imagine for a moment we had an oracle that when given a family of configurations \mathcal{F} would immediately provide the (correct) value of $\text{forb}(m, \mathcal{F})$ and perhaps an extremal matrix satisfying the bound. What could we do with such a wonder? In particular, among other things, this oracle would solve all Erdős-Stone-Simonovits type results, as one could imagine forbidding $(1, 1, 1)^T$ together with the incidence matrix of a graph H . When forbidding these configurations the extremal matrix one would obtain would be $[\mathbf{0} \mid I \mid G]$ where G is the incidence matrix of a graph that does not have H as a configuration.

Because of this, the idea of having a general theory for families of forbidden configurations looks tantalizing, but perhaps a little daunting. Results about small families of small forbidden configurations could very well flourish with a systematic study. If I were to continue research in forbidden configurations, I would certainly start here, as I am sure there are many new results waiting to be discovered.

As of this writing we don't yet have the tools to study configurations with 6 rows or more, but we hope that by chipping away at the problem we are making progress toward perhaps a proof or refutation of Conjecture 1.4.1. Proving this pivotal conjecture would indeed be a milestone for the field, but it is my personal belief that the conjecture will turn out to be false. The reason for this belief, in the face of so much evidence for it, is its constructive nature. Sure, the product constructions described in this Conjecture 1.4.1 have given the best asymptotic bounds for many (small) configurations, but these are not, by any stretch, a random sample of configurations. They are chosen to be small and manageable.

As we've seen, a very small change in a configuration (for example, substituting a 0 by a

1 or adding a column) can lead to wildly different bounds. I find there is not much reason to believe a large, random configuration would satisfy the conjecture that the best (asymptotic) construction will be products of I , I^c and T .

On the other hand, the configurations I , I^c and T do seem to “pop-up” repeatedly while studying configurations. Perhaps they are in some way special, and we have indeed proven many cases in which the conjecture turns out to be true.

Bibliography

- [ABS11] R.P. Anstee, F. Barekat, and A. Sali, *Small forbidden configurations V: Exact bounds for 4×2 cases*, Studia. Sci. Math. Hun. **48** (2011), 1–22. → pages 32, 59, 60
- [AF86] R.P. Anstee and Z. Füredi, *Forbidden submatrices*, Discrete Math. **62** (1986), 225–243. → pages 15, 18, 134
- [AFS01] R.P. Anstee, R. Ferguson, and A. Sali, *Small forbidden configurations II*, Electronic Journal of Combinatorics **8** (2001), R4 25pp. → pages 18, 32
- [AGS97] R.P. Anstee, J.R. Griggs, and A. Sali, *Small forbidden configurations*, Graphs and Combinatorics **13** (1997), 97–118. → pages 18, 32, 33
- [AHK77] Kenneth Appel, Wolfgang Haken, and John Koch, *Every planar map is four colorable*, Illinois Journal of Mathematics (1977). → pages 31
- [AK97] R. Ahlswede and L.H. Khachatrian, *The complete intersection theorem for systems of finite sets*, European Journal of Combinatorics **16** (1997), 125–136. → pages 30
- [AK06] R.P. Anstee and P. Keevash, *Pairwise intersections and forbidden configurations*, European Journal of Combinatorics **27** (2006), 1235–1248. → pages 18, 30, 60, 119, 121
- [AK07] R.P. Anstee and N. Kamoosi, *Small forbidden configurations III*, Electronic Journal of Combinatorics **14** (2007), R79 34pp. → pages 32
- [AK10] R.P. Anstee and S.N. Karp, *Forbidden configurations: Exact bounds determined by critical substructures*, Electronic Journal of Combinatorics **17** (2010), R50 27pp. → pages 32, 59

- [Alo83] N. Alon, *On the density of sets of vectors*, Discrete Math **46** (1983), 199–202.
→ pages 83
- [Ans] R.P. Anstee, *A survey of forbidden configuration results*, <http://www.math.ubc.ca/~anstee/>. → pages 18, 23, 103
- [Ans85] R.P. Anstee, *General forbidden configuration theorems*, Journal of Combinatorial Theory A **40** (1985), 108–124. → pages 13, 32
- [Ans90] R.P. Anstee, *Forbidden configurations, determinants and discrepancy*, European Journal of Combinatorics **11** (1990), 15–19. → pages 88
- [AR11] R.P. Anstee and Miguel Raggi, *Genetic algorithms applied to problems of forbidden configurations*, preprint. <http://www.math.ubc.ca/~anstee/> (2011), 22pp.
→ pages iii, 2
- [ARS10a] R.P. Anstee, Miguel Raggi, and Attila Sali, *Evidence for a forbidden configuration conjecture; one more case solved.*, preprint. <http://www.math.ubc.ca/~anstee/> (2010), 16pp. → pages iii, 2, 18
- [ARS10b] R.P. Anstee, Miguel Raggi, and Attila Sali, *Forbidden configurations and product constructions*, preprint. <http://www.math.ubc.ca/~anstee/> (2010), 22pp. → pages iii, 2
- [ARS11] R.P. Anstee, Miguel Raggi, and Attila Sali, *Forbidden configurations: Quadratic bounds*, preprint. <http://www.math.ubc.ca/~anstee/> (2011), 21pp. → pages iii, 2, 18
- [AS05] R.P. Anstee and A. Sali, *Small forbidden configurations IV*, Combinatorica **25** (2005), 503–518. → pages 15, 18, 33, 95
- [BB05] J. Balogh and B. Bollobás, *Unavoidable traces of set systems*, Combinatorica **25** (2005), 633–643. → pages 33, 134
- [BBM06] J. Balogh, B. Bollobás, and R. Morris, *Hereditary properties of partitions, ordered graphs and ordered hypergraphs*, European Journal of Combinatorics **8** (2006), 1263–1281. → pages 124, 125, 128, 129

- [BEHW89] A. Blumer, A. Ehrenfeucht, D. Haussler, and M. K. Warmuth, *Learnability and the Vapnik–Chervonenkis dimension*, Journal of the ACM **36(4)** (1989), 929–865. → pages 32
- [BKS05] J. Balogh, P. Keevash, and B. Sudakov, *Disjoint representability of sets and their complements*, Journal of Combinatorial Theory B **95** (2005), 12–28. → pages 32
- [EKR61] P. Erdős, C. Ko, and R. Rado, *Intersection theorems for systems of finite sets*, Quarterly Journal of Mathematics, Oxford Series **series 2 12** (1961), 313–320. → pages 30
- [Erd64] P. Erdős, *On extremal problems of graphs and generalized graphs*, Israel Journal of Mathematics **2** (1964), 183–190. → pages 140
- [ES46] P. Erdős and A.H. Stone, *On the structure of linear graphs*, Bulletin of the American Mathematical Society **52** (1946), 1089–1091. → pages 3, 30
- [ES66] P. Erdős and M. Simonovits, *A limit theorem in graph theory.*, Studia Scientiarum Mathematicarum Hungarica **1** (1966), 51–57. → pages 3, 30
- [FH92] Z. Füredi and P. Hajnal, *Davenport-Schinzel theory of matrices*, Discrete Math **103** (1992), 233–251. → pages 123, 128
- [FQ83] Z. Füredi and F. Quinn, *Traces of finite sets*, Ars Combin. **18** (1983), 195–200. → pages 83
- [Für83] Z. Füredi, *personal communication*, 1983. → pages 15
- [Für96] Z. Füredi, *Combinatorics, Probability and Computing* **5** (1996), 29–33. → pages 30, 123
- [Gro80] H.O.F. Gronau, *An extremal set problem*, Studia Sci.Math. Hungar **15** (1980), 29–30. → pages 83
- [Juk01] S. Jukna, *Extremal combinatorics with applications in computer science*, Springer-Verlag, 2001. → pages 29
- [KM07] M. Klazar and A. Marcus, *Extensions of the linear bound in the Füredi-Hajnal conjecture*, Advances in Applied Mathematics **38** (2007), 258–266. → pages 123, 124, 125, 128, 129

- [KST54] Kövari, V. Sós, and P. Turán, *On a problem of K. Zarankiewicz*, Colloq. Math. **3** (1954), 50–57. → pages 30, 123, 140
- [Man07] *Problem 28 soln by H. Gouwentak, W. Mantel, J. Teixeira de Mattes, F. Schuh, and W.A. Wythoff*, Wiskundige Opgavem **10** (1907), 60–61. → pages 29
- [Mat02] Jiri Matousek, *Lectures in discrete geometry*, Springer, 2002. → pages 32
- [MT04] A. Marcus and G. Tardos, *Excluded permutation matrices and the Stanley-Wilf Conjecture*, Journal of Combinatorial Theory Ser. A **107** (2004), 153–160. → pages 123, 124, 125
- [MZ07] D. Mubayi and Y. Zhao, *Forbidding complete hypergraphs as traces*, Graphs and Combinatorics **23** (2007), 667–679. → pages 32
- [O’R87] Joseph O’Rourke, *Art gallery theorems and algorithms*, Oxford University Press, 1987. → pages 32
- [Ram30] F. P. Ramsey, *On a problem of formal logic*, Proceedings London Mathematical Society s2 **30** (1) (1930), 264–286. → pages 31, 118
- [Sau72] N. Sauer, *On the density of families of sets*, Journal of Combinatorial Theory. Series A **13** (1972), 145–147. → pages 15
- [She72] S. Shelah, *A combinatorial problem: Stability and order for models and theories in infinitary languages*, Pacific Journal of Mathematics **4** (1972), 247–261. → pages 15
- [Spe28] Emanuel Sperner, *Ein satz über untermengen einer endlichen menge*, Mathematische Zeitschrift **27** (1) (1928), 544–548. → pages 31
- [Ste78] J. M. Steele, *Empirical discrepancies and subadditive processes*, Annals of Probability **6**(1) (1978), 118–227. → pages 32
- [Tar05] G. Tardos, *On 0-1 matrices and small excluded submatrices*, Journal of Combinatorial Theory Series A **111** (2005), 266–288. → pages 123
- [Tur41] P. Turán, *On an extremal problem in graph theory*, Matematikai és Fizikai Lapok **48** (1941), 436–452. → pages 29

- [TV06] Terence Tao and Van H. Vu, *Additive combinatorics*, Cambridge Studies in Advanced Mathematics, Cambridge University Press, 2006. → pages 32
- [Vap00] V.N. Vapnik, *The nature of statistical learning theory. information science and statistics*, Springer-Verlag, 2000. → pages 32
- [VC68] V. N. Vapnik and A. Ya. Chervonenkis, *On the uniform convergence of relative frequencies of events to their probabilities*, Soviet Math. Dokl. **9** (1968), 915–918. → pages 32
- [VC71] V.N. Vapnik and A.Ya. Chervonenkis, *On the uniform convergence of relative frequencies of events to their probabilities*, Theory of Probability and Applications (1971), 264–280. → pages 15
- [Ver05] R. Vershynin, *Integer cells in convex sets*, Advances in Mathematics **197** (2005), 248–273. → pages 32
- [WD81] R. S. Wencour and R. M. Dudley, *Some special Vapnik–Chervonenkis classes*, Discrete Mathematics **33(3)** (1981), 313–318. → pages 32

Index

- (p, q) -split, 124
- $\{0, 1\}$ -complement, 6
- p_1, p_2, \dots, p_d split, 128
- s -boundary case, 18
- $\text{proj}_i(B)$, 130

- absent, 38
- avoids, 11

- boundary case, 18

- column sum, 5
- concatenation, 6
- configuration, 10
- contain F as a configuration, 11
- critical substructure, 82

- has no configuration, 11

- identity complement I_m^c , 8
- identity matrix I_m , 8
- implication, 40
- impure, 40
- inductive child, 35

- laminar, 13
- long supply, 38

- multiplicity, 7

- non-essential, 93

- pattern, 124
- power set, 4
- predicted s -boundary case, 18
- predicted boundary case, 18
- pure, 40

- restriction, 7

- set system, 4
- shifted matrix, 83
- short supply, 38
- simple, 5
- simple configuration, 10
- simple hypergraph, 4
- standard decomposition, 34
- subconfiguration, 11
- subtraction, 7

- tower matrix T_m , 9

- violates an implication, 40