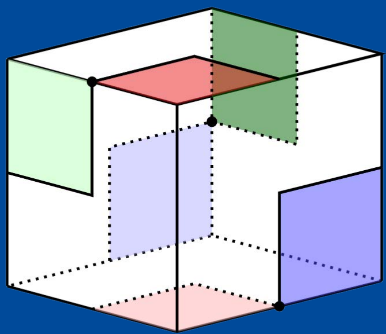
Matija Bucić

Local to Global Phenomenon and Other Topics in Probabilistic and Extremal Combinatorics.



Diss. ETH No. 27626

MATIJA BUCIĆ

LOCAL TO GLOBAL PHENOMENON AND OTHER TOPICS IN PROBABILISTIC AND EXTREMAL COMBINATORICS

DISS. ETH NO. 27626

LOCAL TO GLOBAL PHENOMENON AND OTHER TOPICS IN PROBABILISTIC AND EXTREMAL COMBINATORICS

A dissertation submitted to attain the degree of

DOCTOR OF SCIENCES of ETH ZURICH
(Dr. sc. ETH Zurich)

presented by

MATIJA BUCIĆ
M.Math, University of Cambridge

born on 04. October 1993
citizen of Croatia

accepted on the recommendation of

Prof. Dr. B. Sudakov, examiner
Prof. Dr. M. Krivelevich, co-examiner

2021

DOI: 10.3929/ethz-b-000491057

To my amazing family and friends
and to our Lord creator

ABSTRACT

Extremal combinatorics is one of the fundamental fields of study in modern combinatorics, tracing its origins at the very least to the work of Euler in the 18-th century. In a very general sense, extremal combinatorics is concerned with questions of the form how large or how small a collection of finite objects can be, provided it has to satisfy certain restrictions. Another fundamental branch of modern combinatorics is probabilistic combinatorics which was first systematically studied by Erdős in the 1940s. Probabilistic combinatorics entails the use of probability theory to obtain combinatorial results, which in many cases have no inherent randomness at all. In this thesis, we explore a number of different topics in extremal and probabilistic combinatorics. The first half of the discussed topics have a common theme. Namely, they are concerned with instances of the local to global principle, which states that one can obtain a global understanding of a structure from having a good understanding of its local properties, or vice versa. This phenomenon has been ubiquitous in many areas of mathematics and beyond, with profound consequences.

Our first topic is concerned with the classical, extensively studied, Erdős-Rogers problem dating back to 1962, which can be restated as follows. How large must the independence number $\alpha(G)$ of a graph G be, whose every m vertices contain an independent set of size r? This restatement is due to Erdős and Hajnal and independently Linial and Rabinovich from the early '90s. It arose from a change in perspective, which shifts the focus from fixing $\alpha(G)$ and r to fixing the local parameters m and r instead. We develop two new approaches to attack this problem which allow us to significantly improve previously best known bounds due to Linial and Rabinovich, Erdős and Hajnal, Alon and Sudakov, Krivelevich, and Kostochka and Jancey, depending on the regime. We also discuss a related topic of connecting a local Ramsey property to its global counterpart, introduced by Erdős and Hajnal, which we exploit to obtain new examples of Ramsey graphs.

The next topic is concerned with the following type of questions. How many monochromatic paths, cycles, or general trees does one need to cover all vertices of a given r-edge-coloured graph G? These problems date back to the 1960s and were intensively studied by various researchers over the last 50 years. We establish a connection between this problem and the following natural Helly-type, local to global question for hypergraphs. What is the maximum number of vertices needed to cover all the edges of a hypergraph H if it is known that any collection of a few edges of H has a small cover? This problem was raised by Erdős, Hajnal, and Tuza about 30 years ago. We

obtain quite accurate bounds for the hypergraph problem and use them to give some unexpected answers to several questions about covering graphs by monochromatic trees raised by Bal and DeBiasio; Kohayakawa, Mota, and Schacht; Lang and Lo; and Girão, Letzter, and Sahasrabudhe.

We will also discuss concepts of locally and globally forcing tournaments. Here, for example, a tournament H is said to be locally forcing if a big tournament T having "correct" counts (approximately the same as in the random tournament) of H as a subtournament in any large induced subtournament forces T to be quasirandom. We develop an analogous theory to the one introduced by Chung and Graham for usual graphs with some surprising results. One highlight is a result which states that in order for H to be locally forcing H itself needs to be very strongly quasirandom.

Rota's conjecture, dating back to 1989, states the following. Given n bases $B_1, \dots, B_n$ in an n-dimensional vector space V, one can always find n disjoint bases of V, each containing exactly one element from each B_i (we call such bases transversal bases). Rota's basis conjecture remains wide open despite its apparent simplicity and the efforts of many researchers over the last 30 years. We obtain the first linear bound on the number of disjoint transversal bases one can find. The conjecture has a natural extension to the setting of matroids, where our results also apply.

The classical Erdős-Szekeres theorem dating back almost a hundred years states that any sequence of $(n-1)^2+1$ distinct real numbers contains a monotone subsequence of length n. It has been generalised to higher dimensions in a variety of ways, but perhaps the most natural one was proposed by Fishburn and Graham about 25 years ago. They defined concepts of a monotone and a lex-monotone array and asked how large an array one needs in order to be able to find a monotone or a lex-monotone subarray of size $n \times \ldots \times n$. They obtained Ackermann-type bounds for their problem in both cases. We significantly improve these results. Regardless of the dimension, we obtain at most a quadruple exponential bound in n in both cases. We also connect the problem to a number of interesting topics considered by various researchers over the years.

We conclude with two topics in extremal set theory. The first one involves the minimum possible size of the intersection spectrum of a 3-chromatic intersecting hypergraph, where we prove in a very strong form a conjecture of Erdős and Lovász from 1973. The second one involves saturated families of sets, where we answer a question of Frankl and Tokushige, making first substantial progress on a conjecture of Erdős and Kleitman from 1974. The main ingredient is a connection we observe between the problem and a famous correlation inequality.

ZUSAMMENFASSUNG

Die extremale Kombinatorik ist eines der grundlegenden Forschungsgebiete der modernen Kombinatorik und geht zumindest auf die Arbeit von Euler im 18. Jahrhundert zurück. Ganz allgemein befasst sich die extremale Kombinatorik mit Fragen der Form, wie gross oder wie klein eine Menge endlicher Objekte sein kann, sofern sie bestimmte Einschränkungen erfüllen muss. Ein weiterer grundlegender Zweig der modernen Kombinatorik ist die probabilistische Kombinatorik, die erstmals in den 1940er Jahren von Erdős systematisch untersucht wurde. Probabilistische Kombinatorik beinhaltet die Verwendung der Wahrscheinlichkeitstheorie, um kombinatorische Ergebnisse zu erhalten, die in vielen Fällen überhaupt keine inhärente Zufälligkeit aufweisen. In dieser Arbeit untersuchen wir verschiedene Themen der extremalen und probabilistischen Kombinatorik. Die erste Hälfte der diskutierten Themen hat ein gemeinsamen Fokus. Sie befassen sich nämlich mit Instanzen des sogenannten Lokal-Global-Prinzips, das besagt, dass man ein globales Verständnis einer Struktur erhalten kann, wenn man ihre lokalen Eigenschaften gut versteht, oder umgekehrt. Dieses Phänomen ist in vielen Bereichen der Mathematik und darüber hinaus allgegenwärtig, mit tiefgreifenden Konsequenzen.

Unser erstes Thema befasst sich mit dem klassischen, ausführlich untersuchten Erdős-Rogers-Problem aus dem Jahr 1962, das wie folgt formuliert werden kann. Wie gross muss die Unabhängigkeitszahl $\alpha(G)$ eines Graphen G sein, dessen alle m Eckpunkte eine unabhängige Menge der Grösse r enthalten? Diese alternative Formulierung ist Erdős und Hajnal sowie unabhängig Linial und Rabinovich aus den frühen 90er Jahren zu verdanken. Es entstand aus einem Perspektivwechsel, der den Fokus von der Festlegung von $\alpha(G)$ und r auf die Festlegung der lokalen Parameter m und r verlagerte. Wir entwickeln zwei neue Ansätze, um dieses Problem anzugehen, die es uns ermöglichen, die bisher bekannten Schranken von Linial und Rabinovich, Erdős und Hajnal, Alon und Sudakov, Krivelevich sowie Kostochka und Jancey je nach Regime signifikant zu verbessern. Wir diskutieren auch ein verwandtes Thema der Verbindung einer lokalen Ramsey-Eigenschaft mit ihrem globalen Gegenstück, das von Erdős und Hajnal eingeführt wurde und das wir nutzen, um neue Beispiele für Ramsey Graphen zu erhalten.

Das nächste Thema befasst sich mit den folgenden Fragen. Wie viele monochromatische Pfade, Zyklen oder allgemeine Bäume benötigt man, um alle Ecken eines bestimmten r-kanten-gefärbten Graphen G abzudecken? Diese Probleme stammen aus den 1960er Jahren und wurden in den letzten 50 Jahren von verschiedenen Forschern intensiv untersucht. Wir stellen eine Verbindung zwischen diesem Problem und der folgenden natürlichen, lokalen bis globalen Frage vom Helly-Typ für Hypergraphen her. Was ist die maximale Anzahl von Ecken, die benötigt werden, um alle Kanten eines Hypergraphen H abzudecken, wenn bekannt ist, dass jede Sammlung einiger Kanten von H eine kleine Abdeckung hat? Dieses Problem wurde vor etwa 30 Jahren von Erdős, Hajnal und Tuza angesprochen. Wir erhalten ziemlich genaue Schranken für das Hypergraphen-Problem und verwenden sie, um einige

unerwartete Antworten auf verschiedene Fragen zur Abdeckung von Graphen durch monochromatische Bäume zu geben, die von Bal und DeBiasio; Kohayakawa, Mota und Schacht; Lang und Lo; und Girão, Letzter und Sahasrabudhe aufgeworfen wurden.

Wir werden auch Konzepte für das lokale und globale Erzwingen von Turnieren diskutieren. Hier wird beispielsweise ein Turnier H als lokal erzwungen bezeichnet, wenn ein grosses Turnier T mit korrekten Zählungen von H (ungefähr das gleiche wie im Zufallsturnier) als Subturnier in allen grossen induzierten Subturnieren zwangsläufig quasi-zufällig sein muss. Wir entwickeln eine analoge Theorie zu der von Chung und Graham eingeführten für übliche Graphen mit einigen überraschenden Ergebnissen. Ein Highlight ist ein Ergebnis, das besagt, dass H, um H selbst lokal zu erzwingen, sehr stark quasi-zufällig sein muss.

Rotas Vermutung aus dem Jahr 1989 besagt Folgendes. Gegeben n Basen $B_1, \ldots, B_n$ in einem n-dimensionalen Vektorraum V, kann man immer n disjunkte Basen von V finden, die jeweils genau ein Element von jedem B_i enthalten (wir nennen solche Basen transversale Basen). Rotas Basis-Vermutung bleibt trotz ihrer offensichtlichen weit offen Einfachheit und der Bemühungen vieler Forscher in den letzten 30 Jahren weit offen. Wir erhalten die erste lineare Schranke für die Anzahl von disjunkte transversale Basen Anzahl disjunkter transversaler Basen, die man finden kann. Die Vermutung hat eine natürliche Erweiterung für Matroide, wo unsere Ergebnisse auch zutreffen.

Das klassische Erdős-Szekeres-Theorem aus fast hundert Jahren besagt, dass jede Folge von $(n-1)^2+1$ verschiedenen reellen Zahlen eine monotone Teilfolge der Länge n enthält. Es wurde auf verschiedene Weise auf höhere Dimensionen verallgemeinert, aber das vielleicht natürlichste wurde vor etwa 25 Jahren von Fishburn und Graham vorgeschlagen. Sie definierten Konzepte eines monotonen und eines lex-monotonen Arrays und fragten, wie gross ein Array sein müsse, um ein monotones oder ein lex-monotones Subarray der Grösse $n \times \ldots \times n$ zu finden. In beiden Fällen erhielten sie für ihr Problem Schranken vom Typ Ackermann. Wir verbessern diese Ergebnisse deutlich. Unabhängig von der Dimension erhalten wir in beiden Fällen höchstens eine vierfache Exponentialgrenze in n. Wir verbinden das Problem auch mit einer Reihe interessanter Themen, die von verschiedenen Forschern im Laufe der Jahre behandelt wurden.

Wir schliessen mit zwei Themen der extremalen-Mengentheorie. Die erste betrifft die minimal mögliche Grösse des Schnittspektrums eines 3-chromatischen sich überschneidenden Hypergraphen, wobei wir in sehr starker Form eine Vermutung von Erdős und Lovász aus dem Jahr 1973 beweisen. Die zweite betrifft gesättigte Familien von Mengen, in denen wir eine Frage von Frankl und Tokushige beantworten und erste wesentliche Fortschritte bei einer Vermutung von Erdős und Kleitman aus dem Jahr 1974 machen. Der Hauptbestandteil ist eine Verbindung, die wir zwischen dem Problem und einer berühmten Korrelationsungleichheit beobachten.

ACKNOWLEDGEMENTS

First and foremost I would like to thank my advisor Benny Sudakov, without whose invaluable, sage advice and guidance this thesis would not exist. I am particularly indebted for the amazing opportunities he made available for me, and for being an excellent collaborator on a number of very successful and enjoyable research projects. On a separate note I am extremely grateful for the amazing working environment he fosters and for plenty of sage advice about life beyond mathematics. I'd also like to express my sincere gratitude to Shoham Letzter, for being a great friend, collaborator and mentor and for the invaluable role she played at the start of my studies. I am very grateful to Imre Leader for being an incredibly encouraging force and for his incredibly kind words at various difficult points of my journey. I am also very grateful to Matthew Kwan, Stefan Glock and Alexey Pokrovskiy for being amazing collaborators and for their help and support. I would like to thank Nina Kamčev for being an amazing friend and for plenty of good advice about both life and mathematics. I would like to thank Michael Krivelevich for a number of encouraging conversations and for taking time from his busy schedule to referee my thesis.

There have been a huge number of other people who have had a positive impact on my time as a doctoral student. In particular, I'd like to thank all my coauthors given with number of successful collaborations in parentheses: Benny Sudakov (20), Shoham Letzter (5), Tuan Tran (4), Alexey Pokrovskiy (3), Noga Alon (2), Nemanja Draganić (2), Matthew Kwan(2), Adam Wagner(2), Jacob Cooper, David Correia, Pat Devlin, Jacob Fox, Lior Gisboliner, Stefan Glock, Sven Heberle, Mo Hendon, Dru Horne, Erik Jahn, Oliver Janzer, Tom Kalvari, Dan Kráľ, Dániel Korándi, Eden Kuperwasser, Bernard Lidicky, Eoin Long, Jason Long, Ben Lund, Samuel Mohr, Melkior Ornik, Asaf Shapira, Ufuk Topcu. In addition I would like to thank all other students and postdocs who have been a part of Benny's close-knit research group over the years (Felix, Igor, Istvan, Jon, Nina, Omri, Pedro, Richard, Tash) and the adopted member, the one and only, Davide.

Last but not least, my heartfelt thanks go to my amazing family: Mum, Dad, Tina, Dino, Ana, Joe, Marko, Danijela, Stric, Strina for always being here for me and to some amazing friends who I consider to essentially be a part of my family, Domagoj, Paula, Nina, Leo, Luka and Tea. I have truly been blessed beyond belief with the people closest to me.

Main parts of this thesis correspond closely to a number of papers completed during my doctorate. Specifically:

- Section 2.1 corresponds to the paper "Large independent sets from local considerations", joint work with Benny Sudakov.
- Section 2.2 corresponds to the paper "Large cliques and independent sets all over the place", joint work with N. Alon and B. Sudakov (published in *Proceedings of AMS*).
- Section 3.1 corresponds to the paper "Covering graphs by monochromatic trees and Helly-type results for hypergraphs", joint work with D. Korándi and B. Sudakov (accepted for publication in *Combinatorica*).
- Section 3.2 corresponds to the paper "Tournament quasirandomness from local counting", joint work with E. Long, A. Shapira and B. Sudakov (accepted for publication in *Combinatorica*).
- Chapter 4 corresponds to the paper "Erdős-Szekeres theorem for multidimensional arrays", joint work with B. Sudakov and T. Tran.
- Chapter 5 corresponds to the paper "Halfway to Rota's basis conjecture", joint work with M. Kwan, A. Pokrovskiy and B. Sudakov (published in *International Mathematics Research Notices*).
- Section 6.1 corresponds to the paper "The intersection spectrum of 3-chromatic intersecting hypergraphs", joint work with S. Glock and B. Sudakov.
- Section 6.2 corresponds to the paper "Minimum saturated families of sets", joint work with S. Letzter, B. Sudakov and T. Tran (published in *Bulletin of the LMS*).

CONTENTS

1 INTRODUCTION 1
2 RAMSEY TYPE PROBLEMS 9
2.1 Large independent sets from local considerations 9
2.1.1 Introduction 9
2.1.2 Local to global independence number 14
2.1.3 2-density and local independence number 32
2.1.4 Concluding remarks 43
2.2 Locally Ramsey Graphs 56
2.2.1 Locally Ramsey graphs and lexicographic products . . 58
2.2.2 Locally Ramsey graphs and scrambling 63
2.2.3 Small cliques and independent sets 68
2.2.4 Concluding remarks 69
3 TOPICS CONCERNING HYPERGRAPHS AND TOURNAMENTS 71
3.1 Covering graphs by monochromatic trees and Helly-type results for hypergraphs 71
3.1.1 Introduction 71
3.1.2 Preliminaries 77
3.1.3 Upper bound near the threshold 79
3.1.4 The connection to hypergraph covering 81
3.1.5 Results for general hypergraphs 85
3.1.6 Results for r-partite hypergraphs 91
3.1.7 Concluding remarks and open problems 98
3.2 Tournament quasirandomness from local counting 102
3.2.1 Introduction 102
3.2.2 Globally forcing tournaments 107
3.2.3 Local forcing and the counting polynomial 109
3.2.4 Finding $\mathcal{T}_{cliq}$-forcing tournaments 115
3.2.5 $\mathcal{T}_{bip}$-forcing tournaments 118
3.2.6 Proving local forcing 121
3.2.7 Concluding remarks 129
4 ERDŐS-SZEKERES THEOREM FOR MULTIDIMENSIONAL ARRAYS 131
4.1 Introduction 131
4.2 Monotone arrays 134
4.2.1 Preliminaries 135
4.2.2 Proofs of the main results on monotonicity 136
4.3 Lexicographic arrays 140

4.3.1 Preliminaries 140
4.3.2 2-dimensional case 141
4.3.3 High-dimensional case 146
4.4 Concluding Remarks 150
4.4.1 Long common monotone subsequence 151
4.4.2 Ramsey type problems for vertex-ordered graphs . . . 152
4.4.3 Canonical orderings of discrete structures 153
5 ROTA'S BASIS CONJECTURE FOR MATROIDS 155
5.1 Introduction 155
5.2 Finding many disjoint transversal bases 157
5.2.1 Simple swaps 158
5.2.2 Cascading swaps 161
5.2.3 Increasing the number of initial addable elements . . . 167
5.3 Concluding remarks 170
6 TWO TOPICS IN EXTREMAL SET THEORY 173
6.1 Intersection spectrum of 3-chromatic intersecting hypergraphs 173
6.1.1 Introduction 173
6.1.2 Intersection spectrum 176
6.1.3 Concluding remarks 183
6.2 Minimum saturated families of sets 184
6.2.1 Introduction 184
6.2.2 The proof 187
6.2.3 Conclusion 190

BIBLIOGRAPHY 193

NOTATION

We use standard graph-theoretic notation throughout. We denote the vertex set and edge set of a graph or a hypergraph G by $V(G)$ and $E(G)$ respectively. We will denote by $|G| = |V(G)|$ the number of vertices of G and by $e(G) = |E(G)|$ the number of edges in G. For $v \in G$ we denote by $N_G(v)$ the neighbourhood of v and by $d_G(v) = |N_G(v)|$ its degree in G. When the graph in question is clear we omit it from the index. For a subset $S \subseteq V(G)$ we denote the *induced subgraph* of G on this subset by $G[S]$, that is the subgraph with vertex set S which inherits the edge structure from G. An r-graph or r-uniform hypergraph is a hypergraph with all edges of size r, so for example 2-graphs are the usual graphs. Whenever we work with r-graphs, we tacitly assume $r \geq 2$. An r-graph H is said to be *r-partite* if there is a partition of $V(G) = V_1 \cup \cdots \cup V_r$ such that any $e \in E(H)$ satisfies $|e \cap V_i| = 1$ for all i. The independence number of a graph G, denoted $\alpha(G)$, is defined as the largest size of an independent set (set of vertices spanning no edges) in G. A clique is a graph with all edges present, we denote it by K_r when it consists of r vertices.

We denote by $\mathcal{G}(n, p)$ the binomial random graph obtained by including every possible edge with probability p, independently between different edges. Given a probability distribution $\mathcal{D}$ on X we write $x \sim \mathcal{D}$ to denote an element $x \in X$ selected randomly according to $\mathcal{D}$. We simply write $x \sim X$ when $\mathcal{D}$ is taken to be the uniform distribution on X.

We use standard set-theoretic notation throughout as well. Let $[n] = \{1, 2, \ldots, n\}$, let $2^{[n]}$ be the family of all subsets of $[n]$ and let $(n)_m$ denote the falling factorial $(n)_m = n(n-1)\cdots(n-m+1)$. Given a set X we write $\binom{X}{k}$ for the collection of k-element subsets of X. We denote by $\emptyset$ the empty set.

For a real number x we denote by $\lfloor x \rfloor$ and $\lceil x \rceil$ the largest integer smaller and larger or equal than x, respectively and refer to these as floor and ceiling functions of x. We will often omit these, with the goal of improving clarity of presentation, whenever their use is not crucial for the argument. We use standard asymptotic notation throughout the paper. We also use standard asymptotic notation throughout, as follows. For functions $f = f(n)$ and $g = g(n)$, we write $f = O(g)$ to mean that there is a constant C such that $|f| \leq C|g|$, $f = \Omega(g)$ to mean that there is a constant $c > 0$ such that $f \geq c|g|$, $f = \Theta(g)$ to mean that $f = O(g)$ and $f = \Omega(g)$, and $f = o(g)$ or $f \ll g$ to mean that $f/g \to 0$ as $n \to \infty$. When saying "with high probability", often abbreviated as "w.h.p.", we mean that the probability of an event is $1 - o(1)$.

Finally, we note that we will introduce some further notation, specific to each section as the need for it arises.

1

INTRODUCTION

Extremal and probabilistic combinatorics are two fundamental areas of modern combinatorics. These two closely related subjects have seen spectacular development over the last couple of decades. This has been in large part due to the development of deep and well-developed methods, which solidified the position of combinatorics as a fundamental mathematical discipline with connections to many other areas.

One of the most classical fields of study in extremal combinatorics, which has in addition been central to the development of probabilistic combinatorics, is Ramsey theory. Ramsey theory refers to a large body of deep results, which all roughly speaking say that any sufficiently large structure is guaranteed to have a large, well-organised substructure. Its inception dates back to 1929 and the celebrated theorem of Ramsey, [209] which states that *any* sufficiently large graph must contain a clique or an independent set of arbitrarily large size. Determining the maximum size of such a set that we are guaranteed to find in any graph on n vertices frames this as an extremal problem. Here, already in 1935, Erdős and Szekeres [97] showed that any graph on n vertices contains a clique or an independent set of size $0.5 \log n$. On the other hand, in what was one of the first applications of the now indispensable probabilistic method Erdős [84] has shown that almost all graphs on n vertices contain neither a clique nor an independent set of size $2 \log n$. As an introductory example showcasing the basic probabilistic method let us include a proof of this result.

Theorem 1.0.1. *Almost all graphs on n vertices contain neither a clique nor an independent set of size larger than $2 \log n$.*

Proof. Let us consider the random graph $\mathcal{G}(n, 1/2)$, obtained by choosing every possible edge to exist with probability $1/2$, independently of all other edges. Since every graph is equally likely to be an outcome of this experiment, our task is to show that $\mathcal{G}(n, 1/2)$ contains a clique or an independent set of size $k = \lfloor 2 \log n \rfloor$ with vanishing probability. For any subset K consisting of k out of our n vertices, we define an event E_K to occur if it spans a clique or an independent set. Then $\mathbb{P}(E_K) = 2 \cdot 2^{-\binom{k}{2}}$. Since there are $\binom{n}{k}$ different events E_K the probability that one of them occurs is at most $2 \cdot 2^{-\binom{k}{2}} \cdot \binom{n}{k}$, which can easily be verified to approach 0 as n grows. $\square$

Over the years, Ramsey theory has developed into a full-fledged area of research with a number of deep results and applications to many other fields. Yet, some of the most classical problems remain intractable essentially since the early days. One of them is most certainly what happens in the asymmetric case of the Ramsey problem mentioned above. In other words, what happens if we forbid a clique of fixed size m and are looking for as large an independent set as possible. The asymptotics for this problem have been resolved for $m = 3$ (see [11, 36, 106]) only recently, thanks in large part due to the development of our understanding of the triangle-free process, but already for $m = 4$ the problem is wide open. Another classical problem is concerned with removing randomness from the proof given above. Namely, since almost all graphs fail to contain cliques or independent sets of size $2 \log n$, certainly it should be possible to exhibit an explicit (non-random) one. Despite attention from many researchers and some major recent breakthroughs coming from the theoretical computer science community [26, 49, 61] this problem is still very much open, even if one only wants to find graphs without cliques or independent sets of size $O(\log n)$.

In Chapter 2 we discuss two topics that take a so-called local-global perspective on these two problems. The local-global principle states that one can obtain a global understanding of a structure from having a good understanding of its local properties or vice versa. This phenomenon has been ubiquitous in many areas of mathematics and beyond. See [20, 133, 140, 177, 233] for some examples.

Large independent sets from local considerations

The first of the above-mentioned topics will be discussed in Section 2.1. This section is concerned with the classical, extensively studied [40, 76, 77, 96, 147, 167, 227, 229, 240], Erdős-Rogers problem, dating back to 1962, can be restated as follows. How large must the independence number $\alpha(G)$ of a graph G be if every m vertices of G contain an independent set of size r? Note here that forbidding a clique of size m is equivalent to setting $r = 2$, which recovers the classical asymmetric Ramsey problem mentioned above. There are further, more fundamental links between these problems in the case of general r as well.

This restatement of the Erdős-Rogers problem is due to Erdős and Hajnal [82] and independently Linial and Rabinovich [178] from the early '90s. It arose from a change in perspective, which shifts the focus from fixing $\alpha(G)$ and r to fixing the local parameters m and r instead. In other words, they were interested in how large must the independence number of the whole graph be if we know that any small number of vertices m contain an independent set of size r. In fact, one can define an m-local independence number of a graph G to be the minimum independence number we can find among subgraphs of G on m vertices, and the problem becomes relating the local independence number to the independence number of G itself, the "global" independence number.

Our contribution to this area is a development of two new methods for attacking the local to global independence number problem, which allowed us to significantly improve previously best known bounds due to, depending on the regime, Linial and Rabinovich [178], Erdős and Hajnal [82], Alon and Sudakov [17], Krivelevich [168], and Kostochka and Jancey [163].

As an example, our methods allowed us to prove a conjecture of Erdős and Hajnal from 1991, stating that any graph G on n vertices with the property that any 7 vertices contain an independent set of size 3 must have $\alpha(G) \geq \Omega(n^{1/3+\varepsilon})$. In particular, we show such graphs must have $\alpha(G) \geq \Omega(n^{5/12-o(1)})$. This is an improvement in the exponent over the previously best bound of $n^{1/3}$. This is illustrative of our improvements in general, which also, in most cases, improve exponents when compared to previously best bounds. This kind of improvements are exceedingly rare in this type of Ramsey problems. For example, the above-mentioned closely related classical asymmetric Ramsey number problem has not seen such an improvement in the exponent since the initial result of Erdős and Szekeres [97] from 1935.

Our methods connect various regimes of the local to global independence number problem to a number of other well-studied problems in combinatorics. In addition to the above-mentioned Ramsey problem for a fixed graph versus a large independent set, these include the stability problem for Turán's theorem, the independence number in sparse triangle-free graphs and a 2-density variant of Turán's theorem, corresponding to a natural random process.

Locally Ramsey graphs

The second, closely related topic discussed in Section 2.2 is concerned with local properties of *Ramsey graphs*. A graph is said to be *k-Ramsey* if it contains neither a clique nor an independent set of size k. There has been a lot of work towards better understanding and constructing new Ramsey graphs, with several recent, major breakthroughs coming from the theoretical computer science community [26, 49, 61]. In order to better understand local properties of Ramsey graphs, in 1988 Erdős and Hajnal [82] introduced the concept of an *m-locally k-Ramsey* graph, defined as a graph in which any subset consisting of m vertices contains *both* a clique and an independent set of size k.

There are strong links between the concepts of locally Ramsey graphs and (global) Ramsey graphs. In particular, any *m-locally k-Ramsey* graph, upon minor modification, is also Ramsey with an appropriate parameter. Given this, the natural candidate for a good locally Ramsey graph is the random graph, the best-known example of usual Ramsey graphs. Indeed, prior to our work, this was the best-known example of a locally Ramsey graph.

Our main contribution here is a new way of constructing locally Ramsey graphs which produces examples with significantly better local Ramsey properties compared to the previously best-known examples coming from random graphs. Another interesting aspect of our graphs is that, as already mentioned, the fact they are good locally Ramsey graphs implies they are very good usual Ramsey graphs as well. They are better Ramsey graphs than some of the recent, far more involved constructions coming from the theoretical computer science community [26]. While, unfortunately, we do use randomness in our construction, making them non-explicit, they are still (even on a local level) very different from random graphs.

In Chapter 3 we discuss two further topics showcasing instances of the local to global principle in extremal and probabilistic combinatorics.

Covering graphs by monochromatic trees and Helly-type results for hypergraphs

In Section 3.1 we consider the following type of questions. How many monochromatic paths, cycles, or general trees does one need to cover all vertices of a given r-edge-coloured graph G? These problems were introduced in the 1960s [124] and were intensively studied by various researchers over the last 60 years. For some examples, see [87, 135, 136, 138, 144, 181, 200] and references therein.

One might object that this question at first glance has no local-global flavour. However, one of our main contributions to the area is a connection between this problem and the following natural local to global, Helly-type question for hypergraphs.

What is the maximum number of vertices needed to cover all the edges of a hypergraph H if it is known that any collection of a few edges of H has a small cover? This specific problem was raised by Erdős, Hajnal and Tuza [100] about 30 years ago.

We obtain quite accurate bounds for the hypergraph problem and use them to give some unexpected answers to several questions about covering graphs by monochromatic trees raised and studied by Bal and DeBiasio [21], Kohayakawa, Mota and Schacht [158], Lang and Lo [174], and Girão, Letzter and Sahasrabudhe [125].

As a specific example, it was generally believed that the threshold for the random graph $\mathcal{G}(n, p)$ to have a cover using r monochromatic trees in any r-colouring of its edges is around the point at which any $r + 1$ vertices have a common neighbour. Using our connection to the hypergraph problem, we show that this is exponentially far from the truth and that the correct threshold is only at the point when any $2^{r(1+o(1))}$ vertices have a common neighbour. This answered a question of Kohayakawa, Mota and Schacht.

Tournament quasirandomness from local counting

The second topic discussed in this chapter is concerned with quasirandom tournaments, and we discuss it in Section 3.2. A well-known theorem of Chung and Graham [57] states that if $h \geq 4$ then a tournament T is quasirandom if and only if T contains each h-vertex tournament the "correct number" of times as a subtournament. We discuss the relationship between quasirandomness of T and the count of a *single* h-vertex tournament H in T. We consider two types of counts, the global one and the local one.

We first observe that if T has the correct *global* count of H and $h \geq 7$, then quasirandomness of T is only forced if H is transitive. The next natural question when studying quasirandom objects asks whether possessing the correct *local* counts of H is enough to force quasirandomness of T. A tournament H is said to be locally forcing if it has this property. Variants of the local forcing problem have been studied before in both the graph and hypergraph settings. Perhaps the closest analogue of our problem was considered by Simonovits and Sós [222] who looked at whether having "correct counts" of a fixed graph H as an induced subgraph of G implies G must be quasirandom, in an appropriate sense. They proved that this is indeed the case when H is regular and conjectured that it holds for all H (with the exception of the path on three vertices).

Contrary to the Simonovits-Sós conjecture, in the tournament setting, we prove that a constant proportion of all tournaments are not locally forcing. In fact, any locally forcing tournament must itself be strongly quasirandom. On the other hand, unlike in the globally forcing case, we construct infinite families of non-transitive locally forcing tournaments.

Erdős-Szekeres theorem for multidimensional arrays

In the subsequent Chapter 4 we will focus on another cornerstone result in Ramsey theory, namely the Erdős-Szekeres theorem [97] dating back almost a hundred years. It states that any sequence of $(n-1)^2+1$ distinct real numbers contains a monotone subsequence of length n. This theorem has been generalised to higher dimensions in a variety of ways [47, 48, 152, 171, 179, 187, 219, 231], but perhaps the most natural one was proposed by Fishburn and Graham [105] about 25 years ago. They defined the concept of a monotone and a lex-monotone array and asked how large an array one needs in order to be able to find a monotone or a lex-monotone subarray of size $n \times \ldots \times n$.

Their definitions are as follows. A d-dimensional array is said to be *monotone* if, for each dimension, all the 1-dimensional subarrays along the direction of this dimension are increasing or are all decreasing. A d-dimensional array is *lex-monotone* if it is possible to permute the coordinates and reflect the array along some dimensions to obtain the lexicographically ordered array.

Fishburn and Graham obtained Ackermann-type bounds for their problem in both cases. Our main contribution here is a significant improvement over these bounds. Regardless of the dimension, we obtain at most a triple exponential bound in n in the monotone case and a quadruple exponential one in the lex-monotone case, as well as prove an intermediate conjecture they pose. The key ingredient of our approach is a connection we observe between the high-dimensional Erdős-Szekeres problem and classical Ramsey theory.

We also connect the problem to a number of interesting topics considered by various researchers over the years, including Ramsey type problems for vertex-ordered graphs, a canonical ordering of discrete structures and the long common subsequence of permutations problem.

Rota's basis conjecture for matroids.

Chapter 5 is concerned with Rota's basis conjecture [146] dating back to 1989 and stating the following. Given n bases $B_1, \ldots, B_n$ in an n-dimensional vector space V, one can always find n disjoint bases of V, each containing exactly one element from each B_i (we call such bases *transversal bases*). Rota's basis conjecture remains wide open despite its apparent simplicity and the efforts of many researchers [8, 37, 51, 74, 75, 121, 127, 196, 239] over the last 30 years. For example, the conjecture was recently the subject of a collaborative "Polymath" project [202] in which amateur and professional mathematicians from around the world collaborated on the problem. Our contribution here is a "halfway" result towards this conjecture. Namely, we show that one can always find $(1/2 - o(1))\, n$ disjoint transversal bases. This is the first linear bound on the number of disjoint transversal bases and improves on the previously best bound of $\Omega(n/\log n)$ established using a beautiful probabilistic argument by Dong and Geelen [74].

The conjecture, as well as our results, extend to the setting of matroids. Whitney introduced matroids [238] in 1935 as generalisations of vector spaces. They have found applications in geometry, topology, combinatorial optimisation, network theory and coding theory. Rota's conjecture itself has surprising connections to apparently unrelated subjects. Specifically, there are implications (see [146]) between Rota's basis conjecture, the Alon–Tarsi conjecture [18] concerning the enumeration of even and odd Latin squares, and a conjecture concerning the supersymmetric bracket algebra.

In Chapter 6 we discuss two old problems in extremal set theory. Extremal set theory is one of the main branches of extremal combinatorics and studies extremal questions involving families of sets, which one might view as hypergraphs. It has close ties to a number of other areas, including, in particular, probability theory.

Intersection spectrum of 3-chromatic intersecting hypergraphs

For a hypergraph H, define its intersection spectrum $I(H)$ as the set of all intersection sizes $|E \cap F|$ of distinct edges $E, F \in E(H)$. In their seminal paper [94] from 1973, which introduced the local lemma, Erdős and Lovász asked: how large must the intersection spectrum of a k-uniform 3-chromatic intersecting hypergraph be?

They showed that such a hypergraph must have at least three intersection sizes and conjectured that the size of the intersection spectrum tends to infinity with k. Despite the problem being reiterated several times over the years by Erdős and other researchers [59, 78, 83, 99, 208], the lower bound of three intersection sizes has remarkably withstood any improvement until our work. Our main contribution here is a proof of the Erdős–Lovász conjecture in a very strong form. Namely, we show that there are at least $k^{1/2-o(1)}$ intersection sizes. Our proof consists of a delicate interplay between Ramsey type arguments and a density increment approach.

Minimum saturated families of sets

One of the most studied concepts in extremal set theory is that of an intersecting family, defined as a family of sets in which any two sets have a non-empty intersection, with the classical Erdős-Ko-Rado theorem [93] from 1961 being perhaps the most famous example.

A folklore observation says that any maximal intersecting (in the sense that we can not add a new set to the family and keep it intersecting) family of subsets of an n-element set has size exactly 2^{n-1}. This motivated a number of extensions and generalisations over the years, for some examples see [9, 10, 14, 41, 73, 93, 104, 116, 128, 155, 220].

Observing that an equivalent definition of being intersecting is for the family to not contain two sets which are disjoint, Erdős and Kleitman [92] gave the following generalisation of the concept of a maximal intersecting family. A family of subsets of an n-element set is said to be *s-saturated* if it contains no s pairwise disjoint sets, and moreover, no set can be added to the family while preserving this property.

Almost 50 years ago, Erdős and Kleitman conjectured that an s-saturated family of subsets of an n-element set has size at least $(1-2^{-(s-1)})2^n$. It is easy to show that every s-saturated family has size at least $\frac{1}{2} \cdot 2^n$, but, as was mentioned by Frankl and Tokushige [112], even obtaining a slightly better bound of $(1/2+\varepsilon)2^n$, for some fixed $\varepsilon > 0$, seems difficult. Our main contribution is a proof of such a result. We show that every s-saturated family of subsets of $[n]$ has size at least $(1-1/s)2^n$. We actually resolve completely a certain multipartite version of this problem, from which the above bound follows.

Our proof makes use of algebraic techniques, which arose through a connection that we established between the problem of Erdős and Kleitman and a classical problem in probability theory involving correlation inequalities, namely the Van den Berg-Kesten-Reimer inequality [211, 237].

We give more detailed introductions, including the history of each of the topics mentioned here in their corresponding sections.

2

RAMSEY TYPE PROBLEMS

2.1 LARGE INDEPENDENT SETS FROM LOCAL CONSIDERATIONS

2.1.1 *Introduction*

In this section we study the following classical problem. If we know that any m vertices of a graph contain an independent set of order r how large can the independence number of the whole graph be? The study of this problem for specific choice of parameters dates back almost 60 years, with the first published result being due to Erdős and Rogers [96] in 1962.

Over the years this problem attracted a lot of attention. Originally the focus was on the instance of the problem in which we keep the sizes of independent sets we want to find locally and in the whole graph to be fixed and small. In other words if we forbid in G an independent set of size s how big a subset of vertices one can find without an independent set of size r? Choosing $r = 2$ precisely recovers the usual Ramsey problem and was in fact the original motivation behind the general question. This question became known as the Erdős-Rogers problem and has been extensively studied, for some examples see [40, 76, 96, 147, 167, 227, 229, 240] and a recent survey [77] due to Dudek and Rödl.

In the early 90's Erdős and Hajnal [82] and independently Linial and Rabinovich [178] propose changing the perspective and fixing the local parameters m and r instead. In other words, asking what can be said about the independence number of the whole graph if we know that any small number of vertices m contain an independent set of size r. This frames the problem squarely under the so called local-global principle, stating that one can obtain global understanding of a structure from having a good understanding of its local properties, or vice versa. This phenomenon has been ubiquitous in many areas of mathematics and beyond, see e.g. [20, 133, 140, 177]. In fact one can define an m-local independence number $\alpha_m(G)$ of a graph G to be the minimum independence number we can find among subgraphs of G on m vertices and the problem becomes relating the local independence number to the independence number of G itself, the "global" independence number. In particular, we are interested in the smallest possible size of $\alpha(G)$ in an n-vertex graph satisfying $\alpha_m(G) \geq r$.

In this section we discuss two new approaches for attacking this problem, which allow us to significantly improve previously best known bounds due to Linial and Rabinovich [178], Erdős-Hajnal [82], Alon and Sudakov [17], Krivelevich [168] and Kostochka and Jancey [163]. In the case of lower bounds we improve their results for at least half of the possible choices of m and r and in the case of upper bounds for essentially all choices. Moreover, we believe that both approaches have potential for further improvements.

The initial approach of Linial and Rabinovich [178] and independently Alon and Sudakov [17] reduces the lower bound problem to the question of bounding from above Ramsey numbers of a clique of size $k = \lceil \frac{m}{r-1} \rceil$ vs a large independent set. Our new idea is that one can find other "forbidden" graphs whose Ramsey numbers perform better. For this to work we need to obtain upper bounds on the Ramsey numbers of our new graphs vs a large independent set, which often turns out to be an interesting problem in its own right.

The above introduced parameter k controls in large part the known lower bounds for $\alpha(G)$ among all graphs satisfying $\alpha_m(G) \geq r$. Linial and Rabinovich [178] determine the answer precisely if $k \leq 2$, i.e. for $m \leq 2r-2$. For $k = 3$, they show that an n vertex graph satisfying $\alpha_m(G) \geq r$ must have $\alpha(G) \geq n^{1-\frac{2}{r-1}-o(1)}$ if $m = 2r-1$ and $\alpha(G) \geq \Omega(n^{1/2})$ for the rest of the range $m \leq 3r-3$. Our first result improves the exponent in their bounds for the first half of this range. Moreover, the improvement in the exponent is by a constant factor independent of r, unless $m = 2r-1$.

Proposition 2.1.1. *Let $m = 2r-2+t$ for $1 \leq t \leq r-1$. Then any n-vertex graph G satisfying $\alpha_m(G) \geq r$ has $\alpha(G) \geq \Omega(n^{1-1/\ell})$, where $\ell = \left\lfloor \frac{r-1}{t} \right\rfloor + 1$.*

In the general case of $k \geq 4$, Linial and Rabinovich and independently Alon and Sudakov show that an n-vertex graph satisfying $\alpha_m(G) \geq r$ must have $\alpha(G) \geq \Omega(n^{\frac{1}{k-1}})$. We improve the exponent in these bounds for the first half of the range for any k.

Theorem 2.1.2. *Let $k = \lceil \frac{m}{r-1} \rceil$ and let us assume $m \leq (k-\frac{1}{2})(r-1)$. Then any n-vertex graph G satisfying $\alpha_m(G) \geq r$ has $\alpha(G) \geq \Omega(n^{\frac{1}{k-3/2}})$.*

Going beyond $k-2$ in the denominator of the exponent in the above theorem seems likely to require an improvement over the best known upper bounds on Ramsey numbers, which have not seen an improvement in the exponent since the initial paper of Erdős and Szekeres [97] from 1935. This means our result is in some sense half-way between previously best bound and the Ramsey barrier.

The key part of the above result is actually the special case of $r = 3$. This is due to an easy observation which allows us to generalise any improvement in this case to the first half of the range as above, for any r. The first interesting instance here, which actually lead us to the general improvements above, is $m = 7$ and $r = 3$ in which case we can obtain an even better bound. Studying this case was explicitly proposed by Erdős and Hajnal [82] who observed that any graph G on n vertices with $\alpha_7(G) \geq 3$ must have $\alpha(G) \geq \Omega(n^{1/3})$ and that such a graph G exists with $\alpha(G) \leq O(n^{1/2})$. They conjectured that neither of these bounds is tight. Our next result confirms their first conjecture.

Theorem 2.1.3. *Any n-vertex graph G with $\alpha_7(G) \geq 3$ has $\alpha(G) \geq n^{5/12-o(1)}$.*

By the aforementioned observation this actually gives the same improved bound for first half of the range for any instance with $k = 4$, i.e. for $3r - 2 \leq m \leq 3.5(r-1)$.

To prove an upper bound on the minimum possible $\alpha(G)$ among all graphs with $\alpha_m(G) \geq r$, one needs to find a graph which has small independence number while having big independent sets spread around everywhere. Given the close relation of our problem to Ramsey numbers, random graphs are natural candidates for such examples. Understanding $\alpha_m(G)$ in $G \sim \mathcal{G}(n,p)$ turns out to be an interesting problem in its own right. Observe that the requirement $\alpha_m(G) \geq r$ may be rephrased as stating that G contains no copy of an m-vertex graph H with $\alpha(H) \leq r-1$ as a subgraph. A standard application of Lovász local lemma tells us that if we are only forbidding a single graph H then the largest p we can take is controlled by the 2-density of H. If we are instead forbidding a family of graphs the correct parameter turns out to be the minimum of the 2-densities over all graphs in our family. This reduces our problem to the following natural extremal question, which we propose to study. What is the minimum value of the 2-density of an m-vertex graph H with $\alpha(H) \leq r-1$? If we denote the answer to this question by $M(m,r)$ the above discussion leads us to the following reduction.

Proposition 2.1.4. *Let m, r be fixed, $m \geq 2r-1 \geq 3$ and $M = M(m,r)$. Then for any n there exists an n-vertex graph G with $\alpha_m(G) \geq r$ and $\alpha(G) \leq n^{1/M+o(1)}$.*

The value of $M(m,r)$, and hence also our upper bounds for the local to global independence number problem, are mostly controlled by the same parameter $k = \lceil \frac{m}{r-1} \rceil$ as before. Some intuition behind this, suggested by Linial and Rabinovich [178], is that a natural example of an m vertex graph with independence number at most $r-1$ is a vertex disjoint union of $r-1$ cliques with sizes as equal as possible (in other words complement of a

Turán graph on m vertices with no clique of size r). This graph clearly has no independent set of size r and we picked the clique sizes as equal as possible in order to minimise the 2-density. Our parameter k is simply the size of a largest clique in this example.

Turning to the results, we start once again with the range $2r-1 \leq m \leq 3r-3$, i.e. $k=3$. Here Linial and Rabinovich show that there exist n-vertex graphs G satisfying $\alpha_m(G) \geq r$ and $\alpha(G) \leq n^{1-1/(8r-4)}$. We improve the exponent in this bound for the whole range. Moreover, the improvement in the exponent is by a constant factor independent of r, towards the end of the range.

Proposition 2.1.5. *Let $m \geq 2r-1 \geq 3$, for any n there exists an n-vertex graph G satisfying $\alpha_m(G) \geq r$ with $\alpha(G) \leq n^{1-\frac{1}{2r-2}+o(1)}$ and if $m \geq 3r-4$ with $\alpha(G) \leq n^{\frac{3}{5}-\frac{2}{5r-13}+o(1)}$.*

These bounds follow from our results on $M(2r-1,r)$, which we determine precisely and $M(3r-4,r)$ which we determine up to lower order terms. This means that in terms of using random graphs as examples these bounds are essentially best possible for $m = 2r-1, 3r-4$. We can obtain a constant factor improvement in the exponent for about $1/3$ of the range, but since we believe our current argument does not give the best possible answer, in terms of $M(m,r)$, for the whole range we leave this open for future research.

The problem of determining $M(3r-4,r)$ is closely related to a well-studied problem of finding large independent sets in sparse triangle-free graphs. Perhaps the most famous result in this direction is due to Ajtai-Komlós-Szemerédi [11] and Shearer [217], but for our problem earlier results of Staton [225] and Jones [149] turn out to be more relevant. These results are part of a very active research area of studying graphs having no cliques of size k nor independent sets of size r but which have potentially much fewer vertices than the corresponding Ramsey number $R(k,r)$. Our problem of lower bounding $M(m,r)$ falls under this framework since we can always assume that our graphs, in addition to having no independent sets of size r, are also K_k-free or the 2-density is already large. We point the interested reader to classical papers [12, 225] and numerous papers citing them.

We now turn to the general case of $k \geq 4$. Let us begin with the initial instance, when $r=3$. Unfortunately, here the results for $r=3$ do not immediately generalise as they did in the case of lower bounds. They do however provide a starting point, which serves as a basis for more general results. Here we determine $M(m,3)$ precisely for all m, which allows us to improve exponents in the previously best bounds of Linial and Rabinovich [178]. They showed there are n-vertex graphs G with $\alpha_m(G) \geq 3$ and $\alpha(G) \leq n^{\frac{4+o(1)}{m-4/(m-2)}}$ if m is even and $\alpha(G) \leq n^{\frac{4+o(1)}{m-3/(m-2)}}$ if m is odd.

Theorem 2.1.6. *For any n there exists an n-vertex graph G satisfying $\alpha_m(G) \geq 3$ with $\alpha(G) \leq n^{\frac{4+o(1)}{m+2}}$ if m is even and $\alpha(G) \leq n^{\frac{4+o(1)}{m+3-13/\sqrt{m}}}$ when $m \geq 5$ is odd (here both terms $o(1) \to 0$ as $n \to \infty$).*

We remark that in the even case any improvement of our exponent, in terms of m, provably leads to improvement over the best known lower bounds on Ramsey numbers and in the odd case without improving the Ramsey numbers one can only improve the term $13/\sqrt{m}$.

Once again our arguments in this particular regime show that the problem of determining $M(m,r)$ is related to yet another well-studied problem. Namely, the stability problem for Turán's theorem first considered by Erdős, Győri and Simonovits in [88]. While one can use their results to obtain good bounds on $M(m,3)$ obtaining precise answers requires a different, more careful argument.

In the fully general case we determine $M(m,r)$ up to lower order terms (where r is considered fixed and m large), giving us the following result.

Theorem 2.1.7. *Let us assume m is sufficiently larger than r and set $k = \lceil \frac{m}{r-1} \rceil$. Then for any n there exists an n-vertex graph G satisfying $\alpha_m(G) \geq r$ with $\alpha(G) \leq n^{\frac{2+o(1)}{k+1-c_r/\sqrt{k}}}$.*

This improves previous upper bounds of Linial and Rabinovich from roughly $n^{\frac{2}{k-1}}$ to roughly $n^{\frac{2}{k+1}}$ and is once again essentially best possible assuming lower bounds on Ramsey numbers are tight.

While the above result requires m to be large compared to r some of our ideas apply for any choice of the parameters. To illustrate this we consider the case $m = 20, r = 5$ which was used by various researchers as a benchmark to compare their methods. Here Linial and Rabinovich show there are graphs G having $\alpha_{20}(G) \geq 5$ and $\alpha(G) \leq n^{18/39+o(1)}$. Krivelevich [168] improved this to $\alpha(G) \leq n^{14/33+o(1)}$. He obtains this as an application of his result on minimum number of edges in colour-critical graphs. The best possible bound using this approach was later obtained by Kostochka and Jancey [163] who showed $\alpha(G) \leq n^{18/43+o(1)}$. For comparison $39/18 \approx 2.17, 33/14 \approx 2.36$ and $43/18 \approx 2.39$, while our methods allow us to improve this to 3. That is, there exists a graph G with $\alpha_{20}(G) \geq 5$ which has $\alpha(G) \leq n^{1/3+o(1)}$ and once again this is best possible (up to $o(1)$ term) without improving the lower bounds on Ramsey numbers $R(5,s)$.

In addition to above applications and connections another reason which makes the study of $M(m,r)$ interesting is its relation to a random graph process. For a graph property $\mathcal{P}$ the random graph process with respect to $\mathcal{P}$ starts with an empty graph and iteratively adds a new uniformly random edge for as long as this does not violate $\mathcal{P}$. Random graph processes have been extensively studied for a variety of properties and have found numerous applications (see e.g. [35, 36, 106, 154, 169, 197] and references therein). In our setting $M(m,r)$ controls the final density of the random process w.r.t. m-local r-independence property $\alpha_m \geq r$. So $M(m,r)$ essentially controls the behaviour of this random process.

Notation. Whenever working with graphs satisfying $\alpha_m(G) \geq r$ all our asymptotics are with respect to $n = |G|$ and we treat m and r as constants unless otherwise specified. When working with directed graphs $N^{\pm}(v)$ denotes the in/out neighbourhood of v and $d^{\pm}(v)$ in/out degree.

2.1.2 *Local to global independence number*

In this section we prove our lower bounds on $\alpha(G)$ for a graph G satisfying $\alpha_m(G) \geq r$. We begin with our lower bound result for $k = \lceil \frac{m}{r-1} \rceil = 3$, namely Proposition 2.1.1.

Proposition 2.1.1. *Let $m = 2r-2+t$ for $1 \leq t \leq r-1$. Then any n-vertex graph G satisfying $\alpha_m(G) \geq r$ has $\alpha(G) \geq \Omega(n^{1-1/\ell})$, where $\ell = \left\lfloor \frac{r-1}{t} \right\rfloor + 1$.*

Proof. Let us first assume we can find a vertex disjoint collection of $a = t\ell - (r-1)$ cycles $C_{2\ell-1}$ and $t-a$ cycles $C_{2\ell+1}$. Their union makes a subgraph of G of order $a(2\ell-1) + (t-a)(2\ell+1) = (2\ell+1)t - 2a = 2r-2+t = m$ and has no independent set of size larger than $a(\ell-1) + (t-a)\ell = t\ell - a = r-1$. This gives us a contradiction to $\alpha_m(G) \geq r$, which means such a union does not exist in G.

If we can find a cycles $C_{2\ell-1}$ in G then the remainder of the graph can not contain $t-a$ cycles $C_{2\ell+1}$, this means that by removing at most m vertices from G we can find a subgraph which is $C_{2\ell+1}$-free. If there are less than a cycles $C_{2\ell-1}$ in G then we find a subgraph, again with at least $n-m$ vertices which is $C_{2\ell-1}$-free. In either case a classical result from [85] (see also [176, 228] for slight improvements) on cycle-complete Ramsey numbers tells us there is an independent set of size at least $\Omega((n-m)^{1-1/\ell}) = \Omega(n^{1-1/\ell})$, as desired. $\square$

We now show how to generalise any improvement made in the $r = 3$ case to half of the range for any k. It will be convenient to denote by $f(n, m, r)$ the smallest possible size of $\alpha(G)$ in an n vertex graph with $\alpha_m(G) \geq r$.

Lemma 2.1.8. *Let $k = \lceil \frac{m}{r-1} \rceil$ and $\ell = m - (k-1)(r-1)$. Provided $\ell \leq \frac{r-1}{2}$ we have $f(n, m, r) \geq \min(f(n-m, 2k-1, 3), f(n-m, k-1, 2))$.*

Proof. Let G be a graph on n vertices with $\alpha_m(G) \geq r$. Let us first assume that we can find a vertex disjoint union consisting of ℓ subgraphs on $2k-1$ vertices, each having no independent set of size 3, and $r-1-2\ell$ copies of K_{k-1}. This union is a subgraph on $\ell(2k-1) + (r-1-2\ell)(k-1) = (r-1)(k-1) + \ell = m$ vertices which has no independent set larger than $2\ell + r - 1 - 2\ell = r - 1$, contradicting $\alpha_m(G) \geq r$.

If we can find ℓ such subgraphs on $2k-1$ vertices this means that the remainder of the graph can not contain $r-1-2\ell$ copies of K_{k-1}. Removing our subgraphs and a maximal collection of K_{k-1}'s in the remainder we obtain a subgraph on at least $n-m$ vertices which is K_{k-1}-free, or in other words has $\alpha_{k-1} \geq 2$ implying $f(n, m, r) \geq f(n-m, k-1, 2)$. If there are less than ℓ such subgraphs we may remove a maximal collection and obtain a subgraph on at least $n-m$ vertices which contains no subgraphs on $2k-1$ vertices without an independent set of size 3. In other words our new subgraph has $\alpha_{2k-1} \geq 3$ so $f(n, m, r) \geq f(n-m, 2k-1, 3)$ as claimed. $\square$

Note that $f(n, k-1, 2) \geq \alpha$ is equivalent to $R(k-1, \alpha) \leq n$ implying that the best known bound is $f(n, k-1, 2) \geq n^{1/(k-2)-o(1)}$. This represents a natural barrier for our results since it seems very likely that $f(n, 2k-1, 3) \geq f(n, k-1, 2)$. On the other hand the results of [178] and [17] may be stated as $f(n, m, r) \geq f(n-m, k, 2)$. Therefore, obtaining a lower bound for $f(n, 2k-1, 3)$, better than $f(n, k, 2)$, immediately improves their bound whenever the above lemma applies, i.e. $\ell \leq \frac{r-1}{2}$. Our result in the next section gives a bound which is half-way (in terms of exponents) between the above bounds coming from Ramsey numbers of K_{k-1} and K_k vs large independent set.

2.1.2.1 *Independent sets of size three everywhere.*

In this subsection we will show how to find big independent sets in graphs satisfying $\alpha_{2k-1}(G) \geq 3$. This is inherently a Ramsey question in the following sense. How big a graph do we need take in order to guarantee that we can find an independent set of size α or a subgraph H with $2k-1$ vertices and no independent set of size 3? The approach of [178] and [17] is to always look for only a single graph H, namely a vertex disjoint union of K_k and K_{k-1}.

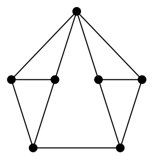


Figure 2.1: H_7

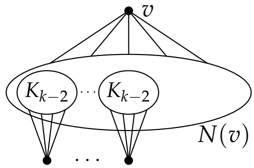


Figure 2.2: $\mathcal{M}$

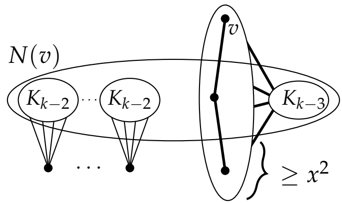


Figure 2.3: Extending $\mathcal{M}$

This H clearly has no independent set of size 3 and their approach actually further reduces to finding a copy of K_k. The reason is that if we can find a single copy of K_k then the rest of the graph has lost barely any vertices and yet must be K_{k-1}-free or we find H. Hence, this would give us an even better bound than just using K_k-freeness. Our key new ingredient is to in addition look for a different $2k-1$-vertex graph which we call H_{2k-1} and define to be a blow-up of C_5 with parts of sizes $1, k-2, 1, 1, k-2$ appearing in that order around the cycle, with cliques placed inside of parts (see Figure 2.1 for an illustration). Since the complement of this graph is an actual blow-up of C_5 it is triangle-free, implying that H_{2k-1} has no independent set of size 3 and is hence forbidden in any graph satisfying $\alpha_{2k-1}(G) \geq 3$.

We start by explaining the general idea behind our argument. As argued above our goal is to find in an arbitrary K_k-free and H_{2k-1}-free n-vertex graph G a large independent set. We will do so by finding a vertex v and a large collection of vertex disjoint K_{k-1}'s with $k-2$ vertices inside $N(v)$. We know that the remaining vertex of any such K_{k-1} lies outside $N(v)$ as our graph is K_k-free. Furthermore, we know that the set of these last vertices spans an independent set as otherwise any edge between such vertices together with their K_{k-1}'s and v make a copy of H_{2k-1} (see Figure 2.2 for an illustration). This gives us our desired large independent set.

The more difficult part of the argument is to actually find such a collection of K_{k-1}'s. The following two easy lemmas will help us control how many K_{k-1}'s we can find with $k-2$ vertices inside a neighbourhood of v and how many such K_{k-1}'s can intersect another one, respectively. Let us denote by $t_i(G)$ the number of copies of K_i in G, (we omit G when it is clear from context).

Lemma 2.1.9. *Let G be a graph with $\alpha(G) \leq \alpha$ and let $k \geq 2$. Provided $t_{k-1}(G) > 0$, we have*

$$\frac{t_k(G)}{t_{k-1}(G)} \geq \frac{|G|}{k\alpha^{k-1}} - 1.$$

Proof. We will prove the claim by induction on k. For the base case of $k = 2$ (note that $t_1 = |G|$) the claim follows from Turán's theorem which gives $e(G) \geq \frac{n^2}{2\alpha} - \frac{n}{2}$ (see [16]). Let us now assume $k \geq 3$ and that the claim holds for $k-1$. Given $S \subseteq V(G)$ let us denote by $e(S)$ the number of edges in the common neighbourhood of S and by $d(S)$ the number of common neighbours of S. Then we have

$$\begin{aligned}\binom{k}{2} t_k &= \textstyle\sum_{G[S]=K_{k-2}} e(S) \geq \sum_{G[S]=K_{k-2}} \left(\frac{(d(S))^2}{2\alpha} - \frac{d(S)}{2}\right) \\ &\geq \frac{\left(\sum_{G[S]=K_{k-2}} d(S)\right)^2}{2t_{k-2}\alpha} - \textstyle\sum_{G[S]=K_{k-2}} \frac{d(S)}{2} \\ &= \frac{(k-1)^2 t_{k-1}^2}{2t_{k-2}\alpha} - \frac{(k-1)t_{k-1}}{2}\end{aligned}$$

Where we used Turan's theorem within common neighbourhood of S in the first inequality and Cauchy-Schwarz inequality for the second. Dividing by $(k-1)t_{k-1}/2$ and using the induction assumption we obtain

$$\frac{kt_k}{t_{k-1}} \geq \frac{(k-1)t_{k-1}}{\alpha t_{k-2}} - 1 \geq \frac{|G|}{\alpha^{k-1}} - \frac{k-1}{\alpha} - 1 \geq \frac{|G|}{\alpha^{k-1}} - k.$$

□

Lemma 2.1.10. *Let G be a K_k-free graph with $\alpha(G) < \alpha$. Then for any $i \leq k$ we have*

$$t_i(G) \leq \frac{1}{i!} \cdot \alpha^{\binom{k}{2} - \binom{k-i}{2}}.$$

Proof. We prove the claim by induction on i. For the base case of $i = 1$ the claim is equivalent to $t_1 = |G| \leq \alpha^{k-1}$ which holds by the classical bound on the Ramsey number $R(k, \alpha)$. Let us now assume the claim holds for $i - 1$. Given a subset of vertices S we denote by $N(S)$ the set of common neighbours of S and by $d(S) = |N(S)|$. If $G[S] = K_{i-1}$ then $N(S)$ is K_{k-i+1}-free in addition to having no independent set of size α. So the same classical bound on Ramsey numbers as above implies $d(S) \leq \alpha^{k-i}$. Taking a sum and using the inductive assumption we obtain:

$$it_i = \sum_{G[S]=K_{i-1}} d(S) \leq \alpha^{k-i} t_{i-1} \leq \frac{1}{(i-1)!} \cdot \alpha^{k-i+\binom{k}{2}-\binom{k-i+1}{2}} \tag{2.1}$$

$$= \frac{1}{(i-1)!} \cdot \alpha^{\binom{k}{2}-\binom{k-i}{2}}. \tag{2.2}$$

□

We are now ready to prove our general result for $r = 3$.

Proof of Theorem 2.1.2. Our task is to show that in an n-vertex graph G which contains an independent set of size 3 among any $2k-1$ vertices we can find an independent set of size $\alpha = \Omega(n^{1/(k-3/2)})$. With this choice of α we may assume $n \geq C\alpha^{k-3/2}$ for an arbitrarily large constant C. Since the result for $k = 3$ holds by Proposition 2.1.1 we may assume $k \geq 4$.

As discussed above if G contains a K_k then the remainder of the graph has no K_{k-1}, so by the classical Ramsey bound we get $\alpha(G) \geq (n-k)^{1/(k-2)} \geq \Omega(n^{1/(k-3/2)})$. So we may assume G is K_k-free. Our goal is to find a vertex v and a collection $\mathcal{M}$ of α vertex disjoint K_{k-1}'s each with $k-2$ vertices in $N(v)$. Then, as we already explained above, any such K_{k-1} has exactly one vertex outside $N(v)$ as our graph is K_k-free and these vertices form an independent set as otherwise any edge between such vertices together with their K_{k-1}'s and v make a copy of H_{2k-1}, a contradiction. In order to do so we will analyse common neighbourhoods of cliques of size $k-3$ inside $N(v)$. Any edge we find inside such a common neighbourhood gives rise to a copy of K_{k-1} and if we find there a path of length 2 starting with v then the last edge of this path gives rise to a copy of K_{k-1} with exactly $k-2$ vertices in $N(v)$, which we are looking for.

Let us first give a lower bound on $T := (k-2)t_{k-2}$ (recall that t_{k-2} counts the number of K_{k-2}'s in G) which counts the number of extensions of a K_{k-3} into a K_{k-2}, i.e. the sum of sizes of common neighbourhoods of K_{k-3}'s in our graph. It is a simple application of Lemma 2.1.9.

$$T = (k-2)t_{k-2} \geq t_{k-3} \cdot \left(\frac{n}{\alpha^{k-3}} - k + 2\right) \geq t_{k-3} \cdot \frac{n}{2\alpha^{k-3}}, \tag{2.3}$$

where in the second inequality we used $n \geq 2k\alpha^{k-3}$ which holds for our choice of α. Using Lemma 2.1.9 repeatedly in a similar way we get the following lower bound on T which will be useful later in the argument.

$$\begin{aligned} T = (k-2)t_{k-2} &= (k-2) \cdot \frac{t_{k-2}}{t_{k-3}} \cdots \frac{t_2}{t_1} \cdot n \\ &\geq (k-2) \cdot \frac{n}{2(k-2)\alpha^{k-3}} \cdots \frac{n}{4\alpha} \cdot n = \frac{n^{k-2}}{2^{k-3}(k-3)!\alpha^{\binom{k-2}{2}}} \end{aligned} \tag{2.4}$$

In order to carry out the above proof strategy, for every clique of size $k-3$ we are going to restrict our attention only to a large part of its common neighbourhood where independent sets expand (meaning they have many vertices adjacent to some vertex of the set). This will achieve our goal since G

being K_k-free means that the common neighbourhood of a K_{k-3} is triangle-free, so part of this common neighbourhood inside $N(v)$ is an independent set. Hence, expansion of this independent set precisely means there are many endpoints of a path of length 2 starting with v, giving many choices for a K_{k-1} with $k-2$ vertices in $N(v)$.

To obtain such an expansion we are going to restrict attention to K_{k-3}'s which have common neighbourhood of order at least half the average. By (2.3), this means it is at least $d := n/(4\alpha^{k-3})$. We will call such a K_{k-3} *average* and we know that altogether there are at least $T/2$ ways of extending average K_{k-3}'s into a K_{k-2}.

We now proceed to obtain independent set expansion inside the common neighbourhood of each average K_{k-3}. Let S be such a K_{k-3}. As long as we can find a subset X of the common neighbourhood N of S which has less than $\frac{d-2\alpha}{2\alpha} \cdot |X|$ neighbours inside N we add X to our current independent set, remove its neighbours and repeat. We either find an independent set of size α or we have removed at most $\alpha + \frac{d-2\alpha}{2\alpha} \cdot \alpha = \frac{d}{2}$ vertices from N. Any remaining vertex in N is called an *expanding neighbour* of S. Since S was arbitrary, every average K_{k-3} has at least $d/2$ expanding neighbours and inside its expanding neighbourhood independent sets expand by a factor of $\frac{d-2\alpha}{2\alpha} \geq \frac{d}{4\alpha} = \frac{n}{16\alpha^{k-2}} =: d'$. In particular, every vertex in N (being an independent set of size one) has degree at least d' in this set. Furthermore, there are still at least $T/4$ ways to extend an average K_{k-3} into a K_{k-2} using an expanding neighbour.

We now pick our v to be a vertex which is an expanding neighbour of $g_v \geq T/(4n)$ average K_{k-3}'s (such v exists by double counting and the above bound). Let $\mathcal{S}$ denote the collection consisting of all such K_{k-3}'s. Let us first observe some properties of an $S \in \mathcal{S}$. We denote by N_S its expanding neighbourhood and by $D_S := N(v) \cap N_S$ the set of its expanding neighbours inside $N(v)$. By definition, $v \in N_S$ for all $S \in \mathcal{S}$. Also, as was explained above, we know that v has at least d' neighbours within N_S, i.e. $|D_S| \geq d'$. These neighbours span an independent set (since they belong to the common neighbourhood of $k-2$ vertices in $v \cup S$) of size at least d', so they expand inside N_S. This gives us d'^2 different vertices which together with S and one of the vertices in D_S make a K_{k-1} with exactly $k-2$ vertices inside $N(v)$. To find many such disjoint K_{k-1}'s we will use the fact that there are in total $\sum_{S \in \mathcal{S}} |D_S| \geq g_v \cdot d'$ ways to extend K_{k-3}'s in $\mathcal{S}$ into a K_{k-2} using an expanding neighbour belonging to $N(v)$.

Let us now consider a maximal collection $\mathcal{M}$ of vertex disjoint K_{k-1}'s each with exactly $k-2$ vertices inside $N(v)$. Let us assume towards a contradiction that $|\mathcal{M}| < \alpha$. We will show below that if $|\mathcal{M}| < \alpha$ we can still find some

$S \in \mathcal{S}$ and a set $D'_S \subseteq D_S$ of at least $d'/2$ of its expanding neighbours in $N(v)$ such that both S and D'_S are vertex disjoint from all cliques in $\mathcal{M}$. For now, suppose we found such S and D'_S. Since $D'_S \subseteq D_S$ is an independent set (as we explained above), it expands within neighbourhood of S meaning that there are at least $d'^2/2 = \frac{n^2}{2^9\alpha^{2k-4}} \geq \alpha$ vertices which together with S and some vertex in D'_S make a K_{k-1} with $k-2$ vertices in $N(v)$. Moreover, note that all these $d'^2/2$ vertices lie outside of $N(v)$ or we get a K_k in G. Since we removed less than α vertices outside of $N(v)$ (recall that each $K_{k-1} \in \mathcal{M}$ has exactly one vertex outside $N(v)$) one of these vertices is disjoint from all cliques in $\mathcal{M}$ and gives rise to the desired copy of K_{k-1}, which is vertex disjoint from any clique in $\mathcal{M}$ (since both S and D'_S are chosen disjoint from any clique in $\mathcal{M}$) and hence contradicts its maximality.

Therefore, it remains to be shown that there is an $S \in \mathcal{S}$ and $d'/2$ of its expanding neighbours inside $N(v)$, all disjoint from any clique in $\mathcal{M}$. Note that cliques in $\mathcal{M}$ cover at most $\alpha(k-2)$ vertices inside $N(v)$. On the other hand note that any vertex $u \in N(v)$ can belong to at most $\alpha^{\binom{k-2}{2}}/(k-3)!$ copies of K_{k-2} inside $N(v)$. This follows from Lemma 2.1.10 since any such copy of K_{k-2} amounts to a copy of K_{k-3} in the common neighbourhood of v and u which spans a K_{k-2}-free graph with no independent set of size α (or we are done). On the other hand, any K_{k-2} can be an extension of at most $k-2$ different copies of K_{k-3} in $\mathcal{S}$ so there are at least $g_v d' - \alpha^{\binom{k-2}{2}+1}(k-2)^2/(k-3)!$ extensions of a K_{k-3} from $\mathcal{S}$ to a K_{k-2} inside $N(v)$ both disjoint from any clique in $\mathcal{M}$. Note that

$$\begin{aligned}\frac{g_v d'}{\alpha^{\binom{k-2}{2}+1}(k-2)^2/(k-3)!} &\geq \frac{n^{k-3}/(2^{k-1}\alpha^{\binom{k-2}{2}}) \cdot n/(16\alpha^{k-2})}{\alpha^{\binom{k-2}{2}+1}(k-2)^2} \\ &= \frac{n^{k-2}}{2^{k+3}(k-2)^2\alpha^{(k-2)^2+1}} \geq \frac{1}{2}\end{aligned}$$

where in the first inequality we used $g_v \geq T/(4n)$ and (2.4) to bound T, while in the last inequality we used that $n \geq C\alpha^{k-3/2} \geq C\alpha^{k-2+1/(k-2)}$. This means that there are at least $g_v d'/2$ such extensions and since $g_v = |\mathcal{S}|$ there must be a K_{k-3} with $d'/2$ extensions, as desired. □

2.1.2.2 *The Erdős-Hajnal (7,3) case.*

For $k = 4$ the result from the previous section implies that graphs with $\alpha_7 \geq 3$ have $\alpha \geq \Omega(n^{2/5})$ which already suffices to confirm the conjecture of Erdős and Hajnal [82]. In this section we show how to further improve this bound to $\alpha \geq n^{5/12-o(1)}$, i.e. we prove Theorem 2.1.3.

The general idea will be similar as in the previous subsection. Here since $k = 4$ we may assume our G satisfying $\alpha_7(G) \geq 3$ is K_4 and H_7-free. In fact in this case, $\alpha_7(G) \geq 3$ is essentially (up to removal of a few vertices) equivalent to G being K_4 and H_7-free. Since we do not need the non-obvious direction here, we prove it as Lemma 2.1.19 in the following section where it will be useful.

Unlike in the previous section, since desired α is bigger we will not be able to find a large enough set of vertex disjoint triangles (not containing v) with an edge in $N(v)$. However, these triangles will still play a major role in the argument. We will call them *v-triangles* and the vertex of a v-triangle not in $N(v)$ is going to be called a v-extending vertex (in other words any non-neighbour of v which belongs to a K_4 minus an edge together with v). While we can not find a large enough collection of disjoint v-triangles the fact our graph is H_7-free imposes many restrictions on the subgraph induced by v-extending vertices, which we call $N_\triangle(v)$. The following lemma establishes the properties of $N_\triangle(v)$ that we will use in our argument.

Lemma 2.1.11. *Let G be a K_4 and H_7-free graph and $v \in G$. Then*

a) $N_\triangle(v) \cap N(v) = \emptyset$.

b) If $u, w \in N_\triangle(v)$ belong to vertex disjoint v-triangles then $u \nsim w$.

c) $N_\triangle(v)$ is triangle-free.

d) Let C be a connected subgraph of $N_\triangle(v)$ consisting only of vertices belonging to at least 5 different v-triangles. If $|C| \geq 2$ then there exists $u \in N(v)$ such that for any $w \in C$ $N(v) \cap N(w)$ induces a star centred at u in G.

Proof. Part a) is immediate since G is K_4-free and part b) since it is H_7-free.

For part c) assume to the contrary that there is a triangle x, y, z in $N_\triangle(v)$ and that f, g, h are edges in $N(v)$ completing a v-triangle with x, y, z respectively. By part b) any two of f, g, h need to intersect. This is only possible if they make a triangle (in which case together with v they make a K_4) or if they make a star, in which case the centre of the star together with x, y, z makes a K_4, either way we obtain a contradiction.

For part d) let $u, w \in N_\triangle(v)$ be adjacent and belong to at least 5 different v-triangles. We claim that then $(N(w) \cap N(v)) \cup (N(u) \cap N(v))$ make a star in G. Indeed if $N(u) \cap N(v)$ would contain two disjoint edges then by part b) any edge in $N(w) \cap N(v)$, of which there are at least 5 by assumption, must intersect them both. Since there can be at most 4 edges which intersect each of two disjoint edges, we conclude there can be no disjoint edges in $N(u) \cap N(v)$. This means that they span either a star or a triangle. Since there are at least 5 edges, the former must occur. We can repeat for w in place of u and observe that the only way for each pair of edges, one per star, to intersect is that they share the centre, as claimed. Propagating along any path in C we deduce that the same holds for any pair of vertices in C. □

The following corollary, based mostly on part d) of the above lemma, allows us to partition $N_\triangle(v)$ into three parts which we will deal with separately in our argument.

Corollary 2.1.12. *Let G be a K_4 and H_7-free graph and $v \in G$. Then there exists a partition of $N_\triangle(v)$ into three sets L, I and C with the following properties:*

a) L consists only of vertices belonging to at most 4 different v-triangles.

b) I is an independent set.

c) C can be further partitioned into $C_1, \dots, C_m$ such that there are no edges between different C_i's and for every C_i there is a distinct $v_i \in N(v)$ such that any v-triangle containing a vertex from C_i must contain v_i as well.

Proof. We chose L to consist of all v-extending vertices belonging to at most 4 v-triangles. We chose I to consist of isolated vertices in $G[N_\triangle(v) \setminus L]$ and $C = N_\triangle(v) \setminus (L \cup I)$. Note that C is a union of connected components of $G[N_\triangle(v) \setminus L]$, which we denote by $C'_1, C'_2, \dots$ each of order at least 2 and consisting entirely of vertices belonging to at least 5 different v-triangles. In particular, Lemma 2.1.11 part d) implies that there exists a vertex $v'_i \in N(v)$ such that any v-triangle containing a vertex from C'_i must also contain v'_i. Finally, we merge any C'_is which have the same vertex for their v'_i to obtain the desired partition $C_1, \dots, C_m$ of C. □

We are now ready to prove Theorem 2.1.3.

Proof of Theorem 2.1.3. Let G be an n-vertex graph with $\alpha_7(G) \geq 3$. Our task is to show it has an independent set of size $\alpha = n^{5/12-o(1)}$. If G contains a K_4 the remainder of the graph must be triangle-free so has an independent set of size $\Omega(\sqrt{n}) > \alpha$. Hence, we may assume G is K_4-free as well as H_7-free.

We begin by ensuring minimum degree is high, so that we can ensure good independent set expansion inside neighbourhoods. We repeatedly remove any vertex with degree at most $2n/\alpha$. Observe that if we remove more than half of the vertices, then the removed vertices induce a subgraph with at least $n/2$ vertices and at most $n \cdot 2n/\alpha$ edges. So, Turán's theorem implies there is an independent set of size at least $\frac{n/2}{8n/\alpha+1} \geq \Omega(\alpha)$ and we are done. Let us hence assume that G has minimum degree at least $2n/\alpha$ (we technically need to pass to a subgraph on at least $n/2$ vertices but this only impacts the constants).

Let us fix a vertex v and do the following. As long as we can, we find an independent set X inside $N(v)$, which has less than $n/\alpha^2 \cdot |X|$ neighbours inside $N(v)$, we add X to our current independent set, remove its neighbours from $N(v)$ and continue. We either find an independent set of size $\alpha/2$ and are done or we have removed at most $\alpha/2 + n/\alpha^2 \cdot \alpha/2 \leq n/\alpha$ vertices from the neighbourhood. We direct[1] all edges from v towards the remaining vertices in $N(v)$. Note that $d^+(v) \geq d(v)/2 \geq n/\alpha$ and inside $N^+(v)$ independent sets expand by at least a factor of $x := n/\alpha^2 \geq \Omega(n^{2/12})$. We repeat for every vertex v, and note that some edges of G might be assigned both directions, while some none.

Our first goal is to show there are in total many v-triangles for which their edge in $N(v)$ belongs to $N^-(v)$ and the remaining two edges of the triangle are directed away from this edge. We will call such a v-triangle a *directed v-triangle*. In other words this counts the number of K_4's minus an edge with the 4 edges incident to the missing edge all being directed towards vertices of the missing edge. We denote this count by T_4.

Claim. *Unless $\alpha(G) \geq \Omega(n^{5/12})$ we have $T_4 \geq \Omega(n^2)$.*

Proof. We will show that for any vertex v with in-degree at least half the average we get at least $\Omega(d^-(v)^2/n^{2/12})$ directed v-triangles. Let us for now assume this holds. Note that v with lower in-degree contribute at most half to the value of $\sum_{v \in G} d^-(v) = \sum_{v \in G} d^+(v) \geq n^2/\alpha \geq \Omega(n^{19/12})$ (recall that $d^+(v) \geq n/\alpha$). An application of Cauchy-Schwarz implies there are at least $\Omega((\sum_{v \in G} d^-(v))^2/n \cdot n^{-2/12}) \geq \Omega(n^2)$ directed triangles.

Let us now fix a vertex v with in-degree at least half the average, this in particular means $d^-(v) \geq n/(2\alpha) \geq n^{7/12}$. We know that any vertex $u \in N^-(v)$ has v as an out-neighbour which means that v, as a single vertex independent set, expands inside $N^+(u)$. So v has at least x neighbours inside $N^+(u)$, i.e. u has at least x out-neighbours inside $N(v)$. Since u was an arbitrary vertex in $N^-(v)$ this means there are at least $d^-(v)x \geq \Omega(n^{9/12})$ edges inside $N(v)$ which are directed away from a vertex in $N^-(v)$ (note that if an edge of G has both directions and is inside of $N^-(v)$ it is counted twice and considered as 2 directed edges). Let us call the set of such directed edges M, so in particular $|M| \geq d^-(v)x \geq \Omega(n^{9/12})$. Note further that given such a directed edge $uw \in M$, since u has w as a single vertex independent set in its out-neighbourhood w expands there. This means that our edge lies in at least x v-triangles with both edges incident to u directed away from

1 The assignment of directions is simply a convenient way to encode the information about in what part of $N(v)$ we know independent sets expand. An out-neighbour corresponds to an expanding neighbour in the previous argument.

u. We call the third vertex of any such triangle uw-extending, note that it belongs to $N_\triangle(v)$. We are now going to assign types to edges in M according to where, inside $N_\triangle(v)$, we find the majority of their extending vertices.

Let us fix a partition of $N_\triangle(v)$ into C, L and I, provided by Corollary 2.1.12. Given a directed edge $uw \in M$ we say it is of type L or C if it has at least $x/3$ extending neighbours in L or C, respectively. We say it is of type I if it was not yet assigned a type. In particular, an edge of type I also has at least $x/3$ extending vertices in I (since we have shown above it has at least x in total), but we also know it has at most $2x/3$ extending vertices in other parts of $N_\triangle(v)$.

First case: at least a third of the edges in M are of type C.
Any vertex in $N(v)$ is adjacent to less than α other vertices inside $N(v)$ (since these vertices make an independent set as G is K_4-free). This means there is a matching M' of at least $|M|/(6\alpha)$ edges of type C, as otherwise vertices making a maximal matching are incident to less than $|M|/3$ edges so it can be extended. Let us denote by N_e the set of extending vertices of an edge $e \in M'$ inside C. By assumption $|N_e| \geq x/3$ and note that each N_e spans an independent set (being in a common neighbourhood of e). On the other hand we also claim N_e's are disjoint. To see this suppose $x \in N_e \cap N_{e'}$ for two distinct edges $e, e' \in M'$. By definition of N_e there is some i such that $x \in C_i$ and by Corollary 2.1.12 part c) we know that $v_i \in e$ and $v_i \in e'$, a contradiction. Note also that there can be no edges between distinct $N_e, N_{e'}$ or we find an H_7. This means that $\bigcup_{e \in M'} N_e$ is an independent set of size at least $x/3 \cdot |M|/(6\alpha) = \Omega(n^{6/12})$.

Second case: at least a third of the edges in M are of type L.
Since any vertex in L belongs to at most 4 distinct v-triangles it can in particular be an extending vertex of at most 4 edges from M. Since every edge of type L has at least $x/3$ extending vertices in L this means $|L| \geq x|M|/12 \geq n^{11/12}/12$. Now Lemma 2.1.11 part c) implies $L \subseteq N_\triangle(v)$ is triangle-free so there is an independent set of size at least $\sqrt{|L|} \geq \Omega(n^{5.5/12})$.

Third case: at least a third of the edges in M are of type I.
Let $u \in N^-(v)$. Let us denote by T_u the set of out-neighbours of u which together with u make an edge of type I. We know by the case assumption that $\sum_{u \in N^-(v)} |T_u| \geq |M|/3$. Note that $T_u \subseteq N(v) \cap N^+(u)$ so must span an independent set (or we find a K_4 in G) and hence expands inside $N^+(u)$. This means there are at least $x|T_u|$ distinct vertices extending an out-edge of u of type I. Note however that we do not know that all of them must be in I. But since any edge of type I has at most $2x/3$ extending neighbours outside

of I this means that there are are least $x|T_u|/3$ vertices in I extending an out-edge of u. This in particular means that u sends at least $x|T_u|/3$ edges directed towards I. By taking the sum over all u we obtain that the number of edges directed from $N^-(v)$ to I is at least $xM/9 \geq d^-(v)x^2/9 \geq \Omega(n^{11/12})$.

Since I spans an independent set we may assume $|I| < \alpha$. Let $S_u = N^-(v) \cap N^-(u)$ for any $u \in I$. Let $s_u = |S_u|$ so we know that $\sum_{u \in I} s_u \geq d^-(v)x^2/9 \geq \Omega(n^{11/12})$. Let I' be the subset of I consisting of vertices u with $s_u \geq 2\alpha$. Since vertices of $I \setminus I'$ contribute at most $2|I|\alpha < n^{10/12}$ to the above sum, we still have $\sum_{u \in I'} s_u \geq d^-(v)x^2/18$. By Turán's theorem for any $u \in I'$ there needs to be at least $s_u^2/(4\alpha)$ (using that $s_u \geq 2\alpha$) edges inside S_u, or we find an independent set of size α. Each such edge gives rise to a directed v-triangle. Hence, using Cauchy-Schwarz, there are at least

$$\sum_{u \in I'} s_u^2/(4\alpha) \geq \Omega(d^-(v)^2 x^4/\alpha^2) \geq \Omega(d^-(v)^2/n^{2/12})$$

directed v-triangles, as desired. $\square$

Let us give some intuition on how we are going to use the fact that T_4 is big. Let us denote by $\overrightarrow{d_e}$ the number of common out-neighbours of vertices making an edge e. Recall that T_4 can be interpreted as the number of K_4's minus an edge with all its edges oriented towards the missing edge. We call the remaining edge (for which we are not insisting on the direction) the *spine*. Note that the number of our K_4's minus an edge having some fixed edge e as spine is precisely $\binom{\overrightarrow{d_e}}{2}$. This means that $T_4 = \sum_{e \in E(G)} \binom{\overrightarrow{d_e}}{2}$. Our bound on T_4 obtained above tells us that in certain average sense $\overrightarrow{d_e}$'s should be big.

Observe now that if one finds a star centred at v consisting of s edges each with $\overrightarrow{d_e} \geq t$ then this means that the s leaves each have t out-neighbours inside $N(v)$ (note that we are disregarding the information that they are in fact inside $N^+(v)$). This will allow us to play a similar game as we did in the previous claim, indeed there we tackled the same problem with $s = d^-(v)$ and $t = x$ (which we obtained through expansion) with an important difference, namely that the star was in-directed. The bound on T_4 tells us that there is such a star with larger (in certain sense) parameters s and t and the next claim shows how this gives rise to many of our K_4's minus an edge, which we find in a different place (in particular they do not use the centre of the star since we do not know the direction of centre's edges).

For any $v \in G$ let S_v be the star consisting of a centre v and all its edges in G. Let us also denote by s_v the total number of our K_4's minus an edge with an edge of S_v as their spine, i.e. $s_v = \sum_{e \in S_v} \binom{\overrightarrow{d_e}}{2}$. Summing over v we also have $2T_4 = \sum_v \sum_{e \in S_v} \binom{\overrightarrow{d_e}}{2} = \sum_v s_v$.

Claim. *Unless $\alpha(G) \geq n^{5/12-o(1)}$ and provided $s_v \geq \frac{T_4}{n}$ there exist $\frac{s_v^2}{n^{7/12+o(1)}}$ of our K_4's minus an edge with their spine inside $N(v)$.*

Proof. Let us first "regularise" $\overrightarrow{d_e}$'s for $e \in S_v$. Let us partition these edges into at most $\log n$ sets with all edges belonging to a single set having $\overrightarrow{d_e} \in [t, 2t]$ for some t. Since $s_v = \sum_{e \in S_v} \binom{\overrightarrow{d_e}}{2}$ and the edges in S_v are split into at most $\log n$ sets, we conclude that one set contributes at least $\frac{s_v}{\log n}$ to this sum. I.e. edges in this set make a substar S of S_v consisting of s edges, each having $t \leq \overrightarrow{d_e} \leq 2t$ for some s, t satisfying $2st^2 \geq s\binom{2t}{2} \geq \frac{s_v}{\log n} \geq \frac{T_4}{n \log n} \geq n^{1-o(1)}$.[2] We may assume that $t \leq n^{5/12}$ since $\overrightarrow{d_e}$ counts the number of certain common neighbours of a fixed edge which must span an independent set (or there is a K_4).

Similarly as in the previous claim we define M to be the set of directed edges inside $N(v)$ with source vertex being a leaf of S. Now the fact that any edge $e \in S$ has $\overrightarrow{d_e} \geq t$ means that any leaf u of S is a source of at least t edges in M. Let us remove all but exactly t such edges from M, so in particular $|M| = st$ (as before while some edges might be oriented both ways we treat this as two distinct directed edges). Let us denote by T_u the set of t out-neighbours of u which together with u make an edge in M. We know T_u is an independent set (it consists of common neighbours of the edge vu). In particular, both T_u and any of its subsets expand inside $N^+(u)$. Let us consider an auxiliary bipartite graph with the left part being T_u and the right part being $N_\triangle(v) \cap N^+(u)$. We put an edge between two vertices if together with u they make a triangle in G. The expansion property translates to the fact that any subset of size t' of the left part has at least $t'x$ distinct neighbours on the right. A standard application of Hall's theorem (to the graph obtained by taking x copies of every vertex on the left) tells us we can find t disjoint stars each of size x in this graph. Translating back to our graph, for each of the t edges incident to u in M we have found a set of x out-neighbours of u which extend it into a triangle. Moreover these sets are disjoint for distinct edges. We call these x vertices extending for the corresponding edge[3].

Once again let us take a C, L, I partition provided by Corollary 2.1.12 and assign types C, L and I to edges in M, which have at least $x/3$ extending

2 We decided to pay the $\log n$ factor here for simplicity, it is possible to do the same argument more carefully and avoid it.

3 Note that this is a subset of what we considered to be extending vertices in the previous claim. Here it is important for us to fix the number of extending neighbours for every edge for certain regularity considerations.

neighbours in C, L and I, respectively. This is similar as in the previous argument except that we are using our new, slightly modified definition of extending vertices. We again split into three cases according to which type is in the majority.

First case: at least $st/3$ edges in M are of type C.
As in 1. case of the previous claim we can find a matching $\mathcal{M}$ of at least $st/(6\alpha)$ edges of type C (since vertices of $\mathcal{M}$ are incident to at most $2|\mathcal{M}|\alpha$ edges inside $N(v)$). Let us denote by N_e the set of extending neighbours of an edge $e \in \mathcal{M}$ inside C, so $|N_e| \geq x/3$. Again, as before, each N_e is independent (neighbours of the same edge), they are disjoint for different e (otherwise if a vertex belonging to two N_e's belongs to C_i we get a contradiction to uniqueness of v_i) and there are no edges between distinct N_e's (or we find an H_7). So their union makes an independent set of size at least $stx/(18\alpha)$. If $t \leq n^{4/12}$ then $st^2 \geq n^{1-o(1)}$ implies $st \geq n^{8/12-o(1)}$ and our independent set is of size at least $n^{5/12-o(1)}$, as desired. So let us assume $t \geq n^{4/12}$.

Note that by Corollary 2.1.12, given an edge $e \in M$ and its extending neighbour which belongs to some C_i we know that $v_i \in e$ and v_i can be either source or sink of e. In the former case we say the neighbour is source-extending and in the latter sink-extending. We say e is of type C-source if it has at least $x/6$ source-extending neighbours in C and of type C-sink if it has at least $x/6$ sink-extending neighbours in C.

If there are at least $st/6$ edges in M of type C-sink that means there is a leaf u of S which is a startpoint of at least $t/6$ such edges. If uw is one of these $t/6$ edges it has a set of at least $x/6$ sink-extending neighbours which span an independent set (being common neighbours of an edge) and all belong to the same C_i (namely the one for which $v_i = w$) and these C_i's are distinct between edges. This means that these sets are disjoint between ones corresponding to distinct edges and span an independent set (there are no edges between distinct C_i's) so we found an independent set of size at least $tx/36 \geq \Omega(n^{6/12})$.

If there are at least $st/6$ edges in M of type C-source we need to be able to find $s/12$ leaves of S incident to at least $t/12$ such edges each (since we know that every leaf of S is incident to exactly t edges in M so otherwise, there would be less than $s/12 \cdot t + s \cdot t/12 = st/6$ such edges in total). For any such leaf u this means we find $\frac{t}{12} \cdot \frac{x}{6} \geq \frac{tx}{72}$ source-extending neighbours of its edges (using our preprocessing fact that extending neighbours of distinct edges incident to u are disjoint). By Lemma 2.1.11 c) we know there is an independent set of size $\sqrt{tx}/9$ among these neighbours. Since these are source-extending neighbours we know they all belong to C_i for which $v_i = u$.

In particular, for distinct u they belong to distinct C_i's, meaning we obtain an independent set of size

$$\Omega(s\sqrt{tx}) = \Omega(st^2\sqrt{x}/t^{3/2}) \geq n^{13/12-o(1)}/t^{3/2} \geq n^{5.5/12-o(1)},$$

using $t \leq n^{5/12}$.

Second case: at least $st/3$ edges in M are of type L.
Let us first assume $s \geq t$. We again find a matching of size $st/(6\alpha)$ of edges of type L in M. Each edge e in the matching gives rise to a set N_e of $x/3$ extending neighbours in L. N_e spans an independent set (being inside common neighbourhood of an edge) and there can be no edges between distinct N_e's (or we find an H_7). This means that union of N_e's spans an independent set. Since any extending neighbour in this union can be extending for at most 4 edges (so belongs to at most 4 different N_e's) this gives $\alpha(G) \geq stx/(72\alpha) \geq (st^2)^{2/3}x/(72\alpha) \geq n^{5/12-o(1)}$, using $s \geq t$ and $st^2 \geq n^{1-o(1)}$.

Let us now assume $t \geq s$. We can find at least $s/6$ leaves of S each being a start vertex of at least $t/6$ edges in M of type L (otherwise there would be less than $s/6 \cdot t + s \cdot t/6 = st/3$ edges in total). Given a directed edge $uw \in M$ of type L with u being one of these $s/6$ leaves we define A_{uw} as the set of extending vertices of uw belonging to L. We will now state some properties of these sets A_{uw} which will allow us to find a big independent set. Since this is the most technical part of the proof and once the appropriate properties are identified is independent of the rest of the argument we prove it as a separate lemma afterwards.

1.) No vertex belongs to more than 4 different A_{uw}'s. Since $A_{uw} \subseteq L$ and by Corollary 2.1.12 part a) any vertex in L belongs to at most 4 different v-triangles, this means it can belong to at most 4 different A_{uw}'s.

2.) $|A_{uw}| \leq x$. This follows since, by our definition, there are exactly x uw-extending vertices.

3.) $\sum_w |A_{uw}| \geq tx/18$. This follows since for any u there are at least $t/6$ edges uw for which A_{uw} is defined and each such edge being of type L means there are at least $x/3$ extending vertices, meaning $|A_{uw}| \geq x/3$.

4.) If uw and $u'w'$ are independent then there can be no edges between A_{uw} and $A_{u'w'}$. Else, we find H_7.

This precisely establishes the conditions of Lemma 2.1.13, which provides us with an independent set of size $\min(\Omega(s\sqrt{tx}), \Omega(s^{1/2}t^{3/4}x^{1/4}), \Omega(s^{3/5}t^{3/5}x^{2/5}))$. Each of the three expressions is minimised when t is as large as possible (under the assumption $st^2 \geq n^{1-o(1)}$) so we may plug in

$t = n^{5/12}$ and $s = n^{2/12-o(1)}$ in which case the first expression evaluates to $n^{5.5/12-o(1)}$, the second to $n^{5.25/12-o(1)}$ and the third to $n^{5/12-o(1)}$.

Third case: there are at least $st/3$ edges in M of type I.
Since I spans an independent set we know $|I| < \alpha$. Any directed edge uw in M of type I has at least $x/3$ extending neighbours in I. Since these are distinct for different w's by our definition of an extending neighbour and since we insist that extending neighbours are out-neighbours of u this means that overall there are at least $stx/9$ edges directed from leaves of S to I.

This means that the average out-degree from S to I is at least $tx/9$. This together with a standard application of Cauchy-Schwarz implies there are $\Omega(s(tx)^2) = n^{16/12-o(1)}$ (recall that $st^2 \geq s_v/(2\log n) \geq n^{1-o(1)}$) out-directed cherries ($K_{1,2}$'s) with the centre in S. If we denote by $\mathcal{P}$ the set of pairs of vertices in I then $|\mathcal{P}| = \binom{|I|}{2} \leq \alpha^2$, let us also denote by d_p the number of common in-neighbours of a pair of vertices $p \in \mathcal{P}$. So in particular, $\sum_{p\in\mathcal{P}} d_p = \Omega(s(tx)^2) = n^{16/12-o(1)}$. Pairs p with $d_p < 2\alpha$ contribute at most $|\mathcal{P}| \cdot 2\alpha \leq n^{15/12-o(1)}$ to this sum so if $\mathcal{P}' \subseteq \mathcal{P}$ denotes the set of pairs which have $d_p \geq 2\alpha$ then also $\sum_{p\in\mathcal{P}'} d_p = \Omega(s(tx)^2) = n^{16/12-o(1)}$. Applying Turán's theorem inside a common neighbourhood of $p \in \mathcal{P}'$ we find there $d_p^2/(4\alpha)$ edges (using that $d_p \geq 2\alpha$) or there is an independent set of size α. Note that any edge we find inside this common in-neighbourhood gives rise to our desired K_4 minus an edge. In particular, using Cauchy-Schwarz we find at least

$$\sum_{p\in\mathcal{P}'} \frac{d_p^2}{4\alpha} \geq \frac{(\sum_{p\in\mathcal{P}'} d_p)^2}{4\alpha|\mathcal{P}'|} \geq \Omega(s^2(tx)^4/\alpha^3) \geq s_v^2/n^{7/12+o(1)}$$

copies of our desired K_4 minus an edge, as claimed. □

Recall that $2T_4 = \sum_v s_v$. Note also that stars with $s_v \geq T_4/n$ contribute at least T_4 to this sum. Hence, taking the sum over v of the number of copies of our K_4's minus an edge with spine in $N(v)$ we obtain at least

$$\sum s_v^2/n^{7/12+o(1)} \geq T_4^2/n^{19/12+o(1)} \geq T_4 n^{5/12-o(1)},$$

where we used Cauchy-Schwarz in the first inequality and our bound $T_4 \geq \Omega(n^2)$, from the first claim in the second. Note however that certain copies of our K_4's minus an edge got counted multiple times. But, for every v that counted our K_4 minus an edge we know it had its spine inside $N(v)$. This means that a single copy could be counted at most $\alpha(G)$ times since the spine (being an edge in G) can have at most $\alpha(G)$ neighbours, as they span an independent set. In particular, unless $\alpha(G) \geq n^{5/12-o(1)}$, this shows that there are more than T_4 distinct copies of our K_4's minus an edge, contradicting the definition of T_4 and completing the proof. □

We now prove the lemma we used in the proof above. Let us first attempt to help the reader parse the statement. It says that if we can partition vertices of G into a grid of subsets each of size at most x (so each cell of the grid contains at most x vertices), such that G only has edges between vertices in the same row or column of the grid and we additionally know that there is a large number of vertices in each row then we can find a big independent set in the whole graph.

Lemma 2.1.13. *Let G be a triangle-free graph with vertex set $\bigcup_{i,j} A_{ij}$ where $i \in [s], j \in \mathbb{N}$. If*

1. *no vertex appears in more than 4 different A_{ij}'s;*
2. *$|A_{ij}| \leq x$, for any i, j;*
3. *for some $t \geq s$ and any $i \in [s]$ there are at least tx vertices in $\cup_j A_{ij}$;*
4. *there are no edges of G between A_{ij} and $A_{k\ell}$ for any $i \neq k$ and $j \neq \ell$*

then $\alpha(G) \geq \min(\Omega(s\sqrt{tx}), \Omega(s^{1/2}t^{3/4}x^{1/4}), \Omega(s^{3/5}t^{3/5}x^{2/5}))$.

Proof. Let us first replace any vertex which appears in multiple A_{ij}'s with distinct copies of itself, one per A_{ij} it appears in. Our new graph has all A_{ij} disjoint and satisfies the same conditions as the original. In addition the independence number went up by at most a factor of 4 so showing the result for our new graph implies it for the original. So let us assume that sets A_{ij} actually partition the vertex set of G.

Let $\alpha = \alpha(G)$. We will call $\cup_j A_{ij}$ a row of our grid, $\cup_i A_{ij}$ a column and each A_{ij} a cell. Let us first clean-up the graph a bit. As long as we can find an independent set I of size more than $2\alpha/s$ using vertices from at most t/s cells inside some row, we take I, delete rest of the row and all the columns containing a vertex of I from G. If we repeat this at least $s/2$ many times we obtain an independent set of size larger than α which is impossible. This means that upon deleting at most $s/2$ many rows and at most $(t/s) \cdot s/2 = t/2$ many columns we obtain a subgraph for which in any row any t/s cells don't contain an independent set of size at least $2\alpha/s$. This subgraph still satisfies all the conditions of the lemma with $t := t/2$ and $s := s/2$. The only non-immediate condition is 3, it holds since we deleted at most $t/2$ cells in any of the remaining rows, so in total at most $tx/2$ vertices in that row altogether, using that any cell contains at most x vertices. From now on we assume our graph G satisfies the property that in any row any t/s cells don't contain an independent set of size $2\alpha/s$

Let us delete vertices from our graph until we have exactly tx in every row. Let n denote the number of vertices of G, so $n = stx$. Observe that at least half of the vertices of G have degree at least $n/(4\alpha)$ as otherwise vertices with degree lower than this induce a subgraph which has an independent set of size at least α by Turán's theorem. Condition 4 ensures each such vertex either has at least $n/(8\alpha)$ neighbours in its row or $n/(8\alpha)$ neighbours in its column. In particular, at least a quarter of vertices of G fall under one of these cases. Since G is triangle-free, neighbourhood of any vertex is an independent set. We conclude that either there are at least $s/4$ rows containing an independent set of size $n/(8\alpha)$ or there are $t/4$ columns containing an independent set of size $n/(8\alpha)$.

Let us first consider the latter case. Let U be the union of our $t/4$ independent sets of size at least $n/(8\alpha)$, belonging to distinct columns, so consisting of at least $tn/(32\alpha)$ vertices. Let a_i denote the number of vertices of U in row i. Then $\sum_{i=1}^{s} a_i \geq tn/(32\alpha)$ and each $a_i \leq tx$ (since we removed all but tx vertices in any row). On the other hand since G is triangle-free we know that in each row we can find an independent subset of U of size $\sqrt{a_i}$. In particular, since U was constructed as a union of independent set in columns (and all edges of G are either within columns or within rows) this means that U contains an independent set of size $\sum_{i=1}^{s} \sqrt{a_i} \geq \frac{tn}{32\alpha \cdot tx} \cdot \sqrt{tx} = \Omega(st^{3/2}x^{1/2}/\alpha)$ (where we used $n = stx$ and the standard fact that sum of roots is minimised, subject to constant sum, when as many terms as possible are as large as possible). In other words we showed $\alpha \geq \Omega(st^{3/2}x^{1/2}/\alpha)$ giving us the second term of the minimum.

Moving to the former case let us again take a union U of our $s/4$ independent sets of size at least $n/(8\alpha)$, belonging to distinct rows, so again $|U| \geq sn/(32\alpha)$. Call a cell A_{ij} full if it contains at least $4\alpha/t$ vertices of U. There are less than $t/(2s)$ full cells in any row, since otherwise, U restricted to $\lceil t/(2s) \rceil$ full cells gives us an independent set of size at least $2\alpha/s$ using at most $\lceil t/(2s) \rceil \leq t/s$ cells (using $t \geq s$), which contradicts our property from the beginning. Using this and once again the property from the beginning we conclude there can be at most $2\alpha/s$ vertices of U in full cells of any fixed row. If $2\alpha/s \geq n/(16\alpha)$ then $\alpha^2 \geq \Omega(ns) \geq \Omega(s^2tx)$, so first term of the minimum is satisfied. So we may assume $2\alpha/s \leq n/(16\alpha)$. Hence, by removing from U any vertex belonging to a full cell we remove at most half the vertices of U (since U had at least $n/(8\alpha)$ vertices in every row). Now, finally denote by a_i the number of vertices of U belonging to the column i. So $\sum_{i=1}^{s} a_i = |U| \geq sn/(64\alpha)$ and $a_i \leq s \cdot 4\alpha/t$, since all the remaining vertices of U belong to non-full cell. As before $\alpha \geq \sum_{i=1}^{s} \sqrt{a_i} \geq \frac{sn/(64\alpha)}{4s\alpha/t} \cdot \sqrt{4s\alpha/t} \geq \Omega(s^{3/2}t^{3/2}x/\alpha^{3/2})$ giving us the third term of the minimum. $\square$

2.1.3 *2-density and local independence number*

In this section we show our upper bounds on the maximum possible $\alpha(G)$ in a graph G satisfying $\alpha_m(G) \geq r$. In order to do this we need to exhibit a graph with no large independent set in which any m-vertex subgraph contains an independent set of size r. As discussed in the introduction the natural candidates are random graphs and the answer is controlled by $M(m,r)$ which is defined to be the minimum value of the 2-density over all graphs H on m vertices having $\alpha(H) \leq r-1$. It will be convenient to define $d_2(H) = \frac{e(H)-1}{|H|-2}$ so that the 2-density is simply the maximum of $d_2(H')$ over subgraphs H' of order at least 3, from now on whenever we consider 2-density we will implicitly assume the subgraphs we take have at least 3 vertices. We begin by proving Proposition 2.1.4.

Proposition 2.1.4. *Let m, r be fixed, $m \geq 2r-1 \geq 3$ and $M = M(m,r)$. Then for any n there exists an n-vertex graph G with $\alpha_m(G) \geq r$ and $\alpha(G) \leq n^{1/M+o(1)}$.*

Proof. A graph H is said to be *strictly* 2*-balanced* if $m_2(H) > m_2(H')$ for any proper subgraph H' of H. I.e., if H itself is the maximiser of $d_2(H)$ among its subgraphs and in particular $m_2(H) = d_2(H) = \frac{e(H)-1}{|H|-2}$.

Let $\mathcal{H} = \{H_1, \ldots, H_t\}$ be a collection of strictly 2-balanced graphs such that any m-vertex H with $\alpha(H) \leq r-1$ contains some H_i as a subgraph. We can trivially obtain it by replacing any H in our family which is not strictly 2-balanced by its subgraph H' which maximises $d_2(H')$. In particular, $\mathcal{H}$ is a family of strictly balanced 2-graphs H satisfying $m_2(H) = d_2(H) \geq M$, with at most m vertices and with the property that if a graph is $\mathcal{H}$-free then it satisfies $\alpha_m \geq r$. Note that $t \leq 2^{\binom{m}{2}}$, so is in particular bounded by a constant (depending on m).

Let $G \sim \mathcal{G}(n,p)$, where we choose $p := 1/(48tn^{1/M})$ and will be assuming n to be large enough throughout. Let A_K denote the event that a subset $K \subseteq V(G)$, consisting of $k := \frac{8\log n}{p} + 2 = O(n^{1/M}\log n)$ vertices, spans an independent set. In particular, we have $\mathbb{P}(A_k) = (1-p)^{\binom{k}{2}}$. Let B_j^i denote the event that we find a copy of H_i at j-th possible location (so fixing the subset of vertices of G where we could find H_i and the labellings of vertices). In particular, $\mathbb{P}(B_j^i) = p^{e(H_i)}$. Our goal is to show that with positive probability none of the events A_K or B^i_j occur, which implies that there is an $\mathcal{H}$-free graph with no independent set of size $O(n^{1/M}\log n)$ as desired. We will do so by using the asymmetric version of the Lovász local lemma (see Lemma 5.1.1 in [16]). In order to apply the lemma we first need to understand how many dependencies there are between different types of events. In particular, given A_K it depends only on $\binom{k}{2}$ edges of G, so in particular it is mutually independent of all events B^j_i which do not contain one of these edges. In particular, it is mutually independent from all but at most $k^2n^{|H_i|-2}$ events B^j_i and at most n^k other events $A_{K'}$ (since there are at most this many such events in total). Similarly, any B^i_j is mutually independent of all but at most $e(H_i)n^{|H_{i'}|-2} \le m^2n^{|H_{i'}|-2}$ events $B^{i'}_j$ for a fixed i' and at most n^k events A_k.

We now need to choose parameters x (corresponding to events of type A_K) and y_i (corresponding to events of type B^j_i) such that

$$\mathbb{P}(A_K) \le x\cdot(1-x)^{n^k}\cdot\prod_i(1-y_i)^{k^2n^{|H_i|-2}} \quad \text{and}$$
$$\mathbb{P}(B^j_i) \le y_i\cdot(1-x)^{n^k}\cdot\prod_{i'}(1-y_{i'})^{m^2n^{|H_{i'}|-2}}$$

which will complete the proof. We choose $x = 1/n^k$ so that in particular $(1-x)^{n^k} \ge 1/3$ (as n is large) and $y_i = p/(8tn^{|H_i|-2}) \le 1/2$ so that in particular $(1-y_i)^{n^{|H_i|-2}} \ge e^{-p/(4t)}$ (using $1-a \ge e^{-2a}$ for $a \le 1/2$). With these choices we obtain

$$\begin{aligned}
x\cdot(1-x)^{n^k}\cdot\prod_i(1-y_i)^{k^2n^{|H_i|-2}} &\ge \frac{1}{n^k}\cdot\frac{1}{3}\cdot e^{-pk^2/4} \ge e^{-2k\log n - pk^2/4}\\
&= e^{-p\binom{k}{2}} \ge (1-p)^{\binom{k}{2}} = \mathbb{P}(A_K) \quad \text{and}\\
y_i\cdot(1-x)^{n^k}\cdot\prod_{i'}(1-y_{i'})^{m^2n^{|H_{i'}|-2}} &\ge \frac{p}{8tn^{|H_i|-2}}\cdot\frac{1}{3}\cdot e^{-m^2p/4}\\
&\ge \frac{p}{48t\cdot n^{\frac{|H_i|-2}{e(H_i)-1}\cdot(e(H_i)-1)}}\\
&\ge \frac{p}{(48tn)^{\frac{1}{M}\cdot(e(H_i)-1)}} \ge p^{e(H_i)} \ge \mathbb{P}(B^j_i)
\end{aligned}$$

where in the second inequality we used the fact that m is a constant while $p \to 0$ so $m^2p \to 0$ and in particular $e^{-m^2p/4} \le 1/2$ (since n is large) in the third inequality we used $n^{-\frac{|H_i|-2}{e(H_i)-1}} \ge n^{-\frac{1}{M}}$ which follows since M is equal to the minimum of $\frac{e(H_i)-1}{|H_i|-2}$ over H_i (and we used $|H_i| \geq 3$ to put the $48t$ factor under the exponent). □

Remark. This result seems to be the best one can get using random graphs, up to the polylog factor. The polylog factor can likely be slightly improved compared to the above argument by using the $\mathcal{H}$-free process (see e.g. [197] for more details about this process).

If we replace $m_2(H)$ in the definition of $M(m,r)$ with $d_2(H) = \frac{e(H)-1}{|H|-2}$ the problem of determining $M(m,r)$ would reduce to the classical Turán's theorem. Indeed, since the number of vertices is fixed, minimising $d_2(H)$ is tantamount to minimising the number of edges in an m-vertex graph with $\alpha(H) < r$ and upon taking complements we reach the setting of the classical Turán's theorem. This is why it is natural to call our problem of determining $M(m,r)$ the 2-density Turán problem. Note that since $m_2(H) \geq d_2(H)$ the proposition also holds if we replace M with $\min d_2(H)$. This essentially recovers the argument of Linial and Rabinovich [178]. However, it turns out one can in many cases do much better by using the actual 2-density.

2.1.3.1 *The 2-density Turán problem*

In this subsection we show our results concerning the 2-density Turán problem of determining $M(m,r)$ which together with Proposition 2.1.4 give upper bounds in the local to global independence number problem mentioned in the introduction.

Triangle-free case

Here we show our bounds for the case $k = 3$. This means that m and r satisfy $2r-1 \leq m \leq 3r-3$ and as expected the behaviour will be very different at the beginning and end of the range. Our first observation determines $M(2r-1,r)$.

Proposition 2.1.14. *Let $r \geq 2$. Then $M(2r-1,r) = m_2(C_{2r-1}) = 1 + \frac{1}{2r-3}$.*

Proof. Since C_{2r-1} is a $2r-1$ vertex graph with no independent set of size r we obtain $M(2r-1,r) \leq m_2(C_{2r-1})$. For the lower bound let G be a graph on $2r-1$ vertices with $\alpha(G) \leq r-1$, our goal is to show $m_2(G) \geq m_2(C_{2r-1})$. If G contains a cycle of length ℓ then $m_2(G) \geq m_2(C_\ell) \geq m_2(C_{2r-1})$ where in the last inequality we used $\ell \leq 2r-1$ since G has only $2r-1$ vertices. If G contains no cycles it is a forest so in particular it is bipartite. One part of the bipartition must have at least r vertices giving us an independent set of size at least r, which is a contradiction. □

Turning to the other end of the range we show.

Theorem 2.1.15. *For $r \geq 2$ we have $M(3r-4, r) \geq \frac{5}{3} - \frac{1}{r-2}$.*

Proof. Let G be a graph on $m = 3r-4$ vertices with $\alpha(G) \leq r-1$. If G contains a triangle then $m_2(G) \geq 2$ and we are done. If G contains a subgraph G' on m' vertices with minimum degree at least 4 then $m_2(G) \geq d_2(G') \geq \frac{2m'-1}{m'-2} > 2$ so again we are done. In particular, we may assume that G is 3-degenerate. These conditions allow us to apply a modification of a result of Jones [149] (see Appendix C for more details about the modification) which tells us that $e(G) \geq 6m - 13(r-1) - 1 = 5r - 12$. This implies $m_2(G) \geq d_2(G) \geq \frac{5r-13}{3r-6} = \frac{5}{3} - \frac{1}{r-2}$ as claimed. □

This is close to best possible, for example the chain graph H_r (see [149] for more details) has $3r-4$ vertices, no independent set of size r and $m_2(H_r) = \frac{5}{3} - \frac{1}{9} \cdot \frac{1}{r-2}$. We believe that, as in the problem of [149], these graphs should be optimal, it is not hard to verify that this is indeed the case for first few values of r and one can improve our result above by repeating more carefully the stability type argument from [149] for our graphs.

In the above result we did not look at the very end of the range for $k = 3$, namely $m = 3r-3$. The reason is that it seems to behave differently. Of course $M(3r-3, r) \geq M(3r-4, r)$ so the same bound as above applies, however it seems possible that a stronger bound is the actual truth, it is even possible that the answer jumps to $M(3r-3, r) \geq 2$.

Independence number two.

In this subsection we solve the 2-density Turán problem for graphs with no independent sets of size 3. The behaviour depends on parity of m, we begin with the easier case when m is even.

Lemma 2.1.16. *For any $k \geq 2$ we have $M(2k, 3) \geq (k+1)/2$.*

Proof. Let G be a graph on $2k$ vertices with $\alpha(G) \leq 2$. This condition implies that for any vertex v of G the set of vertices not adjacent to v must span a clique, since otherwise the missing edge together with v makes an independent set in G of size 3. On the other hand, if we can find $K_k \subseteq G$ then $m_2(G) \geq m_2(K_k) = \frac{k+1}{2}$ and we are done. So we may assume G is K_k-free. Combining these two observations implies every vertex has at most $k-1$ non-neighbours and in particular $\delta(G) \geq 2k-1-(k-1) = k$. This in turn implies $m_2(G) \geq \frac{e(G)-1}{|G|-2} \geq \frac{k^2-1}{2k-2} = (k+1)/2$ completing the proof. □

We now turn to the more involved case of odd $m = 2k - 1$. The increase in difficulty is partially due to the fact that the answer becomes very close (but not equal) to $m_2(K_k)$ which we have seen above is the answer for graphs with one more vertex. So the bound we need to show is much stronger in the odd case. We begin with the following lemma which is at the heart of our argument. We state it for the complement of our actual graphs for convenience.

Lemma 2.1.17. *Let $k \geq 5$ and $1 \leq t < \sqrt{(k-1)/2}$. Let G be a triangle-free graph on $2k-1$ vertices with the property that any k of its vertices span at least $t+1$ edges. Then $e(G) \leq (k-1)^2 - t^2 + 1$.*

Proof. The following easy claim will be used at various points in the proof. It also provides an illustration for the flavour of the more involved arguments we will be using later.

Claim. *If there are 2 vertex disjoint independent sets of order $k-1$ then $e(G) \leq (k-1)^2 - t^2 + 1$.*

Proof. Let v be the (only) vertex not belonging to either of the independent sets, which we call L and R. Let i denote the number of neighbours of v in L and j in R. Observe first that $v \cup L$ and $v \cup R$ are both sets of k vertices so need to span at least $t+1$ edges, by our main assumption on G. Since all edges in these sets are incident to v (L and R are both independent sets) we conclude that $i, j \geq t+1$. Note further that since G is triangle-free there can be no edges between neighbours of v, which means that there can be at most $(k-1)^2 - ij$ edges in $L \cup R = G \setminus v$. Adding the $i+j$ edges incident to v we obtain $e(G) \leq i + j + (k-1)^2 - ij = (k-1)^2 + 1 - (i-1)(j-1) \leq (k-1)^2 + 1 - t^2$. $\square$

We now proceed to obtain some information on structure of G. Observe first that by our assumption on t we have $(k-1)^2 - t^2 + 1 > (k-1)^2 - (k-1)/2 + 1 = k^2 - 5k/2 + 5/2$ so if we can show $e(G) \leq k^2 - 5k/2 + 5/2$ we are done. So let us assume $e(G) \geq k^2 - 5k/2 + 3$, which will suffice to give us some preliminary information about G.

Since G is triangle-free, neighbours of any vertex span an independent set. By our main assumption on G there can be no independent set of order k so $\Delta(G) \leq k-1$. On the other hand, we have $\Delta(G) \geq 2e(G)/(2k-1) \geq (2k^2 - 5k + 6)/(2k-1) > k-2$, so $\Delta(G) = k-1$. In particular, there exists a vertex with $k-1$ neighbours, which means that there is an independent set R of size $k-1$ in G. If every vertex in R has degree at most $k-3$ then the sum of degrees in G is at most $k(k-1) + (k-1)(k-3) = 2k^2 - 5k + 3 < 2e(G)$.

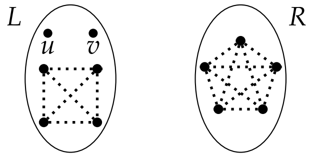


Figure 2.4: Initial structure for $k = 6$, dotted lines depict missing edges.

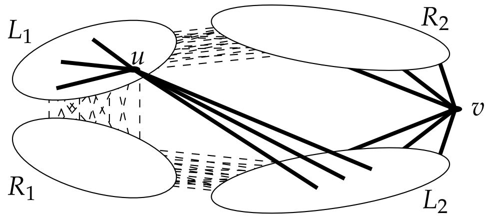


Figure 2.5: Blow-up of C_5 structure, dashed lines denote only possible locations of edges, there can be no edges between non-adjacent parts or inside parts, apart from those belonging to S.

So there is a vertex in R with degree at least $k-2$ and in particular there exist an independent set of size at least $k-2$ disjoint from R. In other words $L := G \setminus R$ contains an independent set of size $k-2$ and two remaining vertices, say v and u (see Figure 2.4 for the illustration of the current state). Let us w.l.o.g. assume that v has at most as many neighbours in L as u.

This almost gives us the situation in the claim above. In particular by the claim, we may assume that both u and v have at least one neighbour in $L \setminus \{u, v\}$.

We now proceed to obtain more detailed information on how G should look like. Denote by $L_2 := N(v) \cap L, R_2 := N(v) \cap R, L_1 = L \setminus (L_2 \cup \{v\})$ and $R_1 := R \setminus R_2$. Note that all edges within $L_1 \cup L_2$ must touch u so $L_1 \cup L_2$ induces a star S with a centre at u, say of size s. Note further that no edges in $R_2 \cup L_2 = N(v)$ can exist as G is triangle-free. Putting these observations together we conclude that G without edges of S is a subgraph of a blow-up of C_5 with parts $\{v\}, R_2, L_1, R_1$ and L_2 in order (see Figure 2.5 for an illustration in the case when $u \in L_1$).

This means that v contributes $|L_2| + |R_2|$ edges while the remaining edges all come between $L_1 \cup L_2$ and $R_1 \cup R_2$ and S. Since $|R_1 \cup R_2| = |L_1 \cup L_2| = k-1$ and there can be no edges between R_2 and L_2 there are $(k-1)^2 - |L_2||R_2| - X$ edges between $L_1 \cup L_2$ and $R_1 \cup R_2$ where X counts the number of non-edges between R_2 and L_1, L_1 and R_1, and R_1 and L_2. In total we have $e(G) = |L_2| + |R_2| + (k-1)^2 - |L_2||R_2| - X + s$.

Let us denote by $i = |R_2|$ and $j = |L_2|$ so $k-1-j = |L_1|$; $k-1-i = |R_1|$ and $e(G) = i + j + (k-1)^2 - ij - X + s = (k-1)^2 + 1 - (i-1)(j-1) - X + s$. Since by our main assumption on G there needs to be at least $t+1$ edges among the k vertices $v \cup R_1 \cup R_2$ and we know there are exactly $|R_2| = i$ edges in this set ($R_1 \cup R_2$ is an independent set) we conclude that $i \geq t+1$.

Similarly, we know $s + j \geq t + 1$ as otherwise $v \cup L_1 \cup L_2$ make a k-vertex subset with $|L_2| + s \leq t$ edges in total. Note that since we observed by the claim that both v and u need to have a neighbour inside $L \setminus \{u, v\}$ we must have $j, s \geq 1$, while by our choice of v as having less neighbours in L than u we have $j \leq s$.

We distinguish two cases depending on whether $v \sim u$ or not (i.e. whether $u \in L_1$ or $u \in L_2$). Let us deal with the case $u \in L_2$ first. There can be no edges within $L_2 = N(v)$ so all leaves of S must be in L_1. u must have at least $t + 1$ neighbours within $R_1 \cup R_2$ (otherwise $u \cup R_1 \cup R_2$ are k vertices with at most t edges) so there must be at least $s(t+1)$ edges missing between $R_1 \cup R_2$ and L_1, i.e. $X \geq s(t+1)$. In particular, $e(G) \leq (k-1)^2 + 1 - (i-1)(j-1) + s - s(t+1) \leq (k-1)^2 + 1 - t(j-1) - ts \leq (k-1)^2 + 1 - t^2$, (where we used $j \geq 1$ and $i \geq t+1$ in the second inequality and $j - 1 + s \geq t$ in the third), as desired.

In the remaining case $u \in L_1$. Let's say s_1 leaves of S are in L_1 and s_2 in L_2. Let $x \geq t + 1$ be the number of neighbours of u in $R_1 \cup R_2$ we know as before there must be $xs_1 + (k - 1 - x)$ non-edges between L_1 and $R_1 \cup R_2$ and at least $(x - |R_2|)s_2 = (x - i)s_2$ non-edges between R_1 and L_2. In total we have $X \geq xs_1 + (k - 1 - x) + \max(x - i, 0) \cdot s_2 = k - 2 + s + (x-1)(s-1) - \min(x, i) \cdot s_2$. If we denote by $m := \min(x, i)$ we get

$$
\begin{aligned}
e(G) &\leq (k-1)^2 + 1 - (k-2) - (i-1)(j-1) - (x-1)(s-1) + m \cdot s_2 \\
&\leq (k-1)^2 + 1 - (k-2) - (m-1)(s+j-2) + m \cdot s_2 \\
&\leq (k-1)^2 + 1 - (k-2) - (m-1)(s+s_2-2) + m \cdot s_2 \\
&= (k-1)^2 + 1 - (k-4) - (m-2)(s-2) - s_1
\end{aligned}
$$

Where we used $i, x \geq m$ in the second inequality and $j \geq s_2$ (since s_2 leaves live inside L_2 of size j) in the third. The term $(m-2)(s-2) + s_1$ is non-negative provided $s \geq 2$, since $m = \min(x, i) \geq t + 1 \geq 2$. So if $s \geq 2$ we have $e \leq (k-1)^2 + 1 - (k-4) \leq (k-1)^2 + 1 - t^2$ where we used $t < \sqrt{(k-1)/2} \implies t^2 \leq k - 4$, which holds for $k \geq 5$ (using integrality of t for $k = 5, 6$). If $s = 1$ we must also have $j = 1$ (since $j \leq s$ and $j \geq 1$) and in turn $j + s \geq t + 1$ implies $t = 1$. If $s_2 = 0$ the first inequality above (and $k \geq 3$) gives $e(G) \leq (k-1)^2 = (k-1)^2 - t + 1$ and we are done. If $s_2 = |L_2| = 1$ then $s_1 = 0$ and removing the single vertex in L_2 removes all neighbours of v, u in L and gives us again the situation from the claim. $\square$

We are now ready to deduce our bound on $M(2k-1, 3)$.

Theorem 2.1.18. *Let $k \geq 4$ then*

$$M(2k-1,3) \geq \frac{k+1}{2} - \max_{1 \leq t \leq k-2} \min\left(\frac{t}{k-2}, \frac{(k+1)/2-(t-1)^2}{2k-3}\right).$$

Proof. Let G be a graph with $2k-1$ vertices and $\alpha(G) \leq 2$. If $k = 4$ the desired bound evaluates to 2. Since in this case G does not satisfy the 7-local 3-independence property by Lemma 2.1.19 it contains either a K_4 or an H_7 as a subgraph. Since $m_2(K_4) = 5/2 > 2$ and $m_2(H_7) \geq \frac{e(H_7)-1}{|H_7|-2} = 2$ (since $e(H_7) = 11$ and $|H_7| = 7$) we deduce that in either case $m_2(G) \geq 2$ as desired. Let us now assume $k \geq 5$.

Let $m(t) = \min\left(\frac{t}{k-2}, \frac{(k+1)/2-(t-1)^2}{2k-3}\right)$. Observe that if we choose $t = \sqrt{(k-1)/2}$ we obtain $\frac{(k+1)/2-(t-1)^2}{2k-3} = \frac{t^2+1-(t-1)^2}{2k-3} = \frac{2t}{2k-3} < \frac{t}{k-2}$. Let a be the maximiser of $m(t)$.

If $m(a) = \frac{a}{k-2}$ then by above observation we must have $a < \sqrt{(k-1)/2}$ (since the first term of the minimum is increasing and the second is decreasing in t). We may assume that among any k vertices there are at least $a+1$ missing edges, as otherwise the induced subgraph on these k vertices implies $m_2(G) \geq (k+1)/2 - a/(k-2) = (k+1)/2 - m(a)$ and we are done. This allows us to apply Lemma 2.1.17 to the complement of G to deduce G must have at least $\binom{2k-1}{2} - ((k-1)^2 + 1 - a^2) = k(k-1) + a^2 - 1$ edges. This in turn implies $m_2(G) \geq \frac{k(k-1)+a^2-2}{2k-3} = (k+1)/2 - \frac{(k+1)/2-a^2}{2k-3} \geq (k+1)/2 - m(a)$, where the last inequality follows since otherwise $m(a) < \frac{(k+1)/2-a^2}{2k-3}$ which implies $m(a) < m(a+1)$ (since also $\frac{a+1}{k-2} > \frac{a}{k-2} = m(a)$), which contradicts maximality of $m(a)$ (note that $a+1 < \sqrt{(k-1)/2}+1 \leq k-2$ for $k \geq 5$).

If on the other hand $m(a) = \frac{(k+1)/2-(a-1)^2}{2k-3}$ we may assume $a \geq 2$ (since $\frac{1}{k-2} < \frac{(k+1)/2}{2k-3}$ for $k \geq 5$). So we must have $m(a) \geq \frac{a-1}{k-2}$ (otherwise since the second term in the definition of $m(t)$ is decreasing in t we conclude $m(a-1) > m(a)$ and get a contradiction). As in the previous case this means that any k vertices must miss at least a edges, or we are done. We also know that $m(a-1) = (a-1)/(k-2)$ (as otherwise again $m(a-1) > m(a)$) so again as in the previous case by our initial observation we must have

$a - 1 < \sqrt{(k-1)/2}$ and we may apply Lemma 2.1.17 with $t = a - 1$ to complement of G to obtain $e(G) \geq k(k-1) + (a-1)^2 - 1$. This implies

$$m_2(G) \geq \frac{k(k-1) + (a-1)^2 - 2}{2k-3} = \frac{k+1}{2} - \frac{(k+1)/2 - (a-1)^2}{2k-3} = \frac{k+1}{2} - m(a),$$

as desired. □

We now prove the lemma which we used for $k = 4$ case in the above theorem and mentioned in Section 2.1.2.

Lemma 2.1.19. *Any K_4 and H_7-free graph G satisfies $\alpha_7(G) \geq 3$.*

Proof. It is enough to show that any 7 vertex graph G which is K_4-free and has no I_3 (independent set of size 3) must contain H_7. First observe that $\delta(G) \geq 3$ as otherwise there is a vertex with 4 non-neighbours who must span a K_4, in order to avoid I_3. Note also that $\Delta(G) \leq 5$ as if a vertex v had degree 6 then by $R(3,3) = 6$ in its neighbourhood we find an I_3 or a K_3 which together with v makes a K_4. Since G has odd size it must contain a vertex v of degree exactly 4.

Our goal is to find two vertex disjoint triangles in G. If some vertex v has degree 3 then its non-neighbours span a triangle and since its neighbours don't span an I_3 the edge among them together with v give us our second triangle. If all vertices have degree at least 4 then by pigeonhole principle any two adjacent vertices lie in a triangle. If we take v as our guaranteed vertex of degree 4, let u, w be its non-neighbours. Then $u \sim w$ and u and w lie in some triangle. Removing it leaves us with v and 3 of its neighbours, so we again find a second triangle.

So we can always find a pair of vertex disjoint triangles xyz and abc. Let v be the remaining vertex of G. v can send at most 2 edges towards each of the triangles. If it sends exactly 2 to both then the non-neighbours of v must be adjacent and we found our H_7. If it sends 2 to xyz, say $v \sim x, y$ of them but only one to abc say $v \sim a$ then replacing v with z we find disjoint triangles xyv and abc such that z sends 2 edges towards each ($v \nsim z, b, c$ implies $z \sim b, c$) so we are back in the first case and are done. □

Combining Lemma 2.1.16 and Theorem 2.1.18 with Proposition 2.1.4 we obtain Theorem 2.1.6. Both above results are tight. Since we only needed above lower bounds for Theorem 2.1.6 we will only describe our tightness examples here and postpone (the somewhat tedious) computation of their 2-density to Appendix A.

If $m = 2k$ then our example is simply a vertex disjoint union of 2 cliques on k vertices. This graph clearly has no independent set of size 3 and it is not hard to see that its 2-density is equal to $m_2(K_k) = \frac{k+1}{2}$ (see Lemma 2.1.31). If $m = 2k - 1$ the answer is more complicated since it needs to match the somewhat messy bound of Theorem 2.1.18. The examples however still arise naturally from looking at the proof and will be a blow-ups of C_5 with cliques placed into parts which we choose to have sizes $1, a, k-1-a, k-1-a, a$ in order around the cycle, where a is the optimal choice of t in Theorem 2.1.18. Complement of any such graph is an actual blow-up of C_5 so is triangle-free and for the computation of its 2-density see Lemma 2.1.32.

General Turán 2-density problem

In this section we show our general bounds on $M(m, r)$. Combining the following proposition with Proposition 2.1.4 we obtain Theorem 2.1.7.

Proposition 2.1.20. *Let $k = \lceil m/(r-1) \rceil$. Provided m is sufficiently larger than r we have $M(m, r) \geq \frac{k+1}{2} - \frac{c_r}{\sqrt{k}}$, where $c_r > 0$ is a constant depending only on r.*

Proof. Let G be an m-vertex graph with $\alpha(G) \leq r - 1$, our task is to show $m_2(G) \geq \frac{k+1}{2} - \frac{c_r}{\sqrt{k}}$, where $k = \lceil \frac{m}{r-1} \rceil$, under the assumption that m is large. $\alpha(G) \leq r - 1$ implies that G's complement $\overline{G}$ is K_r-free. Let $t = t_{r-1}(m) - e(\overline{G})$, where $t_{r-1}(m)$ denotes the Turán number for K_r-free graphs on m vertices. If $t \geq \frac{3}{2}m$ then we get

$$m_2(G) \geq \frac{e(G)-1}{|G|-2} = \frac{\binom{m}{2} - t_{r-1}(m) + t}{m-2} \geq \frac{\frac{m^2}{2(r-1)} + t - m/2}{m}$$
$$\geq \frac{m}{2(r-1)} + 1 \geq \frac{k-1}{2} + 1 = \frac{k+1}{2},$$

where in the second inequality we used the standard bound $t_{r-1}(m) \leq \left(1 - \frac{1}{r-1}\right) \frac{m^2}{2}$.

So we are done unless $t < \frac{3}{2}m$. Since $\overline{G}$ is K_r-free and has $t_{r-1}(m) - t$ edges a stability theorem (see Theorem 1.3 in [23]) implies that $\overline{G}$ can be made $r-1$-partite by removing at most $\frac{rt^{3/2}}{2m}$ edges (being crude and using that m is sufficiently larger than r). Translating this to G we conclude G is a vertex disjoint union of $r - 1$ cliques missing a few edges, in total at most $\frac{rt^{3/2}}{2m} \leq r\sqrt{m}$ edges. At least one of these "cliques" needs to have size $s \geq k$. In particular if take a subset of k vertices of this "clique" it still misses at most $r\sqrt{m}$ edges. In particular, it has 2-density at least $\frac{k+1}{2} - \frac{r\sqrt{m}}{k-2} \geq \frac{k+1}{2} - \frac{c_r}{\sqrt{k}}$. $\square$

Remark. The stability result we used above was also independently discovered in [212] (in an asymptotic form), we used the variant from [23] since it is explicit. Our problem seems to be closely related to this type of stability problems for Turán's theorem. For example the bipartite variant, which was precisely solved in [88], has the same form of optimal examples as we found for $M(2k-1,3)$. This was recently generalised to r-partite graphs in [160] which might be helpful for studying $M(m,r)$ for larger r.

Note that for the special case of $r = 3$ and m odd this result matches (up to a constant factor in front of the lower order term) our bound in Theorem 2.1.18 and is hence almost best possible in this case by Lemma 2.1.32. On the other hand if m is even it is some way off. This seems to happen in general, we found examples (disjoint unions of our examples for the $r = 3$ case) which show that above bound is tight up to the constant factor in front of the lower order term provided $m \pmod{r-1}$ is between 1 and $(r-1)/2$. This condition ensures that in the Turán K_r-free graph on m vertices there is more small parts (of size $k-1$) which allows us to pair up small and big parts and place there a copy of our example from the $r = 3$ case, we once again relegate the details to Appendix A. It seems that as $m \pmod{r-1}$ approaches r stronger bounds should hold and ultimately if $m \mid (r-1)$ the lower order term disappears completely as it did in the $r = 3$ case.

Theorem 2.1.7, while being close to best possible, unfortunately requires m to be somewhat large (compared to r) which misses many interesting instances of the problem. The following result illustrates some of our ideas for obtaining results which hold for any choice of parameters. We restrict attention to the divisible case $m = k(r-1)$ to keep the argument as simple as possible.

Proposition 2.1.21. *Let $r \geq 3$ and $m = k(r-1)$ then we have $M(m,r) \geq \frac{k}{2} + \frac{k-1}{m-2}$.*

Proof. We will prove the result by induction on r while keeping k fixed. For the base case of $r = 3$ we have $k = m/2$ and the statement matches precisely Lemma 2.1.16.

Let G be a graph on m vertices having $\alpha(G) \leq r-1$. If G has a vertex v of degree less than k then removing $v \cup N(v)$ from G we obtain a graph G' on at least $m-k = k(r-2)$ vertices which has $\alpha(G') \leq r-2$, since v extends any independent set we can find in G'. This implies by the inductive assumption for $r-1$ that $m_2(G) \geq m_2(G') \geq M(m-k, r-1) \geq \frac{k}{2} + \frac{k-1}{m-k-2} \geq \frac{k}{2} + \frac{k-1}{m-2}$ with room to spare. Hence, we may assume $\delta(G) \geq k$ which implies $m_2(G) \geq d_2(G) \geq \frac{mk/2-1}{m-2} = \frac{k}{2} + \frac{k-1}{m-2}$. $\square$

The above proof clearly leaves quite some room for improvement. However, it (combined with $M(m, r)$ being increasing in m to capture the non-divisible cases) already suffices to improve the bound of Linial and Rabinovich for all values of m and r with $k \geq 4$. It also suffices to obtain a significant improvement in the benchmark case $m = 20, r = 5$ of $M(20, 5) \geq 49/18$ over the previously best bound of 43/18 of Kostochka and Yancey [163]. We have more involved ideas which allow one to improve on the above bound quite substantially. In particular, we manage to resolve the benchmark case and show $M(20, 5) = 3$. Since that argument is somewhat more involved and its generalisations become even more complicated, while ultimately still falling short of the asymptotic result of Theorem 2.1.7, we relegate it to Appendix B.

2.1.4 *Concluding remarks*

In this section we studied the local to global independence number problem, i.e. how big an independent set one finds in a graph with the property that any m vertices contain an independent set of size r. While many of our results break previous barriers on this problem, there is still room for improvement and we believe we have not fully exhausted the potential of our ideas.

In terms of lower bounds we improve previously best bounds for about half of the possible choices of m and r. It would be interesting to obtain a similar improvement for the whole, or at least most of the range. Our argument here relied on improving the bounds for $r = 3$ which then generalised through Lemma 2.1.8. One can follow our approach for $r \geq 4$ as well. Lemma 2.1.8 easily generalises so for example if one improves bounds say for $m = 3k - 1, r = 4$ this leads to improvement for about 2/3 of the possible values in general. For $r = 3$ our arguments relied on a Ramsey result for graphs H_{2k-1} which were certain blow-ups of C_5. We believe similar story should happen for larger r, in the initial cases role of C_5 seems to be taken by the chain graphs (see graphs H_k in [149]) and to obtain a general result for fixed r one should prove a Ramsey bound for appropriate blow-ups of these chain graphs. This should lead to an improvement for essentially all values of m and r except when $r - 1 \mid m$ which seems more difficult. In fact using a minor modification of Lemma 2.1.8 if one improves the bounds in such a "divisible" case, say $m = 2k$ and $r = 3$ this immediately improves the bounds for any choice of m and r with $r - 1 \nmid m$ (and most divisible cases as well). Here a good starting point seems to be the case $m = 8, r = 3$.

Question 2.1.22. *Does any graph with an independent set of size* 3 *among any* 8 *vertices have* $\alpha(G) \geq n^{1/3+\varepsilon}$*?*

The reason we raise the $(8,3)$ case instead of $(6,3)$ is that the latter is easily seen to be essentially equivalent to the problem of how large independent sets we find in triangle-free graphs. Since the answer to this classical problem is known up to a constant factor [35, 106, 217] the same holds for our problem. This raises the possibility that the $(8,3)$ case is essentially equivalent to the same problem for K_4-free graphs which is open and believed hard. It turns out however that this is not the case since for example square of C_8 is an 8 vertex K_4-free graph with no independent set of size 3 and could play the role of our H_{2k-1}'s as the intermediate forbidden graph in this case. In fact it is the only possible candidate, as can be seen by looking at optimal examples for $R(4,3)$ [183] which show that property $\alpha_8 \geq 3$ is essentially equivalent to graph being K_4-free and C_8^2-free.

In terms of improving our new bounds a good starting place are graphs in which every 7 vertices have an independent set of size 3. We showed such graphs must have $\alpha(G) \geq n^{5/12-o(1)}$ proving a conjecture of Erdős and Hajnal. Here the natural limit for our methods is actually $n^{3/7}$ and most of our argument works up to this point. It should be possible to push our methods at least beyond $5/12$. On the other hand breaking $3/7$ seems to require new ideas. The main question here is whether it is possible to reach $1/2$, namely whether the second conjecture of Erdős and Hajnal holds.

Question 2.1.23. *Does any graph with an independent set of size* 3 *among any* 7 *vertices have* $\alpha(G) \geq n^{1/2-o(1)}$*?*

Lemma 2.1.19 shows that this is in some sense equivalent to a Ramsey problem of our graph H_7 vs an independent set, with the added benefit that we know the graph is K_4-free, which however seems to be a weaker condition than being H_7-free, so it is unclear if it is actually needed at all. The above bound for $m = 7, r = 3$ is stronger than our general bound which makes it likely that the general bound can be further improved.

Let us now turn to the upper bounds. Our bounds all arise from our results on the Turán 2-density problem and the main open problem is to solve this problem precisely for all choices of parameters.

Question 2.1.24. *What is the minimum value of the* 2*-density of a graph on* m *vertices having no independent set of size* r*?*

We defined the answer to be $M(m,r)$ and determine it precisely for $r = 3$, for ends of the range with $k = 3$ (for $m = 2r - 1$ and up to lower order term for $m = 3r - 4$), for certain small cases such as $m = 20, r = 5$ and determine it up to $O_r(1/\sqrt{m})$ in general. While the parameter $k = \lceil \frac{m}{r-1} \rceil$ seems to control the rough behaviour of $M(m,r)$, in order to obtain precise results one needs to take into account the residue of m modulo $r-1$. We have seen this in the $r=3$ case with the distinction between even and odd cases. This is also evident in the $k=3$ case from our results for the ends of this range. The behaviour for $k=3$ across the whole range also seems interesting and may be a good starting point for obtaining precise general results.

In general we can show that for the smaller half of the non-zero residues the $O_r(1/\sqrt{m})$ term is needed. It could be interesting to determine what happens for the remaining half of the residues and in particular when $r-1 \mid m$. Based on our results for $r=3$ and $m=20, r=5$, it seems plausible that $M(m,r) = m_2(K_k) = \frac{k+1}{2}$ for any r and $m = k(r-1)$.

Acknowledgments. We would like to thank Alexandr Kostochka for helping us find [149] and Michael Krivelevich for useful comments.

Appendix A: Independence number in sparse triangle-free graphs

Theorem 2.1.25. *Any 3-degenerate, triangle-free graph G with m vertices and no independent set of size r has $e(G) \geq 6m - 13r - 1$.*

Proof. Our proof is by induction on m. We will actually prove a slightly stronger result. We will show that $e(G) \geq 6m - 13r - 1 + I$, where I is equal to 1 if $\delta(G) < 3$ and equal to 0 otherwise.

For the base case we show the result for $m \leq 5$ and any r. If $m \leq 2$ the result is immediate since the desired bound follows from $e(G) \geq 0$ if $m \geq 3$ since the graph is triangle-free we must have $r \geq 2$ so the bound again follows for $m \leq 4$ immediately and if $r \geq 3$ also for $m = 5$, while if $r = 2$ then G must be C_5 and the bound again holds.

Now let us assume $m \geq 6$ and that the result holds for any graph on at most $m-1$ vertices satisfying our conditions. If there is an isolated vertex in G then removing it we obtain a graph on $m-1$ vertices with independence number at most $r-1$ which implies $e(G) \geq 6(m-1) - 13(r-1) - 1 \geq 6m - 13r + 6 \geq 6m - 13r - 1 + I$. Similarly if there is a vertex of degree 1 removing it and its neighbour leaves us with a graph on $m-2$ vertices with independence number at most $r-1$ giving us the bound $e(G) \geq 6(m-2) - 13(r-1) - 1 \geq 6m - 13r \geq 6m - 13r - 1 + I$. So we may assume $\delta(G) \geq 2$.

If $\delta(G) = 2$ let v be a vertex of degree 2, with neighbours u, w. If there are at most 2 vertices, other than v, adjacent to one of u or w then removing them and v, u, w leaves us with a graph on at least $m-5$ vertices with independence number at most $r-2$. This leftover graph has at least $6(m-5) - 13(r-2) - 1 = 6m - 13r - 5$ edges. On the other hand G in addition has at least 5 edges touching these 5 removed vertices since minimum degree is 2 giving us $e(G) \geq 6m - 13r \geq 6m - 13r + I - 1$ as desired. So there are at least 3 vertices (other than v) adjacent to u or w and in particular v, u, w are incident to at least 5 edges. By removing v, u and w we obtain a graph on $m - 3$ vertices with independence number at most $r - 1$ which hence has at least $6(m-3) - 13(r-1) - 1 = 6m - 13r - 6$ edges. Since v, u, w are incident to at least 5 edges we need to gain one more. Note first that $G \setminus u, v, w$ must have minimum degree 3 or we gain one from its I term. Note also that either u or w must have degree 2 or v, u, w touch at least 6 edges and we gain. Say u is of degree 2, and w' is its neighbour other than w. w' must have degree at least 3 in $G \setminus u, v, w$ so at least 4 in G. This means that u, v, w' touch at least 6 edges and repeating the above argument with u in place of v we are done.

Final case is if $\delta(G) = 3$. We may assume G is connected as otherwise we may apply induction on each of the components and are done. If G is 3-regular then the number of edges is $e(G) = 3m/2$ and a result of Staton (see Theorem 6 in [225]) on graphs with maximum degree 3 implies $r \geq \frac{5}{14}m$ this implies $e(G) = 3m/2 \geq 6m - 13r$ as desired. Hence, we may assume there is a vertex of degree at least 4 and in particular, since G is 3-degenerate that there exists a vertex v of degree 3 adjacent to a vertex u of degree 4. Let w, q be the remaining neighbours of v. Since $\delta(G) = 3$ and $d(u) = 4$ we know there are at least 10 edges touching v, u, w or q. Removing these 4 vertices we get a graph on $m - 4$ vertices with independence number at most $r - 1$ so by induction it has at least $6(m-4) - 13(r-1) + I' - 1 = 6m - 13r - 1 - 11 + I'$ edges, where $I' = 1$ if the remainder graph has minimum degree at most 2. This means that we are done unless there are exactly 10 edges touching v, u, w, q and the remainder graph has minimum degree at least 3. w is a vertex of degree 3 with two neighbours in the remainder graph, so each having degree at least 3 there and at least 4 in G (since they are adjacent to w). This means that w and its neighbours touch at least 11 edges so repeating the argument as above with w in place of v we obtain the desired bound. $\square$

Appendix B: The M(20,5) case

In order to show $M(20, 5) \geq 3$, we will need a few intermediate results. Let us define $e(m, r)$ to be the minimum possible number of edges in an m-vertex graph with independence number at most $r - 1$, provided it has 2-density less than $\frac{k+1}{2}$ (where $k = \lceil m/(r-1) \rceil$). It can be thought of as the variant of determining $M(m, r)$ where we care about the final number of edges instead of the 2-density, but we impose a restriction of not having too dense parts.

For determining $M(20, 5)$ we will need bounds on $e(14, 4)$ and $M(15, 4)$ for which in turn we will need $e(9, 3)$.

Lemma 2.1.26. $e(9, 3) \geq 19$.

Proof. Let G be a graph with 9-vertices with $\alpha(G) \leq 2$ and $m_2(G) < 3$. If G has a vertex v of degree at most 3 then v has at least 5 non-neighbours which must span a clique (since $\alpha(G) \leq 2$) which implies $m_2(G) \geq m_2(K_5) = 3$, a contradiction. If $\delta(G) \geq 5$ then $e(G) \geq 9 \cdot 5/2 > 19$, and we are done. So $\delta(G) = 4$ and there exists a vertex v with degree exactly 4. Let $L = v \cup N(v)$ and $R = G \setminus L$, so $|L| = 5, |R| = 4$. Since $|L| = 5$ there must be a missing edge in L (or we find a K_5 and have $m_2(G) \geq 3$). Observe that vertices making a missing edge in L can not both be non-adjacent to the same vertex in R (or $\alpha(G) \geq 3$) meaning they need to send at least 4 edges towards R. We obtain that there must be at least $4 + 3 \cdot 4 - 3 = 13$ edges touching L, since there are 4 cross edges touching the missing edge, the remaining 3 vertices inside L each have degree at least $\delta(G) = 4$ and we double counted only the edges between these 3 vertices, so at most 3. Now since R consists of non-neighbours of v it spans a K_4 so there are 6 edges within R and in total we have the claimed $13 + 6 = 19$ edges. $\square$

Lemma 2.1.27. $e(14, 4) \geq 33$.

Proof. Let G be a graph with 14-vertices with $\alpha(G) \leq 3$ and $m_2(G) < 3$. If G has a vertex v of degree at most 3 then v has at least 10 non-neighbours which contain no independent set of size 2 so must have 2-density at least 3 by Lemma 2.1.16. If $\delta(G) \geq 5$ then $e(G) \geq 14 \cdot 5/2 > 33$ as desired. So $\delta(G) = 4$ and there exists a vertex v with degree exactly 4. Let again $L = v \cup N(v)$ and R be the rest of the graph. Note now that a missing edge in L must touch at least 5 edges going to R, as otherwise there are $9 - 4 \geq 5$ common non-neighbours of the missing edge and they must span a clique. Similarly as in Lemma 2.1.26 this means there needs to be 14 edges touching L. On the other hand $G[R]$ is a 9-vertex graph with $\alpha(G[R]) \leq 3$ (since v is a non neighbour to anyone in R) and it has $m_2(G[r]) \leq m_2(G) < 3$ so Lemma 2.1.26 applies implying there are at least 19 edges inside $G[R]$ and giving us our claimed total. $\square$

Backtracking along the above proofs it is not hard to show that they are both optimal and even deduce a lot about the structure of the optimal examples. We are now ready to prove our final intermediate result which determines $M(15, 4)$.

Lemma 2.1.28. $M(15, 4) = 3$.

Proof. Let G be a graph with 15-vertices with $\alpha(G) \le 3$ and $m_2(G) < 3$. If there is a vertex of degree at most 4 removing it and its neighbourhood, similarly as above reduces to the case of $M(10,3)$ which by Lemma 2.1.16 is equal to 3. So $\delta(G) \ge 5$. If $\delta(G) \ge 6$ then $e(G) \ge 45$ and $d_2(G) \ge \frac{45-1}{15-2} > 3$. So let v be a vertex with degree exactly 5. Let $L = v \cup N(v)$ and R be the rest of the graph. $G[R]$ is a 9-vertex graph with $\alpha(G[R]) \le 2$ and it has $m_2(G[r]) \le m_2(G) < 3$, so R spans at least 19 edges by Lemma 2.1.26. We claim there needs to be at least 21 edges touching L. If there were a triangle of missing edges inside L its vertices need to send at least 9 edges across (or we get $\alpha(G) \ge 4$) so our usual calculation tells us there are $9 + 3 \cdot 5 - 3 = 21$ edges touching L as desired. Note also that missing edges inside L can not span a star (or we can remove its centre and be left with a K_5 inside L so there need to exist 2 disjoint missing edges. Each missing edge sends at least 5 edges across (or we find a K_5 in their common non-neighbourhood) and since we are assuming there is no missing triangle inside L there needs to be 2 actual edges inside L between our missing pair. Putting these together we get $2 \cdot 5 + 2 + 2 \cdot 5 - 1 = 21$ edges touching L as desired. So there are at least $21 + 19 = 40$ edges in G and $m_2(G) \ge 3$ as desired. □

Theorem 2.1.29. $M(20,5) = 3$.

Proof. Let G be a graph with 20-vertices with $\alpha(G) \le 4$ and $m_2(G) < 3$. If there is a vertex of degree at most 4 removing it and its neighbourhood, reduces to the case of $M(15,4)$ solved in Lemma 2.1.28. So $\delta(G) \ge 5$. If there are more than 8 vertices of degree greater than 5 then $e(G) \ge 55$ and $d_2(G) \ge \frac{55-1}{20-2} = 3$. So there are at least 12 vertices of degree 5. Let v be such a vertex and Let $L = v \cup N(v)$ and R be the rest of the graph. $G[R]$ is a 14-vertex graph with $\alpha(G[R]) \le 3$ and it has $m_2(G[r]) \le m_2(G) < 3$, so R spans at least 33 edges by Lemma 2.1.27. If we find 22 edges touching L we obtain $e(G) \ge 55$ and are done.

Note that any missing triangle in L sends at least 10 edges across (or the ≥ 5 common non-neighbours need to make a clique). So as in the previous lemma we obtain at least 22 edges touching L as desired. Once again the missing edges can not span a star or we find a K_5 inside L so again we need to be able to find a disjoint pair of missing edges in L. Since again any missing edge must send at least 5 edges to R (or we reduce to the case of $M(10,3) = 3$) and at least 2 actual edges must exist between the missing edges (or we find a missing triangle) so again we obtain $2 \cdot 5 + 2 + 2 \cdot 5 - 1 = 21$ edges touching L.

This already implies $M(20,5) \geq 53/18$ but to obtain our best bound we need to work a little bit harder. In particular, if we do find only 21 edges touching L above we obtain a lot of structural information about L.

Claim. *If there are* 21 *edges touching L then* $G[L]$ *induces either:*

a) A K_3 *and a* K_4 *intersecting in a single vertex and consisting entirely of vertices of degree* 5 *in G or*

b) Two K_4*'s intersecting in an edge, one* K_4 *consists of vertices of degree* 5 *and the remaining* 2 *vertices are of degree* 6.

Proof. We have already observed that there are no missing triangles in L and that we can find 2 disjoint missing edges, which gave us at least 21 edges touching L. In order for this bound to be tight there needs to be exactly two actual edges in between them and in order to avoid missing triangles this means that the 4 vertices making these two edges span a missing C_4 with both cross edges present. Each of the 4 missing edges making this C_4 must send exactly 5 edges across (or we again gain) and every vertex of this C_4 has at most 3 neighbours in L so has at least 2 neighbours outside. The only way this can happen is if 2 diagonal (so adjacent in G) vertices of L send 3 edges out (call them w_1, w_2) and remaining 2 send 2 edges out (call them s_1, s_2). By minimum degree in G being at least 5 we know that both s_i must be joined to both remaining vertices in L (v and call the final vertex u). Since v is adjacent to everyone the only edges we don't have information about are uw_1, uw_2. If only uw_1 is a missing edge then u, w_1 must send at least 5 edges across and since w_1 sends 3 u must send 2. But this would imply u is adjacent to v, w_2, s_1, s_2 and these 2 outer vertices so have degree at least 6 and improve our bound. Therefore, there are only 2 options both uw_1 and uw_2 are edges, giving us the case b) or neither are edges, giving us case a). Note that we know u, v must be of degree 5 or we gain and since we know how many edges s_i's and w_i's send outside and how $G[L]$ looks like we know degree's of everyone in G. □

Observe that v was an arbitrary vertex of degree 5 so the above must hold for any such vertex.

Claim. *Case b) happens for all vertices of degree* 5.

Proof. Let us assume case a) happens for vertex v. In other words $v \cup N(v)$ induces a triangle v, u, w and a $K_4 = v \cup U_v$. Observe that $v \in N(u)$ and v is not adjacent to 3 vertices in $N(u)$ outside $N(v)$ (u has degree 2 in $v \cup N(v)$ and degree 5 in G). This means $u \cup N(u)$ must also fall under case a) since in

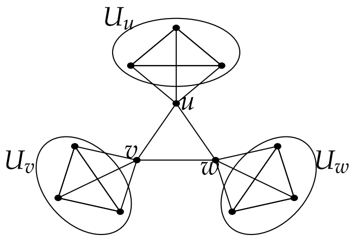


Figure 2.6: Graph W.

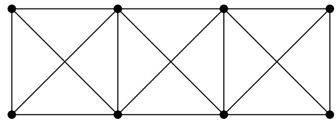


Figure 2.7: Graph H.

case b) any vertex in $G[u \cup N(u)]$ has at most 2 non-neighbours. So $u \cup N(u)$ makes a triangle u, v, w plus $K_4 = u \cup U_u$ (we know the triangles match since only the vertices of the triangle miss 3 edges in the neighbourhood, which means v is in the triangle and is only adjacent to u, w). Analogously we obtain U_w with the same picture. So we find a graph W consisting of 3 vertex disjoint K_4's each with a singled out vertex such that the singled out vertices make a triangle (see Figure 2.6 for an illustration).

We know u, v, w have no further edges among vertices of W as they have degree 5 both in G and W. But we claim that also there are no edges between distinct U_v, U_u, U_w, meaning we find W as an induced subgraph. To see this if say $x \in U_v, y \in U_u$ are adjacent then they can send at most 1 edge outside W (they have degree 5 in G and 4 in W). Since u and x are independent and there are at most 8 vertices adjacent to one of them (w, U_v, U_u and the potential neighbour of x outside W) this means that removing $u, N(u), x, N(x)$ removes at most 10 vertices and leaves us with a 10 vertex graph which has no independent set of size 3 (or together with u and x we get a size 5 one in G) so this reduces to the $M(10, 3)$ case and we are done.

Observe that every vertex of U_v, U_w, U_u has degree 3 in W and 5 in G so sends exactly 2 edges to $G \setminus W$. If there is $x \in U_u$ and $y \in U_w$ which have a common neighbour outside W then there are at least 5 vertices outside of W which are not adjacent to either x or y. In particular, there is a missing edge among such vertices, which together with x, y and v makes an independent set of size 5. This means that $N(U_u), N(U_w)$ and $N(U_u)$ when restricted outside of W must be disjoint. In particular, since there are 8 remaining vertices there is one neighbourhood, say of U_v consisting only of 2 vertices. Since every vertex of U_v sends 2 edges outside W this means that U_v and these 2 vertices span a $K_{3,2}$. Finally, this means that for any vertex v' of U_v we have $v' \cup N(v')$ induces a K_6 minus a triangle, which falls under neither of our cases. This is a contradiction since v' has degree 5 in G. $\square$

Let H denote the graph consisting of two vertex disjoint K_4's with two pairs of vertices, one pair per part joined by a $K_{2,2}$. See Figure 2.7 for an illustration.

Claim. *Any vertex of degree* 5 *lies in an induced copy of* H *in* G *where vertices of degree* 3 *in* H *have degree* 6 *in* G.

Proof. Let v be a vertex of degree 5, since case b) occurs we obtain two K_4's sharing an edge uv. One of the K_4's consists only of vertices of degree 5 so u, v and call its remaining 2 vertices w, s. So by looking at the neighbourhood of w, since case b) must occur we deduce there is a K_4 which intersects our two original cliques only in ws (since v, u would have too high degree if they were in the shared edge). This gives us a desired copy of H and it remains to be shown H is induced.

u, v, s, w all have degree 5 in H and 5 in G so they have no additional edges. If we had an edge between x and y, both of which are of degree 6 and which satisfy $x \sim w, y \sim v$ then x has 4 edges in $G[H]$ so sends at most 2 edges outside. Deleting these two edges and our copy of H we deleted at most 10 vertices among which we deleted $x, N(x), v, N(v)$ and x and v are independent, so we again reduce to the $M(10, 3)$ case and are done. □

Claim. *No two copies of* H *as in the previous claim can intersect.*

Proof. The central K_4 in any such copy of H consists of vertices of degree 5 in G and in H so they have to be disjoint between copies of H.

If a vertex x of degree 6 is shared by two copies H_1, H_2 of H then x has either 6 distinct edges among $H_1 \cup H_2$ or 5, but then $|H_1 \cap H_2| \geq 2$. Let us delete H_1, H_2 and the potential external neighbour of x. In either case we deleted at most 15 vertices. Let us choose $v_i \in H_i$ such that $d(v_i) = 5$ and $v_i \nsim x$. This means x, v_1, v_2 make an independent set and we deleted them and their neighbourhoods in G. This means that among remaining 5 vertices there can be no missing edges as any such edge together with x, v_1, v_2 would span a K_5. □

Finally this means that the copies of H we find should be vertex disjoint and since there are 12 vertices of degree 5 and each H contains 4 this means there should be at least 3 copies so at least 24 vertices in total giving us a contradiction and completing the proof. □

Appendix C: Upper bounds for the Turán 2-density problem

In this section we show our upper bounds on $M(m, r)$. We begin with some basic observations which will prove useful in computing 2-densities. We say two vertices v and u are *equivalent* if there is a transposition swapping vertices v and u in the automorphism group of G.

Lemma 2.1.30. *Let H be the subgraph of G maximising $d_2(G)$ which has as many vertices as possible. For any pair of equivalent vertices, H contains either both or none of them.*

Proof. Let v and u be equivalent vertices and assume for the sake of contradiction that H contains v but not u. Let d denote the number of neighbours of v in H, in particular by equivalence u also has d neighbours in $H \setminus \{v\}$, so in particular at least d neighbours in H.

By maximality of H we must have $\frac{e(H)-1}{|H|-2} \geq \frac{e(H)-d-1}{|H|-1-2} \implies d \geq \frac{e(H)-1}{|H|-2}$ or $H \setminus \{v\}$ would have larger 2-density. Similarly, by comparing with $H \cup \{u\}$ we must have $\frac{e(H)-1}{|H|-2} > \frac{e(H)+d-1}{|H|+1-2} \implies d < \frac{e(H)-1}{|H|-2}$ giving us a contradiction. □

Lemma 2.1.31. *For any $k \geq 2$ we have $M(2k, 3) \leq \frac{k+1}{2}$.*

Proof. We take G to be a disjoint union of two K_k's. It is immediate that $\alpha(G) \leq 2$ so we only need to show $m_2(G) = m_2(K_k) = \frac{k+1}{2}$. To this end let us take the subgraph H of G as in Lemma 2.1.30. Since all the vertices within a single clique are equivalent we deduce that H is either K_k or the whole graph. In the first case we are done immediately. In the second case we have $m_2(G) = d_2(G) = \frac{k(k-1)-1}{2k-2} < \frac{k}{2} < \frac{k+1}{2}$ giving us a contradiction. □

Remark. There are more interesting examples, for example one can find tight examples which are K_k-free. If k is even one can take $k/2$-th power of the cycle on $2k$ vertices while for odd k one can find a different Cayley graph of the Cyclic group of order $2k$.

Lemma 2.1.32. *For any $k \geq 4$ the bound in Theorem 2.1.18 is tight.*

Proof. Let $m(t) = \min\left(\frac{t}{k-2}, \frac{(k+1)/2-(t-1)^2}{2k-3}\right)$ and let a be the maximiser of this expression (over $1 \leq t \leq k-2$), as in the proof of Theorem 2.1.18. Our goal is to find a graph G on $2k-1$ vertices with $\alpha(G) \leq 2$ which has $m_2(G) \leq \frac{k+1}{2} - m(a)$. We take G to be the blow-up of C_5 with cliques placed inside of parts of size $1, k-1-a, a, a, k-1-a$, in order around the cycle. Since complement of G is an actual blow-up of C_5 it is triangle-free so $\alpha(G) \leq 2$.

Observe first that $e(G) = 2(k-1-a) + 2\binom{k-1}{2} + a^2 = k(k-1) + (a-1)^2 - 1$, so

$$\frac{e(G)-1}{|G|-2} = \frac{k+1}{2} - \frac{(k+1)/2 - (a-1)^2}{2k-3} \leq \frac{k+1}{2} - m(a).$$

Note that for $k \geq 4$ we have $\frac{(k+1)/2-(t-1)^2}{2k-3} \leq 1/2$ for any $t \geq 1$. This implies $m(a) \leq 1/2$ and in particular since $m_2(K_{k-1}) = k/2 \leq (k+1)/2 - m(a)$ it is enough to consider only proper subgraphs of size at least k or more when computing $m_2(G)$ (or the desired bound holds). Let H be the subgraph of G maximising $d_2(H)$ as in Lemma 2.1.30. Since all vertices within a part are equivalent we know H must be a union of full parts. Above considerations tell us that H is not the whole graph and has at least k vertices.

Let us also observe that if $\frac{a-1}{k-2} \geq 1/2$ the second term is smaller (or equal) in the minimum for both $m(a-1)$ and $m(a)$ as well as being decreasing in t which would make $m(a-1)$ larger than $m(a)$, a contradiction. Hence, $a-1 < (k-2)/2$, implying $a \leq (k-1)/2$ and in particular $k-1-a \geq a$. This implies that $\delta(G) = \min(2(k-1-a), k-1) \geq k-1$. If H has $2k-1-t$ vertices with $1 \leq t \leq k-2$ we would have

$$\begin{aligned}
d_2(H) = \frac{e(H)-1}{|H|-2} &\leq \frac{e(G) - t\delta(G) + \binom{t}{2} - 1}{2k-t-3} \\
&\leq \frac{k(k-1) + (a-1)^2 - 1 - t(k-1) + \binom{t}{2} - 1}{2k-t-3} \\
&= \frac{k+1}{2} + \frac{(a-1)^2 - (k+1)/2 - t(k-3)/2 + \binom{t}{2}}{2k-t-3} \\
&\leq \frac{k+1}{2} - \frac{(k+1)/2 - (a-1)^2}{2k-3} \leq \frac{k+1}{2} - m(a),
\end{aligned}$$

where in the first inequality we used the fact that H is obtained from G by removing a set of t vertices and hence, these t vertices are incident to at least $\delta(G)t - \binom{t}{2}$ edges.

Finally if H has exactly k vertices it is easy to see that it misses at least a edges (since it is union of parts of the blow-up and every part except the single vertex one has size at least a) which implies $\frac{e(H)-1}{k-2} \leq \frac{\binom{k}{2}-a-1}{k-2} = \frac{k+1}{2} - \frac{a}{k-2} \leq \frac{k+1}{2} - m(a)$, completing the proof. □

The following lemma will be useful in determining when it makes sense to take vertex disjoint unions of graphs in computation of $m_2(G)$. For two graphs G and H we will denote by $G \sqcup H$ the graph obtained by taking a vertex disjoint union of G and H.

Lemma 2.1.33. *Assume $2e(H) > |H|$ and $2e(G) > |G|$. Then $d_2(G \sqcup H) < \max(d_2(G), d_2(H))$.*

Proof. Assume otherwise, so $d_2(G \sqcup H) \geq d_2(G), d_2(H)$. We have $d_2(G \sqcup H) = \frac{e(G)+e(H)-1}{|G|+|H|-2} \geq \frac{e(G)-1}{|G|-2} = d_2(G) \Leftrightarrow \frac{e(G)+e(H)-1}{e(G)-1} \geq \frac{|G|+|H|-2}{|G|-2} \Leftrightarrow \frac{e(H)}{e(G)-1} \geq \frac{|H|}{|G|-2} \Leftrightarrow \frac{e(H)}{|H|} \geq \frac{e(G)-1}{|G|-2}$. Observe now that since $|H| < 2e(H)$ then $\frac{e(H)}{|H|} < \frac{e(H)-1}{|H|-2} = d_2(H)$ so $d_2(G \sqcup H) \geq d_2(G) \implies d_2(H) > d_2(G)$. Repeating with H in place of G we obtain $d_2(G) > d_2(H)$ which is a contradiction. □

The following result shows that bound of Theorem 2.1.7 is tight (up to constant factor in the constant term) for the first half of the modulae.

Theorem 2.1.34. *Let $k = \lceil \frac{m}{r-1} \rceil$ and $\ell = m - (k-1)(r-1)$. Provided $\ell \leq \frac{r-1}{2}$ and that k is sufficiently larger than r there exists $c_r > 0$ such that $M(m, r) \leq \frac{k+1}{2} - \frac{c_r}{\sqrt{k}}$.*

Proof. Let us first describe our graph G. We let G consist of ℓ vertex disjoint copies of our graph G' with $2k-1$ vertices from Lemma 2.1.32 and $r-1-2\ell \geq 0$ copies of K_{k-1}. The only information we will use about G' is that $m_2(G') \leq \frac{k+1}{2} - \frac{1}{2\sqrt{k+1}}$ and that $\alpha(G') \leq 2$.

Observe first that $\alpha(G) < r$ as any set of r vertices of G must have either 3 vertices in some copy of G' or 2 vertices in a copy of K_{k-1}, either way there is an edge and the set is not independent.

On the other hand observe that any graph H with $\alpha(H) < r$ and more than $3r$ vertices must satisfy $2e(H) > |H|$, since there can be at most $(1 - 1/r)|H|^2/2$ non-edges by Turán's theorem so $e(H) \geq |H|^2/(2r) - |H|/2$.

Let us now take H to be a subgraph of G maximising $d_2(H)$ and our goal is to show that, upon minor modification this H must live inside a single copy of G' or K_{k-1} which will let us complete the proof since we know both have small enough m_2.

We know that H must have at least k vertices or we obtain $m_2(G) = d_2(H) \leq m_2(K_{k-1}) = k/2$. This means that if H has at most $3r$ vertices inside some copy of G' or K_{k-1} we can remove them and this only changes the 2-density by $O(\frac{1}{k}) \leq o(\frac{1}{\sqrt{k}})$, where we are taking k sufficiently larger than r. Let us call the graph H_1 obtained from H by deleting all vertices of H belonging to a copy of G' or K_{k-1} if there are at most $3r$ of them inside that copy. Our new graph H_1 by above observation has $d_2(H_1) \geq d_2(H) - O(\frac{1}{k})$ and has the property that inside any copy of G' or K_{k-1} it has either none or more than $3r$ vertices.

Applying the preceding lemma iteratively to restrictions of H_1 to a single copy of G' or K_{k-1} we can find a subgraph H_2 of either G' or K_{k-1} which satisfies $d_2(H_2) > d_2(H_1) \geq d_2(H) - O(\frac{1}{k})$. Since we know $d_2(H_2) \leq \max(m_2(K_{k-1}), m_2(G'))$ and we know $m_2(K_{k-1}) = k/2$ and $m_2(G') \leq \frac{k+1}{2} - \frac{1}{2\sqrt{k+1}}$ so either way $m_2(G) = d_2(H)$ is small enough. $\square$

2.2 LOCALLY RAMSEY GRAPHS

Ramsey theory refers to a large body of deep results, which roughly say that any sufficiently large structure is guaranteed to have a large well-organised substructure. Its inception dates back to 1929 and the celebrated theorem of Ramsey [209] which states that any sufficiently large graph must contain a clique or an independent set of arbitrarily large size. In terms of quantitative results in 1935 Erdős and Szekeres [97] showed that any graph on n vertices contains a clique or an independent set of size $0.5 \log n$. On the other hand in what was one of the first applications of the now indispensable probabilistic method Erdős [84] has shown that in a random graph $\mathcal{G}(n, 1/2)$ w.h.p. there are no cliques or independent sets of size $2 \log n$. Despite considerable effort [26, 49, 61, 113] there are still no known non-probabilistic constructions which match the random graph.

We say a graph is *k-Ramsey* if it contains neither a clique nor an independent set of size k. In general an n-vertex graph is said to be a *Ramsey graph* if it is k-Ramsey for some k "close" to $\log n$. Over the years there has been a wide body of work studying properties of Ramsey graphs. In particular, based on the apparent difficulty of finding non-probabilistic Ramsey graphs, it is widely believed that with an appropriate definition of "close" any Ramsey graph must be random-like. While there is a vast number of results (see [15, 98, 108, 148, 172, 173, 204, 218] and references within) showing that indeed Ramsey graphs need to satisfy, to an extent, various properties usually associated with random graphs, our understanding of Ramsey graphs is still far from sufficient to consider this claim in any way settled.

Given an integer k let G be a k-Ramsey graph with the largest number of vertices. Observe that G must contain *both* a clique and an independent set of size $k-1$, as otherwise we can add a new vertex joined to all or none of the vertices of G to find a larger k-Ramsey graph. This shows that if we have a good Ramsey graph the largest clique and largest independent set should be of similar size. For example, if we consider the random graph, which is the best known Ramsey graph, it will with high probability contain both a clique and an independent set of size a 1 or 2 less than the largest k for which it is k-Ramsey. Furthermore, Ramsey graphs satisfy a similar property locally as well. Given a k-Ramsey graph, since it has no clique or independent set of size k, we know that any subset consisting of $R(k, \ell)$[4] vertices contains *both* a clique and an independent set of size ℓ.

4 $R(k, \ell)$ denotes the off-diagonal Ramsey number, defined as the minimum number of vertices in a graph needed to guarantee there is either a clique of order k or an independent set of order ℓ.

The so called local-global principle, stating that one can obtain global understanding of a structure from having a good understanding of its local properties, or vice versa, has been ubiquitous in many areas of mathematics and beyond for many years [20, 140, 156, 177, 233]. Keeping this in mind the following problem of Erdős and Hajnal [82] seems to be very relevant to understanding Ramsey graphs. Given a k (which might be a function of n) and an n-vertex graph G they ask what is the smallest m for which any m vertex subset of G contains both a clique and an independent set of size k? We denote the answer by $m_G(k)$ and say that a graph is (m,k)*-locally Ramsey* if $m \geq m_G(k)$. To see the relation to Ramsey graphs first observe that being $(m,2)$-locally Ramsey is equivalent to being m-Ramsey. Secondly, more interestingly if we can find an $n \geq 3m$ vertex $(m,k+1)$-locally Ramsey graph it can contain at most $k-1$ vertex disjoint cliques of size m/k, as otherwise they would give us a set of m vertices in which there is no independent set of size $k+1$. The same clearly applies for independent sets. So if we remove a maximal collection of such cliques and independent sets we are left with a graph on at least $n/3$ vertices which is m/k-Ramsey in addition to still being $(m,k+1)$-locally Ramsey.

This means that understanding the behaviour of $m_G(k)$ very well could lead us to better understanding of Ramsey graphs, as well as interesting new examples of Ramsey graphs. Let us first consider what happens with $m_G(k)$ for the random graph $G \sim \mathcal{G}(n,1/2)$. If k is small compared to n we have that w.h.p. $m_G(k) = \Theta(k \log n)$ (see Section 2.2.3) which as we will see, and as one might expect since for $k=2$ this is the standard Ramsey problem, is actually smallest possible among all graphs. On the other hand we also have by Erdős's results [84] that $m_G(k) \geq 2^{k/2}$ so as k becomes larger than $\log \log n$ the bound deteriorates quickly. For example, one needs at least $\sqrt{n}$ size sets to guarantee to be able to find both cliques and independent sets of size $\log n$.

A natural question is whether one can do better. In fact, Erdős [82] singled out the case of $k = \log n$ and asked if such graphs exist with $m_G(k) = (\log n)^3$. If the answer were positive this would give rise to $(\log n)^2$-Ramsey graphs which are very different from $\mathcal{G}(n,1/2)$, since they would still be $\left((\log n)^3, \log n\right)$-locally Ramsey which is very far from being true in the random graph. This question remains open. However, the Alon and Sudakov [17] showed that $m_G(\log n) \geq \Omega((\log n)^3/\log\log n)$, which perhaps validates Erdős' intuition behind asking the question with the parameters he chose.

On the other hand, in terms of upper bounds nothing better than the one mentioned above, coming from random graphs, namely that there is a graph

G for which $m_G(\log n) \leq O(\sqrt{n})$, was known, leaving the possibility that no significantly different Ramsey graphs arise this way. Perhaps surprisingly our main result shows this is not the case, giving a twofold sub-polynomial improvement over the above bound.

Theorem 2.2.1. *There exists an n-vertex graph G for which*

$$m_G(\log n) \leq 2^{2^{(\log\log n)^{1/2+o(1)}}}.$$

As discussed above this gives rise to Ramsey graphs which, while being worse than the random graph, are significantly better than the classical explicit construction of Frankl and Wilson [113] and even the recent breakthrough explicit construction of Barak, Rao, Shaltiel and Wigderson [26]. While unfortunately our construction does use randomness, it still gives rise to somewhat weaker Ramsey graphs which are very different from $\mathcal{G}(n, 1/2)$ in the sense that they are $\left(2^{2^{(\log\log n)^{1/2+o(1)}}}, \log n\right)$-locally Ramsey.

So far we have restricted attention to the case of $k = \log n$ for simplicity and to allow for easier comparison between results. We do find examples in the general case as well.

Theorem 2.2.2. *For any $n \geq 4$ and $k \geq \log n$ there exists an n-vertex graph G with*

$$\log\log m_G(k) \leq 6\sqrt{\log\log n \log\log k}.$$

Finally, we prove a simple proposition which determines $m_n(k)$, defined as the minimum of $m_G(k)$ over all n vertex graphs G, up to a constant factor, provided k is small enough compared to n.

Proposition 2.2.3. *Provided n is sufficiently large compared to $k \geq 2$ we have $m_n(k) = \Theta(k \log n)$.*

Notation. We denote by I_k the independent set consisting of k vertices. We denote by $\omega(G)$ the clique number of G. All our logarithms are in base 2. We note that when saying a graph G is (m, r)-locally Ramsey we do not require either m or r to be integers, we want that any set of at least m vertices contains a clique and an independent set of size at least r.

2.2.1 *Locally Ramsey graphs and lexicographic products*

We begin with a proposition which provides us with a starting point for our further constructions.

Proposition 2.2.4. *For any n there exists an n-vertex graph which is $(2^{r+8}\log n, r)$-locally Ramsey for all r.*

Proof. Let us first fix an $r \geq 2$ and set $m = \lceil 2^{r+8}\log n \rceil$. The chance that an m vertex induced subgraph of $\mathcal{G}(n,1/2)$ does not contain K_r (or I_r) is equal to the chance that $\omega(\mathcal{G}(m,1/2)) < r$. Using Janson's inequality (as in Section 10.3 in [16]) implies that this probability is at most $e^{-\frac{\mu^2}{2(\mu+\Delta)}}$ where $\mu = \binom{m}{r}2^{-\binom{r}{2}}$ and $\Delta = \mu^2 \cdot \sum_{i=2}^{r-1}\binom{r}{i}\binom{m-r}{r-i}2^{\binom{i}{2}}/\binom{m}{r}$. In our case $\Delta \leq \mu^2\binom{m}{r}^{-1}\binom{r}{2}\binom{m-r}{r-2}\cdot 2\sum_{i=2}^{r-1}2^{2-i} \leq 2\mu^2 r^4/m^2$. It is easy to check that $\mu > m^2/r^4$ and hence $\mu + \Delta \leq 2\Delta$. Thus, $\frac{\mu^2}{2(\mu+\Delta)} \geq \frac{m^2}{8r^4}$ and therefore $\mathbb{P}(\omega(\mathcal{G}(m,1/2)) < r) \leq e^{-\frac{m^2}{8r^4}}$. Finally, by a union bound, the probability that there exists a set of m vertices in $\mathcal{G}(m,1/2)$ which does not have K_r or I_r is at most

$$\binom{n}{m}\cdot 2e^{-\frac{m^2}{8r^4}} \leq 2e^{m(\log n - m/(8r^4))} \leq 2e^{-m\log n/8} \leq \frac{1}{n^2}.$$

Here in the second to last inequality we used $2^{r+8} \geq 9r^4$. Now taking a union bound over all $r \leq \log n$ we deduce that the desired graph exists. $\square$

Note that in fact we proved that $\mathcal{G}(n,1/2)$ is $(2^{r+8}\log n, r)$-locally Ramsey with high probability. This bound can be slightly improved (see Section 2.2.3) but we would gain little in our applications since $m_G(r)$ is going to "go within a log" so we chose for simplicity to show the above bound.

As already mentioned in the introduction, the random graph performs close to best possible when r is very small. Our next construction already does much better when $r \geq \log n$, it serves as a basis and an illustration for our main construction presented in the following section.

Lemma 2.2.5. *For any integer $N \geq 4$ there exists an N-vertex (m,r)-locally Ramsey graph for any m, r which satisfy $\log r \leq \frac{(\log m)^2}{2^9\log N}$.*

Upon inverting we obtain a graph G for which $m_G(r) \leq 2^{16\sqrt{2\log N\log r}}$. In particular, when $r = \log N$ this is already significantly better compared to about $\sqrt{N}$ in case of the random graph.

The example we use to prove the lemma is the lexicographic product of a random graph $\mathcal{G}(n,1/2)$ with itself multiple times. The lexicographic product $G \times H$ of two graphs G and H is defined as the graph on the vertex set $V(G) \times V(H)$ in which two vertices (v,u) and (x,y) are adjacent iff $v \sim_G x$ or $v = x$ and $u \sim_H y$. We write G^ℓ for the lexicographic product of G with itself ℓ times. The main property of the lexicographic product which makes them natural candidates for our graphs is that clique and independence numbers are multiplicative (see [123]). Let us give some intuition as to why this is useful. Let $G \sim \mathcal{G}(n, 1/2)$ and let us compare G^ℓ with the random graph on the same number of vertices $G' \sim \mathcal{G}(n^\ell, 1/2)$. If we take an induced subgraph H of G on m vertices then H^ℓ gives us a subset of m^ℓ vertices of G^ℓ which by the multiplicative property above contains both a clique and an independent set of size about $(2\log m)^\ell$. On the other hand a subset of m^ℓ vertices of G' w.h.p. does not have cliques (or independent sets) of size $2\log(m^\ell) = 2\ell \log m$. This means that, at least if we restrict our attention to subsets of G^ℓ arising in this product fashion, G^ℓ is a much better (m,r)-locally Ramsey graph than G' for most choices of the parameters. Of course one may not just restrict attention to such sets. The following lemma allows one to show that even in arbitrary subsets of G^ℓ it is possible to find big cliques (and independent sets).

The statement of the next lemma is somewhat technical, one of the reasons for this is that we want to state it in a very general form since we want to use it twice with very different choices of parameters. Second reason is that we believe it might be useful in improving other constructions people might come-up with in the future, as well as possibly for other problems involving subgraphs of lexicographic products.

Let us sketch the proof idea. The lemma starts with a graph G on n vertices in which for some $2 \le r_2 < r_3 < \ldots < r_k$ we know that any m_t vertices contain both K_{r_t} and I_{r_t} for all $t \ge 2$. We now take a subset S of G^ℓ in which we want to find a big clique (the argument for an independent set will be analogous). For every vertex $v \in G$ we denote by S_v the subset of S consisting of all elements having v as their first coordinate. We then look at m_t vertices with highest $|S_v|$. For some t all these vertices need to actually have a reasonably large $|S_v|$, say at least m', as we know that $\sum_{v \in G} |S_v| = |S|$. We now use the information that any m_t vertices in G have a clique of size r_t, so in particular among our top m_t vertices some r_t make a clique, say $1, \ldots, r_t$. Now for any two elements of S if their first coordinates are adjacent in G then they are also adjacent in G^ℓ. So if we look at the sets $S_i, i \le r_t$ all the edges between S_i and S_j for $i \neq j$ exist. In particular, if we find a clique in each of S_i we may take a union of these cliques to obtain a clique in S. Since all vertices in S_i share the first vertex and $|S_i| \ge m'$, finding a clique reduces to finding a clique in a subset of size m' of $G^{\ell-1}$ for which we may use induction. For an example, if we work with the assumption that G is $(2\log n, 2)$-locally Ramsey, which we can get from the random graph, then $r_2 = 2, m_2 = 2\log n$ and let $|S| = m$. We split in 2 cases, either some vertex has $|S_v| \ge m/(4\log n)$

or there are $2\log n$ vertices which all have $|S_v| \geq m/(2n)$, since otherwise $|S| = \sum_{v\in G}|S_v| < (2\log n)\cdot m/(4\log n) + n\cdot m/(2n) = m$. In the former case we take a vertex v with $|S_v| \geq m/(4\log n)$ and look for a clique in S_v. This reduces the task to looking for a clique in a subset of size $m/(4\log n)$ of $G^{\ell-1}$ which we do by induction. In the latter case by our assumption on G among $2\log n$ vertices there must exist an edge vu of G. This means that we can find a clique of twice the size we are guaranteed in a subset of size $m/(2n)$ of $G^{\ell-1}$, which we once again do by induction. Optimising the choice of parameters will already bring us close to the bound in Lemma 2.2.5.

It will be more convenient to work with the inverse of $m_{G^\ell}(r)$. So, let $\beta_G(m,\ell)$ denote the largest r such that in any m-vertex subset of G^ℓ we can always find both K_r and I_r.

Lemma 2.2.6. *Let G be an n-vertex graph. Suppose that for some $2 \leq r_2 < \ldots < r_k$ we know that $m_G(r_t) \leq m_t$, for $2 \leq t \leq k$. Then*

$$\log \beta_G(m,\ell) \geq \frac{\log m - \ell\log(2m_2)}{\max\limits_{2\leq t\leq k}\left(\frac{\log(m_{t+1}/m_2)+t}{\log r_t}\right)}, \tag{2.5}$$

for any choice of $m_{k+1} \geq \min(n,m)+1$.

Proof. First note that if for some $i<j$ we have $m_i > m_j$ we may decrease m_i to be equal to m_j, since $r_i < r_j$ implies $m_j \geq m_G(r_j) \geq m_G(r_i)$ and doing this can only increase the target function. So we may assume that m_i is increasing.

We will prove the claim by induction on ℓ for every m satisfying $\min(n,m)+1 \leq m_{k+1}$ (where we are treating G, r_i's and m_i's as fixed parameters). For the base case of $\ell = 1$ since $G^\ell = G$ we have $m \leq n$ so $m_{k+1} \geq m+1$. We also have $m \geq 2m_2$ as otherwise $\beta(m,\ell) < 0$ and the claim is trivial. So there exists some $2 \leq t \leq k$ such that $m_t \leq m < m_{t+1}$. Since $m_t \geq m_G(r_t)$ it is sufficient to show that $\beta_G(m,1) \leq r_t$. This indeed holds since $\log m - \log(2m_2) < \log m_{t+1} - \log m_2 = \log(m_{t+1}/m_2)$.

Now assume $\ell \geq 2$ and that the claim holds for $\ell-1$ and any m for which $\min(n,m)+1 \leq m_{k+1}$. Let S be a set of m vertices in G^ℓ. For any vertex $v \in G$ let us denote by $s(v)$ the number of elements in S which have v as their first coordinate. Let us also set $m_1 = m_G(1) = 1$ and $r_1 = 1$ for convenience.

Claim. *For some $1 \leq t \leq k$ there are m_t vertices $v \in G$ with $|S_v| \geq m/(2^t m_{t+1})$.*

Proof. Let $v_1, \ldots, v_s$ be the vertices of G which have $s(v_i) > 0$ ordered so that $s(v_i)$ is decreasing in i. Note that $s \leq \min(|G|,|S|) \leq m_{k+1}-1$. If $s(v_{m_i}) \geq m/(2^i m_{i+1})$ for some $i \leq k$ we may take $t = i$ and $\{v_1, \ldots, v_{m_t}\}$

all have $s(v_i) \geq m/(2^t m_{t+1})$ so we are done. Hence, we may assume that $s(v_{m_i}) < m/(2^i m_{i+1})$ for all $i \leq k$. But this implies that

$$|S| = \sum_{i=1}^{s} s(v_i) = \sum_{i=1}^{k} \sum_{j=m_i}^{m_{i+1}-1} s(v_j) \leq \sum_{i=1}^{k} (m_{i+1} - m_i) \cdot \frac{m}{2^i m_{i+1}} < m \sum_{i=1}^{k} \frac{1}{2^i} < m,$$

which is a contradiction. $\square$

Since $m_t \geq m_G(r_t)$ among the m_t vertices given by the claim there must be r_t forming a K_{r_t} in G. Let us denote by $S_1, \ldots, S_{r_t} \subseteq S$ the sets of elements in S whose first coordinate is the i-th vertex of this clique. By definition of the lexicographic product if 2 elements in S have adjacent first coordinates they are adjacent in G^ℓ as well. This means that we can combine the cliques we find in each of S_i into a single clique in S. Furthermore, since all elements in S_i have the same first coordinate we can delete their first coordinate when looking for a clique, which leaves us with a subset of at least $m' = m/(2^t m_{t+1})$ elements of $G^{\ell-1}$ to which we can apply induction[5] to find a clique. This means we can find a clique in G^ℓ of size at least $r_t \cdot \beta_G(m', \ell - 1)$. Repeating the argument for independent sets we can find an independent set of this size as well. Therefore:

$$\begin{aligned}
\log \beta(m, \ell) &\geq \log\left(r_t \cdot \beta\left(m', \ell - 1\right)\right) = \log r_t + \frac{\log m' - (\ell - 1)\log(2m_2)}{C} \\
&= \frac{C \log r_t + \log m - \log(2^t m_{t+1}) - \ell \log(2m_2) + \log(2m_2)}{C} \\
&\geq \frac{\log m - \ell \log(2m_2)}{C}
\end{aligned}$$

where $C = C(m_2, \ldots, m_{k+1}, r_2, \ldots, r_k) = \max_{2 \leq t \leq k} \left(\frac{\log(m_{t+1}/m_2) + t}{\log r_t}\right)$ and the last inequality follows since we get an equality if $t = 1$ and since $C \log r_t \geq \log m_{t+1} - \log m_2 + t$ if $t \geq 2$. As this is precisely the RHS of (2.5) this completes the proof. $\square$

Lemma 2.2.5 now follows as a corollary upon making the appropriate choice for the parameters.

Proof of Lemma 2.2.5. Let G be the n-vertex graph which is $\left(2^{t+8} \log n, t\right)$-locally Ramsey for all t provided by Proposition 2.2.4. This implies we can apply Lemma 2.2.6 with $r_t = t$ and $m_t := 2^{t+8} \log n \geq m_G(t)$ for $2 \leq t \leq$

5 Note that $m' < m$ so $\min(n, m') \leq \min(n, m) \leq m_{k+1} - 1$ as required by the assumption.

$k \le \log n$, where k is the largest integer such that $m_k \le n$, and $m_{k+1} := 2^{k+9}\log n > n$. The lemma implies that

$$\log \beta_G(m,\ell) \ge \frac{\log m - \ell \log\log n - 11\ell}{2\log n/\log\log n}, \tag{2.6}$$

since $\log m_2 = 10 + \log\log n$, $\log(m_{t+1}/m_2) = t-1$ and $(2t-1)/\log t$ is increasing in t.

We claim that G^ℓ with an appropriate choice of parameters n and ℓ (in terms of N and m) provides us with the desired graph. Note that we may assume that $m \le N$ (as otherwise the claim is vacuous) and that $\log m \ge 2^4\sqrt{2\log N} \ge 2^4 \log\log N$ as otherwise the claimed bound holds trivially. Let n be the smallest integer for which $\log n/\log\log n \ge 32\log N/\log m$ and let $\ell := \lceil \log m/(32\log\log n)\rceil$. With this choice of n and ℓ we get a graph on $n^\ell \ge N$ vertices, since

$$\ell \log n \ge \frac{\log m}{32\log\log n}\cdot\frac{32\log N\log\log n}{\log m} = \log N.$$

Furthermore, by (2.6) any set of size m contains both cliques and independent sets of size at least 2 to the power

$$\frac{\log m - 12\ell\log\log n}{2\log n/\log\log n} \ge \frac{\log m - \frac{12}{16}\log m}{2\cdot 33\log N/\log m} = \frac{(\log m)^2}{33\cdot 8\log N}$$

where in the denominator we used that $\log n/\log\log n < 33\log N/\log m$ (which holds since $\log n/\log\log n$ grows slower than n and $\log N/\log m \ge 1$) and in the numerator $\ell \le \log m/(16\log\log n)$ (which holds since $\log m/(32\log\log n) \ge 1/2$ and $\lceil x\rceil \le 2x$ for any $x \ge 1/2$). $\square$

2.2.2 *Locally Ramsey graphs and scrambling*

It might be tempting to try to reiterate the argument used in the previous section by starting with our better construction in place of the random graph. Notice however, that all our examples are in fact already powers of the random graph so doing this would only provide us with higher powers of the random graph which are already considered by our argument. This idea however has some merit when combined with a further twist. If we start with a high power of the random graph it will be a much better (m,r)-locally Ramsey than the random graph for some fixed value of r but perform comparatively poorly for small values of r. If we now scramble this graph a little bit, in the sense that we flip every edge and non-edge with some small probability this will improve the performance of our graph when r is small while only slightly decreasing performance for larger r. Taking the lexicographic powers of this graph in place of the random graph is how we obtain our improved construction.

Let us define the p-scramble $\mathcal{G}(G,p)$ of a graph G to be the graph obtained by independently removing every edge of G and adding every non-edge of G with probability p. The following lemma makes formal the above idea that by taking a p-scramble of an (m,r)-locally Ramsey graph G we obtain a graph which is close to being as good a Ramsey graph as $\mathcal{G}(n,p)$, meaning it has no cliques or independent sets of size about $\log n/p$ or in other words is $(\log n/p, 2)$-locally Ramsey but is in addition still close to being (m,r)-locally Ramsey.

Lemma 2.2.7. *If there exists an n-vertex (m,r)-locally Ramsey graph then there exists a graph which is both $\left(m, \frac{r}{17\log n}\right)$-locally Ramsey and $(r/2,2)$-locally Ramsey, provided $r \geq 16\log n$.*

Proof. Let G be an (m,r)-locally Ramsey graph on n vertices. Let $G' \sim \mathcal{G}(G,p)$ with $p := \frac{8\log n}{r} \leq 1/2$.

By an immediate coupling, the probability that G' contains a $K_{r/2}$ (or $I_{r/2}$) is at most the probability that $\mathcal{G}(n, 1-p)$ contains a $K_{r/2}$. This probability is, by a union bound, at most $\binom{n}{r/2}(1-p)^{\binom{r/2}{2}} \leq 2^{r/2\cdot\log n - pr^2/8} = n^{-r/2} \leq \frac{1}{4}$ since $\frac{r}{2} = \frac{4\log n}{p}$. So G' contains neither $K_{r/2}$ nor $I_{r/2}$, or in other words is $(r/2,2)$-locally Ramsey with probability more than $1/2$.

Given a set of r vertices forming a clique in G the probability that in G' this set still contains a clique of size t is equal to the probability that $\mathcal{G}(r, 1-p)$ has a K_t. Note that the expected number of missing edges is $\mu = \binom{r}{2}p = 4(r-1)\log n$ so by Chernoff's inequality (see Appendix A of [16]) the probability that there are more than 2μ edges is at most $e^{-\mu/3}$. If we have less edges then by Turan's theorem (see [16]) there is a clique of size at least $\frac{r^2}{4\mu + r} \geq \frac{r}{16r\log n + r} \geq \frac{r}{17\log n}$. This means that with probability $1 - \binom{n}{r}e^{-\mu/3} \geq 1 - n^{r - \frac{4}{3}(r-1)} \geq 1/4$ any clique of size r in G contains a clique of size at least $\frac{r}{17\log n}$ in G'. Repeating for the independent sets we conclude that G' is $\left(m, \frac{r}{17\log n}\right)$-locally Ramsey with probability at least $1/2$. Therefore, with positive probability the desired graph exists. $\square$

Being more careful one can improve $\frac{r}{17\log n}$ to $\Omega\left(\frac{r\log r}{\log n}\right)$ but not more (since $\omega(\mathcal{G}(r, 1-p)) = \Theta\left(\frac{\log r}{p}\right)$ w.h.p.) However, this improvement seems to be negligible in our applications so we opted for the above simpler argument. The following lemma gives our main construction. We obtain it by starting with our construction from the previous section, scrambling it using Lemma 2.2.7 then taking an appropriate lexicographic power using Lemma 2.2.6 and repeating with this new graph. The parameter t will control the number of iterations that we do. We also, for now, add an assumption that the clique/independent set size r we are looking for is not too small.

Theorem 2.2.8. *For any $t \geq 2$ there exists a (m,r)-locally Ramsey graph on $N \geq 4$ vertices, provided $\log m \geq t^{2t}(\log r)^t(\log N)^{1/t}$ and $\log r \geq t \log\log N$.*

Proof. Let us define $\log m(N,r,t) := t^{2t}(\log r)^t(\log N)^{1/t}$. We will prove by induction on t that for any $N \geq 4$ and $\log r \geq t \log\log N$ there exists an $(m(N,r,t),r)$-locally Ramsey graph on N vertices. The base case of induction for $t = 2$ follows (with room to spare) from Lemma 2.2.5.

Let us take an (m',r')-locally Ramsey graph G on $n \geq 4$ vertices (with parameters r', n satisfying $r' \geq 16\log n$, $\log r' \geq t \log\log n$ and $n \geq 4$, to be chosen later) given by the inductive assumption for some $t \geq 2$, so with $m' := m(n,r',t)$. Let G' be the scrambled graph given by Lemma 2.2.7 applied to G. So in particular G' is $(r'/2,2)$-locally Ramsey and (m',r'')-locally Ramsey where we write $r'' := \lceil r'/(17\log n) \rceil$, note that we are using $r' \geq 16\log n$ so that the lemma applies.

We now take lexicographic products of G'. Lemma 2.2.6 allows us to use these locally Ramsey properties of G' to give bounds on the locally Ramsey properties of G'^{ℓ} which will be our actual example for iteration t. So in particular we may apply Lemma 2.2.6 with $r_2 = 2, m_2 = r'/2 \geq m_{G'}(2); r_3 = r'', m_3 = m' \geq m_{G'}(r'')$ and $m_4 = m+1$. The lemma implies

$$\log \beta_{G'}(m,\ell) \geq \frac{\log m - \ell \log r'}{\max(\log m', \log m/\log r'')}, \tag{2.7}$$

since $\log(2m'/r') + 2 \leq \log m'$ (since $r' \geq 8$) and $\log(2(m+1)/r') + 3 \leq m$ (since $r' \geq 16\log n \geq 32$).

We note that at this point what remains to be done is to choose the parameters and use (2.7) to show the induction step holds. The rest of the proof is somewhat technical and it might help the reader to at first ignore various constants and floors and ceils. Let us now choose all our parameters in terms of N, r and t. Let

$$m := m(N,r,t+1), \qquad \log r' := \left(1+\frac{2}{t}\right)\log r,$$

$$\ell := \left\lfloor \frac{\log m}{(t+1)\log r'} \right\rfloor \text{ and } \qquad \log n := \left\lceil \frac{\log N}{\ell} \right\rceil.$$

Our goal is to show that with this choice of parameters the RHS of (2.7) is at least $\log r$. This would give us a graph on $n^\ell \geq N$ (by definition of n) vertices which is (m,r)-locally Ramsey so we obtain the desired graph by taking a subgraph consisting of exactly N vertices. We first show the following easy inequalities.

Claim. *We have* $16 \leq n \leq N$, $\ell \geq 64$ *and*

$$\log r' \geq (1 + 1/t)(\log r + \log(17 \log n)).$$

Proof. Note that since

$$\begin{aligned}\log m = \log m(N, r, t+1) &= (t+1)^{2(t+1)} (\log r)^{t+1} (\log N)^{1/(t+1)} \\ &> (t+1)^{2(t+1)} \log r\end{aligned}$$

so in particular

$$\frac{\log m}{(t+1)\log r'} = \frac{\log m}{(t+1)(1+2/t)\log r} \geq (t+1)^{2(t+1)-1}/(1+2/t) > 64,$$

which in turn implies $\ell = \left\lfloor \frac{\log m}{(t+1)\log r'} \right\rfloor \geq 64$. This together with $N \geq 4$ and the definition of n imply $n < N$. If $m > N$ then there are no subsets of size at least m in G'^ℓ so the induction step is vacuously true, therefore we may assume $m \leq N$. Using this we get $\log n \geq \frac{\log N}{\ell} \geq \frac{\log N \cdot (t+1) \log r'}{\log m} \geq (t+1)\log r' > 3$, (where we used $r' > 1$ and $t \geq 2$). This in particular implies that $\log n \geq 4$ and $\log n \leq 4/3 \cdot \log N/\ell \leq \log N/32$ (using $\ell \geq 64$). This in turn implies $\log r' = (1+2/t)\log r \geq (1+1/t)\log r + \log\log N \geq (1+1/t)\log r + \log(17 \log n)$, where we are using $1/t \cdot \log r \geq \log\log N \geq \log(32 \log n)$. $\square$

This immediately implies the required inequalities on n, r' and t, indeed $r' \geq 17 \log n$ and $n \geq 4$ while $\log r \geq (t+1)\log\log N$ implies $\log r' \geq \log r \geq (t+1)\log\log N \geq t \log n$.

Let us now turn to the main inequalities. Observe that

$$\begin{aligned}\frac{\log m - \ell \log r'}{\log m / \log r''} &\geq \left(\frac{\log m - \log m/(t+1)}{\log m}\right) \log r'' \\ &\geq \left(1 - \frac{1}{t+1}\right)(\log r' - \log(17 \log n)) \geq \log r, \qquad (2.8)\end{aligned}$$

where in the last inequality we used the main inequality from the claim. This shows that one of the two desired inequalities that we need to show to conclude that RHS of (2.7) is at least $\log r$ holds. The second inequality we need is equivalent to $\log m - \ell \log r' \geq \log m' \log r$ and is implied by $\log m \geq (1+1/t) \log m' \log r$ (since $\ell \leq \log m/((t+1)\log r'))$. Let us now show this inequality holds (recall that we have chosen $m' = m(n, r', t)$):

$$\begin{aligned}(1+1/t)\log r \log m' &= (1+1/t)\log r \cdot t^{2t}(\log r')^t(\log n)^{1/t} \\ &\leq (t+1)t^{2t-1}\log r \cdot (\log r')^{t+1/t} \cdot 2\left(\frac{\log N}{\log m}\right)^{1/t} \\ &\leq (t+1)^{2(t+1)}(\log r)^{t+1+1/t}(\log N)^{1/t} \cdot (\log m)^{-1/t} \\ &\leq (\log m)^{1+1/t} \cdot (\log m)^{-1/t} = \log m,\end{aligned}$$

where in the first inequality we used $\log n \leq \frac{4\log N}{3\ell} \leq \frac{4(t+1)\log r' \log N}{3\log m}$ (following since $n \geq 16$ so $\log n \geq 4$ and definitions of n and ℓ) and $(4(t+1)/3)^{1/t} \leq 2$ (since $t \geq 2$). In the second inequality we used $\log r' = (1+2/t)\log r$ and $(1+2/t)^{t+1/t} \leq (1+1/t)^{2(t+1/t)} \leq (1+1/t)^{2t+1}$. The third inequality follows as $m = m(N,r,t+1) = (t+1)^{2(t+1)}(\log r)^{t+1}(\log N)^{1/(t+1)}$ since $(t+1+1/t)\cdot t/(t+1) \leq t+1$. Together with (2.8) this shows that the RHS of (2.7) is at least $\log r$ completing the proof. □

We were relatively lax with various estimates in the argument above for the sake of simplifying the inequalities as much as possible. For example, the optimal exponent of $\log r$ (which one may obtain using exact same parameters as we did above) is $(t+1)/2 - 2/t$. We also note that the assumption $r \geq (\log n)^t$ was also made for the sake of simplicity, since for smaller values of r the above argument would give barely any improvement over just using the above bound with $r = (\log n)^t$ and monotonicity of $m_G(r)$ in r. Let us now optimise over t and obtain Theorem 2.2.2 as a corollary. Recall the statement of Theorem 2.2.2.

Theorem 2.2.2. *For any $n \geq 4$ and $k \geq \log n$ there exists an n-vertex graph G with*

$$\log\log m_G(k) \leq 6\sqrt{\log\log n \log\log k}.$$

Proof. If $\log N \leq k \leq (\log N)^t$ we use Theorem 2.2.8, with $r = (\log N)^t$ to obtain a graph G with

$$\begin{aligned}\log m_G(k) &\leq \log m_G((\log N)^t) \leq t^{3t}(\log\log N)^t(\log N)^{1/t} \\ &\leq t^{3t}(\log k)^t(\log N)^{1/t}.\end{aligned}$$

If $k > (\log N)^t$ we may use Theorem 2.2.8 directly with $r = k$ to conclude that for any $k \geq \log N$ there is a graph G with $\log m_G(k) \leq t^{3t}(\log k)^t(\log N)^{1/t}$.

We now choose $t = \left\lfloor \sqrt{\frac{\log\log N}{\log\log k}} \right\rfloor$. If $t < 2$ we obtain that the desired inequality requires $m > N$ making the claim vacuous. Hence the above inequality gives us

$$\begin{aligned}\log\log m_G(k) &\leq 3t\log t + t\log\log k + \frac{1}{t}\cdot\log\log N\\ &\leq 4t\log\log k + \frac{1}{t}\cdot\log\log N \leq 6\sqrt{\log\log N\log\log k}\end{aligned}$$

where we used $t \leq \log\log N \leq \log k$ and $t \geq 2$. □

Theorem 2.2.1 follows by simply plugging in $k = \log N$ in Theorem 2.2.2.

Remark. By following the argument used in Proposition 2.2.4 it is not hard to show that $\mathcal{G}(n,p)$ is w.h.p. $\left((1/p)^{r+8}\log n, r\right)$-locally Ramsey for all r, assuming $p \leq 1/2$. Using this in Lemma 2.2.7 would give us a graph which is $\left((1/p)^{r+8}\log n, r\right)$-locally Ramsey for all r in addition to being $\left(m, \frac{1}{3p}\right)$-locally Ramsey, provided $1/p \geq r/(8\log n)$. We can then use this extra information for small values of r similarly as we did in the proof of Lemma 2.2.5 to obtain an improvement in (2.7). Ultimately, this would lead to an improvement in Theorem 2.2.8 in which we divide by roughly a $(\log\log N)^{(t-3)/2}$ factor which would only slightly improve the $o(1)$ term in Theorem 2.2.1.

2.2.3 *Small cliques and independent sets*

In this section we show Proposition 2.2.3. We begin with the lower bound.

Proposition 2.2.9. *Provided $n \geq 4r\log n$ and $r \geq 2$ we have $m_n(r) \geq (0.5 + o(1))r\log n$.*

Proof. Let us start with the lower bound. Given an n-vertex graph G, by the standard bound on Ramsey numbers (see e.g. [97]), G must contain a clique or an independent set of size at least $0.5\log n$. If we remove this set and repeat $2r-3$ times we get either $r-1$ vertex disjoint cliques or $r-1$ vertex disjoint independent sets of size at least $0.5\log(n/2)$ as at each step we are left with at least $n - 2r\log n \geq n/2$ vertices. The union of these sets give us a set of $(0.5+o(1))r\log n$ vertices in which we can not find a clique (if the sets were independent) or an independent set (if the sets were cliques) of size r. This shows that $m_G(r) \geq (0.5+o(1))r\log n$. □

Remark. The above bound applies for essentially the whole range but is beaten by the approach in [17] as soon as r is bigger $\log\log n$. They show that provided there is an I_r in every subset of size s then one can find an independent set of size $\Omega(r\log(n/s)/\log(s/r))$. Using this for our graphs and finding r copies of this big independent set as in the proof above would show $m_n(r) \geq \Omega(r^2\log n/\log\log n)$, provided r is at most polylogarithmic in n which beats the bound in Proposition 2.2.9 when $r \gg \log\log n$.

Let us now turn to the upper bound. Perhaps not too surprisingly since we are working with "small" values of r the example is the random graph $G \sim \mathcal{G}(n,1/2)$. Since n is much bigger than r the argument we used to prove Proposition 2.2.4, even when done more carefully, would only give us $m_G(r) \leq O(r^4\log n)$ so we make use of a slightly different approach.

Proposition 2.2.10. *Provided n is sufficiently large compared to r we have for $G \sim \mathcal{G}(n,1/2)$ that w.h.p. $m_G(r) = \Theta(r\log n)$.*

Proof. The lower bound follows from the previous proposition. So we focus on the upper bound. It is well-known (see [159]) that almost all graphs without K_r are $r-1$ colourable. Hence, provided $m \to \infty$ as $n \to \infty$ we have $\mathbb{P}(\omega(\mathcal{G}(m,1/2)) < r) \leq (1+o(1))\mathbb{P}(\chi(\mathcal{G}(m,1/2)) < r) \leq (1+o(1))r^m 2^{-(1+o(1))m^2/(2r)}$. Here, the last inequality follows since there are $(r-1)^m$ many ways to assign $r-1$ colours to m vertices and given a colouring there are at least $(r-1)\binom{m/(r-1)}{2} \geq (1+o(1))m^2/(2r)$ pairs of vertices assigned the same colour which are not allowed to appear as edges. The same estimate holds for the probability that a graph on m vertices contains no independent set of size r. Thus in G the expected number of sets of size m which contain no clique of size r or no independent set of size r is at most

$$\binom{n}{m} \cdot 2(1+o(1))r^m 2^{-(1+o(1))m^2/(2r)}$$

which tends to 0 for $m = (2+o(1))r\log n$, completing the proof. $\square$

The fact that almost all m-vertex K_r-free graphs are $r-1$ colourable has recently been shown to be true for r up to $\log m/(10\log\log m)$ in [25]. Since our sets have size m which is roughly $r\log n$ this means that n sufficiently large in the above result may be replaced with $r \leq O\left(\frac{\log\log n}{\log\log\log n}\right)$.

2.2.4 *Concluding remarks*

In this section we studied the function $m_G(r)$ with particular interest in how small it can be. The function $m_n(r)$ defined as the minimum of $m_G(r)$ over all n-vertex graphs G was introduced by Erdős and Hajnal almost 30 years ago. Combined with the lower bound obtained in [17] we obtain

$$\frac{(\log n)^3}{\log\log n} \le m_n(\log n) \le 2^{2^{(\log\log n)^{1/2+o(1)}}}.$$

In general it would be very interesting to get better bounds on $m_n(\log n)$ and in particular answer Erdős' question of whether $m_n(\log n) > (\log n)^3$. In fact, we suspect that $m_n(\log n)$ may be bigger than any fixed power of $\log n$. Our initial examples in Section 2.2.1 are essentially the classical examples of explicit Ramsey graphs due to Naor [190], following-up on the idea of using lexicographic products to build Ramsey graphs due to Abbott [1]. There has recently been some major progress on finding better explicit Ramsey examples [26, 49, 61]. It would be interesting to see if one can combine these graphs with our ideas to improve our upper bound.

Another possibly interesting perspective arises if we consider a colouring restatement of our problem. Note that $m_n(r) - 1$ may be defined as the largest number m such that in any 2-colouring of K_n we can find m-vertices not containing a monochromatic K_r in one of the colours. With this in mind one can define the m-local Ramsey number $LR_m(G)$ of a graph G as the smallest n for which in any 2-colouring of K_n there are m vertices not containing a monochromatic copy of G in one of the colours. For example if m is sufficiently larger than r Proposition 2.2.3 implies that $LR_m(K_r) = 2^{\Theta(m/r)}$. Natural generalisations to more colours or asymmetric graphs might hold some interest as well.

Acknowledgements. We would like to thank David Conlon for useful conversations and remarks.

3

TOPICS CONCERNING HYPERGRAPHS AND TOURNAMENTS

3.1 COVERING GRAPHS BY MONOCHROMATIC TREES AND HELLY-TYPE RESULTS FOR HYPERGRAPHS

3.1.1 *Introduction*

Given an r-edge-coloured graph G, how many monochromatic paths, cycles or general trees does one need to cover all vertices of G? The study of such problems has a very rich history going back to the 1960's when Gerencsér and Gyárfás [124] showed that for any 2-colouring of the edges of the complete graph, there are two monochromatic paths that cover all the vertices. Gyárfás [135] later conjectured that the same is true for more colours, i.e. in any r-edge-colouring of K_n, there are r monochromatic paths covering the vertex set. This conjecture was solved recently for $r = 3$ by Pokrovskiy [200], but it is still open for all $r \geq 4$. The best known bound in general is that $O(r \log r)$ monochromatic paths suffice, and is due to Gyárfás, Ruszinkó, Sárközy and Szemerédi [138], improving on results by Gyárfás [135] and Erdős, Gyárfás and Pyber [87]. A similar type of question can be asked if we want to cover with cycles instead of paths. In fact, most results mentioned above also hold for disjoint cycles instead of paths. For further examples, generalisations and detailed history, we refer the reader to a recent survey by Gyárfás [136].

We will study the problem of covering graphs using monochromatic connected components. Let us denote by $\mathrm{tc}_r(G)$ the minimum m such that in any r-edge-colouring of G, there is a collection of m monochromatic trees that cover the vertices of G. Since each connected graph contains a spanning tree, we may replace "tree" in this definition by "connected subgraph" or "component", which we do without further mention throughout the section. The question of covering graphs with monochromatic components was first considered by Lovász in 1975 [181] and Ryser in 1970 [144], who conjectured that $\mathrm{tc}_r(K_n) = r - 1$, or in other words, given any r-edge-colouring of K_n we can cover its vertices using at most $r - 1$ monochromatic components. It is easy to see that $\mathrm{tc}_r(K_n) \leq r$ by fixing a vertex and taking the r monochromatic components containing it in each of the colours. On the other hand, it is not hard [21] to construct classes of graphs that only miss very few edges but admit no cover with a number of monochromatic components that is even bounded by a function of r.

Given a graph G it is not clear how to determine $\mathrm{tc}_r(G)$ and in particular if it can be bounded by a function of r only. Here we develop a framework which allows one to translate this question to a covering problem for hypergraphs. We illustrate the merits of this approach by obtaining answers to various well-studied problems in the area. The first set of these problems is about covering random graphs using monochromatic components.

3.1.1.1 *Covering random graphs*

A common theme in the combinatorics of recent years is to obtain sparse random analogues of extremal or Ramsey-type results. For some examples, see Conlon and Gowers [62] and Schacht [214] and references therein. With this in mind, Bal and DeBiasio [21] initiated the study of covering random graphs by monochromatic components. Following this, Korándi, Mousset, Nenadov, Škorić and Sudakov [162] and Lang and Lo [174] studied a version of this problem in which one uses cycles instead of components, and Bennett, DeBiasio, Dudek and English [32] looked at a related problem for random hypergraphs.

Here, we focus on the original problem of covering random graphs with monochromatic components considered by Bal and DeBiasio [21]. They proved that the number of components needed becomes bounded when p is somewhere between $\left(\frac{r\log n}{n}\right)^{1/r}$ and $\left(\frac{r\log n}{n}\right)^{1/(r+1)}$.

Theorem 3.1.1 (Bal, DeBiasio)**.** *Let r be a positive integer. Then for $G \sim \mathcal{G}(n,p)$,*

(a) if $p \ll \left(\frac{r\log n}{n}\right)^{1/r}$, then w.h.p. $\mathrm{tc}_r(G) \to \infty$, and

(b) if $p \gg \left(\frac{r\log n}{n}\right)^{1/(r+1)}$, then w.h.p. $\mathrm{tc}_r(G) \leq r^2$.

They also made the conjecture that w.h.p. $\mathrm{tc}_r(\mathcal{G}(n,p)) \leq r$ when $p \gg \left(\frac{r\log n}{n}\right)^{1/r}$. This was subsequently proved by Kohayakawa, Mota and Schacht [158] for $r = 2$.[1] On the other hand, [158] presents a far from trivial construction, due to Ebsen, Mota, and Schnitzer, showing that $\mathrm{tc}_r(\mathcal{G}(n,p)) \geq r+1$ for $p \ll \left(\frac{r\log n}{n}\right)^{1/(r+1)}$, which disproves the conjecture for $r \geq 3$. Since this example forces just one additional component and only applies for slightly

1 In fact they solve the partitioning version of this problem.

larger values of edge probability, it was still generally believed that the conjecture is close to being true. In fact, Kohayakawa, Mota and Schacht ask if r components are enough to cover $\mathcal{G}(n,p)$ when p is slightly larger than $\left(\frac{r\log n}{n}\right)^{1/(r+1)}$.

We show that the answer to the above question is quite different from what was expected. We also obtain a good understanding of the behaviour of $\mathrm{tc}_r(\mathcal{G}(n,p))$ throughout the probability range. In particular, we find that $\mathrm{tc}_r(G)$ only becomes equal to r when the density is exponentially larger than conjectured.

Theorem 3.1.2. *Let r be a positive integer. There are constants c, C such that for $G \sim \mathcal{G}(n,p)$,*

(a) if $p < \left(\frac{c\log n}{n}\right)^{\sqrt{r}/2^{r-2}}$, then w.h.p. $\mathrm{tc}_r(G) > r$, and

(b) if $p > \left(\frac{C\log n}{n}\right)^{1/2^r}$, then w.h.p. $\mathrm{tc}_r(G) \le r$.

It is easy to see that $\mathrm{tc}_r(G) \ge r$ holds whenever $\alpha(G) \ge r$ (see e.g. [21]). So the second part of the theorem actually implies that $\mathrm{tc}_r(\mathcal{G}(n,p)) = r$ for all larger values of p, so long as $\alpha(\mathcal{G}(n,p)) \ge r$.

Moreover, we show that near the threshold where $\mathrm{tc}_r(G)$ becomes bounded (and for quite some time after that), its value is not linear, as had been conjectured, but is of order $\Theta(r^2)$.

Theorem 3.1.3. *Let r be a positive integer, $d > 1$ a constant and $G = \mathcal{G}(n,p)$. There are constants c, C such that if $\left(\frac{C\log n}{n}\right)^{\frac{1}{r}} < p < \left(\frac{c\log n}{n}\right)^{\frac{1}{d(r+1)}}$ then w.h.p. $\mathrm{tc}_r(G) = \Theta(r^2)$ (the asymptotics depending on r only).*

Note that Theorems 3.1.1 and 3.1.3 together establish a threshold of $\left(\frac{\log n}{n}\right)^{1/r}$ for the property of having tc_r bounded by a function of r. A slightly weaker understanding of this threshold also follows from the result of Korándi et al. [162] that an r-edge-coloured $\mathcal{G}(n,p)$ can be covered with $O(r^8\log r)$ monochromatic cycles whenever $p > n^{-1/r+\varepsilon}$. In fact, our upper bound on $\mathrm{tc}_r(G)$ in this regime borrows some ideas from [162].

The lower bound in Theorem 3.1.3 answers a question of Lang and Lo [174], who considered the problem of partitioning $\mathcal{G}(n,p)$ into cycles, and ask if $o(r^2)$ monochromatic cycles are enough for $p = \Omega\left(\left(\frac{\log n}{n}\right)^{1/r}\right)$. Our result shows that not only is the answer no, it is not even possible for larger values of p, even if we only need to *cover* the vertices and are allowed to use trees instead of cycles.

The above two theorems describe the value of $\mathrm{tc}_r(\mathcal{G}(n,p))$ when p is quite small or quite large. We also obtain the following theorem, which tracks the behaviour of $\mathrm{tc}_r(\mathcal{G}(n,p))$ in the range between.

Theorem 3.1.4. *Let $k > r \geq 2$ be integers, there exist constants c, C such that given $G \sim \mathcal{G}(n,p)$ if $\left(\frac{C\log n}{n}\right)^{1/k} < p < \left(\frac{c\log n}{n}\right)^{1/(k+1)}$ then w.h.p.* $\frac{r^2}{20\log k} \leq \mathrm{tc}_r(G) \leq \frac{16r^2\log r}{\log k}$.

In fact, we obtain slightly better bounds when k approaches either extreme (i.e. when k is linear or exponential in r). This connects the bounds of this theorem to the ones in Theorems 3.1.2 and 3.1.3. For example, we obtain that if k is exponential in r then $\mathrm{tc}_r(G) = \Theta(r)$.

3.1.1.2 *The connection to covering partite hypergraphs.*

As mentioned before, our proofs of the above results rely on an interesting connection to a natural problem about hypergraph covers. Loosely speaking, the question asks how big a cover of a hypergraph H can be if any subgraph of H with few edges has a small cover. Here by a (vertex) cover of a hypergraph H, we mean a set of vertices that has a non-empty intersection with all edges of H. The minimum size of such a cover is called the cover number of H, and is denoted by $\tau(H)$.

In our case, we consider a variant for r-partite r-graphs (r-uniform hypergraphs). A *transversal cover* is then defined as a cover containing exactly one vertex in each part of the r-partition.

Definition. We say that an r-partite r-graph H has the *r-partite k-covering property* if any subgraph of H with at most k edges has a transversal cover.

We define $\mathrm{hp}_r(k)$ to be the largest possible cover number of an r-partite hypergraph satisfying the r-partite k-covering property if such a maximum exists, and set $\mathrm{hp}_r(k) = \infty$ otherwise. The connection between the two problems is particularly striking in the case of random graphs where we get the following result:

Theorem 3.1.5. *Let $k > r \geq 2$ be integers, and let $G \sim \mathcal{G}(n,p)$. There are constants $C, c > 0$ such that:*

1. *If $np^k > C\log n$ then w.h.p.* $\mathrm{tc}_r(G) \leq \mathrm{hp}_r(k)$.
2. *If $np^{k+1} < c\log n$ then w.h.p.* $\mathrm{tc}_{r+1}(G) \geq \mathrm{hp}_r(k) + 1$.

What this is saying is that for $\left(\frac{\log n}{n}\right)^{1/k} \ll p \ll \left(\frac{\log n}{n}\right)^{1/(k+1)}$ the value of $\mathrm{tc}_r(\mathcal{G}(n,p))$ is essentially determined by $\mathrm{hp}_r(k)$.

Let us point out that the $k = r+1$ (first non-trivial) case of estimating $\mathrm{hp}_r(k)$ is directly related to Ryser's famous conjecture [144] which claims $\tau(H) \le (r-1)\nu(H)$ for r-partite r-graphs. Here the *matching number* $\nu(H)$ is defined as the largest number of pairwise disjoint edges of H. Indeed, it is easy to check that the r-partite $(r+1)$-covering property for H is equivalent to $\nu(H) \le r$. So in the special case when $\nu(H) = r$, Ryser's conjecture is equivalent to $\mathrm{hp}_r(r+1) \le r(r-1)$.

3.1.1.3 *Covering by components of different colours*

The second setting in which our method works very well is for graphs of large minimum degree.

Covering such graphs using monochromatic components was first considered by Bal and DeBiasio [21]. A particularly nice conjecture they raised is that any graph G with minimum degree $\delta(G) \ge (1 - 1/2^r)n$ can be covered by monochromatic components of *distinct* colours. They gave an example showing that if true, this conjecture is best possible. Girão, Letzter and Sahasrabudhe [125] proved the conjecture for $r \le 3$. Our methods enable us to completely resolve this conjecture for all values of r.

Theorem 3.1.6. *Let G be an r-coloured graph on n vertices with $\delta(G) \ge (1 - 1/2^r)n$. Then the vertices of G can be covered by monochromatic components of distinct colours.*

3.1.1.4 *Covering general hypergraphs*

For our connection discussed in Section 3.1.1.2, we needed the considered hypergraphs to be partite. As it turns out, the behaviour does not change much if we drop this condition. This gives rise to the following, independently interesting and perhaps even more natural, question: What is the largest possible cover number of an r-graph in which any k edges have a cover of size at most ℓ? Is it even finite? We denote this maximum by $\mathrm{h}_r(k,\ell)$ if it exists, and write $\mathrm{h}_r(k,\ell) = \infty$ otherwise. The question of determining the parameter $\mathrm{h}_r(k,\ell)$ was raised by Erdős, Hajnal and Tuza [100] more than 25 years ago and was later studied by Erdős, Fon-Der-Flaass, Kostochka and Tuza; Fon-Der-Flaass, Kostochka and Woodall; Kostochka [86, 107, 165]. They were all interested in a setting when r is big and ℓ is fixed and small. Here we develop an approach which gives good estimates on $\mathrm{h}_r(k,\ell)$ in general and in particular when $\ell = r$, which is the case relevant to our problem.

Let us now define the covering property in this setting, as well.

Definition. We say that a hypergraph H has the (k,ℓ)*-covering property* if any subgraph of H with at most k edges has a cover of size at most ℓ.

Note that a hypergraph H has the $(k,1)$-covering property if and only if any k edges of H have a non-empty intersection. This property is the central concept in the study of Helly families, dating all the way back to 1913: Helly's theorem [143] states that if in a collection $\mathcal{F}$ of convex sets in $\mathbb{R}^d$, any $d+1$ sets have a non-empty intersection, then all the sets in $\mathcal{F}$ intersect. There is a vast literature on studying what other families exhibit such Helly-type properties. For some classical examples from geometry, see [70], for Helly-type results for hypergraphs, see [38, 117, 175]. For example, a folklore analogue for hypergraphs states that if any $r+1$ edges of an r-graph intersect, then all the edges have a non-empty intersection (see [117]). With our notation, this corresponds to $\mathrm{h}_r(r+1,1)=1$. Studying $\mathrm{h}_r(k,\ell)$ leads to a natural generalisation of this result.

An interesting special case is to determine when $\mathrm{h}_r(k,\ell)=\ell$ holds, i.e. for what k we know that if any k edges of a hypergraph have a cover of size ℓ, then the whole hypergraph has a cover of size ℓ. As observed by Füredi [117], a result of Bollobás [39] (generalising earlier work of Erdős, Hajnal and Moon [90] for graphs), implies that the answer to this question is $k \geq \binom{\ell+r}{r}$.

We obtain good bounds on $\mathrm{h}_r(k,\ell)$ in general, but for the sake of clarity, and because this case illustrates the general behaviour well, we only present our results for $r=\ell$ here, summarised in the next theorem. Our general bounds are presented in Section 3.1.5.

Theorem 3.1.7. *The following table describes the behaviour of* $\mathrm{h}_r(k,r)$ *for fixed* r *as* k *varies. For arbitrary constants* $c>1$ *and* $d>0$ *we have*

Range of k	$[1,r]$	$r+1$	$(r,cr]$	$\left(r,e^{r/2}\right]$	$\left[e^{dr},\infty\right)$	$\left[\binom{2r}{r},\infty\right)$
Value of $\mathrm{h}_r(k,r)$	∞	r^2	$\Theta\left(r^2\right)$	$\left[\frac{r^2}{4\log k},\frac{16r^2\log r}{\log k}\right)$	$\Theta(r)$	r

We can actually prove slightly stronger bounds in the middle range, when k is close to either extreme, see Section 3.1.5 for more details.

Another reason why $r=\ell$ is a case of particular interest for us is its relation to the parameter $\mathrm{hp}_r(k)$, which we need for our connection result. Note that $\mathrm{hp}_r(k)\leq \mathrm{h}_r(k,r)$ follows immediately from the definitions. On the other hand, examples giving lower bounds for $\mathrm{h}_r(k,r)$ need not be r-partite, so we need to work a bit harder to obtain lower bounds for $\mathrm{hp}_r(k)$. Nevertheless, we can establish essentially the same lower bounds for $\mathrm{hp}_r(k)$ as the ones above for $\mathrm{h}_r(k,r)$ (see Theorem 3.1.31).

3.1.1.5 *Organisation of the section*

In the next section we present some preliminary results and tools, mostly simple properties of random graphs. In Section 3.1.3 we will prove the upper bound of Theorem 3.1.3. In Section 3.1.4 we describe the connection between the problem of covering by monochromatic trees and hypergraph covering problems in more detail and deduce Theorem 3.1.5 from it. In Section 3.1.5 we show our results for general hypergraphs, in particular, Theorem 3.1.7. Finally, in Section 3.1.6 we show our bounds on $\mathrm{hp}_r(k)$ which, through our connection to monochromatic covering, yield Theorems 3.1.2 to 3.1.4 and 3.1.6. In Section 3.1.7 we give some open problems. We conclude the section with an appendix in which we give an alternative perspective to the general hypergraph setting considered in Section 3.1.5.

3.1.2 *Preliminaries*

Throughout the section all colourings considered are *edge*-colourings. Whenever we have an r-colouring, we label the colours by the first r positive integers.

We now present several properties of random graphs that play a role in our arguments. Let us begin with the following simple Chernoff-type bound on the tail of the binomial distribution.

Lemma 3.1.8 ([16], Appendix A)**.** *Let $X \sim Bin(n, p)$ be a binomial random variable. Then*

$$\mathbb{P}(X < np/2) \leq e^{-np/8}.$$

The following result is a simple consequence of union bound and the above lemma. For a proof, see e.g. [162](Lemma 3.8).

Lemma 3.1.9. *Let $p = p(n) > 0$. The random graph $G \sim \mathcal{G}(n, p)$ satisfies the following property w.h.p.: for any two disjoint subsets $A, B \subseteq V(G)$ of size at least $\frac{10 \log n}{p}$, there is an edge between A and B.*

The following property allows us to give a lower bound on the number of common neighbours of a small set that send at least one edge to another small set. For a set of vertices A, we denote the set of common neighbours of all vertices in A by $N(A)$. (For convenience, we define $N(\emptyset) = V(G)$.)

Lemma 3.1.10. *Let $r \geq 1$ be an integer, $D \geq 1$ and $p = p(n) \geq \left(\frac{64Dr\log n}{n}\right)^{1/r}$. The random graph $G \sim \mathcal{G}(n, p)$ satisfies the following property w.h.p.: for any two disjoint subsets $A, B \subseteq V(G)$ such that $|A| \leq D/p$ and $|B| = r - 1$, there are at least $|A| \log n$ vertices in $N(B) \setminus A$ that have a neighbour in A.*

Proof. Let A, B be two fixed disjoint subsets of $V(G)$ of sizes $|A| = k \leq D/p$ and $|B| = r-1$. The number of vertices of $N(B) \setminus A$ that have a neighbour in A follows the binomial distribution
$\mathrm{Bin}\left(n-k-r+1, p^{r-1}(1-(1-p)^k)\right)$. Notice that for large enough n we have $n' = n-k-r+1 > n/2$, while $pk \leq D$ implies $p' = p^{r-1}(1-(1-p)^k) \geq p^{r-1}(1-e^{-pk}) \geq \frac{p^r k}{2D}$ using the fact that $1-e^{-x} \geq \frac{x}{2D}$ for $0 \leq x \leq D$ and $1 \leq D$. Applying Lemma 3.1.8, we obtain that the probability that there are fewer than $|A| \log n \leq k \cdot \frac{p^r n}{64Dr} < \frac{n}{2} \cdot \frac{p^r k}{4D} < \frac{n'p'}{2}$ vertices in $N(B) \setminus A$ that have a neighbour in A is at most

$$e^{-n'p'/8} \leq e^{-knp^r/(32D)} \leq n^{-2kr}.$$

Applying the union bound over all $1 \leq k \leq D/p$ and the $\binom{n}{k}\binom{n-k}{r-1}$ possibilities for A and B, we obtain that the probability that there are sets A, B failing the conditions of the problem is at most

$$\sum_{k=1}^{D/p} n^k n^{r-1} n^{-2kr} \to 0$$

as $n \to \infty$, as claimed. $\square$

Taking $D = 1$ and a set A of size one in the above lemma gives us the following corollary

Corollary 3.1.11. *Let $r \geq 1$ be an integer and $G \sim \mathcal{G}(n,p)$ for $p > \left(\frac{64r\log n}{n}\right)^{1/r}$. Then w.h.p. any r vertices in G have at least $\log n$ common neighbours.*

As usual, an independent set in G is a subset of $V(G)$ where no two vertices are connected by an edge, and the independence number $\alpha(G)$ is the largest size of an independent set in G. The following lemma is the only property we are going to use when proving lower bounds on $\mathrm{tc}_r(\mathcal{G}(n,p))$. It is a special case of Lemma 6.4 (ii) in [21].

Lemma 3.1.12. *Let $m > k \geq 2$ be integers. There is a $c > 0$ such that for $G \sim \mathcal{G}(n,p)$ with $p \leq \left(\frac{c\log n}{n}\right)^{1/k}$, w.h.p. there is an independent set S in G of size m such that no k vertices in S have a common neighbour in G.*

The following statement is a simple generalisation of the observation that $\mathrm{tc}_r(K_n) \leq r$.

Proposition 3.1.13. *Let r be an integer and G be a multigraph. Then $\mathrm{tc}_r(G) \leq r\alpha(G)$.*

Proof. Let I be a maximal independent set, and take all monochromatic components containing a vertex of I. This is at most $r\alpha(G)$ components, and as I is maximal, every other vertex is adjacent to I in some colour, hence covered by some component. $\square$

We finish the section with results of a somewhat different flavour. The following theorem, due to Alon [13], is a far-reaching generalisation of the Bollobás set pairs inequality [39] and several of its variants (see [117] for more details).

Theorem 3.1.14 (Alon). *Let $V_1, \ldots, V_s$ be disjoint sets, $a_1, \ldots, a_s$ and $b_1, \ldots, b_s$ be positive integers, and let $A_1, \ldots, A_m$ and $B_1, \ldots, B_m$ be finite sets satisfying the following properties:*

- $|A_i \cap V_j| \leq a_j$ *and* $|B_i \cap V_j| \leq b_j$ *for all* $1 \leq i \leq m$ *and* $1 \leq j \leq s$,
- $A_i \cap B_i = \emptyset$ *for all* $1 \leq i \leq m$ *and*
- $A_i \cap B_j \neq \emptyset$ *for all* $1 \leq i < j \leq m$.

Then $m \leq \prod_{i=1}^{s} \binom{a_i+b_i}{a_i}$.

Definition 3.1.15. A hypergraph H is said to be *critical* if every proper subgraph of H has a strictly smaller cover number.

Corollary 3.1.16 (Bollobás [39]). *A critical r-graph with cover number $t+1$ has at most $\binom{r+t}{t}$ edges.*

Proof. Let H be a critical r-graph with cover number $t+1$, and denote the edges of H by $A_1, \ldots, A_m$. By the criticality of H, we know that if we remove any edge A_i, the remaining hypergraph admits a cover of size at most t. Denote this cover by B_i. Notice that $A_i \cap B_i = \emptyset$, as otherwise B_i would be a cover of H of size at most t, which is impossible. Furthermore, $A_i \cap B_j \neq \emptyset$ whenever $i \neq j$, because B_j is a cover of $H \setminus A_j$, so in particular, it covers the edge A_i.

This means that the sets of vertices A_i, B_i satisfy $|A_i| = r$, $|B_i| \leq t$, $A_i \cap B_i = \emptyset$ and $A_i \cap B_j \neq \emptyset$ whenever $i \neq j$. Therefore, Theorem 3.1.14 with $s = 1$, $V_1 = V(H), a_1 = r$ and $b_1 = t$ implies $m \leq \binom{r+t}{r}$. $\square$

3.1.3 *Upper bound near the threshold*

We start by showing an upper bound on $\mathrm{tc}_r(\mathcal{G}(n,p))$ just above the probability threshold where it becomes bounded (w.h.p.). This is the only regime in which we will not use the relation between the problem of covering with monochromatic trees and the hypergraph covering problem.

As mentioned in Theorem 3.1.1, Bal and DeBiasio [21] obtained an upper bound of r^2 when $p \gg \left(\frac{\log n}{n}\right)^{1/(r+1)}$. Their argument proceeds as follows. First, one can define the *transitive closure multigraph* G^{tc} of G on the same vertex set, where two vertices are connected by an edge of colour i whenever they are in the same colour-i component of G. Clearly, G and G^{tc} have the same set of monochromatic components. Now when $G \sim \mathcal{G}(n,p)$ with $p \gg \left(\frac{\log n}{n}\right)^{1/(r+1)}$, it is not hard to check that $\alpha(G^{\mathrm{tc}}) \leq r$. Indeed, in this probability range, any $r+1$ vertices of G have a common neighbour v w.h.p., so some two of them will be connected by a monochromatic path through v. Proposition 3.1.13 then yields $\mathrm{tc}_r(G) = \mathrm{tc}_r(G^{\mathrm{tc}}) \leq r^2$.

In some sense, our upper bounds through the hypergraph covering problem are far-reaching generalisations of this argument. However, all of these arguments break down for $\left(\frac{\log n}{n}\right)^{1/r} \ll p \ll \left(\frac{\log n}{n}\right)^{1/(r+1)}$, when we only know that any r vertices have a common neighbour. This issue was resolved in [162] by embedding "cascades" in $\mathcal{G}(n,p)$ to show that among any $4r-2$ vertices, some two are connected by monochromatic paths in a robust way. The argument needed there is quite technical and only works for $p > n^{-1/r+\varepsilon}$. We present a much simpler adaptation of the ideas to prove $\alpha(G^{\mathrm{tc}}) \leq 3r-2$ for the whole probability range. This will readily imply the upper bound on $\mathrm{tc}_r(\mathcal{G}(n,p))$ for Theorem 3.1.3.

Theorem 3.1.17. *Let r be a positive integer and $G \sim \mathcal{G}(n,p)$. There is a $C > 0$ such that if $p > \left(\frac{C\log n}{n}\right)^{1/r}$, then w.h.p. $\mathrm{tc}_r(G) \leq (3r-2)r$.*

Proof. Fix an r-edge-colouring of G. By Proposition 3.1.13, it is enough to show that $\alpha(G^{\mathrm{tc}}) \leq 3r-2$.

So suppose for contradiction that there is an independent set S of $3r-1$ vertices. Notice first that no vertex of G can send two edges of the same colour to S, because such edges would belong to the same monochromatic component of G, so G^{tc} would have an edge in S. Notice that this implies that every vertex of G has at most r neighbours in S.

Let us say that a vertex set A is *rainbow to* $\{v_1, \ldots, v_r\}$ if A is contained in the colour-i component of v_i for every i. Now let $X \subseteq S$ be any set of size $2r-1$, and let A be a largest vertex set in G^{tc} that is rainbow to some r-subset $\{v_1, \ldots, v_r\} \subseteq X$. Note that $|A| \geq 1$ since by Corollary 3.1.11 we know that any r vertices in S have a common neighbour, which can not send two edges of the same colour towards S. The heart of our proof is the following bootstrapping argument, which immediately gives $|A| \geq \frac{10 \log n}{p}$.

Claim 1. *If A_0 is a set of size at most $\frac{20r!}{p}$ that is rainbow to some r-set in X, then there is a set of size at least $\frac{|A_0| \log n}{r!}$ that is rainbow to some other r-set in X.*

Proof. Suppose A_0 is rainbow to $X' = \{w_1, \ldots, w_r\} \subseteq X$, and let $w_{r+1}, \ldots, w_{2r-1}$ denote the vertices of $B = X \setminus X'$. Note that A_0 is disjoint from B because otherwise some vertex of B is in the colour-1 component of w_1, contradicting the independence of X in G^{tc}.

This means that we can apply Lemma 3.1.10 with $D = 20r!$ to A_0 and B to find a set A' of at least $|A_0| \log n$ common neighbours of B that each have some neighbour in A_0. Take some vertex $u \in A'$. As we have noted before, all edges uw_j for $j > r$ must have distinct colours. There are $r!$ possible assignments of $r-1$ distinct colours to the edges $uw_{r+1}, \ldots, uw_{2r-1}$, so there is a subset $A'' \subseteq A'$ of at least $|A'|/r!$ vertices such that for every $j > r$, all edges between A'' and w_j have the same colour. Let c be the colour not used by these edges between A'' and B.

By definition, every vertex $u \in A''$ is adjacent to some $v \in A_0$. Let c' be the colour of the edge uv. As A_0 is rainbow to X', u is in the colour-c' component of $w_{c'}$. But X is independent, so c' cannot be the colour of any edge uw_j for $j > r$. This means that $c' = c$ for every such u, and hence, A'' is rainbow to $\{w_c, w_{r+1}, \ldots, w_{2r-1}\}$ (possibly reordered), as needed. □

Let $Y = S \setminus \{v_1, \ldots, v_r\}$ and let B be a largest vertex set in G^{tc} that is rainbow to some r-subset $\{u_1, \ldots, u_r\} \subseteq Y$. The above claim also shows that $|B| \geq \frac{10 \log n}{p}$. Note also that A and B must be disjoint since otherwise $v_1, u_1 \in S$ would belong to the same colour-1 component of G, contradicting the independence of S. This means that we can apply Lemma 3.1.9 to get an edge ab with $a \in A, b \in B$. If ab has colour c, then this implies that v_c and u_c belong to the same colour-c component, so $v_c u_c$ is an edge within S in G^{tc}, a contradiction. □

3.1.4 *The connection to hypergraph covering*

In this section we establish a connection between covering graphs using monochromatic components and a covering problem for hypergraphs.

Given an r-edge-colouring c of a graph G we build the following auxiliary r-partite r-graph $H = H(G, c)$. The vertices of H are taken to be the monochromatic components of G (including singleton vertices). Note that we treat components of distinct colours as different vertices of H even if they consist of the same set of vertices in G. For every $v \in V(G)$, we make an edge $m(v)$ in H consisting of all the monochromatic components that v belongs to. Note that each vertex belongs to exactly one component of each colour, so H is r-uniform and r-partite, where the parts of H are formed by the monochromatic components of the same colour.

Proposition 3.1.18. *Given an edge-colouring c of a graph G, the minimum number of monochromatic components needed to cover $V(G)$ equals $\tau(H(G,c))$. Moreover, if $H(G,c)$ has a transversal cover, then $V(G)$ can be covered using monochromatic components of distinct colours.*

Proof. Let $C_1, \dots, C_t$ be some monochromatic components that cover $V(G)$. For any edge $m(v)$ of H, we know that $v \in C_i$ for some i, so $C_i \in m(v)$. This implies that $C_1, \dots, C_t$ when viewed as vertices of $H = H(G,c)$ make a cover of H. On the other hand, if monochromatic components $C_1, \dots, C_t$ make a cover of H then for any $v \in V(G)$, its corresponding edge $m(v)$ must contain some C_i. But this means that $v \in C_i$, and as v was arbitrary, $C_1, \dots, C_t$ cover $V(G)$. Combining these observations implies the first part of the proposition. If $C_1, \dots, C_r$ make a transversal cover of H then there is at most one C_i in any part of the r-partition, or in other words at most one C_i of any fixed colour. Since we know by the first part that $C_1 \cup \cdots \cup C_r = V(G)$, the second part of the proposition follows. □

Given a colouring of G, this proposition allows us to translate the problem of determining the number of monochromatic components needed for covering $V(G)$ to determining the cover number of the auxiliary hypergraph. However, to determine $\mathrm{tc}_r(G)$, one needs to consider all colourings of G, and this translates to determining the minimum cover number among a class of r-partite r-graphs which arise from different colourings of G. As it turns out, the crucial parameter in determining $\mathrm{tc}_r(G)$ is the largest k such that any k vertices of G have a common neighbour. This also works nicely with the auxiliary hypergraphs in the sense that if any k vertices in G have a common neighbour then for *any colouring* of G, the auxiliary hypergraph will satisfy the r-partite k-covering property. Recall that an r-partite r-graph has the r-partite k-covering property if any k of its edges have a transversal cover, and $\mathrm{hp}_r(k)$ is the maximum cover number of such a hypergraph. This will allow us to prove the following upper bound on $\mathrm{tc}_r(G)$.

Lemma 3.1.19. *Let $k > r \geq 2$ be integers, and let G be a graph in which any k vertices have a common neighbour.*[2] *Then $\mathrm{tc}_r(G) \leq \mathrm{hp}_r(k)$. Moreover, if we can*

2 For this argument, we consider a vertex to be adjacent to itself, so a common neighbour might be one of the k vertices.

find a transversal cover of any hypergraph having the r-partite k-covering property, then $V(G)$ can be covered using components of distinct colours.

Proof. Our goal is to show that in any r-colouring of G, we can cover the vertices of G using at most $\text{hp}_r(k)$ monochromatic components. So fix a colouring c of G, and let $H = H(G,c)$ be the corresponding auxiliary r-partite r-graph.

Now take any set S of k edges in H. These edges correspond to k vertices in G, so they have some common neighbour v. The key observation is that $m(v)$ is a transversal cover of S. Indeed, every hyperedge in S is of the form $m(u)$ for some neighbour u of v, and so the component of colour $c(uv)$ containing u, v is in both $m(u)$ and $m(v)$. This means that H satisfies the r-partite k-covering property, so $\tau(H) \leq \text{hp}_r(k)$. By Proposition 3.1.18 we can cover $V(G)$ using at most $\tau(H) \leq \text{hp}_r(k)$ monochromatic components, and if H has a transversal cover we can do it using components of distinct colours. □

Combining Corollary 3.1.11 and this lemma, we obtain part 1 of Theorem 3.1.5. It might be worth mentioning that the above proof gives an upper bound in terms of a slightly stronger r-partite k-covering property, where the transversal cover of the k edges is required to be another edge of the family. However, we are not aware of any improved bound on $\text{hp}_r(k)$ using this extra condition.

Somewhat unexpectedly it turns out that one can also lower bound $\text{tc}_r(G)$ in terms of hp_r. Here the relevant property of G is that it contains a big enough independent set in which no $k+1$ vertices have a common neighbour.

Lemma 3.1.20. *Let $k > r \geq 2$ be integers. There is a $C = C(r)$ such that if G is a graph with an independent set of size C in which no $k+1$ vertices have a common neighbour in G then* $\text{tc}_{r+1}(G) \geq \text{hp}_r(k) + 1$.

Before turning to the proof let us give a definition of an auxiliary hypergraph and establish some of its properties, which we will use in the proof. Given an r-partite r-graph H we construct an auxiliary r-coloured multigraph $\mathcal{A}(H)$. Let us denote by $P_1, \ldots, P_r$ the parts of H. For the vertex set of $\mathcal{A}(H)$ we take $E(H)$ and for any $e \in E(H)$ we denote the corresponding vertex in $\mathcal{A}(H)$ by $a(e)$. As for the edge set, we place an edge of colour i between two vertices of $\mathcal{A}(H)$ whenever the corresponding edges in $E(H)$ intersect in P_i. Note that if these edges intersect in several different parts, then we add multiple edges (which will be of distinct colours) between the same vertices. The following proposition establishes the properties of $\mathcal{A}(H)$ that we will use.

Proposition 3.1.21. *Let H be an r-partite r-graph satisfying the r-partite k-covering property and let $A = \mathcal{A}(H)$ be defined as above. Then*

1. *We need at least $\tau(H)$ monochromatic components to cover A.*
2. *We can partition any set S of at most k vertices in A into r parts $S_1, \dots, S_r$ (some possibly empty) such that for every i, S_i is contained in a single component of colour i.*

Proof. For part 1, if $a(e), a(f), a(g)$ are vertices of A and $a(e)$ is joined to both $a(f)$ and $a(g)$ with an edge of same colour in A, say i, then e, f and g had to contain the same vertex in part i of H. In particular, propagating this along any monochromatic path in A we conclude that for all vertices of A belonging to a monochromatic component C, their corresponding edges in H need to contain the same vertex $v(C)$ (the one in the part corresponding to the colour of C). This in particular means that if we could cover all vertices of A using $t < \tau(H)$ monochromatic components $C_1, \dots, C_t$ then $v(C_1), \dots, v(C_t)$ would give a cover of all edges of H using less than $\tau(H)$ vertices, a contradiction.

For part 2, let $S = \{a(e_1), \dots, a(e_k)\}$ be a set of k vertices in A. By the r-partite k-covering property, there is a transversal cover of $\{e_1, \dots, e_k\}$ in H. This means that we can partition $\{e_1, \dots, e_k\}$ into r sets such that all edges in the ith set contain the vertex of the transversal cover belonging to part i. Now let S_i consist of the vertices $a(e_j)$ such that e_j is in the ith set to obtain our desired partition of S. $\square$

We are now ready to prove Lemma 3.1.20.

Proof of Lemma 3.1.20. Consider an r-partite r-graph H with the property that any k edges of H have a transversal cover and $\tau(H) = \mathrm{hp}_r(k)$. As we will see in Observation 3.1.32, parts (2) and (3), $k > r$ implies that $\mathrm{hp}_r(k) \le r^2$, in particular, $\mathrm{hp}_r(k)$ is finite. By repeatedly removing edges from H, we may also assume that H is critical, and hence, by Corollary 3.1.16, has at most $\binom{r+r^2}{r}$ edges. Let us number the parts of H by $[r]$, and let $m \le \binom{r+r^2}{r}$ be the number of edges in H.

Let $C = \binom{r+r^2}{r} + 1$ and let G be a graph with an independent set X of size C such that no $k+1$ vertices in X have a common neighbour. Let $S \subseteq X$ be a subset of size m. Since $C \ge m+1$ there exists a vertex w that is not adjacent to S. Let us now take our auxiliary r-coloured graph $A = \mathcal{A}(H)$ and identify its m vertices with the vertices of S.

Using this colouring we will define an $(r+1)$-edge-colouring of G where the vertices cannot be covered with fewer than $\mathrm{hp}_r(k) + 1$ monochromatic components. Let $T = V \setminus S$, so $w \in T$. We colour all edges of G induced by T with colour $r+1$. Since S is independent, all uncoloured edges are between S and T. Since no $k+1$ vertices of S have a common neighbour in G, any vertex $v \in T$ sends at most k edges towards S. Let $S_v = N(v) \cap S$, since $|S_v| \leq k$ by property 2 of A we know we can split it into r parts $S_1(v), \dots, S_r(v)$ such that each $S_i(v)$ is contained in a monochromatic component of A of colour i. We colour an edge from v to S in colour i if its other endpoint belongs to $S_i(v)$. This completes our colouring.

Let us now show that we need $\mathrm{hp}_r(k)+1$ monochromatic components to cover the vertex set of G in our colouring. Note first that all edges touching w have colour $r+1$, so we need one component of this colour. However, this component is disjoint from S because no edge touching S has colour $r+1$. Note also that any monochromatic component in G, when restricted to S, is contained in a monochromatic component of A. This is because for any two vertices $u_0, u_t \in S$ from a monochromatic component of colour j in G, there is a path $u_0 u_1 \dots u_t$ in S and vertices $v_0, \dots, v_{t-1} \in T$ such that $u_{i-1}, u_i \in S_j(v_i)$. By part 2 of Proposition 3.1.21, this means that u_{i-1} and u_i are in the same monochromatic component of colour j in A, for every i. But then u_0 and u_t belong to the same monochromatic component, as well.

This shows that we need at least as many components to cover S as we need to cover $\mathcal{A}(H)$, which, by part 1 of Proposition 3.1.21, is at least $\tau(H) = \mathrm{hp}_r(k)$. Since we also needed one extra component to cover w we obtain $\mathrm{tc}_{r+1}(G) \geq \mathrm{hp}_r(k)+1$. □

This lemma, coupled with Lemma 3.1.12, implies part 2 of Theorem 3.1.5.

Remark. Note that Corollary 3.1.11 and Lemma 3.1.12 imply that for almost all graphs G of certain fixed density the conditions of Lemmas 3.1.19 and 3.1.20 hold with the same value of k implying that $\mathrm{tc}_r(G)$ is essentially determined by $\mathrm{hp}_r(k)$.

3.1.5 *Results for general hypergraphs*

In this section we prove Theorem 3.1.7. Recall that $\mathrm{h}_r(k,\ell)$ is the largest possible cover number of an r-graph H satisfying the (k,ℓ)-covering property, i.e. any k edges of H have a cover of size ℓ. Let us start with some easy observations, establishing some basic properties of $\mathrm{h}_r(k,\ell)$.

Observation 3.1.22. *For $r \geq 2$ we have:*

(1) $\mathrm{h}_r(\ell,\ell) = \infty$,

(2) $\mathrm{h}_r(\ell+1,\ell) \le r\ell$,

(3) $\mathrm{h}_r(k+1,\ell) \le \mathrm{h}_r(k,\ell)$ *and*

(4) $\ell \le \mathrm{h}_r(k,\ell)$.

Proof. For (1), notice that any hypergraph satisfies the (ℓ,ℓ)-covering property. Indeed, given any ℓ edges, we can choose one vertex from each to obtain a cover of size at most ℓ. So taking a hypergraph with arbitrarily large cover number implies the claim.

For (2), let H be an r-graph satisfying the $(\ell+1,\ell)$-covering property. Note that H cannot contain $\ell+1$ pairwise disjoint edges, because covering them would require at least $\ell+1$ vertices. So any maximal set S of pairwise disjoint edges in H has size at most ℓ. By its maximality, every edge of H intersects some edge in S. But then the set of all vertices contained in the edges of S form a cover of size at most $r\ell$.

For (3), notice that the $(k+1,\ell)$-covering property implies the (k,ℓ)-covering property, because if we can cover any $k+1$ edges with ℓ vertices, then the same holds for any k edges, as well.

For (4), we can take a hypergraph consisting of ℓ disjoint edges. This clearly satisfies the (k,ℓ)-covering property for any k, and has cover number exactly ℓ. □

3.1.5.1 *Upper bounds*

Let us start with characterising when we have equality in Observation 3.1.22 part (4). The following result was observed by Füredi [117] (in a slightly different language), and is a simple consequence of Corollary 3.1.16 by Bollobás. We provide the proof for completeness.

Theorem 3.1.23. *If* $k \ge \binom{r+\ell}{\ell}$ *then*

$$\mathrm{h}_r(k,\ell) = \ell.$$

Proof. We know from Observation 3.1.22 part (4) that $\mathrm{h}_r(k,\ell) \ge \ell$, so it is enough to show the upper bound. Suppose for contradiction that there is an r-graph with cover number at least $\ell+1$ satisfying that any k of its edges have a cover of size at most ℓ. Let H be such a hypergraph with the smallest possible number of edges. Then H is clearly critical, so we can apply Corollary 3.1.16 to see that it has at most $\binom{r+\ell}{r}$ edges. But as $k \ge \binom{r+\ell}{r}$, this means that H admits a cover of size at most ℓ, a contradiction. □

Note that the bound on k is tight: indeed, the complete r-graph on $r+\ell$ vertices satisfies the $\left(\binom{r+\ell}{\ell}-1,\ell\right)$-covering property, but cannot be covered with ℓ vertices.

Obtaining any kind of improvement over the easy inequality $\mathrm{h}_r(k,\ell) \leq r\ell$, for $k \geq \ell+1$, given by Observation 3.1.22 seems far from immediate. The following result obtains a good bound by combining a random sampling argument with the above ideas.

Theorem 3.1.24. *Given integers t,k such that $\ell \leq t \leq r\ell$ and $k \geq \binom{r+t}{t}^{1/\lfloor t/\ell \rfloor} \cdot 2r\ell\log(r\ell)$ we have*

$$\mathrm{h}_r(k,\ell) \leq t.$$

Proof. We proceed by contradiction. Let us assume that there is an r-uniform hypergraph that has no cover of size t, but in which any k edges have a cover of size ℓ. Let H be such a hypergraph with minimum number of edges. Then H is critical and Corollary 3.1.16 implies that it has at most $\binom{r+t}{t}$ edges.

Our strategy goes as follows. We start by showing that for any subgraph G of H there must be a set of ℓ vertices of G covering many edges of G, as otherwise a random sample of k edges in G would not have any cover of size ℓ. Armed with this claim, we take such a set of size ℓ covering many edges of H, remove it from the vertex set, and repeat the process with the remaining graph. The claim ensures that in each step we remove many edges, which in turn will imply that this process cannot repeat more than $\frac{t}{\ell}$ times before we reach an empty graph. This algorithm then yields $\frac{t}{\ell} \cdot \ell = t$ vertices covering H, which is a contradiction.

Let us now make this argument precise. Let m denote the number of edges in H (so $m \leq \binom{r+t}{t}$), and let $x = m^{-1/\lfloor t/\ell \rfloor} < 1$. This x will be the proportion of edges that remain in G after removing the ℓ-set in a step of the procedure described above. Given a hypergraph G and a subset $S \subseteq V(G)$, we let $d_G(S)$ denote the number of edges in G that share a vertex with S, i.e. the number of edges in G covered by the set S[3]. Let us start by showing the above-mentioned claim. We denote by $e(G) = |E(G)|$.

Claim 2. *For every non-empty subgraph G of H, there is a subset $S \subseteq V(G)$ of size at most ℓ such that $d_G(S) > (1-x)e(G)$.*

Proof of Claim 2. Let us assume, to the contrary, that there is a subgraph G with the property that any subset of ℓ vertices in G covers at most $(1-x)e(G)$ edges. We may assume that G contains no isolated vertices, as otherwise we

3 Note that $d_G(S)$ is not the codegree of S, which would count the number of sets *containing* S.

can just remove all of them without violating any of the properties. Then $|V(G)| \leq re(G) \leq rm \leq r\binom{r+t}{t}$.

Let T be a random set obtained by sampling k (not necessarily distinct) edges of G, independently and uniformly at random. Then T consists of at most k edges of H, so it has a cover of size ℓ.

On the other hand, for any given set $S \subseteq V(G)$ of size ℓ, the probability that it covers a single randomly sampled edge is $d_G(S)/e(G) \leq 1 - x$. In particular, the probability that S covers all of the k independently sampled edges in T is at most $(1-x)^k$. Finally, by the union bound, the probability that there is a set of size ℓ which covers all k of the sets in T is at most

$$\binom{|V(G)|}{\ell}(1-x)^k \leq (rm)^\ell e^{-xk} < 1,$$

where the last inequality uses $xk > \ell \log(rm)$, which follows from

$$xk \geq m^{-1/\lfloor t/\ell \rfloor}\binom{r+t}{t}^{1/\lfloor t/\ell \rfloor} 2r\ell \log(r\ell) \geq 2r\ell \log(r\ell) = \ell \log((r\ell)^{2r})$$

and, using the assumption $t \leq r\ell$,

$$rm \leq r\binom{2r\ell}{r} < (2r\ell)^r \leq (r\ell)^{2r}.$$

But this means that there is a set of at most k edges without a cover of size ℓ, a contradiction. $\square$

Let us define a sequence of subgraphs $G_{i+1} \subseteq G_i \subseteq \ldots \subseteq G_1 = H$, by taking a set S_i of ℓ vertices of G_i maximising $d_{G_i}(S_i)$ and setting $G_{i+1} = G_i - S_i$ for every $i \geq 1$. It is now easy to see that $|E(G_{i+1})| < x^i m$ for every $i \geq 1$. Indeed, this is clearly true if G_i is empty, whereas otherwise we can apply Claim 2 to obtain $d_{G_i}(S_i) > (1-x)|E(G_i)|$ and $|E(G_{i+1})| < x|E(G_i)|$, and then use induction.

In particular, $|E(G_{\lfloor t/\ell \rfloor+1})| < x^{\lfloor t/\ell \rfloor} m = 1$ means that $G_{\lfloor t/\ell \rfloor+1}$ is empty. But then the set $S_1 \cup \cdots \cup S_{\lfloor t/\ell \rfloor}$ is a cover of H of size at most $\lfloor t/\ell \rfloor \ell \leq t$, which contradicts our assumption. $\square$

The next statement inverts the inequality between k and t given by the above theorem to show the upper bounds of Theorem 3.1.7 in the middle of the range.

Corollary 3.1.25. *Let $e^r \geq k > r \geq 2$ and let $m = \frac{4r}{\log k}$. Then $\mathrm{h}_r(k,r) \leq 4rm \log m$.*

Proof. Note that the bound is trivial from Observation 3.1.22 parts (2) and (3) when $4rm \log m \geq r^2$. We may therefore assume $4rm \log m < r^2$, which after plugging in the chosen value for m and exponentiating is equivalent to $k(\log k)^{16} > (4r)^{16}$ which is in turn easily seen to imply $k \geq r^6$. Let $t = r \lfloor 4m \log m \rfloor$. We only need to check that these values of k, r, t satisfy the condition of Theorem 3.1.24 on k with $\ell = r$. To see this, observe that $4r^2 \log r \leq r^4 < k^{2/3}$, and also,

$$\binom{t+r}{r}^{1/\lfloor 4m \log m \rfloor} \leq \left(\frac{e(t+r)}{r}\right)^{r/(3m \log m)}$$
$$\leq (e(1+4m \log m))^{\log k/(12 \log m)} \leq k^{1/3}$$

using $m \geq 4$ (from $k \leq e^r$) in the last inequality. □

Note that this corollary, combined with Observation 3.1.22 part (3), gives $\mathrm{h}_r(k,r) = O(r)$ when k is exponential in r, and $\mathrm{h}_r(k,r) \leq \frac{16r^2 \log r}{\log k}$ in general for smaller k (using Observation 3.1.22 part (2) for $\log k < 4$) as claimed by Theorem 3.1.7. The remaining upper bounds of Theorem 3.1.7 follow from Observation 3.1.22 parts (1), (2) and (3) in the range $k \in [1, cr]$ and from Theorem 3.1.23 in the range $k \in [\binom{2r}{r}, \infty)$.

3.1.5.2 *Lower bounds*

Let us start with a simple result when $k = \ell + 1$, which was already observed by Erdős et al. in [86].

Proposition 3.1.26. $\mathrm{h}_r(\ell+1, \ell) = r\ell$.

Proof. The upper bound $\mathrm{h}_r(\ell+1,\ell) \leq r\ell$ is given by Observation 3.1.22 part (2), so it is enough to construct an r-graph H with cover number $r\ell$ where any $\ell+1$ edges can be covered with ℓ vertices.

We can actually choose H to be the complete r-uniform hypergraph on $r\ell - 1 + r$ vertices. Indeed, $\tau(H) = r\ell$ because the complement of any $r\ell - 1$ vertices induces an r-edge. On the other hand, H has fewer than $r(\ell+1)$ vertices, so for any $(\ell+1)$-set of r-edges in H there are two edges that intersect. We can therefore cover this set with ℓ vertices by taking a vertex in this intersection, and one vertex for each of the remaining $\ell - 1$ edges. □

To give a lower bound on $\mathrm{h}_r(k,\ell)$ in general, we need to find an r-graph whose cover number is large, but for which any collection of k edges has a cover of size ℓ. The cover number is usually not difficult to estimate, but the (k,ℓ)-covering property can be hard to grasp. As it turns out, a simple counting trick can give good estimates on the largest k satisfying this property. (See also [86] for a somewhat weaker statement).

Theorem 3.1.27. *Let $r \geq 2$ and $t \geq \ell$ be positive integers. For any $k < \binom{t+r}{\ell}/\binom{t}{\ell}$, we have $\mathrm{h}_r(k,\ell) > t$.*

Proof. Let H be the complete r-uniform hypergraph on $t+r$ vertices. The cover number of H is easily seen to be $t+1$, and we will show that this graph satisfies the (k,ℓ)-covering property for any $k < \binom{t+r}{\ell}/\binom{t}{\ell}$. To see this, observe that every edge of H is covered by all but $\binom{t}{\ell}$ vertex ℓ-sets. But then for any k edges, there are at most $k\binom{t}{\ell} < \binom{t+r}{\ell}$ vertex ℓ-sets that do not cover all of them, which leaves at least one vertex cover of size ℓ. □

It will be convenient to formulate this trick in a more general form for future use in Section 3.1.6.

Lemma 3.1.28. *Let H be an r-uniform hypergraph and let G be an ℓ-uniform hypergraph on the same vertex set. Let δ be the minimum of $d_G(S)$ over $S \in E(H)$. Then any $\left\lfloor \frac{e(G)-1}{e(G)-\delta} \right\rfloor$ edges of H can be covered by an edge of G.*

Proof. Let $k \leq \frac{e(G)-1}{e(G)-\delta}$, and consider a collection of k edges $S_1, \ldots, S_k$ of H. There are at most $e(G)-\delta$ edges in G disjoint from each S_i. This means that there are at most $k(e(G)-\delta) \leq e(G)-1$ edges in G disjoint from *some* S_i. In particular, there is an edge of G that intersects all of the S_i. □

This trick might seem to give weak bounds, but they are actually close to best possible in our applications. For example, we have seen that Theorem 3.1.27 is tight for $t = \ell$. When $\ell = r$, it also implies the following result, which is only a $\log(\frac{r}{\log k})$ factor away from the upper bound of Corollary 3.1.25. We will discuss tightness in general in Section 3.1.7.

Corollary 3.1.29. *Let r, k be integers such that $e^{r/2} > k > r \geq 2$. Then $\mathrm{h}_r(k,r) \geq \frac{r^2}{4\log k}$.*

Proof. Let $t = \left\lfloor \frac{r^2}{2\log k} \right\rfloor$. Then $t \geq r \geq 2$ and $\binom{t+r}{r}/\binom{t}{r} \geq \left(\frac{t+r}{t}\right)^r > e^{r^2/(2t)} \geq k$, so we can apply Theorem 3.1.27 to obtain the result. □

Up to now, all our examples came from complete r-graphs. In the following lemma we show that when k is very close to ℓ we obtain better lower bounds by taking multiple copies of complete graphs instead.

Lemma 3.1.30. *Let $r \geq 2$ and $k > \ell$ be positive integers.* $\mathrm{h}_r(k,\ell) \geq \frac{r\ell^2}{6k}$.

Proof. Let H be the disjoint union of $\lceil \ell/2 \rceil$ copies $H_1, \ldots, H_{\lceil \ell/2 \rceil}$ of the complete r-graph on $\left\lfloor \frac{r\ell}{3k} \right\rfloor + r$ vertices. It is easy to see that the cover number of each H_i is at least $\frac{r\ell}{3k}$, and since there are $\lceil \ell/2 \rceil$ copies, we have $\tau(H) \geq \frac{r\ell^2}{6k}$. We just need to prove that any k edges can be covered with ℓ vertices.

Let us first show that any $\left\lceil \frac{3k}{\ell} \right\rceil$ edges in the same H_i have a common vertex. Indeed, every edge avoids exactly $\left\lfloor \frac{r\ell}{3k} \right\rfloor$ vertices, so for any collection of $\left\lceil \frac{3k}{\ell} \right\rceil$ edges, there are at most $\left\lceil \frac{3k}{\ell} \right\rceil \cdot \left\lfloor \frac{r\ell}{3k} \right\rfloor < r + \left\lfloor \frac{r\ell}{3k} \right\rfloor$ vertices that are not contained in all of them. In particular, some vertex is contained in all of them. When $\ell \leq 2$, this already shows that H satisfies the (k,ℓ)-covering property, so we may assume $\ell > 2$ from now on.

Now let S be any collection of k edges in H. It is easy to see that the edges of S can be split into $t \leq \lceil \ell/2 \rceil + k/\left\lceil \frac{3k}{\ell} \right\rceil$ subsets $S_1 \cup \cdots \cup S_t$ so that each S_j consists of at most $\left\lceil \frac{3k}{\ell} \right\rceil$ edges, all from the same H_i. (With at most one S_j containing fewer than $\left\lceil \frac{3k}{\ell} \right\rceil$ edges of H_i for every i.) When $\ell > 2$, we get $t \leq \ell$, so taking a single vertex in the intersection of each S_i gives a cover of S with at most ℓ vertices. $\square$

The lower bounds in Theorem 3.1.7 now follow from Observation 3.1.22 parts (1) and (3) in the range $k \in [1,r]$, from Proposition 3.1.26 for $k = r+1$, from Lemma 3.1.30 in the range $k \in (r, cr]$, from Corollary 3.1.29 in the range $k \in (r, e^{r/2}]$, and from Observation 3.1.22 part (4) in the remaining range $k \in (e^{r/2}, \infty)$.

3.1.6 *Results for r-partite hypergraphs*

In this section we present our estimates on $\mathrm{hp}_r(k)$, the maximum cover number of an r-partite r-graph in which any k edges have a transversal cover. Using the relationship with the monochromatic tree cover problem that we established in Section 3.1.4, this will allow us to prove Theorems 3.1.2 to 3.1.4 and 3.1.6 in Section 3.1.6.3.

The definitions clearly imply that $\mathrm{hp}_r(k) \leq \mathrm{h}_r(k,r)$. Indeed, most of our upper bounds follow from the upper bounds on $\mathrm{h}_r(k,r)$ obtained in the previous section. On the other hand, we need new examples for the lower bounds, as the complete r-graphs or copies of complete r-graphs that we used there are not r-partite. Nevertheless, our lower bounds for $\mathrm{hp}_r(k)$ are only a small constant factor away from the bounds we obtained for $\mathrm{h}_r(k,r)$ in the previous section. We begin by stating an analogue of Theorem 3.1.7 to summarise the results we are going to show in this section all in one place.

Theorem 3.1.31. *The following table describes the behaviour of* $\mathrm{hp}_r(k)$ *for fixed r as k varies. For arbitrary constant* $c > 1$ *we have*

Range of k	$[1, r]$	$r + 1$	$(r, cr]$	$(r, e^r]$	$\left[\binom{2r}{r}, \infty\right)$
Value of $\mathrm{hp}_r(k)$	∞	$[r(r-4), r^2]$	$\Theta\left(r^2\right)$	$\left[\frac{r^2}{12\log k}, \frac{16r^2\log r}{\log k}\right)$	r

Let us proceed with some easy observations, akin to Observation 3.1.22.

Observation 3.1.32. *For* $r \geq 2$ *we have:*

(1) $\mathrm{hp}_r(r) = \infty$,

(2) $\mathrm{hp}_r(r+1) \leq r^2$,

(3) $\mathrm{hp}_r(k+1) \leq \mathrm{hp}_r(k)$ *and*

(4) $r \leq \mathrm{hp}_r(k)$.

Proof. Part (1) can be seen by taking an arbitrarily large collection of disjoint edges. Part (2) follows from Observation 3.1.22 part (2). Part (3) follows, as before, because if any $k+1$ edges have a transversal cover, then the same is true for any k edges, as well. Part (4) can be shown by taking a hypergraph consisting of exactly r disjoint edges. $\square$

3.1.6.1 *Upper bounds*

Since $\mathrm{hp}_r(k) \leq \mathrm{h}_r(k,r)$, Theorem 3.1.23 implies that $\mathrm{hp}_r(k) = r$ whenever $k \geq \binom{2r}{r}$. In this particular case, we can obtain a slightly better result:

Theorem 3.1.33. *Let* $r \geq 2$ *be an integer. If* $k \geq 2^r$, *then any r-graph satisfying the r-partite k-covering property has a transversal cover. In particular,* $\mathrm{hp}_r(k) = r$.

Proof. We follow a similar approach as in Theorem 3.1.23, but use the fact that the covers involved are transversals to obtain an improvement from Theorem 3.1.14.

Suppose for contradiction that there is an r-uniform r-partite hypergraph without a transversal cover satisfying that any 2^r of its edges have a transversal cover. Let H be such a hypergraph with the smallest possible number of edges. Let $A_1, \ldots, A_t$ be the edges of H. By the minimality assumption, the

r-graph $H \setminus A_i$ has a transversal cover B_i for every i. Note that $A_i \cap B_i = \emptyset$, as otherwise B_i would be a transversal cover of H, and $A_i \cap B_j \neq \emptyset$ whenever $i \neq j$, because B_j is a cover of $H \setminus A_j$, which contains A_i.

If $V_1, \ldots, V_r$ denote the parts of the r-partition of H, then we have $|A_i \cap V_j| = 1$ (as H is r-uniform and r-partite), and $|B_i \cap V_j| = 1$ (as B_i is a transversal cover) for every i and j. Therefore we can apply Theorem 3.1.14 with $s = r$ to get that H has at most $\prod_{i=1}^{r} \binom{1+1}{1} = 2^r$ edges. But then H admits a transversal cover, a contradiction. □

This bound on k is tight if we insist on the cover being transversal, as shown by the complete r-partite r-graph with 2 vertices in each part: this satisfies the r-partite $(2^r - 1)$-covering property, but does not have any transversal cover. As we will see later (see Theorem 3.1.36), Theorem 3.1.33 is very close to being tight, in terms of the bound on k, even if we allow any cover of size r.

In general, Corollary 3.1.25 (or Theorem 3.1.7) gives the following bound using $\mathrm{hp}_r(k) \leq \mathrm{h}_r(k, r)$.

Theorem 3.1.34. *For $e^r \geq k > r \geq 2$, we have $\mathrm{hp}_r(k) \leq \frac{16r^2 \log r}{\log k}$.*

It might be interesting to mention that a slight improvement on this result can be obtained using an argument along the lines of Theorems 3.1.23 and 3.1.33 that is based on a generalisation of Theorem 3.1.14 by Moshkovitz and Shapira [188]. However, the improvement is very minor, so we omit the details.

3.1.6.2 *Lower bounds*

As we mentioned before, our lower bounds on $\mathrm{h}_r(k, r)$ do not carry over to $\mathrm{hp}_r(k)$, because complete r-graphs (our constructions) are not r-partite. In fact, it is not immediately clear how to find r-partite graphs that satisfy the r-partite k-covering property and have large cover number.

One of the main difficulties is that the vertices of the smallest part cover all the edges, so to get a meaningful bound on the cover number, we need to construct an r-graph G, whose edges touch many vertices in each part. On the other hand, G cannot contain any matching of more than r edges, otherwise it fails the r-partite k-covering property. This immediately rules out complete or random r-partite r-graphs as possible candidates for G to improve Observation 3.1.32 part (4). We present a construction that does significantly better.

For some m and $t < r$, let $H_{r,t,m}$ be the following r-partite r-graph. In each part of the r-partition, we fix m vertices that we will call *important*. Now for every set S of $r-t$ important vertices in different parts, we add an edge to $H_{r,t,m}$ containing S and t unique "unimportant" vertices in each of the remaining parts. (So $H_{r,t,m}$ has $\binom{r}{r-t}m^{r-t}$ edges, and each unimportant vertex belongs to exactly one edge.) For example, if $t=r$ then $H_{r,t,m}$ is empty and if $t=0$ it is the complete r-uniform, r-partite graph with m vertices per part.

Proposition 3.1.35. $\tau(H_{r,t,m}) = (t+1)m$

Proof. By taking any $t+1$ parts and choosing all important vertices in these parts, we get $(t+1)m$ vertices that cover $H_{r,t,m}$. Indeed, each edge of $H_{r,t,m}$ contains an important vertex in $r-t$ parts, so it must contain one in a part we selected.

Now let S be a set of at most $(t+1)m-1$ vertices. We will show it does not cover $H_{r,t,m}$. As long as there is an unimportant vertex $v \in S$, we can replace it with any important vertex in the unique edge containing v: all previously covered edges are still covered by S. So we may assume that S only contains important vertices. However, since $|S| \leq (t+1)m-1$, at least $r-t$ parts contain some important vertex that is not in S. Then the edge defined by these $r-t$ important vertices is not covered by S. $\square$

Let us start with our bound for the end of the range (where $\mathrm{hp}_r(k)$ becomes r).

Theorem 3.1.36. *For any $k < \binom{r}{\lfloor (r+1)/2 \rfloor} + \binom{r}{\lceil (r+1)/2 \rceil}$, we have $\mathrm{hp}_r(k) > r$.*

Proof. Let H be the disjoint union of $H_1 = H_{r,\lfloor (r-1)/2 \rfloor,1}$ and $H_2 = H_{r,\lceil (r-1)/2 \rceil,1}$. Then, by Proposition 3.1.35, $\tau(H) = \left\lceil \frac{r-1}{2} \right\rceil + 1 + \left\lfloor \frac{r-1}{2} \right\rfloor + 1 = r+1$. On the other hand, we will show that $H \setminus e$ has a transversal cover for any edge e. As H has $\binom{r}{\lfloor (r+1)/2 \rfloor} + \binom{r}{\lceil (r+1)/2 \rceil}$ edges, this will imply that it has the r-partite $\left(\binom{r}{\lfloor (r+1)/2 \rfloor} + \binom{r}{\lceil (r+1)/2 \rceil} - 1\right)$-covering property.

So let e be an edge in H, and assume that it belongs to H_1 (the case $e \in H_2$ is similar). Let C_1 be the set of $\left\lfloor \frac{r-1}{2} \right\rfloor$ important vertices of H_1 that are not contained in e, and let C_2 consist of the $r - \left\lfloor \frac{r-1}{2} \right\rfloor = \left\lceil \frac{r-1}{2} \right\rceil + 1$ important vertices of H_2 in the parts not represented in C_1. It is easy to see that C_1 covers all edges in H_1 except e, and C_2 covers H_2. So $C_1 \cup C_2$ is indeed a transversal cover of $H \setminus e$. $\square$

We now prove our general lower bound.

Theorem 3.1.37. *For any $k > r \geq 2$, we have $\mathrm{hp}_r(k) \geq \frac{r^2}{12 \log k}$.*

Proof. We may assume that $\frac{r^2}{12\log k} > r$, or equivalently, $e^{r/12} > k$, as otherwise the statement is trivial. Let $t = \left\lfloor \frac{r-1}{2} \right\rfloor$ and $m = \left\lfloor \frac{r+1}{4\log k} \right\rfloor$. Since $e^{r/12} > k$ we have $\frac{r+1}{4\log k} > 2$, which implies $m = \left\lfloor \frac{r+1}{4\log k} \right\rfloor \geq \frac{2}{3} \cdot \frac{r+1}{4\log k}$. Thus, by Proposition 3.1.35, $H = H_{r,t,m}$ satisfies $\tau(H) = (t+1)m \geq \frac{r^2}{12\log k}$. To prove that H has the r-partite k-covering property, we will apply Lemma 3.1.28 with G chosen as the complete r-partite r-graph on the set of important vertices of H. Note that any edge of G is transversal, so Lemma 3.1.28 implies that H satisfies the r-partite $\left(\frac{e(G)-1}{e(G)-\delta}\right)$-covering property, where δ is the minimum number of edges in G intersecting an edge of H. We only need to check that $\frac{e(G)-1}{e(G)-\delta} \geq k$.

To see this, note that G has m^r edges, and every edge of H is disjoint from exactly $(m-1)^{r-t}m^t$ of them. Thus indeed,

$$\frac{e(G)-1}{e(G)-\delta} = \frac{m^r - 1}{(m-1)^{r-t}m^t} > \left(1 + \frac{1}{m}\right)^{r-t} > e^{(r-t)/(2m)} \geq k.$$

□

Let us now consider the beginning of the range. As mentioned in the introduction, the problem of determining $\mathrm{hp}_r(r+1)$ is a special case of Ryser's conjecture [144]. Improvements over the trivial upper bound of r^2 on $\tau(H)$ for H with $\nu(H) \leq r$ are only known for $r \leq 5$ [6, 142], in which case they also provide an improvement over our upper bound Observation 3.1.32 part (2). Tightness examples for Ryser's conjecture are not known for all values of r. In fact, they are only known to exist when $r-1$ is a prime power [236], $r-2$ is a prime [3] or for certain special small values [4, 7, 111]. There are, however, examples for every r that are almost tight: Haxell and Scott [141] constructed intersecting r-partite r-graphs with cover number at least $r-4$ for all values of r. We can use them as a black box to get the following bound.

Theorem 3.1.38. *For every $r \geq 2$, we have* $\mathrm{hp}_r(r+1) \geq r(r-4)$.

Proof. Let H' be an intersecting r-partite r-graph with cover number at least $r-4$, as given by [141], and let H consist of r copies of H'. Then we clearly have $\tau(H) \geq r(r-4)$, and it is easy to see that H also satisfies the r-partite $(r+1)$-covering property. Indeed, among any $r+1$ edges, some two belong to the same copy of H', and hence intersect. A vertex in the intersection can then be extended to a transversal cover by taking one vertex each from the remaining edges. □

Of course, the same argument gives $\text{hp}_r(r+1) \geq r(r-1)$ whenever there is a construction for r matching Ryser's conjecture.

As in the previous section, taking more copies of our general construction is helpful at the beginning of the range. However, there is an added difficulty here in ensuring that the cover we get is transversal.

Theorem 3.1.39. *For every k and $r \geq 2$, we have* $\text{hp}_r(k) \geq \frac{r^3}{50k}$.

Proof. If $\frac{r^3}{50k} \leq r$ or $k \leq r$, then the inequality follows from Observation 3.1.32. We may therefore assume $k > r > 50$. Let $t = \left\lfloor \frac{r^2}{10k} \right\rfloor$, and let H consist of $\left\lfloor \frac{r}{4} \right\rfloor$ disjoint copies of $H_{r,t,1}$. Then Proposition 3.1.35 and $\frac{r}{4} > 12$ imply $\tau(H) = (t+1)\left\lfloor \frac{r}{4} \right\rfloor > \frac{r^3}{50k}$. Note that in a single copy of $H_{r,t,1}$, every part contains one important vertex, and every edge contains $r-t$ important vertices. This means that any set of at most $\frac{r}{2t}$ edges within the same copy contains at least $r - t \cdot \frac{r}{2t} = \frac{r}{2}$ common important vertices.

Now let S be any collection of k edges in H. It is easy to see that the edges of S can be split into $m \leq \left\lfloor \frac{r}{4} \right\rfloor + k/\left\lfloor \frac{r}{2t} \right\rfloor \leq \frac{r}{2}$ subsets $S_1 \cup \cdots \cup S_m$ so that each S_i consists of at most $\frac{r}{2t}$ edges from the same copy. Now for every i, there are at least $\frac{r}{2}$ vertices common to all edges in S_i. But then we can greedily choose one vertex in a different part to cover each S_i. By adding an arbitrary vertex from any part where we have used no vertices so far, we obtain a transversal cover of our arbitrary collection of k edges. $\square$

Remark. Another class of examples which would give comparable lower bounds throughout this section is given as follows. The *(r,s)-sum-hypergraph* $S_{r,s}$ is an r-partite hypergraph with $s+1$ vertices in each part i, denoted by $v_{i,j}$ for $0 \leq j \leq s$. The edge set is given by $\{(v_{1,x_1}, \ldots, v_{r,x_r}) \mid \sum_{i=1}^{r} x_i = s\}$. It is not hard to see that the cover number of $S_{r,s}$ is $s+1$ when $r \geq 2$. Moreover, one can show, using Lemma 3.1.28, that $S_{r,s}$ satisfies the r-partite k-covering property with similarly good bounds on k as obtained above.

3.1.6.3 *Proofs of the main theorems*

Proof of Theorem 3.1.2. (a): Let $G \sim \mathcal{G}(n,p)$ with $p < \left(\frac{c \log n}{n}\right)^{\sqrt{r}/2^{r-2}}$. Our goal is to show that w.h.p. $\text{tc}_r(G) > r$. When $r \leq 2$, our assumption implies $p \ll 1/n$. In this case, $\mathcal{G}(n,p)$ has more than 2 components w.h.p. (see e.g. [16]), so the statement follows. When $r \geq 3$, we have $\binom{r-1}{\lfloor r/2 \rfloor} + \binom{r-1}{\lceil r/2 \rceil} > \frac{2^{r-2}}{\sqrt{r}}$, so Theorem 3.1.5(2) implies $\text{tc}_r(G) \geq \text{hp}_{r-1}\left(\binom{r-1}{\lfloor r/2 \rfloor} + \binom{r-1}{\lceil r/2 \rceil} - 1\right) + 1$

w.h.p. in this probability range. Here $\mathrm{hp}_{r-1}\left(\binom{r-1}{\lfloor r/2\rfloor}+\binom{r-1}{\lceil r/2\rceil}-1\right)>r-1$ follows from Theorem 3.1.36, so indeed, $\mathrm{tc}_r(G)>r$.

(b): Let $G\sim\mathcal{G}(n,p)$ with $p>\left(\frac{C\log n}{n}\right)^{1/2^r}$. Our goal is to show that w.h.p. $tc_r(G)\leq r$. Theorem 3.1.5(1) implies that w.h.p. $\mathrm{tc}_r(G)\leq\mathrm{hp}_r(2^k)$. The result now follows since Theorem 3.1.33 implies $\mathrm{hp}_r(2^k)\leq r$. □

Proof of Theorem 3.1.3. Let $G\sim\mathcal{G}(n,p)$ with

$$\left(\frac{C\log n}{n}\right)^{\frac{1}{r}}<p<\left(\frac{c\log n}{n}\right)^{\frac{1}{d(r+1)}}$$

, where $d>1$. Our goal is to show that w.h.p. $\mathrm{tc}_r(G)=\Theta(r^2)$. Theorem 3.1.17 shows that in this probability range w.h.p. we have $\mathrm{tc}_r(G)\leq 3r^2$. On the other hand, for $r\geq 3$ Theorem 3.1.5(2) shows that in this probability range w.h.p. we have $\mathrm{tc}_r(G)\geq\mathrm{hp}_{r-1}(d(r+1)-1)+1$. We can then apply Theorem 3.1.39 to get $\mathrm{hp}_{r-1}(d(r+1)-1)\geq\frac{r^2}{300d}$. □

Proof of Theorem 3.1.4. Let $G\sim\mathcal{G}(n,p)$ with

$$\left(\frac{C\log n}{n}\right)^{\frac{1}{k}}<p<\left(\frac{c\log n}{n}\right)^{\frac{1}{k+1}}$$

, for $k>r\geq 2$. Our goal is to show that w.h.p. $\frac{r^2}{20\log k}\leq\mathrm{tc}_r(G)\leq\frac{16r^2\log r}{\log k}$. Theorem 3.1.5 shows that w.h.p. $\mathrm{hp}_{r-1}(k)\leq\mathrm{tc}_r(G)\leq\mathrm{hp}_r(k)$. Then Theorem 3.1.34 gives $\mathrm{hp}_r(k)\leq\frac{16r^2\log r}{\log k}$. Also, for $r\geq 10$, Theorem 3.1.37 gives $\mathrm{hp}_{r-1}(k)\geq\frac{(r-1)^2}{12\log k}\geq\frac{r^2}{20\log k}$. For $r<10$, the lower bound is trivial. □

Proof of Theorem 3.1.6. Let G be a graph with $\delta(G)\geq(1-1/2^r)n$. Our goal is to show that in any r-colouring of G we can find a collection of trees of distinct colours which cover the vertex set of G. In order to apply Lemma 3.1.19, we need to show that any $k=2^r$ vertices $\{v_1,\dots,v_k\}$ have a common neighbour. Note that for this application, we can consider a vertex to be adjacent to itself, so if M_i denotes the set of non-neighbours of v_i, then $|M_i|<n/k$. But then $|\bigcup_{i=1}^k M_i|<n$, so indeed there is a common neighbour (possibly one of the v_i). This means that we can apply Lemma 3.1.19 and then Theorem 3.1.33 to get $\mathrm{tc}_r(G)\leq\mathrm{hp}_r(k)=r$. Moreover, Theorem 3.1.33 gives a transversal cover for the hypergraph, so according to Lemma 3.1.19, any r-edge-colouring of G has a desired monochromatic tree cover using distinct colours. □

3.1.7 *Concluding remarks and open problems*

In this section we obtained a very good understanding of how $\mathrm{h}_r(k,\ell)$ behaves as k varies. However, there is a multiplicative gap of about $\log(\frac{\ell}{\log k})$ between our lower and upper bounds in the middle of the range. It would be interesting to determine the behaviour of $\mathrm{h}_r(k,\ell)$ more accurately. Any improvement on the upper bound would translate to improved bounds on $\mathrm{hp}_r(k)$ and tc_r, as well. An improvement on the lower bounds, while still interesting, would only yield such an improvement if the example constructed was r-partite.

Another very interesting problem is to determine what is the smallest k for which $\mathrm{hp}_r(k) = r$. The bounds given by Theorems 3.1.33 and 3.1.36 (2^r vs $2\binom{r}{r/2}$) are only a factor of $\sqrt{r}$ away from each other. We believe that Theorem 3.1.36 is probably closer to the truth. In fact, it is tight for small values of r.

Our Theorem 3.1.3 shows that the tree cover number of a random graph is $\Theta(r^2)$ just above the probability threshold where $\mathrm{tc}_r(\mathcal{G}(n,p))$ becomes bounded. More precisely, we show that $r^2(1-o(1)) \le \mathrm{tc}_r(\mathcal{G}(n,p)) \le 3r^2$ when p is slightly above $(\frac{\log n}{n})^{1/r}$, it would be interesting to know it more accurately.

Very recently, Korándi, Lang, Letzter and Pokrovskiy [161] obtained a related result about partitioning into monochromatic cycles with the same answer as in our Theorem 3.1.3: They proved that r-edge-coloured graphs of minimum degree $\frac{n}{2} + \Theta(\log n)$ can be partitioned into $\Theta(r^2)$ disjoint monochromatic cycles.

Acknowledgements. We thank Louis DeBiasio and Tuan Tran for helpful comments on an earlier version of this manuscript.

Appendix A: An alternative perspective

In this section we present an alternative perspective for bounding $\mathrm{h}_r(k,\ell)$ compared to what we did in Section 3.1.5. One advantage of this reformulation is that it can be used to show that Lemma 3.1.28 is close to optimal in our applications. We believe it could also be useful in improving the upper bounds on $\mathrm{h}_r(k,\ell)$ obtained in Theorem 3.1.24.

Let H be an r-graph. For a set $S \subseteq V(H)$, let $m(\bar{S})$ denote the set of edges in H disjoint from S. We define the *ℓ-covering hypergraph* $\mathrm{ch}_\ell(H)$ of H on vertex set $E(H)$, where for each subset $S \subseteq V(H)$ of size ℓ, we take $m(\bar{S})$

as an edge of $\mathrm{ch}_\ell(H)$. Let us show how some properties of $\mathrm{ch}_\ell(H)$ relate to those of H.

Proposition 3.1.40. *Let H be an r-graph.*

1. *The smallest k for which H does* not *satisfy the (k, ℓ)-covering property is equal to $\tau(\mathrm{ch}_\ell(H))$.*
2. *$\tau(H) > t\ell$ if and only if $\mathrm{ch}_\ell(H)$ is t-intersecting (i.e. any t edges in $\mathrm{ch}_\ell(H)$ have a common vertex).*

Proof. We start with part 1. Let $t = \tau(\mathrm{ch}_\ell(H))$ and let T be a minimal cover of $\mathrm{ch}_\ell(H)$. Note that T is a set of t edges in H. This set cannot be covered with ℓ vertices in H, because if S was such a cover, then $m(\bar{S})$ would be disjoint from T in $\mathrm{ch}_\ell(H)$. Therefore, H does not satisfy the (t, ℓ)-covering property.

Let T' be a set of at most $t-1$ edges of H. Since T' is not a cover of $\mathrm{ch}_\ell(H)$, there is an edge $m(\bar{S})$ in $\mathrm{ch}_\ell(H)$ that is not covered by T'. This means that S intersects every edge in T', so T' can be covered by ℓ vertices of H. Since T' was arbitrary, H satisfies the $(t-1, \ell)$-covering property.

Let us now turn to part 2. If H has a cover of size at most $t\ell$, then it can be covered with at most t subsets of $V(H)$ of size ℓ. The edges corresponding to these sets in $\mathrm{ch}_\ell(H)$ have no common vertex, so $\mathrm{ch}_\ell(H)$ is not t-intersecting.

Since $\mathrm{ch}_\ell(H)$ is not t-intersecting, $V(H)$ has t subsets of size ℓ such that no edge of H is disjoint to all of them. This means that union of these subsets covers H, and in particular, $\tau(H) \le t\ell$. □

This proposition allows us to translate the two concepts involved in $\mathrm{h}_r(k, \ell)$ into somewhat simpler parameters of covering hypergraphs. It will be useful for us to use an even simpler parameter, the relaxation of the covering number:

A fractional cover of a hypergraph H is an assignment of weights to the vertices of H so that each edge receives a total weight of at least 1. The minimum possible total weight assigned is called the *fractional cover number*, and is denoted by $\tau^*(H)$. Lovász [182] found the following connection between the two cover numbers when H has maximum degree d:

$$\tau^*(H) \le \tau(H) \le (1 + \log d)\tau^*(H). \tag{3.1}$$

As it turns out, several of our tools can be proved by combining Proposition 3.1.40 with (3.1).

Indeed, we can now easily get a new proof of Lemma 3.1.28 (one of our main tools for bounding $\mathrm{h}_r(k, \ell)$ from below), one that actually shows that the lower bound we get on k is close to optimal. Indeed, the edges of G in Lemma 3.1.28 (as ℓ-sets) define $e(G)$ edges in $\mathrm{ch}_\ell(H)$, and the condition of the lemma ensures that every vertex of $\mathrm{ch}_\ell(H)$ touches at most $e(G)-\delta$ of them (recall that δ is the minimum of $d_G(S)$ over $S \in E(H)$). So these edges form a subhypergraph $\mathcal{H} \subseteq \mathrm{ch}_\ell(H)$ of maximum degree at most $e(G)-\delta$. A double-counting argument then immediately implies $\tau^*(\mathrm{ch}_\ell(H)) \geq \tau^*(\mathcal{H}) \geq \frac{e(G)}{e(G)-\delta}$. Using the first inequality of (3.1) and Proposition 3.1.40, we get that H has the (k,ℓ)-covering property for any $k < \frac{e(G)}{e(G)-\delta}$, establishing Lemma 3.1.28. Moreover, the same argument shows that $\tau^*(\mathrm{ch}_\ell(H))$ (which can easily be found with a linear program) approximates, up to a factor of $1+\log d(\mathrm{ch}_\ell(H))$ (here d denotes the maximum degree as in (3.1)), the smallest k such that H satisfies the (k,ℓ)-covering property.

In Theorem 3.1.27, for example, we choose H as the complete r-graph on $t+r$ vertices. As $\mathrm{ch}_\ell(H)$ is regular and uniform, it is not hard to see that $\tau^*(\mathrm{ch}_\ell(H)) = \binom{t+r}{\ell}/\binom{t}{\ell}$ in this case. So Lemma 3.1.28 gives us the smallest k for which this H has the (k,ℓ)-covering property, up to a factor of $1+\log\binom{|V(H)|}{\ell} \approx \ell \log |V(H)|$. A similar analysis gives essentially the same bound for the graph H used in Theorem 3.1.37, as well.

Proposition 3.1.40 can also be used to obtain upper bounds on $\mathrm{h}_r(k,\ell)$. For example, in order to prove $\mathrm{h}_r(k,\ell) \leq t\ell$, it is enough to show that if $\tau(\mathrm{ch}_\ell(H)) = k+1$ for some hypergraph H, then $\mathrm{ch}_\ell(H)$ cannot be t-intersecting. This raises the following natural question: What is the largest possible cover number of a t-intersecting hypergraph?

A standard upper bound is $\frac{r-1}{t-1}+1$ for hypergraphs with hyperedges of size at most r (see [180]), which is tight for certain complete r-graphs. However, it can be quite far from the truth for sparser hypergraphs with high uniformity. The following theorem works better in our case.

Theorem 3.1.41. *Let G be a t-intersecting hypergraph with n vertices and maximum degree d. Then $\tau(G) \leq n^{1/t}(1+\log d)$.*

Proof. The proof is secretly the same as the one for Theorem 3.1.24, but translated into the language of this setting. We do, however, benefit from this new setting by essentially replacing Claim 2 with (3.1).

Claim. *Let $S \subseteq V(G)$ be a nonempty set. If $|e \cap S| \geq |S|/n^{1/t}$ for every edge $e \in E(G)$, then $\tau(G) \leq (1+\log d(G))n^{1/t}$.*

Proof. Let us assign weight $n^{1/t}/|S|$ to every vertex of S, and weight 0 to every other vertex of G. This is a fractional cover because the total weight of every edge e in G is at least $|e \cap S| \cdot n^{1/t}/|S| \geq 1$. In particular, $\tau^*(G) \leq n^{1/t}$, and using (3.1) we get $\tau(G) \leq (1 + \log d(G))n^{1/t}$. □

We may now assume that for every nonempty subset $S \subseteq V(G)$, there is an edge $e \in E(G)$ such that $|e \cap S| < |S|/n^{1/t}$, as otherwise we are done by the claim. Let us denote this edge by $m(S)$. We construct a series of nested subsets of $V(G)$ as follows. We set $S_1 = V(G)$, and define $e_i = m(S_i)$ and $S_{i+1} = S_i \cap e_i$ for $i = 1, \ldots, t$. This is possible because $S_{i+1} = e_1 \cap \cdots \cap e_i$ and H is t-intersecting. Moreover, it is easy to see by induction that $|S_{i+1}| < n^{1-i/t}$. But then $1 \leq |e_1 \cap \cdots \cap e_t| = |S_{t+1}| < 1$ gives us a contradiction. □

Finally, let us show how Theorem 3.1.24 can be deduced from Theorem 3.1.41.

Alternative proof of Theorem 3.1.24. As in the original proof, let us assume for contradiction that there is an r-uniform hypergraph H satisfying the (k, ℓ)-covering property with $\tau(H) > t$. As before, we may assume that H is critical and has no isolated vertices, and thus has $|E(H)| \leq \binom{r+t}{t}$ and $|V(H)| \leq r\binom{r+t}{t}$. Note that for $G = \mathrm{ch}_\ell(H)$, we have $n = |V(G)| = |E(H)| \leq \binom{r+t}{t}$ and since each edge of G corresponds to an ℓ-subset of $V(H)$ we get $d(G) \leq \binom{|V(H)|}{\ell} \leq |V(H)|^\ell \leq r^\ell \binom{r+t}{r}^\ell \leq (r+t)^{r\ell} \leq (r + r\ell)^{r\ell} \leq e^{2r\ell\log(r\ell)-1}$ using the assumption $t \leq r\ell$. As $\tau(H) > t$, we can apply Proposition 3.1.40 part 2 to deduce that G is $\lfloor t/\ell \rfloor$-intersecting. Thus Theorem 3.1.41 implies that

$$\tau(G) \leq n^{1/\lfloor t/\ell \rfloor}(1 + \log d(G)) \leq \binom{r+t}{t}^{1/\lfloor t/\ell \rfloor} 2r\ell\log(r\ell).$$

Finally, Proposition 3.1.40 part 1 implies $k < \tau(G) \leq \binom{r+t}{t}^{1/\lfloor t/\ell \rfloor} 2r\ell\log(r\ell)$, which is a contradiction. □

3.2 TOURNAMENT QUASIRANDOMNESS FROM LOCAL COUNTING

3.2.1 *Introduction*

A combinatorial structure is said to be "quasirandom" if it behaves in a similar manner to a random structure, where the comparison is made with respect to some deterministic property. The systematic study of quasirandomness was initiated by Thomason [235] [234], and Chung, Graham and Wilson [58], who examined notions of quasirandomness arising from various graph properties. One of the surprising conclusions of these papers is that a wide-range of natural graph properties all lead to essentially the same notion of quasirandomness, in the sense that a graph satisfying one of the properties necessarily satisfies them all. Since then, notions of quasirandomness have been extensively studied in a wide variety of contexts, including hypergraphs [56], [130], [189], [213], permutations [67], [166] and groups [131]. The reader is referred to the survey [170] for an overview of this extensive topic.

In this section we will study notions of quasirandomness for tournaments. The first paper on this topic is due to Chung and Graham [57] who proved that, as in the graph case, a wide range of natural tournament properties give rise to the same notion of quasirandomness. Before stating some of their results we require a little notation. A tournament $T = (V, E)$ consists of a set of vertices $V = V(T)$, together with a set of edges $E = E(T) \subset V \times V$, with the property that (i) $(u, u) \notin E$ for all $u \in V$ and (ii) exactly one of (u, v) and (v, u) lies in E for all distinct $u, v \in V$. We often write $\overrightarrow{uv}$ to denote an edge (u, v). A tournament H appears as a subtournament of T if there is a map $\phi : V(H) \to V(T)$ such that $\overrightarrow{uv} \in E(H)$ if and only if $\overrightarrow{\phi(u)\phi(v)} \in E(T)$. The map ϕ is said to be a labelled embedding of H into T. Let

$$N_T^*(H) = \big|\{\phi : V(H) \to V(T) : \phi \text{ is a labelled embedding of } H \text{ into } T\}\big|.$$

Given $U \subset V(T)$ let $T[U]$ denote the subtournament of T induced by the vertex set U. Also let $N_T^*(H; U) := N_{T[U]}^*(H)$. For $u \in V(T)$ let $d_T^+(u) = |\{v \in V(T) : \overrightarrow{uv} \in E(T)\}|$ and $d_T^-(u) = |\{v \in V(T) : \overrightarrow{vu} \in E(T)\}|$. A tournament T is regular if $d_T^+(u) = d_T^-(u)$ for all $u \in V(T)$. For $U, W \subseteq V(T)$ we denote by $e(U, W)$ the number of edges starting in U and ending in W.

We say that T is an n-vertex tournament if $|V(T)| = n$. An ordering of T is a bijective map $\sigma : V \to [n]$ and the set of all orderings of T is naturally identified with S_n, the symmetric group on n elements. An edge $\overrightarrow{uv} \in E$ is a σ-forward edge if $\sigma(u) < \sigma(v)$ and we write $F_{\sigma,T} \subset E$ to denote the set of σ-forward edges of T. Let $B_{\sigma,T} = E \setminus F_{\sigma,T}$, the set of σ-backward edges. T is said to be transitive if $F_{\sigma,T} = E$ for some $\sigma \in S_n$ and write Tr_n to denote the unique (up to isomorphism) n-vertex transitive tournament. Lastly, we will write $a \pm b$ to denote some value c with $a - b \le c \le a + b$.

Our starting point is a result due to Chung and Graham [57], which gives two equivalences of tournament quasirandomness.

Theorem 3.2.1 (Chung–Graham). *Let $h \in \mathbb{N}$ with $h \ge 4$. Then for any n-vertex tournament T, the following properties are equivalent:*

- $\mathcal{P}_1$: $|F_{\sigma,T}| = \frac{1}{2}\binom{n}{2} \pm o(n^2)$ *for every ordering σ of T.*
- $\mathcal{P}_2(h)$: $N^*_T(H) = 2^{-\binom{h}{2}} n^h \pm o(n^h)$ *for every h-vertex tournament H.*

In fact, there are nine further equivalences given in this paper (see also [132], [153]). Equivalence here is understood in the following sense, focusing on the implication $\mathcal{P}_1 \implies \mathcal{P}_2(h)$: given $\varepsilon > 0$ there is $\delta > 0$ and $n_0 \in \mathbb{N}$ so that if T is an n-vertex tournament which satisfies $|F_{\sigma,T}| = \frac{1}{2}\binom{n}{2} \pm \delta n^2$ for every ordering σ of T and $n \ge n_0$ then $N^*_T(H) = 2^{-\binom{h}{2}} n^h \pm \varepsilon n^h$ for every h-vertex tournament H.

It is easily seen that both properties $\mathcal{P}_1$ and $\mathcal{P}_2(h)$ hold for a random n-vertex tournament T with high probability. We will say T is *quasirandom* if it satisfies $\mathcal{P}_1$, that is $|F_{\sigma,T}| = \frac{1}{2}\binom{n}{2} \pm o(n^2)$. In light of Theorem 3.2.1, this notion is equivalent to the analogous notion arising from $\mathcal{P}_2(h)$.

3.2.1.1 *Globally forcing tournaments*

Although $\mathcal{P}_2(h)$ guarantees that T is quasirandom, it is natural to ask whether this can already be deduced from the count of a *single* tournament H. That is, does quasirandomness of T already follow if $N^*_T(H) = 2^{-\binom{h}{2}} n^h \pm o(n^h)$ for a single h-vertex tournament H?

Definition. Let H be an h-vertex tournament. Given an n-vertex tournament T, consider the following property:

- $\mathcal{P}_2(H)$: $N^*_T(H) = 2^{-\binom{h}{2}} n^h \pm o(n^h)$.

The tournament H is said to be *globally forcing* if $\mathcal{P}_2(H) \implies \mathcal{P}_1$.

It follows easily from exercise 10.44(b) of [180] that for $h \ge 4$ each transitive tournament Tr_h is globally forcing. This statement was recently reproved in the language of flag-algebras by Coregliano and Razborov [69]. Furthermore, Coregliano, Parente and Sato [68] constructed a 5-vertex non-transitive globally forcing tournament T_5. Our first observation is that there are no further globally forcing tournaments.

Proposition 3.2.2. *A tournament H is globally forcing if and only if it is either transitive with 4 or more vertices or if $H = T_5$.*

Here we establish the above proposition for tournaments on 7 or more vertices. After the first version of the source material appeared on arXiv, Hancock, Kabela, Kral, Martins, Parente, Skerman and Volec [139] managed to resolve the remaining cases of 5 and 6 vertex tournaments, building on some of our ideas.

3.2.1.2 *Locally forcing tournaments*

We have seen that having the "correct count" of a fixed tournament H is not enough to guarantee quasirandomness, for essentially any non-transitive H. This is a fairly common situation when studying quasirandom properties in general and the key insight to understand why came from Simonovits and Sós [222] who observed that quasirandomness is a hereditary property, in the sense that any large subgraph of a random structure must also be random-like.

This leads us to the natural question of whether requiring that T contains the "correct count" of H in all large subsets of $V(T)$ is sufficient to guarantee quasirandomness of T. To be more precise, consider the following definition.

Definition. Let H be an h-vertex tournament. Given an n-vertex tournament T, consider the following property:

- $\mathcal{P}_2^*(H)$: $N_T^*(H;U) = 2^{-\binom{h}{2}}|U|^h \pm o(n^h)$ for every set $U \subset V(T)$.

The tournament H is said to be *locally forcing* if $\mathcal{P}_2^*(H) \implies \mathcal{P}_1$.

An analogous property for graphs was studied by Simonovits and Sós for both induced [222] and not-necessarily induced subgraphs [221] as well as for hypergraphs by Conlon, Hàn, Person and Schacht [66] and Dellamonica and Rödl [71]. In the case of graphs the not-necessarily induced case turned out to be much easier and was resolved by Simonovits and Sós [221] who showed that all non-empty graphs must be "locally forcing". This was recently reproved, without use of the regularity lemma, by Conlon, Fox and Sudakov [64]. On the other hand the induced case is still very much open and seems to be the closer analogue to our problem. Indeed, in the case of embedding a tournament one needs to ensure that each pair of vertices get mapped to an edge with one of two possible orientations, while in the induced case each pair needs to be sent to one of two possible states (is an edge or is not an edge). One of the main conjectures in the area, due to Simonovits and Sós [221] says that all graphs on 4 or more vertices must be "locally forcing" in the appropriate induced sense. They prove their conjecture holds for regular graphs and for various other families of graphs.

In the rest of this subsection we present our main results, which give a good understanding of locally forcing tournaments and show a surprisingly different behaviour compared to the one conjectured by Simonovits and Sós in the graph case. Our first result shows that in order for a tournament to be locally forcing it must be quite strongly quasirandom in several ways.

Theorem 3.2.3. *Any non-transitive, locally forcing h-vertex tournament H satisfies*

(i)
$$\sum_{v\in V(H)} \left(d_H^+(v) - d_H^-(v)\right)^2 \le h(h-1).$$

(ii) $|F_{\sigma,H}| = \frac{1}{2}\binom{h}{2} \pm h^{3/2}\sqrt{\log h}$ *for every ordering* σ *of* H.

(iii) For any disjoint subsets $U, W \subseteq H$ *we must have* $e(U,W) \le \frac{1}{2}|U||W| + 2h^{3/2}$

Given that Theorem 3.2.3 says that any locally forcing graph must be strongly quasirandom a natural guess for an example might be to take the random tournament, obtained by orienting each edge of a complete graph uniformly at random, independently between edges. Our next result shows that quite surprisingly with positive probability the random tournament fails condition (i) of Theorem 3.2.3.

Corollary 3.2.4. *The random tournament is not locally forcing with positive probability.*

This tells us that a positive proportion of all tournaments are in fact not locally forcing, in stark contrast to the induced graph problem, where Simonovits-Sós conjecture states that all graphs on 4 or more vertices should be locally forcing. In particular, one can easily find a regular tournament which fails conditions (ii) or (iii) whereas in the induced graph case any regular graph is known to be locally forcing.

This might suggest that there are essentially no non-transitive locally forcing tournaments, as in the global forcing case. We show that this is also not the case and find many non-transitive examples of locally forcing tournaments.

Theorem 3.2.5. *For any large enough h there exist non-transitive, h-vertex, locally forcing tournaments.*

Our examples of locally forcing tournaments come from the following random construction. We take a Steiner triple system on h vertices, that is a partition of the edge set of the complete graph on h vertices into edge disjoint triangles. Now we orient each of these triangles into one of two cycles of length three uniformly at random and independently between triangles. We show that this produces with high probability a locally forcing tournament. It is well known (see [157]) that Steiner triple systems exist provided $h = 1, 3 \pmod 6$ and our arguments allow us a lot of freedom to modify the above procedure in order to obtain examples for other h as well.

All our results are based on a characterisation of locally forcing tournaments in terms of certain graph polynomials. In order to motivate how these polynomials arise let us describe two natural candidates for tournaments with correct local counts of a tournament H, which are not quasirandom.

Definition. Given $\alpha \in [0,1]$, let $\mathcal{T}(n,\alpha)$ denote the distribution on the set of tournaments with vertex set $\{1,\ldots,n\}$ in which each pair $\{i,j\}$ with $1 \le i < j \le n$ appears independently as $\overrightarrow{ij}$ with probability α. We denote by $\mathcal{T}_{cliq}$ the collection of all distributions $\mathcal{T}(n,\alpha)$ as α varies.

Loosely speaking we say that a graph H is $\mathcal{T}_{cliq}$-forcing if no $\mathcal{T}(n,\alpha)$ shows that H is not locally forcing. In other words if there is no $\alpha \neq 1/2$ such that $\mathcal{T}(n,\alpha)$, which is clearly not quasirandom, has the same local counts of H as $\mathcal{T}(n,1/2)$ (which *is* quasirandom).

Definition. Given $\alpha \in [0,1]$, let $\mathcal{T}(n,n,\alpha)$ denote the distribution on the set of tournaments with vertex set $\{1,\ldots,2n\}$ in which both $\{1,\ldots,n\}$ and $\{n+1,\ldots,2n\}$ are oriented according to $\mathcal{T}(n,1/2)$ and each pair $\{i,j\}$ with $1 \le i \le n, n+1 \le j \le 2n$ appears as $\overrightarrow{ij}$ with probability α, with all choices made independently. We denote by $\mathcal{T}_{bip}$ the collection of all distributions $\mathcal{T}(n,n,\alpha)$ as α varies.

Loosely speaking a graph H is $\mathcal{T}_{bip}$-forcing if no $\mathcal{T}(n,n,\alpha)$ shows that H is not locally forcing. I.e. if there is no $\alpha \neq 1/2$ such that $\mathcal{T}(n,n,\alpha)$, which is not quasirandom, has the same local counts of H as $\mathcal{T}(n,n,1/2) = \mathcal{T}(2n,1/2)$.

Both properties $\mathcal{T}_{cliq}$-forcing and $\mathcal{T}_{bip}$-forcing can be expressed in terms of certain polynomial equations depending on the graph H. Our actual definitions of both properties go directly via these polynomial equations and we refer the reader to Section 3.2.3 for exact definitions.

Quite remarkably it turns out that if H is not locally forcing one must be able to find an α such that either $\mathcal{T}(n,\alpha)$ or $\mathcal{T}(n,n,\alpha)$ show H is not locally forcing. In particular, we show:

Theorem 3.2.6. *H is locally forcing if and only if it is $\mathcal{T}_{cliq}$-forcing and $\mathcal{T}_{bip}$-forcing.*

This result allows us to answer the question of whether a fixed graph H is locally forcing or not in terms of whether certain polynomial equations have a common real root in a bounded interval. This can be answered using a combination of Sturm's Theorem (see [27]) and the Euclidean Algorithm for computing the greatest common divisor of a sequence of polynomials. See the concluding remarks for more details.

Organisation of the section: The short proof of Proposition 3.2.2 is given in the next section. In Section 3.2.3 a counting polynomial associated with $\mathcal{T}_{cliq}$-forcing is introduced. In the same section we show that a locally forcing tournament must be $\mathcal{T}_{cliq}$-forcing and that this property imposes many restrictions on H, which gives us a proof of Theorem 3.2.3. Despite this in Section 3.2.4 we find many $\mathcal{T}_{cliq}$-forcing tournaments. In Sections 3.2.5 and 3.2.6 we introduce the polynomials related to the property of $\mathcal{T}_{bip}$-forcing, show that any nearly regular tournament must be $\mathcal{T}_{bip}$-forcing and show Theorem 3.2.6. Using this we prove that many of our examples of $\mathcal{T}_{cliq}$-forcing tournaments are in fact locally forcing.

Remark. All our logarithms in this section are natural, so in base e.

3.2.2 *Globally forcing tournaments*

In this section we will give the short proof of Proposition 3.2.2. We first show that asymptotically the number of transitive subtournaments of order h in a tournament of order n is minimised by the random tournament. This is precisely the content of Exercise 10.44(b) of [180]. We need a variation of this proof, which allows us to also show the fact that any asymptotic minimiser must be quasirandom. This implies that Tr_h is globally forcing for $h \geq 4$ showing the if part of the statement of Proposition 3.2.2.

Proof of Proposition 3.2.2. We will first prove that $N_T^*(h) := N_T^*(Tr_h)$ satisfies $N_T^*(h) \geq f_h(n) = \left(2^{-\binom{h}{2}} - o(1)\right)n^h$ for all n-vertex tournaments T where

$$f_h(n) = \begin{cases} \prod_{j=0}^{h-1} \left(\frac{n+1}{2^j} - 1\right) & \text{if } n \geq 2^{h-1} - 1 \\ 0 & \text{else.} \end{cases}$$

We will prove this by induction on h. For $h = 1$ and 2 the result is immediate. Furthermore,

$$N_T^*(3) = \sum_{v \in V(T)} \binom{d^+(v)}{2} \geq n \binom{\frac{n-1}{2}}{2} = n(n-1)(n-3)/8 = f_3(n). \quad (3.2)$$

Now assume $h \geq 4$ and $n \geq 2^{h-1} - 1$ as otherwise the result is trivial. For each edge $e = \overrightarrow{uv} \in E(T)$ let $N(e) = \{w \in V(T) : \overrightarrow{uw}, \overrightarrow{vw} \in E(T)\}$. Clearly

$$N_T^*(3) = \sum_{e \in E(T)} |N(e)|. \tag{3.3}$$

Letting $T[N(e)]$ be the subtournament of T with vertex set $N(e)$, by induction on h

$$N_T^*(h) = \sum_{e \in E(T)} N_{T[N(e)]}^*(h-2) \geq \sum_{e \in E(T)} f_{h-2}(|N(e)|).$$

It is easy to see that f_h is convex so Jensen's inequality together with (3.3) implies

$$\begin{aligned} N_T^*(h) \geq \binom{n}{2} f_{h-2}\left(\frac{N_T^*(3)}{\binom{n}{2}}\right) &\geq n \cdot \frac{n-1}{2} \cdot f_{h-2}\left(\frac{n-3}{4}\right) \\ &= n \cdot \frac{n-1}{2} \cdot \prod_{j=0}^{h-3}\left(\frac{n+1}{2^{j+2}} - 1\right) \\ &= \prod_{j=0}^{h-1}\left(\frac{n+1}{2^j} - 1\right). \end{aligned} \tag{3.4}$$

Where in the second inequality we used (3.2) and in the first inequality that $\frac{n-3}{4} \geq \frac{2^{h-1}-1-3}{4} = 2^{h-3} - 1$.

We now show that Tr_h is globally forcing for $h \geq 4$. To see this, suppose that $N_T^*(h) \leq (2^{-\binom{h}{2}} + o(1))n^h$. From (3.4) this gives $N_T^*(3) \leq (1 + o(1))n^3/8$. As $h \geq 4$, the function x^{h-2} is strictly convex and by the application of Jensen in (3.4) we find $|N(e)| = (1 \pm o(1))\frac{n}{4}$ for almost all $e \in E(T)$. However, by a result of Chung and Graham (see property P_4 and Theorem 1 in [57]) this property implies that T is quasirandom. Thus Tr_h is globally forcing for $h \geq 4$.

To complete the proof of the proposition, it remains to show that any non-transitive h-vertex tournament H with $h \geq 7$ is not globally forcing. To see this, let $V(H) = \{u_1, \ldots, u_h\}$. For each $n \in \mathbb{N}$ construct an n-vertex tournament T_1 as follows. Let $V(T_1) = V = \{v_1, \ldots, v_n\}$ denote a set of order n and let $V = \cup_{i \in [h]} V_i$ denote a partition with $|V_i| = \lfloor n/h \rfloor$ or $\lceil n/h \rceil$ for all $i \in [h]$. For each edge $\overrightarrow{u_i u_{i'}}$ of H, orient all edges from V_i to $V_{i'}$ in T_1. Orient the remaining edges of T_1 arbitrarily. As each map $\phi : V(H) \to V(T_1)$ with $\phi(u_i) \in V_i$ for all $i \in [h]$ is a labelled embedding, we see that

$$N_{T_1}^*(H) \geq \left(h^{-h} - o(1)\right)n^h. \tag{3.5}$$

Now, starting from T_1, construct a sequence of tournaments $T_1, \dots, T_n$ on vertex set $V(T_1)$, where for $i \geq 1$ each T_{i+1} is obtained from T_i by reorienting all edges to go from v_i to $\{v_{i+1}, \dots, v_n\}$. Using that $h^{-h} > 2^{-\binom{h}{2}}$ for $h \geq 7$, by (3.5) there is $\delta > 0$ such that $N^*_{T_1}(H) > (2^{-\binom{h}{2}} + \delta)n^h$ for large n. On the other hand $T_n = Tr_n$ and as H is not transitive we have $N^*_{T_n}(H) = 0$. Since $N^*_{T_i}(H) = N^*_{T_{i+1}}(H) \pm hn^{h-1}$, by an intermediate value property $N^*_{T_i}(H) = (2^{-\binom{h}{2}} + o(1))n^h$ for some $i \in [n]$. However, T_i is not quasirandom. Indeed as $\left(2^{-\binom{h}{2}} + o(1)\right)n^h = N^*_{T_i}(H) \geq (2^{-\binom{h}{2}} + \delta)n^h - ihn^{h-1}$ we have $i \geq \delta n/2h$. As $|d^+_{T_i}(v_j) - d^-_{T_i}(v_j)| = n - 2j + 1 \geq n/2$ for all $j \in [\min(i-1, n/4)]$, this contradicts the quasirandom property $\mathcal{P}_5$ from [57] which requires that

$$\sum_{v \in V(T_i)} |d^+_{T_i}(v) - d^-_{T_i}(v)| = o(n^2).$$

Thus T_i is not quasirandom and so H cannot be globally forcing. $\square$

Remark: Note that our argument shows that in order for H to be globally forcing $N^*_T(H)$ must be asymptotically maximised when T is the random tournament. In other words if we can find a sequence of n-vertex tournaments T_n such that $N^*_{T_n}(H) \geq (2^{-\binom{h}{2}} + \delta)n^h$ our argument above (switching edges incident to one vertex at a time) shows that H is not globally forcing. In fact it is enough to find a single such T_m for some $m \geq h$ since if we let T'_n be the blow up of T_m with parts of size $\frac{n}{m}$ then $N^*_{T'_n}(H) \geq (2^{-\binom{h}{2}} + \delta)m^h \left(\frac{n}{m}\right)^h = (2^{-\binom{h}{2}} + \delta)n^h$ so T'_n provides us with our sequence.

3.2.3 *Local forcing and the counting polynomial*

To study locally forcing tournaments we introduce the following polynomial.

Definition. The *counting polynomial* of an h-vertex tournament H is given by

$$p_H(x) := \mathbb{E}_{\sigma \sim S_h}\left(x^{|F_{\sigma,H}|}(1-x)^{|B_{\sigma,H}|}\right) = \frac{1}{h!}\sum_{\sigma \in S_h}\left(x^{|F_{\sigma,H}|}(1-x)^{|B_{\sigma,H}|}\right).$$

The following lemma collects a number of useful basic facts concerning $p_H(x)$.

Lemma 3.2.7. *Given an h-vertex tournament H, the following hold:*

(i) *For* $\alpha \in [0,1]$ *we have* $\mathbb{E}_{T\sim\mathcal{T}(n,\alpha)}\big(N^*_T(H)\big) = (n)_h \times p_H(\alpha)$.

(ii) $p_H(1/2) = 2^{-\binom{h}{2}}$ *and* $p_H(1) = 0$ *if* H *is not transitive.*

(iii) *For all* $\alpha \in [0,1/2]$ *we have* $p_H(1/2+\alpha) = p_H(1/2-\alpha)$.

Proof. Let $V(H) = \{u_1, \ldots, u_h\}$. A set $\{i_1, \ldots, i_h\} \subset [n] = V(T)$ with $i_1 < i_2 < \ldots < i_h$ forms a σ-copy of H if $\overrightarrow{i_j i_k} \in E(T)$ if and only if $\overrightarrow{u_{\sigma(j)}u_{\sigma(k)}} \in E(H)$. For $T \sim \mathcal{T}(n,\alpha)$, the probability that a fixed h-set forms a σ-copy of H is $\alpha^{|F_{\sigma,H}|}(1-\alpha)^{|B_{\sigma,H}|}$. Letting $N^*_T(H,\sigma)$ denote the number of σ-copies of H in T gives

$$\mathbb{E}_{T\sim\mathcal{T}(n,\alpha)}(N^*_T(H,\sigma)) = \binom{n}{h}\alpha^{|F_{\sigma,H}|}(1-\alpha)^{|B_{\sigma,H}|}.$$

As $N^*_T(H) = \sum_{\sigma\in S_h} N^*_T(H,\sigma)$ we have

$$\mathbb{E}_{T\sim\mathcal{T}(n,\alpha)}\big(N^*_T(H)\big) = \sum_{\sigma\in S_h} \mathbb{E}_{T\sim\mathcal{T}(n,\alpha)}\big(N^*_T(H,\sigma)\big) = (n)_h \times p_H(\alpha).$$

This gives *(i)*.

To see *(ii)*, note that $p_H(1/2) = 2^{-\binom{h}{2}}$ from the definition of $p_H(x)$ since $|F_{\sigma,H}| + |B_{\sigma,H}| = \binom{h}{2}$ for every $\sigma \in S_h$. Also note that if T is non-transitive then $|B_{\sigma,H}| > 0$ for all σ. So, each term of $p_H(x)$ is zero for $x = 1$ implying $p_H(1) = 0$.

Lastly, we show *(iii)*. For each $\sigma \in S_h$ let $\overline{\sigma}$ denote the reversal of σ, given by $\overline{\sigma}(i) = \sigma(h+1-i)$ for all $i \in [h]$. As this gives a bijection from S_h to itself and $F_{\sigma,H} = B_{\overline{\sigma},H}$, for all $\sigma \in S_h$, we have

$$\begin{aligned} p_H(x) &= \frac{1}{2h!}\sum_{\sigma\in S_h}\left(x^{|F_{\sigma,H}|}(1-x)^{|B_{\sigma,H}|} + x^{|F_{\overline{\sigma},H}|}(1-x)^{|B_{\overline{\sigma},H}|}\right) \\ &= \frac{1}{2h!}\sum_{\sigma\in S_h}\left(x^{|F_{\sigma,H}|}(1-x)^{|B_{\sigma,H}|} + x^{|B_{\sigma,H}|}(1-x)^{|F_{\sigma,H}|}\right) \end{aligned}$$

As $x^a(1-x)^b + x^b(1-x)^a$ is symmetric around $1/2$ for any $a, b \in [0,1/2]$ so is $p_H(x)$. $\square$

We are now ready to give the formal definition of $\mathcal{T}_{cliq}$-forcing which plays a central role in our arguments.

Definition. A tournament H is said to be $\mathcal{T}_{cliq}$*-forcing* if $p_H(x) \neq 2^{-\binom{h}{2}}$ for all $x \in [0,1] \setminus \{\frac{1}{2}\}$.

Note that if H is not locally forcing then there is a tournament T which is not quasirandom but has the same local counts of H as the random tournament $\mathcal{T}(n, 1/2)$. We will see that not being $\mathcal{T}_{cliq}$-forcing implies that there exists an $\alpha \neq 1/2$ such that $\mathcal{T}(n, \alpha)$ has the same local counts of H as the random tournament $\mathcal{T}(n, 1/2)$. Indeed, in the next theorem we prove that any locally forcing tournament must be $\mathcal{T}_{cliq}$-forcing, which will allow us to pass any restrictions imposed on the tournament by being $\mathcal{T}_{cliq}$-forcing to locally forcing tournaments and in particular prove Theorem 3.2.3.

Theorem 3.2.8. *Every locally forcing tournament is $\mathcal{T}_{cliq}$-forcing.*

Proof. Suppose that H is not $\mathcal{T}_{cliq}$-forcing. Then by definition $p_H(\alpha) = 2^{-\binom{h}{2}}$ for some $\alpha \in [0,1] \setminus \{1/2\}$. As $p_H(x)$ is symmetric about $1/2$ by Lemma 3.2.7 *(iii)* we can assume $\alpha > 1/2$. Now select $T \sim \mathcal{T}(n, \alpha)$. For a set $U \subset V(T)$ let X_U denote the random variable $X_U(T) = N^*_T(H; U)$. Noting that for $T \sim \mathcal{T}(n, \alpha)$ we have $T[U] \sim \mathcal{T}(|U|, \alpha)$, by Lemma 3.2.7 *(i)* we have

$$\mathbb{E}_{T \sim \mathcal{T}(n,\alpha)}(X_U) = (|U|)_h \times p_H(\alpha) = (|U|)_h \times 2^{-\binom{h}{2}}.$$

Since each edge belongs to at most $\binom{h}{2} n^{h-2}$ copies of H, X_U is sharply concentrated by Azuma's inequality (see Chapter 7 [16])

$$\mathbb{P}(|X_U - \mathbb{E}(X_U)| > \varepsilon n^h) \leq 2e^{\frac{-\varepsilon^2 n^{2h}}{2\binom{n}{2}\left(\binom{h}{2} n^{h-2}\right)^2}} = e^{-\Omega(n^2)}.$$

In particular, by a union bound with probability $1 - o(1)$ we have $X_U = 2^{-\binom{h}{2}} |U|^h + o(n^h)$ for every $U \subset V(T)$.

On the other hand the number of forward edges of T is distributed as $\mathrm{Bin}(\binom{n}{2}, \alpha)$ so by standard Chernoff type estimates T has $\alpha\binom{n}{2} \pm o(n^2)$ forward edges with probability $1 - o(1)$ (see Appendix A of [16]). Combined with the previous paragraph, we have shown that there is an n-vertex tournament T which satisfies $\mathcal{P}^*_2(H)$ but not $\mathcal{P}_1$. Thus H cannot be locally forcing. $\square$

For the rest of this section it will be more convenient to work with a rescaled and recentered counting polynomial $q_H(x) := 2^{\binom{h}{2}} \times p_H\left(\frac{1+x}{2}\right)$. We now show several restrictions that the $\mathcal{T}_{cliq}$-forcing condition imposes on a tournament. The first one is that a $\mathcal{T}_{cliq}$-forcing tournament should be almost regular.

Lemma 3.2.9. *If H is a non-transitive h-vertex $\mathcal{T}_{cliq}$-forcing tournament then*

$$\sum_{u \in V(H)} \left(d^+_H(u) - d^-_H(u)\right)^2 \leq h(h-1). \tag{3.6}$$

Proof. We start by determining the coefficient of x^2 in q_H. We claim it is equal to

$$\frac{1}{6} \times \Big(\sum_{u \in V(H)} \big(d_H^+(u) - d_H^-(u)\big)^2 - h(h-1)\Big).$$

Given $e \in E(H)$ let I_e denote the function $I_e : S_h \to \{-1, 1\}$ with $I_e(\sigma) = 1$ if $e \in F_{\sigma,H}$ and $I_e(\sigma) = -1$ if $e \in B_{\sigma,H}$. Then we have

$$q_H(x) = \mathbb{E}_{\sigma \sim S_h}\Big((1+x)^{|F_{\sigma,H}|}(1-x)^{|B_{\sigma,H}|}\Big) = \mathbb{E}_{\sigma \sim S_h}\Big(\prod_{e \in E(H)} (1 + I_e(\sigma)x)\Big). \tag{3.7}$$

This implies that the coefficient of x^2 in $q_H(x)$ is

$$\sum_{\{e,e'\} \in \binom{E(H)}{2}} \Big(\mathbb{E}_{\sigma \sim S_h} I_e(\sigma) I_{e'}(\sigma)\Big). \tag{3.8}$$

The contribution of each pair $\{e, e'\} \in \binom{E(H)}{2}$ to this sum is determined by how these edges meet. If $e \cap e' = \emptyset$ then the contribution to the sum is 0. If e and e' are both out-edges of a single vertex, say $e = \overrightarrow{u_1u_2}$ and $e' = \overrightarrow{u_1u_3}$, the contribution to the sum is $1/3$. This is also true if e and e' are both in-edges of a single vertex. Lastly, if e is an out-edge of a vertex v and e' is an in-edge then the contribution to the sum is $-1/3$. By counting the non-zero contributions of pairs $\{e, e'\}$ according to the vertex of $V(H)$ which they intersect in, by (3.8) the coefficient of x^2 in $q_H(x)$ is

$$\sum_{u \in V(H)} \Big(\frac{1}{3}\binom{d_H^+(u)}{2} + \frac{1}{3}\binom{d_H^-(u)}{2} - \frac{1}{3} d_H^+(u) d_H^-(u)\Big) =$$
$$\sum_{u \in V(H)} \frac{(d_H^+(u) - d_H^-(u))^2}{6} - \frac{h(h-1)}{6}.$$

as claimed.

Finally, if (3.6) fails, $q_H(\varepsilon) = 1 + a\varepsilon^2 + O(\varepsilon^4)$ with $a > 0$, where we used Lemma 3.2.7 *(ii)* to get $q_H(0) = 1$ and *(iii)* to get that q_H is even. This implies that $q_H(\varepsilon) > 1$ for sufficiently small $\varepsilon > 0$. As H is not transitive, Lemma 3.2.7 *(ii)* gives $q_H(1) = 0$, and so by the intermediate value theorem $q_H(\varepsilon) = 1$ for some $\varepsilon \in (0, 1)$. But this gives $p_H(\frac{1+\varepsilon}{2}) = 2^{-\binom{h}{2}}$ and so H is not $\mathcal{T}_{cliq}$-forcing. □

The following lemma shows that, in addition to being almost regular, any $\mathcal{T}_{cliq}$-forcing H must be strongly quasirandom.

Lemma 3.2.10. *If H is a non-transitive h-vertex $\mathcal{T}_{cliq}$-forcing tournament then for every ordering σ of H*

$$|F_{\sigma,H}| = \frac{1}{2}\binom{h}{2} \pm h^{3/2}\sqrt{\log h}.$$

Proof. Assume that there is an ordering σ such that $f = |F_{\sigma,H}| \geq \frac{1}{2}\binom{h}{2} + h^{3/2}\sqrt{\log h}$. If we let $b = |B_{\sigma,H}|$ since $f + b = \binom{h}{2}$ we get $b \leq \frac{1}{2}\binom{h}{2} - h^{3/2}\sqrt{\log h}$ and $f - b \geq 2h^{3/2}\sqrt{\log h}$.

We will show that a single term of $q_H(x)$, $(1+x)^f(1-x)^b > h!$ if we choose $x = \frac{3}{4}h^{-1/2}\sqrt{\log h}$. This shows that $q_H(x) > 1$ and we can complete the argument as in the previous lemma. To show the above inequality note that

$$\begin{aligned}(1+x)^f(1-x)^b &> e^{(x-x^2)f+(-x-x^2)b} \\ &> e^{x(f-b)-x^2h^2/2} \\ &> e^{h\log h} = h^h > h!\end{aligned}$$

where in the first inequality we used $1 + t > e^{t-t^2}$ for $|t| \leq 1/\sqrt{3}$. □

By a result of Spencer [223] it is known that for any tournament H there is an ordering of its vertices σ for which $|F_{\sigma,H}| = \frac{1}{2}\binom{h}{2} + \Omega(h^{3/2})$. So the above result is best possible up to the $\sqrt{\log h}$ factor. Fernandez de la Vega [103] showed that for the random tournament $H \sim \mathcal{T}(h, 1/2)$ w.h.p. there is no ordering with $|F_{\sigma,H}| > \frac{1}{2}\binom{h}{2} + 2h^{3/2}$. So it seems quite likely that the $\sqrt{\log h}$ term in the above lemma is not necessary. In fact for a different definition of quasirandomness we do obtain a result which is best possible up to a constant factor.

Lemma 3.2.11. *If H is a non-transitive h-vertex $\mathcal{T}_{cliq}$-forcing tournament then for any disjoint subsets $U, W \subseteq V(H)$ we must have $e(U, W) \leq \frac{1}{2}|U||W| + 2h^{3/2}$.*

Proof. Assume towards a contradiction that there are disjoint sets $U, W \subseteq V(H)$ for which $e(U, W) > \frac{1}{2}|U||W| + 2h^{3/2}$. Using these sets we will find many orderings σ such that $|F_{\sigma,H}| \geq \frac{1}{2}\binom{h}{2} + 2h^{3/2}$. Let $S = V(H) \setminus (U \cup W)$. We will always place all vertices of U before all vertices of W in the orderings σ. Also if $e(S, U \cup W) \geq e(U \cup W, S)$ we place all the vertices of S before all vertices of $U \cup W$ and behind otherwise. Within the sets U, W, S we take the orderings which have more forward edges than backward edges. Note that for any subset of H exactly one out of each of its orderings and its reverse has at least half of the edges going forwards. Therefore, we get at least $\frac{|U|!}{2} \cdot \frac{|W|!}{2} \cdot \frac{|S|!}{2} = \frac{h!}{8\binom{h}{|U|,|W|,|S|}} > h!/(8 \cdot 3^h)$ such orderings (where we used the trinomial expansion). Note that each such ordering σ of H has $|F_{\sigma,H}| \geq \frac{1}{2}\binom{h}{2} + 2h^{3/2}$ since inside each set and between each pair of sets there are at least half of them going forwards and there is a gain of at least $2h^{3/2}$ between U and W.

Assume now σ is an ordering of H with $f = |F_{\sigma,H}| \geq \frac{1}{2}\binom{h}{2} + 2h^{3/2}$. If we let $b = |B_{\sigma,H}|$ similarly as in the previous lemma we obtain that if we choose $x = h^{-1/2}$ then

$$\begin{aligned}(1+x)^f(1-x)^b &> e^{(x-x^2)f+(-x-x^2)b} \\ &> e^{x(f-b)-x^2h^2/2} \\ &> e^{7h/2} > 8 \cdot 3^h.\end{aligned}$$

From the previous two paragraphs we get $q_H(x) > 1$ and complete the proof as in Lemma 3.2.9. □

Spencer showed in [223] that for some constant $c > 0$ in any h-vertex tournament there are disjoint sets of vertices U, W such that $e(U, W) \geq \frac{1}{2}|U||W| + \Omega(h^{3/2})$. So the above lemma is best possible up to the constant factor.

Proof of Theorem 3.2.3. Combining Theorem 3.2.8 with Lemmas 3.2.9, 3.2.10 and 3.2.11 proves Theorem 3.2.3. □

Lemmas 3.2.10 and 3.2.11 show that in order for H to be $\mathcal{T}_{cliq}$-forcing it needs to be quasirandom in 2 different ways introduced by Chung and Graham in [57]. But it shows even more, namely that the error term should not only be qualitatively small (as in the definition of quasirandom properties in the introduction) but should in fact be close to their extremal value.

Since the previous two lemmas require H to be quasirandom and are both satisfied for the random tournament $\mathcal{T}(h, 1/2)$ w.h.p. a natural guess would be that it should be $\mathcal{T}_{cliq}$-forcing w.h.p.. This turns out to be false, due to Lemma 3.2.9.

Proof of Corollary 3.2.4. Let $H \sim \mathcal{T}(h, 1/2)$. We show that with positive probability $\sum_{v \in V(H)} \left(d_H^+(v) - d_H^-(v)\right)^2 > h(h-1)$ which will show the result by Lemma 3.2.9 through Theorem 3.2.8.

Let us set $V(H) = [h]$ and define $I_{ij} = 1$ if $\overrightarrow{ij} \in E(H)$ and $I_{ij} = -1$ if $\overrightarrow{ji} \in E(H)$, so $\mathbb{P}(I_{ij} = 1) = \mathbb{P}(I_{ij} = -1) = 1/2$. Note that $I_{ij} = -I_{ji}$ and that otherwise indicators I_{ij} are mutually independent. We also set $I_{ii} = 0$ for convenience. We have

$$\sum_{i\in[h]} (d_H^+(i) - d_H^-(i))^2 = \sum_{i\in[h]} \left(\sum_{j\in[h]} I_{ij} \right)^2 = h(h-1) + \sum_{i,j,k\in[h],j\neq k} I_{ij} I_{ik}.$$

Let $X = \sum_{i,j,k\in[h],j\neq k} I_{ij} I_{ik}$. Note that $\mathbb{E}(X) = \sum_{i,j,k\in[h],j\neq k} \mathbb{E}(I_{ij})\mathbb{E}(I_{ik}) = 0$ since each I_{ij} is independent of I_{ik} when $j \neq k$ and $\mathbb{E}(I_{ij}) = 0$. Our goal is to show $\mathbb{P}(X > 0) > c$ for some $c > 0$ independent of h. To do this we need to determine some higher moments of X. Note first that

$$\mathbb{E}(X^2) = \sum_{i,j,k\in[h],j\neq k} \quad \sum_{i',j',k'\in[h],j'\neq k'} \mathbb{E}(I_{ij} I_{ik} I_{i'j'} I_{i'k'}) = 4h \binom{h-1}{2}$$

where we used the fact that unless $i = i', \{j,k\} = \{j',k'\}$ at least one of the sets $\{i,j\}, \{i,k\}, \{i',j'\}$, $\{i',k'\}$ is distinct from the others, so its indicator is independent of the others and its contribution vanishes.

It is not hard to estimate $\mathbb{E}X^4$ directly but it requires some case analysis. Instead note that if we replace every occurrence of I_{ji} with $i < j$ with $-I_{ij}$ we get a degree 2 polynomial whose variables are independent indicators. Thus, by a special case of Bonami-Beckner's hypercontractive inequality (see [195] for a simple proof) we have that $\mathbb{E}(X^4) \leq (9\mathbb{E}(X^2))^2 = 81[\mathbb{E}(X^2)]^2$.

Now a simple lemma (see [19]) says that if a random variable Y has expectation 0, $\mathbb{E}Y^2 > 0$ and $\mathbb{E}Y^4/(\mathbb{E}Y^2)^2 \leq b$ then $\mathbb{P}(Y > 0) > 1/(2^{4/3}b)$. So in our case $b = 81$ and $\mathbb{P}(X > 0) \geq 1/205$. □

3.2.4 *Finding $\mathcal{T}_{cliq}$-forcing tournaments*

In the previous section we saw that in order for a tournament to be $\mathcal{T}_{cliq}$-forcing it needs to be strongly quasirandom. We have also seen that the random tournament is not regular enough to be $\mathcal{T}_{cliq}$-forcing w.h.p. It is natural to try to avoid this obstruction by trying next the random regular tournaments. However, the probability space of random regular tournaments is not at all easy to work with so instead we consider a modification of a different probability space on regular tournaments first introduced by Adler, Alon and Ross in [5] to study the maximum number of Hamilton paths in tournaments. Using it we show that there are many $\mathcal{T}_{cliq}$-forcing h-vertex tournaments for any large enough h.

To describe the construction, suppose that triangles $\Delta_1, \ldots, \Delta_L$ and edges $e_1, \ldots, e_F$ form a partition $\mathcal{P}$ of the edge set of the complete graph K_h. We now generate a random tournament on the vertex set $[h]$ as follows. Orient the edges of each Δ_i as a cyclic triangle, each orientation appearing with probability $1/2$ and then orient each e_j uniformly at random as well, so that all the triangles and edges are oriented independently. Write $\mathcal{D}_{\mathcal{P}}$ for the resulting distribution on the set of h-vertex tournaments.

Lemma 3.2.12. *There exists $c > 0$ such that for $h \geq h_0$ and $F \leq ch^2$, the random tournament $H \sim \mathcal{D}_{\mathcal{P}}$ is $\mathcal{T}_{cliq}$-forcing w.h.p..*

Proof. Note that $3L + F = \binom{h}{2}$ so by taking c small enough we may assume $L \geq 10F$, implying $h^2/8 \leq L \leq h^2/6$.

To prove the lemma, we again work with $q_H(x)$, showing that with positive probability $q_H(x) \neq 1$ for all $x \in [-1, 1] \setminus \{0\}$. Note that in any ordering σ of the vertices of H, the triangle corresponding to Δ_i either has two forward edges or two backward edges. Let $J_{i,\sigma}(H) = +1$ in the first case and $J_{i,\sigma}(H) = -1$ in the second case. Similarly let $J_{L+j,\sigma}(H) = \pm 1$ depending on whether edge e_j is forwards or backwards. We have

$$\begin{aligned}
q_H(x) &= \mathbb{E}_{\sigma \sim S_h}\Big((1+x)^{|F_{\sigma,H}|}(1-x)^{|B_{\sigma,H}|}\Big) \\
&= \mathbb{E}_{\sigma \sim S_h}\Big(\prod_{i \in [L]}(1+x)(1-x)(1+J_{i,\sigma}(H)x) \prod_{i \in [L+1, L+F]}(1+J_{i,\sigma}(H)x)\Big) \\
&= (1-x^2)^L \times \mathbb{E}_{\sigma \sim S_h}\Big(\prod_{i \in [L+F]}(1+J_{i,\sigma}(H)x)\Big) \\
&= (1-x^2)^L \times s_H(x).
\end{aligned}$$

As $q_H(x)$ is an even polynomial, so is the (random) polynomial $s_H(x)$. Let $s_H(x) = \sum_{\ell \in [(L+F)/2]} c_{2\ell} x^{2\ell}$, where $\{c_{2\ell}\}_{\ell \in [(L+F)/2]}$ are random variables depending on H.

To complete the lemma it will suffice to prove that with high probability $c_{2\ell} \leq \binom{L}{\ell}$ for all $\ell \in [(L+F)/2]$. Indeed, then

$$s_H(x) = \sum_{\ell \in [(L+F)/2]} c_{2\ell} x^{2\ell} \leq \sum_{\ell \in [(L+F)/2]} \binom{L}{\ell} x^{2\ell} \leq \sum_{\ell \in [L]} \binom{L}{\ell} x^{2\ell} = (1+x^2)^L,$$

which gives $q_H(x) = (1-x^2)^L \times s_H(x) \leq (1-x^2)^L \times (1+x^2)^L = (1-x^4)^L < 1$ for $x \in [-1, 1] \setminus \{0\}$, i.e. H is $\mathcal{T}_{cliq}$-forcing. To obtain the required bound on $c_{2\ell}$, note that

$$c_{2\ell} = \sum_{A \in \binom{[L+F]}{2\ell}} \mathbb{E}_{\sigma \sim S_h}\Big(\prod_{i \in A} J_{i,\sigma}(H)\Big). \tag{3.9}$$

Notice that for any $A \in \binom{[L+F]}{2\ell}$ if there is $i \in A$ such that the triangle or edge corresponding to i is vertex disjoint from all other objects indexed by A then $\mathbb{E}_{\sigma\sim S_h}\Big(\prod_{i\in A} J_{i,\sigma}(H)\Big) = 0$. Indeed, in this case we can cancel the contribution of a permutation σ to the expectation with the permutation $\widetilde{\sigma}$ obtained from σ by reversing the orientation of Δ_i or e_{i-L}. Let $\mathcal{A} \subseteq \binom{[L+F]}{2\ell}$ denote the sets not of this form, i.e. if $A \in \mathcal{A}$ then every object indexed by A shares a vertex with another object indexed by A. We have shown that

$$c_{2\ell} = \sum_{A\in\mathcal{A}} \mathbb{E}_{\sigma\sim S_h}\Big(\prod_{i\in A} J_{i,\sigma}(H)\Big). \tag{3.10}$$

Note further that any $A \in \mathcal{A}$ must have a subset $A' \subset A$ with $|A'| = \ell$ such that any object indexed by A intersects an object indexed by A'. We can choose these A' in $\binom{L+F}{\ell}$ many ways and they span at most 3ℓ vertices. This leaves us with at most $3\ell h$ options for each of the remaining objects. In particular, we have shown that

$$|\mathcal{A}| \leq \binom{L+F}{\ell} \cdot \binom{3\ell h}{\ell}. \tag{3.11}$$

Turning back to upper bounding $c_{2\ell}$ note that (3.10) gives us

$$\mathbb{E}_{H\sim\mathcal{D}_\mathcal{P}}\left(c_{2\ell}\right)^2 = \mathbb{E}_{\sigma,\sigma'\sim S_h}\mathbb{E}_{H\sim\mathcal{D}_\mathcal{P}}\Big(\textstyle\sum_{A,B\in\mathcal{A}} \displaystyle\prod_{i\in A} J_{i,\sigma}(H) \prod_{j\in B} J_{j,\sigma'}(H)\Big). \tag{3.12}$$

We claim that the contribution coming from distinct $A, B \in \mathcal{A}$ vanishes. Indeed let $A \neq B$ and let us fix an $i' \in A \setminus B$ (if $A \subset B$ we can use an analogous argument which takes $i' \in B \setminus A$ instead). For any fixed orderings σ, σ' we have that $J_{i',\sigma}(H)$ is independent from all $J_{i,\sigma}(H)$ for $i \in A \setminus \{i'\}$ and all $J_{j,\sigma'}(H)$ for $j \in B$. This implies

$$\begin{aligned}
&\mathbb{E}_{H\sim\mathcal{D}_\mathcal{P}}\big(\prod_{i\in A} J_{i,\sigma}(H) \prod_{j\in B} J_{j,\sigma'}(H)\big) = \\
&\mathbb{E}_{H\sim\mathcal{D}_\mathcal{P}}(J_{i',\sigma}(H)) \cdot \mathbb{E}_{H\sim\mathcal{D}_\mathcal{P}}\Big(\prod_{i\in A\setminus\{i'\}} J_{i,\sigma}(H) \prod_{j\in B} J_{j,\sigma'}(H)\Big) = 0,
\end{aligned}$$

since $\mathbb{E}_{H\sim\mathcal{D}_\mathcal{P}}(J_{i',\sigma}(H)) = 0$. Hence,

$$\mathbb{E}_{\sigma,\sigma'\sim S_h} \sum_{\substack{A,B\in\mathcal{A}\\ A\neq B}} \mathbb{E}_{H\sim\mathcal{D}_\mathcal{P}}\Big(\prod_{i\in A} J_{i,\sigma}(H) \prod_{j\in B} J_{j,\sigma'}(H)\Big) = 0.$$

Thus, $\mathbb{E}_{H\sim\mathcal{D}_\mathcal{P}}(c_{2\ell})^2 \leq |\mathcal{A}|$ by (3.12). By Markov's inequality the probability that $c_{2\ell} > \binom{L}{\ell}$ is at most $|\mathcal{A}|/\binom{L}{\ell}^2$. For $\ell \geq 2\log L$ we get

$$|\mathcal{A}|/\binom{L}{\ell}^2 < \binom{L+F}{2\ell}/\binom{L}{\ell}^2 < \prod_{i=0}^{\ell-1}\left(\frac{L+F-2i}{L-i}\right)^2\binom{2\ell}{\ell}^{-1}$$
$$\leq \left(\frac{L+F}{L}\right)^{2\ell}/\binom{2\ell}{\ell} \leq 2^{-\ell} \leq \frac{1}{L^2} \tag{3.13}$$

where we used $L \geq F$ and in the second to last inequality we took c small enough and $\ell \geq \ell_0$ since $h \geq h_0$. When $\ell < 2\log L$ using (3.11)

$$|\mathcal{A}|/\binom{L}{\ell}^2 \leq \binom{L+F}{\ell}\cdot\binom{3\ell h}{\ell}/\binom{L}{\ell}^2 \leq \left(\frac{3\ell h(L+F)}{(L-\ell)^2}\right)^\ell \leq \frac{16\log L}{\sqrt{L}}.$$

Here we used $L \geq 10F$ and $h^2/8 \leq L \leq h^2/6$ and assumed h is large enough for $\frac{16\log L}{\sqrt{L}} \leq 1$. Summing over all ℓ we have shown that $c_{2\ell} \leq \binom{L}{\ell}$ for all $\ell \in [L]$ with probability at least $1 - 2\log L \cdot \frac{16\log L}{\sqrt{L}} + \frac{L-2\log L}{L^2} = 1 - o(1)$. $\square$

Remark: It is well-known (see [157]) that for any $h \in \mathbb{N}$ with $h = 1$ or $3 \mod 6$ the complete graph K_h on vertex set $[h]$ admits a Steiner triple decomposition. That is, there is a partition $\mathcal{P}$ of K_h as above for which $F = 0$. Note that in this case $\mathcal{D}_\mathcal{P}$ is always a regular tournament. So there is an infinite family of $\mathcal{T}_{cliq}$-forcing regular tournaments.

3.2.5 $\mathcal{T}_{bip}$-forcing tournaments

In the previous section we have shown that there are many $\mathcal{T}_{cliq}$-forcing tournaments. Our original goal however was to study locally forcing tournaments and Theorem 3.2.8 only says that locally forcing tournaments are necessarily $\mathcal{T}_{cliq}$-forcing. While we believe this to be a sufficient condition, our argument requires a weak additional assumption, which we will call $\mathcal{T}_{bip}$-forcing. We show that being $\mathcal{T}_{bip}$-forcing and $\mathcal{T}_{cliq}$-forcing is equivalent to being locally forcing, which allows us to show that many $\mathcal{T}_{cliq}$-forcing tournaments found in the previous section are in fact locally forcing. We now give the formal definition of $\mathcal{T}_{bip}$-forcing.

Definition. We define the a-th order degree counting polynomial of an h-vertex tournament by:

$$p_{H,a}(x) := \binom{h}{a}^{-1} 2^{-\binom{a}{2}-\binom{h-a}{2}} \sum_{A\in\binom{V(H)}{a}} x^{e(A,V\setminus A)}(1-x)^{e(V\setminus A,A)}.$$

Definition. An h-vertex tournament is *$\mathcal{T}_{bip}$-forcing* if there is no $\alpha \in (1/2, 1]$ such that $p_{H,a}(\alpha) = 2^{-\binom{h}{2}}$ for all $1 \le a \le h-1$ simultaneously.

We have seen that if H fails to be $\mathcal{T}_{cliq}$-forcing then there is an $\alpha \neq 1/2$ such that $\mathcal{T}(n,\alpha)$ has the same local count of H as the random tournament $\mathcal{T}(n,1/2)$. In the following lemma we will see that H not being $\mathcal{T}_{bip}$-forcing means that there is an $\alpha \neq 1/2$ such that $\mathcal{T}(n,n,\alpha)$ does have the same count of H as $\mathcal{T}(n,n,1/2) = \mathcal{T}(2n,1/2)$. In fact H not being $\mathcal{T}_{bip}$-forcing is a seemingly much stronger restriction on H in the sense that it means that there is an $\alpha > 1/2$ such that for *all* $1 \le a \le h-1$, $\mathcal{T}(n,n,\alpha)$ has the correct count of H with a vertices embedded to the left and $h-a$ to the right side for all a simultaneously.

Lemma 3.2.13. *Every locally forcing tournament is $\mathcal{T}_{bip}$-forcing.*

Proof. Suppose towards a contradiction that there is an $\alpha \in (1/2, 1]$ such that $p_{H,a}(\alpha) = 2^{-\binom{h}{2}}$ for all $1 \le a \le h-1$. Now select $T \sim \mathcal{T}(n,n,\alpha)$ with bipartition (L,R). For a set $U \subset L, W \subset R$ let $X_{U,W}$ denote the random variable $X_{U,W}(T) = N_T^*(H; U \cup W)$. The probability that a fixed subset of size a of U and a fixed subset of size $h-a$ of W span an embedding of H with $A \subset V(H)$ being embedded to the left is $2^{-\binom{a}{2}-\binom{h-a}{2}}\alpha^{e(A,V\setminus A)}(1-\alpha)^{e(V\setminus A,A)}$. Summing over all possibilities we obtain

$$
\begin{aligned}
&\mathbb{E}_{T\sim\mathcal{T}(n,n,\alpha)}(X_{U,W}) \\
&\quad = \sum_{a=0}^{h} (|U|)_a(|W|)_{h-a} \cdot \sum_{A\in\binom{V(H)}{a}} 2^{-\binom{a}{2}-\binom{h-a}{2}}\alpha^{e(A,V\setminus A)}(1-\alpha)^{e(V\setminus A,A)} \\
&\quad = (|W|)_h 2^{-\binom{h}{2}} + \sum_{a=1}^{h-1} (|U|)_a(|W|)_{h-a} \cdot \binom{h}{a} p_{H,a}(\alpha) + (|U|)_h 2^{-\binom{h}{2}} \\
&\quad = \sum_{a=0}^{h} \binom{h}{a} (|U|)_a(|W|)_{h-a} \cdot 2^{-\binom{h}{2}} = (|U|+|W|)_h \cdot 2^{-\binom{h}{2}},
\end{aligned}
$$

Where in the last equality we used the identity $\sum_{a=0}^{h}\binom{b}{a}\binom{c}{h-a} = \binom{b+c}{h}$ holding for any b, c.

The rest of the proof proceeds in the same way as the proof of Theorem 3.2.8 except that the number of forwards edges of $T(n,n,\alpha)$ is $(1/4 + \alpha/2)\binom{2n}{2} \pm o(n^2)$ with probability $1 - o(1)$. $\square$

This lemma together with Theorem 3.2.8 shows the only if part of Theorem 3.2.6. We postpone the proof of the if part to the next section, showing first that there are many tournaments which are $\mathcal{T}_{bip}$-forcing, in addition to being $\mathcal{T}_{cliq}$-forcing.

As already mentioned, we believe being $\mathcal{T}_{bip}$-forcing is a very weak additional condition which might be already implied by being $\mathcal{T}_{cliq}$-forcing. It is even possible that all tournaments are $\mathcal{T}_{bip}$-forcing. We now give certain simple conditions that make a tournament $\mathcal{T}_{bip}$-forcing. We say that an h-vertex tournament H is *nearly regular* if $|d_H^+(v) - d_H^-(v)| < \sqrt{h}/2$ for all $v \in V(H)$.

Lemma 3.2.14. *Any nearly regular tournament is $\mathcal{T}_{bip}$-forcing.*

Proof. We are going to show that $p_{H,1}(x) + p_{H,h-1}(x) < 2 \cdot 2^{-\binom{h}{2}}$ for all $x \in (1/2, 1]$. It will be easier to work with the following polynomials

$$\begin{aligned} q_{H,a}(x) &= 2^{\binom{h}{2}} \cdot p_{H,a}\left(\frac{1+x}{2}\right) \\ &= \binom{h}{a}^{-1} \sum_{A \in \binom{V(H)}{a}} (1+x)^{e(A, V\setminus A)}(1-x)^{e(V\setminus A, A)}. \end{aligned}$$

So that

$$\begin{aligned} & q_{H,1}(x) + q_{H,h-1}(x) \\ &\quad = \frac{1}{h} \sum_{v \in V(H)} (1+x)^{d^+(v)}(1-x)^{d^-(v)} + (1+x)^{d^-(v)}(1-x)^{d^+(v)} \\ &\quad = \frac{1}{h} \sum_{v \in V(H)} (1-x^2)^{\min(d^+(v), d^-(v))} \left((1+x)^{|d^+(v)-d^-(v)|} + (1-x)^{|d^-(v)-d^+(v)|}\right) \\ &\quad \leq (1-x^2)^{h/6} \left((1+x)^{\lfloor \sqrt{h}/2 \rfloor} + (1-x)^{\lfloor \sqrt{h}/2 \rfloor}\right) \end{aligned}$$

Where we used that $\min(d^+(v), d^-(v)) \geq (h-1)/2 - \sqrt{h}/2 \geq h/6$, when $h \geq 5$ and that if $h < 5$ near regularity implies regularity so $(h-1)/2 \geq h/6$. Note that since $\binom{\sqrt{h}/2}{2\ell} \leq \frac{(\sqrt{h}/2)^{2\ell}}{(2\ell)!} \leq \left(\frac{h}{6\ell}\right)^{\ell} \leq \binom{h/6}{\ell}$ we have

$$\begin{aligned} (1+x)^{\lfloor \sqrt{h}/2 \rfloor} + (1-x)^{\lfloor \sqrt{h}/2 \rfloor} &\leq 2 \sum_{\ell=0}^{\sqrt{h}/4} \binom{\sqrt{h}/2}{2\ell} x^{2\ell} \leq 2 \sum_{\ell=0}^{\sqrt{h}/4} \binom{h/6}{\ell} x^{2\ell} \\ &\leq 2(1+x^2)^{h/6}. \end{aligned}$$

Combining the above two inequalities we obtain $q_{H,1}(x) + q_{H,h-1}(x) \leq 2(1-x^4)^{h/6} < 2$ as desired. □

Note that we know by Lemma 3.2.9 that any $\mathcal{T}_{cliq}$-forcing tournament must be almost regular but the restriction of the above lemma is slightly stronger. On the other hand our argument only uses a much weaker property than the one given to us by the definition of being $\mathcal{T}_{bip}$-forcing.

Remark: It is not hard to adapt our proof of Lemma 3.2.12 to show that the random tournament $\mathcal{D}_{\mathcal{P}}$ is $\mathcal{T}_{bip}$-forcing provided $\mathcal{P}$ consists of at most ch^2 edges (and the remaining objects are triangles) for sufficiently small c. So in some sense all our examples from the previous section are in fact locally forcing.

To conclude the section we deduce Theorem 3.2.5, assuming the if statement from Theorem 3.2.6 (which will be proven in the next section).

Proof of Theorem 3.2.5. Let $\mathcal{P}$ be a partition of K_h consisting of triangles and edges with every vertex incident to less than $\sqrt{h}/2$ of the edges in $\mathcal{P}$. It is easy to see that such a partition exists for any large enough h (for example by a result of Gustavsson [134]).

By Lemma 3.2.12 $H \sim \mathcal{D}_{\mathcal{P}}$ is $\mathcal{T}_{cliq}$-forcing w.h.p.. Furthermore, any $H \sim \mathcal{D}_{\mathcal{P}}$ is nearly regular and so by Lemma 3.2.14 it is $\mathcal{T}_{bip}$-forcing. Putting these together, Theorem 3.2.6 shows that the tournament $H \sim \mathcal{D}_{\mathcal{P}}$ is w.h.p. locally forcing. $\square$

3.2.6 *Proving local forcing*

Before proceeding to the proof of the if statement of Theorem 3.2.6 in subsection 3.2.6.2, we first recall some results on regularity lemmas for directed graphs in the next subsection.

3.2.6.1 *Regularity and counting lemmas for directed graphs*

A directed graph $D = (V, E)$ consists of a set V of vertices and a set of edges $E \subset V \times V$. Clearly any tournament is also a directed graph. Given disjoint sets $A, B \subset V$ we write $E(A, B)$ to denote the collection of edges $(a, b) \in E \cap (A \times B)$ and $e(A, B) = |E(A, B)|$. We will write $d(A, B)$ to denote the *density* of the pair (A, B), given by

$$d(A, B) = \frac{|E(A, B)|}{|A||B|}.$$

Note that if T is a tournament then $d(A,B) = 1 - d(B,A)$. Given disjoint sets $X, Y \subset V$ we say that (X,Y) is an ε-regular pair if all $X' \subset X$ and $Y' \subset Y$ with $|X'| \geq \varepsilon|X|$ and $|Y'| \geq \varepsilon|Y|$ satisfy

$$|d(X',Y') - d(X,Y)| \leq \varepsilon \qquad \text{and} \qquad |d(Y',X') - d(Y,X)| \leq \varepsilon.$$

A partition of $V = \{V_0, V_1, \ldots, V_K\}$ is said to be an ε-regular partition of D if:

(i) $|V_0| \leq \varepsilon|V|$,

(ii) $|V_1| = \cdots = |V_K|$,

(iii) all but at most $\varepsilon\binom{K}{2}$ of the pairs (V_i, V_j) with $1 \leq i < j \leq K$ are ε-regular.

The set V_0 is called the exceptional set and the sets $V_1, \ldots, V_K$ are clusters. This partition is said to refine a partition $V = U_1 \cup \cdots \cup U_L$ if for all $k \in [K]$ we have $V_k \subset U_l$ for some $l \in [L]$.

The directed regularity lemma of Alon and Shapira from [194], which extends Szemerédi's graph regularity lemma [203], states the following:

Theorem 3.2.15. *Given $m, L \in \mathbb{N}$ and $\varepsilon > 0$, there is $M = M(m, L, \varepsilon)$ with the following property. Given a directed graph $D = (V,E)$ with $|V| \geq M$ and a partition $V = U_1 \cup \cdots \cup U_L$ there is an ε-regular partition $\{V_0, V_1, \ldots, V_K\}$ with $m \leq K \leq M$, which refines $U_1 \cup \cdots \cup U_L$.*

Remark: While the theorem in [194] does not mention refinements, it follows easily from the proof.

A convenient structure associated with a regular partition $\{V_0, V_1, \ldots, V_K\}$ is the reduced graph $\mathcal{R}$, which has vertex set $\{V_1, \ldots, V_K\}$ with the property that V_iV_j is an edge of $\mathcal{R}$ if (V_i, V_j) is an ε-regular pair. Note that by definition $\mathcal{R}$ has at least $(1-\varepsilon)\binom{K}{2}$ edges.

We will also require the following counting lemma.

Lemma 3.2.16. *Let $T = (V,E)$ be a tournament and $V_1, \ldots, V_h$ be disjoint subsets of V. Suppose that for each $1 \leq i < j \leq h$ the pair (V_i, V_j) is ε-regular with density d_{ij}, with $d_{ji} = 1 - d_{ij}$. Then given an h-vertex tournament H with $V(H) = \{u_1, \ldots, u_h\}$, the number of copies of H in T, with u_i sent to V_i for all $i \in [h]$ is*

$$\Big(\prod_{\overrightarrow{u_iu_j} \in E(H)} d_{ij} \pm C_h\varepsilon \Big) \prod_{l \in [h]} |V_l|,$$

where C_h is a constant depending only on h.

Proof. By deleting directed edges of T which are not of the form $\overrightarrow{v_iv_j}$ with $v_i \in V_i$, $v_j \in V_j$ and $\overrightarrow{u_iu_j} \in E(H)$ and ignoring the directions of the remaining edges, this follows immediately from the usual graph counting lemma; see [215]. □

An embedding ϕ of a h-vertex tournament H into a tournament T is said to be partite with respect to the disjoint sets $U_1, \ldots, U_h \subset V(T)$ if each set U_i receives one vertex of the embedding. Let $\mathrm{Emb}_T(H; U_1, \ldots, U_h)$ denote the set of all such ϕ and let $N^*_T(H; U_1, \ldots, U_h) = |\mathrm{Emb}_T(H; U_1, \ldots, U_h)|$.

The following proposition gives a partite version of property $\mathcal{P}^*_2(H)$.

Proposition 3.2.17. *Let H be an h-vertex tournament. Suppose that for some constants ρ, C and an n-vertex tournament T we have $N^*_T(H; U) = \rho|U|^h \pm C$ for all $U \subset V(T)$. Then $N^*_T(H; U_1, \ldots, U_h) = h!\rho \prod_{i\in[h]} |U_i| \pm 2^h C$ for all disjoint sets $U_1, \ldots, U_h \subset V(T)$.*

Proof. By the inclusion-exclusion principle we have

$$N^*_T(H; U_1, \ldots, U_h) = \sum_{r=1}^{h} (-1)^{h-r} \Big(\sum_{I \in \binom{[h]}{r}} N^*_T(H; \bigcup_{i\in I} U_i) \Big).$$

Using $N^*_T(H; \bigcup_{i\in I} U_i) = \rho\big(\sum_{i\in I} |U_i|\big)^h \pm C$ for all $I \subset [h]$ gives

$$\begin{aligned} N^*_T(H; U_1, \ldots, U_h) &= \rho \times \left(\sum_{r=1}^{h} (-1)^{h-r} \Big(\sum_{I \in \binom{[h]}{r}} \big(\sum_{i\in I} |U_i| \big)^h \Big) \right) \pm 2^h C \\ &= \rho \times \big(h! \prod_{i\in[h]} |U_i| \big) \pm 2^h C. \end{aligned}$$

The final equality holds as the summed term in the penultimate equation can be viewed as counting, using inclusion-exclusion, the maps $g : [h] \to \bigcup_{i\in[h]} U_i$ sending each $i \in [h]$ to distinct U_j. Indeed there are $h! \prod_{i\in[h]} |U_i|$ such g which intersect each of the h sets U_i and for each $1 \le r \le h$ there are $\sum_{I\in\binom{[h]}{r}} \big(\sum_{i\in I} |U_i| \big)^h$ such maps which intersect at most r of the U_i's. □

3.2.6.2 *$\mathcal{T}_{cliq}$-forcing and $\mathcal{T}_{bip}$-forcing imply local forcing*

Before proceeding to the main result of this section, we note a simple consequence of Ramsey's theorem.

Lemma 3.2.18. *Given $\alpha > 0$ and $k, \ell \in \mathbb{N}$ there is $\gamma = \gamma(k, \ell, \alpha) > 0$ and $n_0 = n_0(k, \ell, \alpha) \in \mathbb{N}$ with the following property. Suppose that G is an n-vertex* *graph with $n \geq n_0$ and at least $(1-\gamma)\binom{n}{2}$ edges. Then in any k-colouring of $E(G)$ there are vertex disjoint sets $U_1, \dots, U_M \subset V(G)$ so that:*

(i) $G[U_m]$ is a monochromatic clique of order ℓ for all $m \in [M]$;

(ii) $|V(G) \setminus \big(\cup_{m\in[M]} U_m\big)| \leq \alpha n$.

Proof. Set $R := R_k(\ell)$, the k-colour Ramsey number of an ℓ-vertex clique. Also set $\gamma = \alpha^2/2R$ and $n_0 = \lceil 2R/\alpha \rceil$. Suppose we are given a k-colouring of $E(G)$ as in the statement. To prove the lemma it suffices to show that every $W \subset V(G)$ with $|W| \geq \alpha n$ contains a set $U \subset W$ with $|U| = \ell$ so that $G[U]$ is a monochromatic clique. Indeed, using this property we can then greedily find sets $U_1, \dots, U_M$ as in the lemma.

To see that this holds, first note that since $|W| \geq \alpha n \geq 2R$ and $|W|^2 \geq \alpha^2 n^2 = 2\gamma R n^2$, we have

$$\begin{aligned} e(G[W]) &\geq \binom{|W|}{2} - \gamma\binom{n}{2} > \left(1 - \frac{1}{R}\right)\frac{|W|^2}{2} + \left(\frac{|W|^2}{2R} - \frac{|W|}{2} - \frac{\gamma n^2}{2}\right) \\ &\geq \left(1 - \frac{1}{R}\right)\frac{|W|^2}{2}. \end{aligned}$$

By Turán's theorem there is $W' \subset W$ such that $|W'| = R$ and $G[W']$ is complete. As $G[W]$ is k-coloured, from Ramsey's theorem and the definition of R, there is $U \subset W'$ with $|U| = \ell$ and $G[U]$ is monochromatic. This completes the proof. $\square$

We now turn to the proof of the if part of Theorem 3.2.6.

Theorem 3.2.19. *Any $\mathcal{T}_{cliq}$-forcing and $\mathcal{T}_{bip}$-forcing tournament is locally forcing.*

Proof. Let H be a $\mathcal{T}_{cliq}$-forcing and $\mathcal{T}_{bip}$-forcing tournament with $|H| = h$. We are required to show that given $\theta > 0$ there is $\delta > 0$ and n_0 such that the following holds. Suppose that T is an n-vertex tournament with $n \geq n_0$ which satisfies

$$N^*_T(H;U) = 2^{-\binom{h}{2}}|U|^h \pm \delta n^h \tag{3.14}$$

for all $U \subset V(T)$. Then any ordering of $V(T)$ has at most $\frac{1}{2}\binom{n}{2} + \theta n^2$ forward edges.

Let us now give a sketch of the proof before delving into the details. Let $v_1, \dots, v_n$ be an ordering of $V(T)$. We begin by splitting the vertices into consecutive sets U_ℓ of almost equal size. Then we apply the regularity lemma (Theorem 3.2.15) to refine this partition. Taking the reduced graph we define $\mathcal{R}_\ell$ to be its subgraph consisting of clusters contained in U_ℓ. We colour all edges of $\mathcal{R}_\ell$ which join clusters with density between them belonging to the same small interval in the colour indexed by this interval. If we split $[0, 1]$ into finitely many such intervals we obtain a colouring of $\mathcal{R}_\ell$ to which we can apply Lemma 3.2.18 to group most of the clusters inside each $\mathcal{R}_\ell$ into monochromatic cliques. We now show that all edges inside these cliques must have density close to $1/2$. To see this assume the opposite, so that there is a clique C with all edges having density bounded away from $1/2$. Now using Lemma 3.2.16 and H being $\mathcal{T}_{cliq}$-forcing we conclude that there are too few copies of H between the clusters of C, compared to what is guaranteed by Proposition 3.2.17, which holds by (3.14).

We then proceed to upper bound the number of forwards edges of T. The main contribution comes from edges between ε-regular pairs of clusters between different U_ℓ's. To bound this number for a pair of cliques belonging to different U_ℓ's and a fixed $d > 1/2 + \theta$ we build an auxiliary bipartite graph with clusters of the two cliques making the sides of the bipartition and making an edge for any pair of ε-regular clusters which have density roughly d in the forwards direction. We show that this auxiliary graph can not contain $K_{a,h-a}$ for some a, as otherwise by using a similar argument as before, we get h clusters between which we have a wrong count of the copies of H using the fact H is $\mathcal{T}_{bip}$-forcing. By grouping densities and applying the above reasoning for each group we show that there are few forwards edges between the two cliques. Trivially bounding the remaining contributions we show that there are fewer forwards edges than required, completing the proof.

Before beginning we will fix a number of parameters to be used in the proof. Let $\xi = \xi(H, \theta)$ be the minimum of the continuous function $f(x) = \max_{a \in [h-1]}(|2^{\binom{h}{2}} p_{H,a}(x) - 1|)$ on the interval $[1/2 + \theta, 1]$. Note that since H is $\mathcal{T}_{bip}$-forcing we have $f(x) > 0$ for each x in this range so since f is continuous we get $\xi > 0$. Set $\eta = \xi 2^{-\binom{h}{2}-2} h^{-2}$. Since H is $\mathcal{T}_{cliq}$-forcing and $p_H(x)$ is continuous there is $\zeta \in (0, \eta)$ with the property that if $x \in [0, 1]$ and $p_H(x) \geq 2^{-\binom{h}{2}} - \zeta$ then $x = (1 \pm \eta)/2$. Take $\alpha = \theta/64$, $m = \max(\lceil 4h^2/\zeta \rceil, \lceil 2/\eta \rceil)$, $L = \lceil 16\theta^{-1} \rceil$ and $N_1 = \lceil (2(\theta\eta)^{-1}h)^h \rceil$. Also set $m_{min} = 4Ln_0(m, N_1, \alpha)$ and $\gamma = \gamma(m, N_1, \alpha)$ as in Lemma 3.2.18. With C_h as in Lemma 3.2.16, take $\varepsilon > 0$ so that

$$\varepsilon = \min\Big(\frac{1}{4L}, \frac{\gamma}{8L^2}, \frac{\zeta}{4C_h}, \frac{\theta^3}{32}\Big).$$

Lastly, set $n_0 = M = \max(M(m_{min}, L, \varepsilon), 8L)$ as in Theorem 3.2.15 and $\delta = \zeta(4M)^{-h}/2$.

To begin the proof set $U_\ell = \{v_i \in V(T) : i \in ((\ell-1)n/L, \ell n/L]\}$ for all $\ell \in [L]$. Note that $|U_\ell| \geq n/L - 2$. Provided $n \geq n_0$, we may apply the Theorem 3.2.15 to T to obtain an ε-regular partition $\{V_k\}_{k\in[K]} \cup \{V_0\}$ refining $\{U_\ell\}_{\ell\in[L]}$, with $m_{min} \leq K \leq M$. Let $\mathcal{R}$ denote the reduced graph of this partition and $\mathcal{R}_\ell$ denote the subgraph of $\mathcal{R}$ consisting of the clusters contained in U_ℓ. Setting $W_\ell = |\mathcal{R}_\ell|$, we have $W_\ell \geq \frac{K}{2L}$ for all $\ell \in [L]$. Indeed, since $|V_k| = |V_{k'}|$ for all $k, k' \in [K]$ we have

$$n \geq \sum_{k\in[K]} |V_k| = \frac{K}{W_\ell} \sum_{V_k \in V(\mathcal{R}_\ell)} |V_k| \geq \frac{K}{W_\ell}\Big(|U_\ell| - |V_0|\Big) \geq \frac{K}{W_\ell} \times \frac{n}{2L},$$

using $|U_\ell| - |V_0| \geq (n/L - 2) - \varepsilon n \geq \frac{n}{2L}$. Rearranging, we find $W_\ell \geq \frac{K}{2L}$ for all $\ell \in [L]$.

Claim: Each $\mathcal{R}_\ell$ contains a collection of vertex disjoint cliques $\mathcal{C}_\ell$ with the following properties:

(i) Each clique $C \in \mathcal{C}_\ell$ has order N_1,

(ii) $|\mathcal{R}_\ell \setminus (\cup_{C\in\mathcal{C}_\ell} C)| \leq \alpha W_\ell$,

(iii) $d(V_k, V_{k'}) = 1/2 \pm \eta$ for each edge $V_k V_{k'}$ in a clique $C \in \mathcal{C}_\ell$.

To prove the claim, colour the edges of $\mathcal{R}_\ell$ with m colours, where each pair $V_k V_{k'}$ with $k < k'$ gets color $j \in [m-1]$ if $d(V_k, V_{k'}) \in j/m \pm 1/m$ (ties broken arbitrarily). The graph $\mathcal{R}_\ell$ contains at least $\binom{W_\ell}{2} - \varepsilon\binom{K}{2} > (1-\gamma)\binom{W_\ell}{2}$ edges, since $W_\ell \geq \frac{K}{2L}$, $\frac{\gamma}{8L^2} \geq \varepsilon$ and $K \geq m_{min} \geq 4L$. Therefore, from Lemma 3.2.18, since $|\mathcal{R}_\ell| \geq \frac{K}{2L} \geq \frac{m_{min}}{2L} \geq n_0(m, N_1, \alpha)$, the graph $\mathcal{R}_\ell$ contains a collection $\mathcal{C}_\ell$ of vertex disjoint monochromatic cliques which satisfy parts (i) and (ii) from the claim.

It remains to show that part (iii) holds. Let $C \in \mathcal{C}_\ell$ be monochromatic with colour j and $V_{k_1}, \ldots, V_{k_h} \in C$ with $k_1 < \ldots < k_h$. As each pair $(V_k, V_{k'})$ in C is ε-regular with $d(V_{k_i}, V_{k_{i'}}) = (j \pm 1)/m$, by Lemma 3.2.16 we have

$$\begin{aligned}
&N_T^*(H; V_{k_1}, V_{k_2}, \ldots, V_{k_h})\\
&\quad= \sum_{\sigma\in S_h} \left(\Big(\frac{j\pm 1}{m}\Big)^{|F_{\sigma,H}|}\Big(\frac{m-j\pm 1}{m}\Big)^{|B_{\sigma,H}|} \pm C_h\varepsilon\right) \prod_{i\in[h]} |V_{k_i}|\\
&\quad= h!\big(p_H(j/m) \pm (C_h\varepsilon + h^2 m^{-1})\big) \prod_{i\in[h]} |V_{k_i}|\\
&\quad= h!\big(p_H(j/m) \pm \zeta/2\big) \prod_{i\in[h]} |V_{k_i}|.
\end{aligned}$$

In the second equality here we used repeatedly the fact that for any $x, y, t \in [0,1]$ such that $x \pm t \in [0,1]$ we have $(x \pm t)(y \pm t) = (x \pm t)y \pm t = xy \pm 2t$. On the other hand, from (3.14) and Proposition 3.2.17 we also have

$$N_T^*(H; V_{k_1}, V_{k_2}, \dots, V_{k_h}) = h! 2^{-\binom{h}{2}} \prod_{i \in [h]} |V_{k_i}| \pm 2^h \delta n^h$$
$$= h! \left(2^{-\binom{h}{2}} \pm \zeta/2\right) \prod_{i \in [h]} |V_{k_i}|. \tag{3.15}$$

Here we have used that all clusters V_k satisfy $|V_k| \geq (n - |V_0|)/K \geq n/2M$ and $\delta(4M)^h = \zeta/2$. Combined, these bounds give $p_H(j/m) = 2^{-\binom{h}{2}} \pm \zeta$. By our choice of ζ this forces $j/m = (1 \pm \eta)/2$ and so $d(V_k, V_{k'}) = (1 \pm \eta)/2 \pm m^{-1} = 1/2 \pm \eta$ for any pair $V_k V_{k'}$ contained in a clique $C \in \mathcal{C}_\ell$ giving (iii). This completes the proof of the claim.

We can now proceed to prove an upper bound on the number of forward edges of T. To do this, first fix $1 \leq \ell < \ell' \leq L$ and $C \in \mathcal{C}_\ell$ and $C' \in \mathcal{C}_{\ell'}$. We will prove that

$$\sum_{\substack{V_k \in C, V_{k'} \in C': \\ (V_k, V_{k'})\ \varepsilon-\text{reg}}} |E(V_k, V_{k'})| \leq \left(\frac{1}{2} + \frac{3\theta}{2}\right) \Big| \bigcup_{V_k \in C} V_k \Big| \Big| \bigcup_{V_{k'} \in C'} V_{k'} \Big|. \tag{3.16}$$

To see this fix some $d > 1/2 + \theta$ and consider the auxilliary bipartite graph G with vertex set on one side being C and on the other side C'. We put an edge between clusters $V_k \in C$ and $V_{k'} \in C'$ if the pair $(V_k, V_{k'})$ is ε-regular and has density $d(V_k, V_{k'}) = d \pm \eta$. Let a be such that $|2^{\binom{h}{2}} p_{H,a}(d) - 1| \geq \xi$, which exists since $d > 1/2 + \theta$. Our goal is to show that G is $K_{a,h-a}$-free, which by Kővári-Sós-Turán theorem is going to tell us that G is sparse, allowing us to bound the number of forwards edges of T between C and C'.

To see this assume towards a contradiction that G contains a $K_{a,h-a}$. Let $V_{k_1}, \dots, V_{k_a}$ make one side and $V_{k_{a+1}}, \dots, V_{k_h}$ the other of this $K_{a,h-a}$. By part (iii) of the claim, all pairs $(V_{k_i}, V_{k_{i'}})$ are ε-regular with $d(V_{k_i}, V_{k_{i'}}) = 1/2 \pm \eta$ for distinct $i, i' \in [a]$ or distinct $i, i' \in [a+1, h]$. Any $\phi \in \mathrm{Emb}_T(H; V_{k_1}, \dots, V_{k_h})$ embeds $\binom{a}{2} + \binom{h-a}{2}$ edges of H into pairs with density $1/2 \pm \eta$. If $A := \phi^{-1}(\{V_{k_1}, \dots, V_{k_a}\})$ then $e(A, V \setminus A)$ edges of H get embedded into pairs with density $d \pm \eta$, and $e(V \setminus A, A)$ edges into pairs with density $1 - d \pm \eta$. So Lemma 3.2.16 gives

$$\begin{aligned}
&\frac{N_T^*(H; V_{k_1}, \dots, V_{k_h})}{\prod_{i \in [h]} |V_{k_i}|} \\
&\quad = a!(h-a)! \sum_{A \in \binom{V(H)}{a}} \left(\left(\frac{1}{2} \pm \eta\right)^{\binom{a}{2} + \binom{h-a}{2}} (d \pm \eta)^{e(A, V \setminus A)} (1 - d \pm \eta)^{e(V \setminus A, A)} \pm C_h \varepsilon \right) \\
&\quad = h! \binom{h}{a}^{-1} \sum_{A \in \binom{V(H)}{a}} \left(2^{-\binom{a}{2} - \binom{h-a}{2}} d^{e(A, V \setminus A)} (1-d)^{e(V \setminus A, A)} \pm (h^2 \eta + C_h \varepsilon) \right) \\
&\quad = h!(p_{H,a}(d) \pm (h^2 \eta + C_h \varepsilon)). \qquad (3.17)
\end{aligned}$$

As $(h^2\eta + C_h\varepsilon) + \zeta \leq h^2\eta + 2\zeta \leq 3h^2\eta < \xi 2^{-\binom{h}{2}}$ this gives

$$\left| \frac{N_T^*(H; V_{k_1}, \dots, V_{k_h})}{\prod_{i \in [h]} |V_{k_i}| h!} - 2^{-\binom{h}{2}} \right| \geq \zeta,$$

which contradicts (3.15). Thus, there is no $K_{a,h-a}$ in G. The Kővári-Sós-Turán theorem (see [81]) now tells us that G has at most $hN_1^{2-1/h}$ edges (recall that $|C| = |C'| = N_1$.) Since d was arbitrary (provided it is bigger than $1/2 + \theta$) by splitting the interval $[1/2 + \theta, 1]$ into at most η^{-1} subintervals of width at most 2η and building a graph as above for each of these subintervals with d being equal to the center of the interval we obtain that there are at most $\eta^{-1}hN_1^{2-1/h}$ pairs $V_k \in C, V_{k'} \in C'$ such that $(V_k, V_{k'})$ is ε-regular and $d(V_k, V_{k'}) > 1/2 + \theta$. As $|C| = N_1 \geq (2(\eta\theta)^{-1}h)^h$ this gives

$$\begin{aligned}
\sum_{\substack{V_k \in C, V_{k'} \in C': \\ (V_k, V_{k'})\ \varepsilon-reg}} |E(V_k, V_{k'})| &\leq \left(\frac{1}{2} + \theta + \eta^{-1}hN_1^{-1/h}\right) \left| \bigcup_{V_k \in C} V_k \right| \left| \bigcup_{V_{k'} \in C'} V_{k'} \right| \\
&\leq \left(\frac{1}{2} + \frac{3\theta}{2}\right) \left| \bigcup_{V_k \in C} V_k \right| \left| \bigcup_{V_{k'} \in C'} V_{k'} \right|
\end{aligned}$$

i.e. (3.16) holds.

We can now complete the proof, upper bounding the number of forward edges of T. The ε-regular pairs $(V_k, V_{k'})$ with $V_k \in C \in \mathcal{C}_\ell$ and $V_{k'} \in C' \in \mathcal{C}_{\ell'}$ for some $\ell < \ell'$ contribute at most $\big(1/2 + 3\theta/2\big)\binom{n}{2}$ forward edges to T by (3.16). The remaining forward edges either (a) lie entirely in some set U_ℓ, (b) contain a vertex from cluster $V_k \notin \bigcup_\ell \bigcup_{C \in \mathcal{C}_\ell} C$, (c) contain a vertex in V_0, or

(d) lie between pairs $(V_k, V_{k'})$ which are not ε-regular. The number of such edges is at most

$$\sum_{\ell\in[L]} \binom{|U_\ell|}{2} + \sum_{\ell\in[L]} |\mathcal{R}_\ell \setminus (\cup_{C\in\mathcal{C}_\ell} C)| \binom{n^2}{K} + |V_0|n + \varepsilon\binom{K}{2}\left(\frac{n}{K}\right)^2 \leq$$
$$\left(\frac{2}{L} + 8\alpha + 3\varepsilon + \varepsilon\right)\binom{n}{2} \leq \frac{\theta}{2}\binom{n}{2}.$$

Here we have used that $|\mathcal{R}_\ell \setminus (\cup_{C\in\mathcal{C}_\ell} C)| \leq \alpha W_\ell \leq 2\alpha K/L$ by part (ii) of the claim, that $L \geq 16/\theta$, $64\alpha = \theta$ and $16\varepsilon \leq \theta$. Combined, these estimates show that T has at most $\frac{1}{2}\binom{n}{2} + \theta n^2$ forward edges, as required. □

3.2.7 *Concluding remarks*

We have shown that a large tournament H is globally forcing if and only if H is transitive. Our main focus was on the stronger property of being locally forcing. We proved that while many tournaments do not satisfy this property (Lemma 3.2.7) it does hold for many tournaments which we draw from a certain random distribution on nearly regular tournaments. The most natural model of random regular tournaments is to take a uniform distribution over all regular tournaments. We believe that in fact this also gives w.h.p. a $\mathcal{T}_{cliq}$-forcing and hence (by Lemma 3.2.14 and Theorem 3.2.6) a locally forcing tournament.

Another result of this section shows that a tournament is locally forcing if and only if it is $\mathcal{T}_{cliq}$-forcing and $\mathcal{T}_{bip}$-forcing. This in some sense says that in order to check whether a tournament H is locally forcing one only needs to check whether models $\mathcal{T}(n,\alpha)$ or $\mathcal{T}(n,n,\alpha)$ can have the same local counts of H as $\mathcal{T}(n,1/2)$ for some $\alpha \neq 1/2$.

We actually believe that the $\mathcal{T}_{bip}$-forcing assumption may be dropped entirely in Theorem 3.2.6 either because it is implied by $\mathcal{T}_{cliq}$-forcing or because it is satisfied by every tournament. In other words it would be interesting to determine if every $\mathcal{T}_{cliq}$-forcing tournament is locally forcing. We reduced this question to the following problem about degree counting polynomials. Does there exist an h-vertex tournament H such that the rescaled and recentered a-th order degree counting polynomials defined as

$$q_{H,a}(x) := \binom{h}{a}^{-1} \sum_{A\in\binom{V(H)}{a}} (1+x)^{e(A,V\setminus A)}(1-x)^{e(V\setminus A,A)} - 1$$

have a common root in $(0,1]$, for all $1 \leq a \leq h-1$. We note that in a certain sense this is the correct tournament analogue of the Simonovits-Sós conjecture from the induced graph case, discussed in the introduction.

Our arguments rely on the regularity lemma. It is possible that one can find a nice class of tournaments for which one can show the locally forcing property directly, avoiding the use of the regularity lemma. For example, the tournament Tr_3 is not globally forcing but is locally forcing and this is not hard to see directly. Indeed, any tournament T has at most

$$\frac{1}{2}\sum_{v\in T}\left(\binom{d^+(v)}{2}+\binom{d^-(v)}{2}\right)=$$
$$\frac{1}{4}\sum_{v\in T}\frac{1}{2}\left((d^+(v)+d^-(v))^2+(d^+(v)-d^-(v))^2\right)-(n-1)=$$
$$\frac{1}{8}n^3(1+o(1))+\frac{1}{8}\sum_{v\in T}(d^+(v)-d^-(v))^2$$

copies of Tr_3. So in order for the local counts to match that of the random tournament, by Cauchy-Schwarz, we must have $n^{-1}\big(\sum_{v\in U}|d^+_{T[U]}(v)-d^-_{T[U]}(v)|\big)^2 \leq \sum_{v\in U}(d^+_{T[U]}(v)-d^-_{T[U]}(v))^2 = o(n^3)$ for any $U \subseteq V(T)$. This gives condition $\mathcal{P}_5$ from Chung and Graham [57] and implies T must be quasirandom. Thus Tr_3 is locally forcing. Since any three vertices of a tournament either induce Tr_3 or C_3, if any subset of $V(T)$ has the correct count of Tr_3 then it has the correct count of C_3 as well. Thus C_3 is also locally forcing.

We note that deciding whether a fixed tournament is $\mathcal{T}_{cliq}$-forcing is a matter of counting how many zeros of the counting polynomial (minus a constant) one can find in the interval $[1/2,1]$. Since the counting polynomial is of degree at most $\binom{h}{2}$ this can be done by Sturm's algorithm in polynomial time. Deciding whether a fixed tournament is $\mathcal{T}_{bip}$-forcing can be done in a similar fashion, we first find the greatest common divisor of our degree counting polynomials and then find the number of roots of this greatest common divisor in $[1/2,1]$. While all of this can be done in polynomial time it is not clear how to compute the coefficients of our polynomials in polynomial time since they are defined in terms of vertex orderings of which there are $h!$ or subsets, in which case there are potentially as many as $\binom{h}{h/2}$. It could be interesting to determine if these terms could be calculated in polynomial time, as this would give a polynomial time algorithm for deciding whether a given tournament is locally forcing or not.

Acknowledgements. We would like to thank Jan Volec for checking many small cases using a computer and Igor Balla for drawing our attention to [19].

4

ERDŐS-SZEKERES THEOREM FOR MULTIDIMENSIONAL ARRAYS

4.1 INTRODUCTION

A classical paper of Erdős and Szekeres [97] from 1935 is one of the starting points of a very rich discipline within combinatorics: Ramsey theory. A main result of the paper, which has become known as the Erdős-Szekeres theorem, says that any sequence of $(n-1)^2+1$ distinct real numbers contains either an increasing or decreasing subsequence of length n, and this is tight. Among simple results in combinatorics, only few can compete with this one in terms of beauty and utility. See, for example, Steele [226] for a collection of six proofs and some applications.

A very natural question which arises is how does one generalise the Erdős-Szekeres theorem to higher dimensions? The main concept which does not have an obvious generalisation is that of the monotonicity of a subsequence. Several candidates have been proposed [47, 48, 152, 171, 179, 187, 219, 231] but perhaps the most natural one was introduced more than 25 years ago by Fishburn and Graham [105]. A multidimensional array is said to be monotone if for each dimension all the 1-dimensional subarrays along the direction of this dimension are increasing or are all decreasing. To be more formal, a d-dimensional array f is an injective function from $A_1 \times \ldots \times A_d$ to $\mathbb{R}$ where $A_1, \ldots, A_d$ are non-empty subsets of $\mathbb{Z}$; we say f has size $|A_1| \times \ldots \times |A_d|$.

Definition (Monotone array)**.** A d-dimensional array $f\colon A_1 \times \ldots \times A_d \to \mathbb{R}$ is monotone if for each $i \in [d]$ one of the following alternatives occurs:

(i) $f(a_1, \ldots, a_{i-1}, x, a_{i+1}, \ldots, a_d)$ is increasing in x for all choices of $a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_d$;

(ii) $f(a_1, \ldots, a_{i-1}, x, a_{i+1}, \ldots, a_d)$ is decreasing in x for all choices of $a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_d$.

For example, of the following 2-dimensional arrays first and second are monotone, while the third is not (since some rows contain increasing and some rows decreasing sequences).

7	8	9
4	5	6
1	2	3

1	3	6
2	5	7
4	8	9

7	8	9
6	5	4
1	2	3

The higher dimensional version of the Erdős-Szekeres problem introduced by Fishburn and Graham [105] now becomes: given positive integers d and n, determine the smallest N such that any d-dimensional array of size $N \times \ldots \times N$ contains a monotone d-dimensional subarray of size $n \times \ldots \times n$, we denote this N by $M_d(n)$. The Erdős-Szekeres theorem can now be rephrased as $M_1(n) = (n-1)^2 + 1$. Fishburn and Graham [105](Section 3) showed that $M_2(n) \leq \mathrm{towr}_5(O(n))$[1], that $M_3(n)$ is bounded by a tower of height at least a tower in n and that $M_d(n)$ is bounded from above by an Ackermann-type[2] function of order at least d for $d \geq 4$. We significantly improve upon these results.

Theorem 4.1.1.

(i) $M_2(n) \leq 2^{2^{(2+o(1))n}}$,

(ii) $M_3(n) \leq 2^{2^{(2+o(1))n^2}}$,

(iii) $M_d(n) \leq 2^{2^{2^{(1+o(1))n^{d-1}}}}$ *for* $d \geq 4$,

where the terms $o(1)$ *tend to* 0 *as* $n \to \infty$.

Fishburn and Graham introduced another very natural generalisation of the notion of monotonicity of a sequence to higher dimensional arrays. A multidimensional array is said to be *lexicographic* if for any two entries the one which has the larger position in the first coordinate in which they differ is larger. For example, the following array is lexicographic:

3	6	9
2	5	8
1	4	7

An array is said to be *lex-monotone* if it is possible to permute the coordinates and reflect the array along some dimensions to obtain a lexicographic

1 We define the tower function $\mathrm{towr}_k(x)$ by $\mathrm{towr}_1(x) = x$ and $\mathrm{towr}_k(x) = 2^{\mathrm{towr}_{k-1}(x)}$ for $k \geq 2$.

2 The Ackermann function A_k of order k is defined recursively by $A_k(1) = 2, A_1(n) = 2n$ and $A_k(n) = A_{k-1}(A_k(n-1))$. It is an incredibly fast growing function, for example $A_2(n) = 2^n$, $A_3(n) = \mathrm{towr}_n(2)$ and $A_4(n)$ is a tower of height tower of height tower, iterated n times, of 2.

array. To be more formal, for two vectors $\boldsymbol{u} = (u_1, \ldots, u_d)$ and $\boldsymbol{v} = (v_1, \ldots, v_d)$ in $\mathbb{R}^d$, we write $\boldsymbol{u} <_{\text{lex}} \boldsymbol{v}$ if $u_i < v_i$, where i is the smallest index such that $u_i \neq v_i$.

Definition (Lex-monotone array)**.** A d-dimensional array f is said to be lex-monotone if there exist a permutation $\sigma\colon [d] \to [d]$ and a sign vector $\boldsymbol{s} \in \{-1, 1\}^d$ such that

$$f(\boldsymbol{x}) < f(\boldsymbol{y}) \Leftrightarrow (s_{\sigma(1)} x_{\sigma(1)}, \ldots, s_{\sigma(d)} x_{\sigma(d)}) <_{\text{lex}} (s_{\sigma(1)} y_{\sigma(1)}, \ldots, s_{\sigma(d)} y_{\sigma(d)}).$$

Note that a 1-dimensional array is lex-monotone if and only if it is a monotone sequence. The following 2-dimensional arrays are lex-monotone since for the first one the above matrix is obtained by swapping the coordinates, for the second one by reflecting along the first dimension and for the third by performing both of these operations.

7	8	9
4	5	6
1	2	3

9	6	3
8	5	2
7	4	1

9	8	7
6	5	4
3	2	1

Given positive integers d and n, let $L_d(n)$ denote the minimum N such that for any d-dimensional array of size $N \times \ldots \times N$, one can find a lex-monotone subarray of size $n \times \ldots \times n$. Fishburn and Graham [105](Theorem 1) showed that $L_d(n)$ exists. This result has been used to prove interesting results in poset dimension theory [102] and computational complexity theory [34].

Note that any lex-monotone array is monotone, so a very natural strategy to bound $L_d(n)$ is to first find a monotone subarray and then within this subarray find a lex-monotone subarray. This motivates the following problem which is of independent interest. For positive integers d and n, we define $F_d(n)$ to be the minimum N such that any d-dimensional monotone array of size $N \times \ldots \times N$, contains a lex-monotone subarray of size $n \times \ldots \times n$. It is easy to see by the above reasoning that $L_d(n) \leq M_d(F_d(n))$. Fishburn and Graham [105](Lemma 1) showed $F_2(n) \leq 2n^2 - 5n + 4$ and $F_3(n) \leq 2^{2n+o(n)}$, while for $d \geq 4$ their argument gives $F_d(n) \leq \operatorname{towr}_{d-1}(O_d(n))$. We determine $F_2(n)$ completely and significantly improve the bound for all $d \geq 3$.

Theorem 4.1.2.

(i) $F_2(n) = 2n^2 - 5n + 4$,

(ii) $F_d(n) \leq 2^{(c_d+o(1))n^{d-2}}$ *for* $d \geq 3$, *where* $c_d = \frac{1}{2}(d-1)!$ *and the term* $o(1)$ *tends to* 0 *as* $n \to \infty$.

Part (i) of Theorem 4.1.2 answers a question of Fishburn and Graham asking whether $F_2(n) = (1 + o(1))n^2$, in negative. Combining Theorems 4.1.1 and 4.1.2 with the inequality $L_d(n) \leq M_d(F_d(n))$ gives the following upper bounds on $L_d(n)$.

Theorem 4.1.3.

(i) $L_2(n) \leq 2^{2^{(4+o(1))n^2}}$,

(ii) $L_3(n) \leq 2^{2^{2^{(2+o(1))n}}}$,

(iii) $L_d(n) \leq \operatorname{towr}_5\left(O_d(n^{d-2})\right)$ *for* $d \geq 4$,

where the terms $o(1)$ *tend to* 0 *as* $n \to \infty$.

For comparison, the best lower bound on $L_d(n)$, due to Fishburn and Graham [105](Theorem 2) is achieved by taking a random array, is $L_d(n) \geq n^{(1-1/d)n^{d-1}}$ for all $d \geq 2, n \geq 3$.

Notation and organisation. The rest of the chapter is organised as follows. We prove Theorem 4.1.1 in Section 4.2 and Theorem 4.1.2 in Section 4.3. The final section contains some concluding remarks and open problems.

All asymptotic notation in this section is with respect to $n \to \infty$. Given $d \in \mathbb{N}$, we denote the set of all permutations of $[d]$ by $\mathfrak{S}_d$. For real numbers α and β, we employ the interval notation

$$[\alpha, \beta] := \{x \in \mathbb{Z} : \alpha \leq x \leq \beta\}.$$

A set of the form $A_1 \times \ldots \times A_d$, where A_i is a finite subset of $\mathbb{Z}$ for each $i \in [d]$, is called an $|A_1| \times \ldots \times |A_d|$ grid or a grid of size $|A_1| \times \ldots \times |A_d|$. Note that a d-dimensional array $f \colon A_1 \times \ldots \times A_d \to \mathbb{R}$ is equivalent to an ordering of the vertices of the d-dimensional grid $A_1 \times \ldots \times A_d$; we switch between these points of view interchangeably.

We generally use lowercase bolded letters for vectors and uppercase bolded letters for grids. For a vector $\boldsymbol{u}$ we denote by u_i the value of the i-th coordinate.

4.2 MONOTONE ARRAYS

In this section we will prove Theorem 4.1.1. We begin with a few preliminaries.

4.2.1 *Preliminaries*

We collect here several well-known Ramsey-type results, whose simple proofs are included for completeness.

Lemma 4.2.1. *Given $k, n, t \in \mathbb{N}$ in any vertex k-colouring of an $tk\binom{nk}{n} \times kn$ grid there is a monochromatic subgrid of size $t \times n$.*

Proof. Since each row of the grid has kn points, by pigeonhole principle each row has n points with same colour. Fix a choice of such n points for each row. These n points may appear in $\binom{kn}{n}$ different positions and are monochromatic in one of k colours. Since we have $tk\binom{kn}{n}$ rows in total, there are t rows whose chosen points appear at the same positions and use the same colour. The $t \times n$ subgrid formed by the chosen points of these t rows is monochromatic. $\square$

Lemma 4.2.2. *Given integers $d, k \geq 2$ there exists a positive constant $C = C(d,k)$ such that for any positive integers n and N with $N \geq 2^{Cn^{d-1}}$, in any k-colouring of the d-dimensional $N \times \ldots \times N$ grid there is a monochromatic subgrid of size $n \times \ldots \times n$.*

Proof. Choose constants $C(d,k)$ so that $C(2,k) = 2\log_2(ek)$, and $C(d,k) = 2(d-1)C(d-1,k)$ for every $d \geq 2$. We prove the lemma by induction on d. For $d = 2$, since $N \geq 2^{C(2,k)n} = (ek)^{2n} \geq nk\binom{nk}{n}$, the statement follows from Lemma 4.2.1. We proceed to the case $d > 2$, and suppose that the lemma holds for $d-1$. Letting $M = 2^{C(d-1,k)n^{d-2}}$, we will show that any k-colouring of the d-dimensional grid $[N] \times [M] \times \ldots \times [M]$ contains a monochromatic $n \times \ldots \times n$ subgrid. For each $i \in [N]$, define $\boldsymbol{S}_i = \{i\} \times [M] \times \ldots \times [M]$. By the induction hypothesis, each $\boldsymbol{S}_i$ contains a monochromatic subgrid $\{i\} \times \boldsymbol{T}_i$ of size $1 \times n \times \ldots \times n$. There are at most $\binom{M}{n}^{d-1} \leq M^{(d-1)n}$ possibilities for $\boldsymbol{T}_i$, and each $\boldsymbol{T}_i$ uses one of the k colours. Hence, noting that $\frac{N}{kM^{(d-1)n}} \geq \frac{1}{k}2^{(d-1)C(d-1,k)n^{d-1}} \geq n$, there exist a $(d-1)$-dimensional $n \times \ldots \times n$ grid $\boldsymbol{T}$, and a size-n subset $A_1 \subseteq [N]$ such that all $\{i\} \times \boldsymbol{T}$ with $i \in A_1$ have the same monochromatic colour. In particular, $A_1 \times \boldsymbol{T}$ is a monochromatic $n \times \ldots \times n$ subgrid, as desired. $\square$

Remark. Given $d, n_1, \ldots, n_d \in \mathbb{N}$, let $K^{(d)}_{n_1,\ldots,n_d}$ denote the complete d-uniform d-partite hypergraph on vertex sets $V_1 = [n_1], \ldots, V_d = [n_d]$. Edges of $K^{(d)}_{n_1,\ldots,n_d}$ correspond in the obvious way to vertices of the d-dimensional grid $[n_1] \times \ldots \times [n_d]$. Subhypergraphs of $K^{(d)}_{n_1,\ldots,n_d}$ of the form $K^{(d)}_{m_1,\ldots,m_d}$ then correspond to subgrids of $[n_1] \times \ldots \times [n_d]$ of size $m_1 \times \ldots \times m_d$. Using this correspondence, one can derive Lemmas 4.2.1 and 4.2.2 from known results on (hyper)graph Zarankiewicz problem. For example, Lemma 4.2.1 follows from [118](Theorem 2), while Lemma 4.2.2 is a consequence of [65](Theorem 4).

4.2.2 *Proofs of the main results on monotonicity*

We begin with the 2-dimensional case. We will actually prove a stronger version of Theorem 4.1.1 (i) as we will need it for the 3-dimensional case.

Theorem 4.2.3. *For every $n, t \in \mathbb{N}$, any $4n^2 \times (2t)^{2^{2n}}$ array contains an $n \times t$ monotone subarray.*

Proof. Let f be an array indexed by $[N] \times [M]$, where $N = 4n^2$ and $M = (2t)^{2^{2n}}$. By Erdős-Szekeres Theorem we know that in each column of f there is a monotone subsequence of length $2n$. The entries of this subsequence can appear in $\binom{4n^2}{2n}$ different positions so there must be a set $R \subseteq [N]$ of $2n$ positions for which at least $M/\binom{4n^2}{2n}$ columns are monotone when restricted to (rows) R. We take $C \subseteq [N]$ to be the subset of these columns for which the restriction is increasing, we may w.l.o.g. assume that C consists of at least half of these columns. We obtain a subarray $f' = f|_{R\times C}$ which is increasing in each column and has size $2n \times M'$ where $M' \geq \frac{M}{2\binom{4n^2}{2n}} \geq t^{2^{2n}}$.

By applying Erdős-Szekeres Theorem to the sequence given by the first row of f', we can find a subset $C_1 \subseteq C$ of size $|C_1| \geq \sqrt{|C|}$ such that the first row of $f'|_{R\times C_1}$ is monotone. Repeating this argument at step i we find a subset $C_i \subseteq C_{i-1}$ of size $|C_i| \geq \sqrt{|C_{i-1}|} \geq |C|^{1/2^i}$ such that the first i rows of $f'|_{R\times C_i}$ are monotone. Continuing this process until $i = 2n$ we obtain an $2n \times t$ array with each row being either increasing or decreasing. By taking the ones of the type which appears more often we obtain a monotone $n \times t$ subarray as claimed. $\square$

One can easily generalise the above proof to give a bound of the form $M(d, n) \leq \mathrm{towr}_{d+1}((2 + o(1))n^{d-1})$ for any $d \geq 2$, which would already be a substantial improvement over the Ackermann bound due to Fishburn and Graham [105]. However, to prove the desired and better bound $M(d, n) \leq \mathrm{towr}_4((2 + o(1))n^{d-1})$ for $d \geq 4$, we need to consider an intermediate problem which we find interesting in its own right.

Definition (Inconsistently monotone array)**.** An array $f\colon A_1 \times \ldots \times A_d \to \mathbb{R}$ is inconsistently monotone if for each $i \in [d]$, $f(a_1, \ldots, a_{i-1}, x, a_{i+1}, \ldots, a_d)$ is monotone in x for all choices of $a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_d$.

Main difference compared with the definition of a monotone array is that we do not require all the lines along a fixed dimension to be all increasing or all decreasing but allow some to be increasing and some to be decreasing. For positive integers d and n, let $M'_d(n)$ denote the minimum N such that for any d-dimensional array of size $N \times \ldots \times N$, one can find a d-dimensional inconsistently monotone subarray of size $n \times \ldots \times n$. We have $M'_1(n) = (n-1)^2 + 1$ according to Erdős-Szekeres Theorem. When $d \geq 2$ we obtain the following improved version of Theorem 4.1.1 for inconsistently monotone arrays.

Theorem 4.2.4. *For every $d \geq 2$, we have $M'_d(n) \leq 2^{2^{(1+o(1))n^{d-1}}}$.*

Proof. We will prove the following recursive bound

$$M'_d(n) \leq \binom{M'_{d-1}(n)}{n}^{d-1} n^{2^{n^{d-1}}}. \tag{4.1}$$

Let $m = M'_{d-1}(n)$ and $N = \binom{m}{n}^{d-1} n^{2^{n^{d-1}}}$. To prove (4.1), let f be an array indexed by $[m]^{d-1} \times [N]$. For each "height" $h \in [N]$, consider the restriction of f to $[m]^{d-1} \times \{h\}$. As $m = M'_{d-1}(n)$, there exist an $n \times \ldots \times n$ subgrid S_h of $[m]^{d-1}$ such that f is inconsistently monotone on $S_h \times \{h\}$. Given $h \in [N]$, there are at most $\binom{m}{n}^{d-1}$ possibilities for the location of S_h. Hence, by the pigeonhole principle, we can find an $n \times \ldots \times n$ subgrid S of $[m]^{d-1}$ and a subset $H \subset [N]$ of size

$$|H| \geq \frac{N}{\binom{m}{n}^{d-1}} = n^{2^{n^{d-1}}}$$

such that f is inconsistently monotone on $S \times \{h\}$ for every $h \in H$. Let us denote the elements of S by $s_1, \ldots, s_{n^{d-1}}$. By Erdős-Szekeres Theorem, we can construct a nested sequence $H_0 := H \supseteq H_1 \supseteq \ldots \supseteq H_{n^{d-1}}$ such that $|H_i| \geq \sqrt{|H_{i-1}|}$ for every $i \geq 1$, and that $\{s_j\} \times H_i$ is monotone for $j = 1, \ldots, i$. In particular, we have that $|H_{n^{d-1}}| \geq |H|^{1/2^d} \geq n$, and that the restriction of f to $S \times H_{n^{d-1}}$ is inconsistently monotone. This completes the proof of (4.1).

What remains to be shown is that (4.1) implies the desired bound $M'_d(n) \leq \mathrm{towr}_3((1+o(1))n^{d-1})$. We proceed by induction on d, noting that the case

$d = 2$ follows from (4.1) and the fact that $M'_1(n) = (n-1)^2 + 1$. For the induction step, assuming $d \geq 3$. Using (4.1) and the induction hypothesis we find

$$\begin{aligned} M'_d(n) &\leq M'_{d-1}(n)^{(d-1)n} \cdot n^{2^{n^{d-1}}} \\ &\leq \operatorname{towr}_3(O(n^{d-2})) \cdot \operatorname{towr}_3((1+o(1))n^{d-1}) \\ &= \operatorname{towr}_3((1+o(1))n^{d-1}), \end{aligned}$$

finishing the proof. □

The following definition is going to help us find a monotone array inside an inconsistently monotone array.

Definition 4.2.5 (Monotonicity pattern). Let $f \colon \boldsymbol{A} \to \mathbb{R}$ be an inconsistently monotone d-dimensional array. Let $\boldsymbol{a} = (a_1, \ldots, a_d) \in \boldsymbol{A}$. For each $i \in [d]$, let $s_i \in \{-1, 1\}$ be such that $s_i f(a_1, \ldots, a_{i-1}, x, a_{i+1}, \ldots, a_d)$ is increasing in x. The vector $\boldsymbol{s} = (s_1, \ldots, s_d)$ is called the *monotonicity pattern* of f at point $\boldsymbol{a}$.

Notice that if f is a monotone array then f has the same monotonicity pattern at all points, in which case we just call it the monotonicity pattern of f. We now use Theorem 4.2.4 to prove Part (iii) of Theorem 4.1.1.

Theorem 4.2.6. *For every $d \geq 4$, we have $M_d(n) \leq \operatorname{towr}_4\left((1+o(1))n^{d-1}\right)$.*

Proof. Let $N = \operatorname{towr}_4\left((1+o(1))n^{d-1}\right)$, and let $C = C(d, 2^d)$ be the positive constant given by Lemma 4.2.2. It follows from Theorem 4.2.4 that in any d-dimensional array of size $N \times \ldots \times N$, one can find an inconsistently monotone subarray f indexed by $\boldsymbol{A} = A_1 \times \ldots \times A_d$ such that $|A_1| = \ldots = |A_d| = 2^{Cn^{d-1}}$.

Let us colour every point in $\boldsymbol{A}$ with the monotonicity pattern of f at this point. This gives us a vertex-colouring of $\boldsymbol{A}$ with 2^d colours given by $\{-1, 1\}^d$. By Lemma 4.2.2 and the choice of C, there exists a monochromatic $n \times \ldots \times n$ subgrid $\boldsymbol{B}$ of $\boldsymbol{A}$ with colour $(s_1, \ldots, s_d)$. From the definition of the monotonicity pattern $(s_1, \ldots, s_d)$, we can see that $f|_{\boldsymbol{B}}$ is monotone. □

In order to prove $M_3(n) \leq 2^{2^{(2+o(1))n^2}}$, we devise a different argument, not going through the intermediate problem of bounding $M'_d(n)$.

Theorem 4.2.7. *We have $M_3(n) \leq 2^{2^{(2+o(1))n^2}}$.*

Proof. Let $X_1 = [16n^2], X_2 = [2^{2^{6n}}]$ and $X_3 = [2^{2^{(2+o(1))n^2}}]$. To prove Theorem 4.1.1 (ii) it suffices to show that any 3-dimensional array f indexed by $X_1 \times X_2 \times X_3$ contains an $n \times n \times n$ monotone subarray.

For each "height" $h \in X_3$, let $\boldsymbol{C}_h = X_1 \times X_2 \times \{h\}$. As $2^{2^{6n}} \geq (2n2^{2n})^{2^{4n}}$, Theorem 4.2.3 implies that each $\boldsymbol{C}_h$ contains a $2n \times n2^{2n} \times 1$ monotone subarray. There are $\binom{16n^2}{2n}\binom{2^{2^{6n}}}{n2^{2n}}$ different possibilities for a $2n \times n2^{2n} \times 1$ monotone subarray, and the monotonicity pattern of each such subarray is a vector $\boldsymbol{s} \in \{-1,1\}^2$. Since $4\binom{16n^2}{2n}\binom{2^{2^{6n}}}{n2^{2n}} = 2^{2^{o(n^2)}}$, by pigeonhole principle, we can find a vector $\boldsymbol{s} \in \{-1,1\}^2$ and three subsets $S_1 \subseteq X_1, S_2 \subseteq X_2, S_3 \subseteq X_3$ with $|S_1| = 2n$, $|S_2| = n2^{2n}$ and $|S_3| = 2^{2^{(2+o(1))n^2}}$ such that for any $h \in S_3$ the array $f|_{S_1 \times S_2 \times \{h\}}$ is monotone with pattern $\boldsymbol{s}$. Our remaining goal is to find an $n \times n$ subgrid of $S_1 \times S_2$ such that for any pair (a_1, a_2) of this subgrid, $f(a_1, a_2, \cdot)$ is always increasing or always decreasing on some fixed subset of size n of S_3.

For each $h \in S_3$, let $\boldsymbol{L}_h = S_1 \times S_2 \times \{h\}$. We can think of $\boldsymbol{L}_h$'s as "layers" stacked one on top of each other. Given two layers $\boldsymbol{L}_h$ and $\boldsymbol{L}_{h'}$ with $h < h'$, we colour an element $\boldsymbol{v} \in S_1 \times S_2$ in red if $f(\boldsymbol{v}, h) > f(\boldsymbol{v}, h')$, and blue otherwise. This way we obtain a colouring of $S_1 \times S_2$ with two colours, so by Lemma 4.2.1 we can find a monochromatic subgrid $\boldsymbol{B}_{hh'}$ of size $n \times n$. We now consider the following edge-colouring of the complete graph on the vertex set S_3 using k colours. We colour the edge between h and h' by a pair made of $\boldsymbol{B}_{hh'}$ and its monochromatic colour. Since there are at most $\binom{2n}{n}\binom{n2^{2n}}{n}$ possibilities for $\boldsymbol{B}_{hh'}$, we must have $k \leq 2\binom{2n}{n}\binom{n2^{2n}}{n} = 2^{(2+o(1))n^2}$, giving $k^{kn} \leq 2^{2^{(2+o(1))n^2}} = |S_3|$. From this and a result of Erdős and Rado [95](Theorem 1) on the multicolor Ramsey numbers which states that in any k-edge colouring of the complete graph on k^{kn} many vertices contains a monochromatic K_n, we deduce that our colouring contains a monochromatic K_n. Let $H \subseteq S_3$ correspond to the vertices of this K_n and let its colour correspond to an $n \times n$ subgrid $\boldsymbol{B}$ and say w.l.o.g. blue. This means that $f(a_1, a_2, \cdot) : H \to \mathbb{R}$ is increasing for all $(a_1, a_2) \in \boldsymbol{B}$. So by our construction of S_1, S_2 we have that f when restricted to $\boldsymbol{B} \times H$ is a monotone array of size $n \times n \times n$. $\square$

Remark. In the above proof we used the usual Ramsey theorem on our colouring of the complete graph on S_3. However, our colouring is not arbitrary and in fact one can instead use the ordered Ramsey number of a path (see [63]). The third alternative is to only colour an edge according to $\boldsymbol{B}_{hh'}$, and record whether the values increase or decrease between $\boldsymbol{B}_{hh'} \times \{h\}$ and $\boldsymbol{B}_{hh'} \times \{h'\}$ by a directed edge. This gives us a colouring of a tournament in which we want to find a monochromatic directed path (see [60, 137]). Both approaches give slightly better bounds than the one in Theorem 4.1.1 (ii), but unfortunately still give bounds of the form $M_3(n) \leq \text{towr}_3(O(n^2))$.

4.3 LEXICOGRAPHIC ARRAYS

In this section we show our bounds on $F_d(n)$, in particular we prove Theorem 4.1.2.

4.3.1 *Preliminaries*

A monotone array f is said to be *increasing* if restriction of f to any axis parallel line is an increasing sequence (i.e. case (i) of the definition of monotonicity always occurs). More formally:

Definition 4.3.1. A d-dimensional array f is *increasing* if $f(\boldsymbol{x}) \leq f(\boldsymbol{y})$ whenever $x_i \leq y_i$ for all $i \in [d]$.

The following definition generalises the notion of a lexicographic array to allow for a custom priority order of coordinates.

Definition 4.3.2. Given a d-dimensional array f and a permutation $\sigma \in \mathfrak{S}_d$, we say f is *(lexicographic) of type σ* if $f(\boldsymbol{x}) < f(\boldsymbol{y}) \Leftrightarrow (x_{\sigma(1)}, \ldots, x_{\sigma(d)}) <_{\text{lex}} (y_{\sigma(1)}, \ldots, y_{\sigma(d)})$ for all possible $\boldsymbol{x}$ and $\boldsymbol{y}$.

Recall that an array is said to be lex-monotone if it is possible to permute the coordinates and reflect the array along some dimensions to obtain a lexicographic array. The above definition allows us to separate these two actions. In particular, an alternative definition is that an array is lex-monotone if one can reflect the array along some dimensions to obtain a lexicographic array of some type.

Notice that any subarray of a monotone array is itself monotone and moreover has the same monotonicity pattern. This means that when looking for a lex-monotone subarray within a monotone array we can only ever find one with the same monotonicity pattern. In other words we may w.l.o.g. assume that the starting array is increasing. The following immediate lemma makes this statement formal.

Lemma 4.3.3. *For every $d, n \in \mathbb{N}$, $F_d(n)$ equals the minimum N such that any increasing d-dimensional array of size $N \times \ldots \times N$ contains an $n \times \ldots \times n$ subarray of type σ for some $\sigma \in \mathfrak{S}_d$.*

4.3.2 *2-dimensional case*

Notice first that in 2 dimensions there are only two possible types of a (lexicographic) array, namely (1,2) and (2,1). See Figure 4.1 for an illustration of both together with an example of the arrow notation which we found useful when thinking about the problem.

Figure 4.1: Lexicographic arrays of type (1,2) and (2,1), arrows point towards larger points.

We begin with a proof of the upper bound $F_2(n) \leq 2n^2 - 5n + 4$ as it sheds some light to where our lower bound construction is coming from.

Theorem 4.3.4 (Fishburn and Graham [105]). *For $n \in \mathbb{N}$ we have $F_2(n) \leq 2n^2 - 5n + 4$.*

Proof. Let f be an increasing array indexed by $[N] \times [N]$, where $N = (n-1)(2n-3)+1$. For $i \in [2n-2]$, let $a_i = (n-1)(i-1)+1$. Define a red-blue colouring of the grid $\{a_1, \ldots, a_{2n-3}\} \times \{a_1, \ldots, a_{2n-3}\}$ as follows. For every $i, j \in [2n-3]$, we colour (a_i, a_j) red if $f(a_{i+1}, a_j) < f(a_i, a_{j+1})$, and blue otherwise. As $(n-2)(2n-3) + (n-2)(2n-3) < (2n-3)^2$, there exists a row with at least $n-1$ red points or a column with at least $n-1$ blue points. By symmetry, we can assume $(a_i, a_{j_1}), \ldots, (a_i, a_{j_{n-1}})$ are $n-1$ red points in a row a_i with $j_1 < j_2 < \ldots < j_{n-1}$. One can check that the $n \times n$ subarray of f indexed by $[a_i, a_{i+1}] \times \{a_{j_1}, \ldots, a_{j_{n-1}}, a_{j_{n-1}+1}\}$ is of type $(1,2)$. Hence $F_2(n) \leq N = 2n^2 - 5n + 4$, as required. $\square$

The remainder of this subsection is devoted to the proof of the lower bound $F_2(d) \geq 2n^2 - 5n + 4$. We will make ample use of the immediate observation that any subarray of a lexicographic array of type σ which has size at least 2 in each dimension must also be of type σ. We first construct a "building block" for our actual construction showing $F_2(d) \geq 2n^2 - 5n + 4$.

Lemma 4.3.5. *For $n \geq 3$, there exists an increasing array g of size $(n-1)(n-2) \times (n-1)^2$ such that*

(G1) *g does not contain an $(n-1) \times 2$ subarray of type $(1,2)$,* (G2) *g does not contain an $n \times 2$ subarray of type $(2,1)$.*

Proof. For $1 \leq i \leq n-2$, let $\boldsymbol{C}_i = [(n-1)(i-1)+1, (n-1)i] \times [(n-1)^2]$. We choose an array g (see Figure 4.2 for an illustration), indexed by $[(n-1)(n-2)] \times [(n-1)^2]$, such that

- $g|_{\boldsymbol{C}_1} < \ldots < g|_{\boldsymbol{C}_{n-2}}$,
- For each $i \in [n-2]$, $g|_{\boldsymbol{C}_i}$ is of type $(2,1)$.

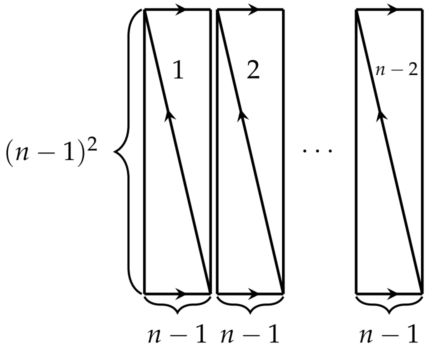


Figure 4.2: An illustration of the array g. Directed arrows point towards a position with a larger value of g. Numbers denote relative order of subarrays.

For (G1), if such a subarray exists, then at least one of $\boldsymbol{C}_i$'s would need to intersect this subarray in some 2×2 subarray. By the second property of g, the 2×2 subarray is of type $(2,1)$, a contradiction.

For (G2), if such a subarray exists it would intersect at least two distinct $\boldsymbol{C}_i$'s, and so it would contain a 2×2 subarray of type $(1,2)$, a contradiction. □

Another building block of our construction is the following.

Lemma 4.3.6. *For $n \geq 3$, there is an increasing array h of size $(n-1)^2 \times (n-1)(n-2)$ such that*

(H1) *h does not contain a $2 \times n$ subarray of type $(1,2)$,*

(H2) *h does not contain a $2 \times (n-1)$ subarray of type $(2,1)$.*

Proof. Let $\boldsymbol{R}_i = [(n-1)^2] \times [(n-1)i - n + 1, (n-1)i]$ for $i \in [n-2]$. Let us define h to be an array indexed by $[(n-1)^2] \times [(n-1)(n-2)]$ so that $h|_{\boldsymbol{R}_i} < h|_{\boldsymbol{R}_j}$ whenever $i < j$ and so that $h|_{\boldsymbol{R}_i}$ is of type $(1,2)$ (see Figure 4.3 for an illustration). This array satisfies the properties (H1) and (H2) by the same argument as in Lemma 4.3.5. □

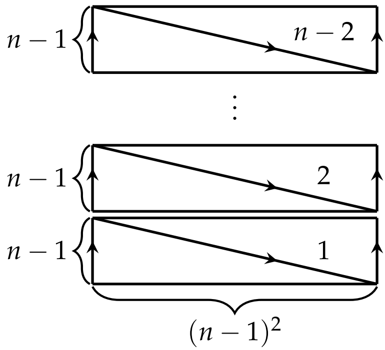


Figure 4.3: An illustration of the array h. Directed arrows point towards a position with a larger value of h. Numbers denote relative order of subarrays.

We are now in a position to prove Theorem 4.1.2 (i).

Theorem 4.3.7. *For $n \in \mathbb{N}$ we have $F_2(n) \geq 2n^2 - 5n + 4$.*

Proof. It is immediate that the statement holds for $n = 1, 2$. We henceforth assume that $n \geq 3$.

Let $N = 2n^2 - 5n + 3 = (n-1)^2 + (n-1)(n-2)$. To prove the statement, it suffices to construct an increasing array $f : [N]^2 \to \mathbb{R}$ which does not contain an $n \times n$ subgrid of type (1,2) or (2,1).

We first split $[N]^2$ into five subgrids $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_5$ (see Section 4.3.2) such that both $\boldsymbol{A}_1$ and $\boldsymbol{A}_5$ have size $(n-1)(n-2) \times (n-1)^2$, both $\boldsymbol{A}_2$ and $\boldsymbol{A}_4$ have size $(n-1)^2 \times (n-1)(n-2)$, while $\boldsymbol{A}_3$ has size $(n-1) \times (n-1)$. Let g and h be arrays given by Lemma 4.3.5 and Lemma 4.3.6, respectively. The array f is chosen so that $f|_{\boldsymbol{A}_1} < f|_{\boldsymbol{A}_2} < \ldots < f|_{\boldsymbol{A}_5}$, $f|_{\boldsymbol{A}_1}$ and $f|_{\boldsymbol{A}_5}$ are copies of g, $f|_{\boldsymbol{A}_2}$ and $f|_{\boldsymbol{A}_4}$ are copies of h, and $f|_{\boldsymbol{A}_3}$ is an arbitrary increasing array. Since $f|_{\boldsymbol{A}_1} < f|_{\boldsymbol{A}_2} < \ldots < f|_{\boldsymbol{A}_5}$ and $f|_{\boldsymbol{A}_i}$ is increasing for every $1 \leq i \leq 5$, f is increasing as well. It remains to show that f does not contain an $n \times n$ subarray of type (1,2) or (2,1).

As $f|_{\boldsymbol{A}_1} < f|_{\boldsymbol{A}_2} < \ldots < f|_{\boldsymbol{A}_5}$ and $f|_{\boldsymbol{A}_i}$ is increasing for every $1 \leq i \leq 5$, we also find that

(P1) Any 2×2 subarray with two vertices in $\boldsymbol{A}_1$ and two vertices to the right of $\boldsymbol{A}_1$ is of type (1,2),

(P2) Any 2×2 subarray with two vertices in $\boldsymbol{A}_5$ and two vertices to the left of $\boldsymbol{A}_5$ is of type (1,2),

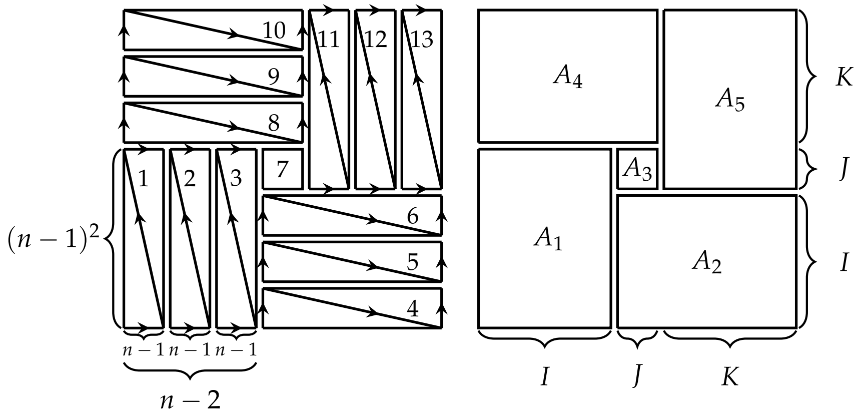


Figure 4.4: Directed arrows point towards a larger value of f and indicate whether the given subarray is of type (1,2) or type (2,1) (as in Figure 4.1). Numbers denote the relative order of subarrays.

(P3) Any 2×2 subarray with two vertices in $\boldsymbol{A}_2$ and two vertices above $\boldsymbol{A}_2$ is of type (2,1),

(P4) Any 2×2 subarray with two vertices in $\boldsymbol{A}_4$ and two vertices below $\boldsymbol{A}_4$ is of type (2,1).

We will show that f has the desired property using properties (P1)–(P4) together with conditions (G1), (G2), (H1), (H2) from Lemmas 4.3.5 and 4.3.6.

Suppose towards a contradiction that $[N]^2$ contains an $n \times n$ subgrid $\boldsymbol{L} = L_1 \times L_2$ such that $f|_{\boldsymbol{L}}$ is of type (1,2) or (2,1). Letting $I = [(n-1)(n-2)]$, $J = [(n-1)(n-2)+1, (n-1)^2]$, $K = [(n-1)^2+1, N]$, we define

$$a = |L_1 \cap I|, \; b = |L_1 \cap J|, \; c = |L_1 \cap K|$$
$$x = |L_2 \cap I|, \; y = |L_2 \cap J|, \; z = |L_2 \cap K|.$$

We will obtain various inequalities involving a, b, c, x, y, z, and eventually reach a contradiction. Since $\boldsymbol{L}$ has size $n \times n$ and $|J| = n-1$ we obtain

$$a+b+c = x+y+z = n, \;\; 0 \le a, c, x, z \le n \;\text{ and }\; 0 \le b, y \le n-1. \quad (4.2)$$

We divide our analysis into two cases.

Case 1: $\boldsymbol{L}$ is of type $(2,1)$.

We have the following series of observations

$$\text{(G2)} \Rightarrow \quad a \leq n-1 \text{ or } x+y \leq 1, \tag{4.3}$$

$$\text{(G2)} \Rightarrow \quad c \leq n-1 \text{ or } y+z \leq 1, \tag{4.4}$$

$$\text{(H2)} \Rightarrow \quad a+b \leq 1 \text{ or } z \leq n-2, \tag{4.5}$$

$$\text{(H2)} \Rightarrow \quad b+c \leq 1 \text{ or } x \leq n-2, \tag{4.6}$$

$$\text{(P1)} \Rightarrow \quad a=0 \text{ or } b+c=0 \text{ or } x+y \leq 1, \tag{4.7}$$

$$\text{(P2)} \Rightarrow \quad a+b=0 \text{ or } c=0 \text{ or } y+z \leq 1. \tag{4.8}$$

Observe that

$$a=0 \text{ or } x+y \leq 1, \tag{4.9}$$

$$c=0 \text{ or } y+z \leq 1. \tag{4.10}$$

To see (4.9) if $b+c=0$ then by (4.2) we have $a=n$ which according to (4.3) implies $x+y \leq 1$ so (4.7) implies (4.9). Similarly, to see (4.10) if $a+b=0$ then by (4.2) we have $c=n$ which according to (4.4) implies $y+z \leq 1$ so (4.8) implies (4.10).

To complete our analysis of Case 1, we show $a=c=0$, giving a contradiction to (4.2). Suppose to the contrary that $a \geq 1$. Then $x+y \leq 1$ by (4.9), and so $z \geq n-1$ by (4.2), which according to (4.5) shows $a+b \leq 1$. Hence

$$c = n-(a+b) \geq n-1 \geq 1 \text{ and } y+z \geq z \geq n-1 \geq 2,$$

giving a contradiction to (4.10). It remains to show that $c=0$. If we instead have $c \geq 1$, then (4.10) implies $y+z \leq 1$, and so $x \geq n-1$ by (4.2), which by (4.6) implies $b+c \leq 1$. Thus

$$a = n-(b+c) \geq n-1 \geq 1 \text{ and } x+y \geq x \geq n-1 \geq 2,$$

contradicting (4.10) and completing the proof in this case.

Case 2: L is of type (1,2).

The analysis of this case is very similar to that of Case 1. We first have the following observations

$$\text{(G1)} \Rightarrow \quad a \leq n-2 \text{ or } x+y \leq 1, \tag{4.11}$$

$$\text{(G1)} \Rightarrow \quad c \leq n-2 \text{ or } y+z \leq 1, \tag{4.12}$$

$$\text{(H1)} \Rightarrow \quad a+b \leq 1 \text{ or } z \leq n-1, \tag{4.13}$$

$$\text{(H1)} \Rightarrow \quad b+c \leq 1 \text{ or } x \leq n-1, \tag{4.14}$$

$$\text{(P4)} \Rightarrow \quad a+b \leq 1 \text{ or } x+y=0 \text{ or } z=0, \tag{4.15}$$

$$\text{(P3)} \Rightarrow \quad b+c \leq 1 \text{ or } x=0 \text{ or } y+z=0. \tag{4.16}$$

We next show

$$a+b \le 1 \text{ or } z=0, \tag{4.17}$$
$$b+c \le 1 \text{ or } x=0. \tag{4.18}$$

To see (4.17) if $x+y=0$ then by (4.2) we have $z=n$ which according to (4.13) implies $a+b \le 1$ so (4.15) implies (4.17). Similarly, to see (4.18) if $y+z=0$ then by (4.2) we have $x=n$ which according to (4.14) implies $b+c \le 1$ so (4.16) implies (4.18).

Finally, we show $x=z=0$, giving a contradiction to (4.2). Suppose $x \ge 1$. Then $b+c \le 1$ by (4.18), and so (4.2) gives $a \ge n-1$, which by (4.11) forces $x+y \le 1$. From this we conclude

$$a+b \ge a \ge n-1 \ge 2 \text{ and } z = n-(x+y) \ge n-1 \ge 1,$$

giving a contradiction to (4.10). To show $z=0$, we suppose $z \ge 1$. Then (4.17) gives $a+b \le 1$, and so (4.2) implies $c \ge n-1$, which by (4.12) results in $y+z \le 1$. Thus

$$b+c \ge c \ge n-1 \ge 2 \text{ and } x = n-(y+z) \ge n-1 \ge 1,$$

contradicting (4.18). This completes our proof of Theorem 4.1.2 (i). □

The example used above is partially motivated by certain examples considered in [44]. This paper also considers higher dimensional examples which may be of some use in higher dimensional instances of our problem as well, but only in terms of optimizing the dependency on d.

4.3.3 *High-dimensional case*

In this subsection we prove Theorem 4.1.2 (ii). Our first ingredient in the proof will be the following lemma.

Lemma 4.3.8 (Dominant coordinate)**.** *Let $d, m, t \ge 2$ be integers, and let f be an increasing array indexed by $[d^2mt]^d$. Then, there exist a dimension $i \in [d]$, sets $B_1, \dots, B_{i-1}, B_{i+1}, \dots, B_d \subseteq [d^2mt]$ of size $m+1$ and t subgrids $\boldsymbol{A}_h := B_1 \times \dots \times B_{i-1} \times \{h\} \times B_{i+1} \times \dots \times B_d$ such that $f|_{\boldsymbol{A}_h} < f|_{\boldsymbol{A}_{h'}}$ whenever $h < h'$.*

One should think of this lemma as saying that there is a dimension i such that one can find a "stack" of subgrids appearing at the same location along the remaining $d-1$ dimensions and different positions along dimension i, which can be thought of as heights of the subgrids. Furthermore, the subgrids are not much smaller than the initial one in the remaining dimensions and our array is always bigger on a higher subgrid. Our proof of this lemma borrows some ideas of [105].

Proof. The proof of the lemma is in some sense a high-dimensional generalisation of the argument used to prove Theorem 4.3.4. We split the grid $[d^2mt]^d$ along each coordinate into td intervals of equal size, obtaining a partition of $[d^2mt]^d$ into translates of $[dm]^d$

$$[d^2mt]^d = \bigcup_{\boldsymbol{u}\in \boldsymbol{T}} \left(\boldsymbol{u} + [dm]^d\right), \tag{4.19}$$

where $\boldsymbol{T} = \{0, dm, 2dm, \ldots, (dt-1)dm\}^d$. The reason behind considering this is that we are now going to compare values taken by the array on certain points in $\boldsymbol{u} + [dm]^d$ for each $\boldsymbol{u} \in \boldsymbol{T}$, and once we find the one with the largest entry, the fact that points of $\boldsymbol{T}$ are suitably spaced apart will allow us to get information about the ordering of a relatively large $(d-1)$-dimensional subarray. For each $i \in [d]$, the aforementioned points are given by $\boldsymbol{x}_i$ and the subarrays by $\boldsymbol{C}_i$ below.

$$\begin{aligned}\boldsymbol{x}_i &= ((i-1)m, (i-2)m, \ldots, m, dm, (d-1)m, \ldots, im) \in [dm]^d,\\ \boldsymbol{C}_i &= [(i-1)m, im] \times \ldots \times [m, 2m] \times \{m\} \times [(d-1)m, dm] \times \ldots \times [im, (i+1)m]\\ &\subset [dm]^d.\end{aligned}$$

Notice that $\boldsymbol{x}_{i+1}$ has every coordinate larger than $\boldsymbol{x}_i$, except i-th, here $\boldsymbol{x}_{d+1} := \boldsymbol{x}_1$. Notice further that $\boldsymbol{x}_{i+1} \in \boldsymbol{C}_i$ is larger in every coordinate than any other point of $\boldsymbol{C}_i$. Similarly $\boldsymbol{x_i}$ with its i-th coordinate reduced by $d(m-1)$ is the point of C_i which is smaller than any other in every coordinate. In other words with respect to the componentwise order of $[dm]^d$:

$$\max \boldsymbol{C}_i = \boldsymbol{x_{i+1}}, \text{ and } \min \boldsymbol{C}_i = \boldsymbol{x_i} - (m-1)d\boldsymbol{e}_i, \tag{4.20}$$

where $\boldsymbol{e}_i$ stands for the i-th unit vector $(0, \ldots, \underset{\substack{\uparrow\\ i\text{-th}}}{1}, \ldots, 0)$.

Now consider a colouring $\chi \colon \boldsymbol{T} \to [d]$ given by:

$$\chi(\boldsymbol{u}) = i \text{ if and only if } f(\boldsymbol{u} + \boldsymbol{x}_i) = \max\{f(\boldsymbol{u} + \boldsymbol{x}_1), \ldots, f(\boldsymbol{u} + \boldsymbol{x}_d)\}.$$

By pigeonhole principle, there is a colour $i \in [d]$ which appears at least $(td)^d/d$ times. This implies that the grid $\boldsymbol{T}$ contains a column in the direction of the i-th coordinate for which at least t vertices of this column have colour i. We list those vertices of $\boldsymbol{T}$ from smallest to largest with respect to their i-th coordinates: $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_t$.

We show that the grids $\boldsymbol{A}_1 = \boldsymbol{u}_1 + \boldsymbol{C}_i, \ldots, \boldsymbol{A}_t = \boldsymbol{u}_t + \boldsymbol{C}_i$ have the desired properties. Indeed, (4.19) implies that $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_t$ are subgrids of $[d^2mt]^t$. Since we have chosen $\boldsymbol{u}_j$'s as in the same column in the direction of the i-th coordinate, all of them have the same coordinates in all other dimensions.

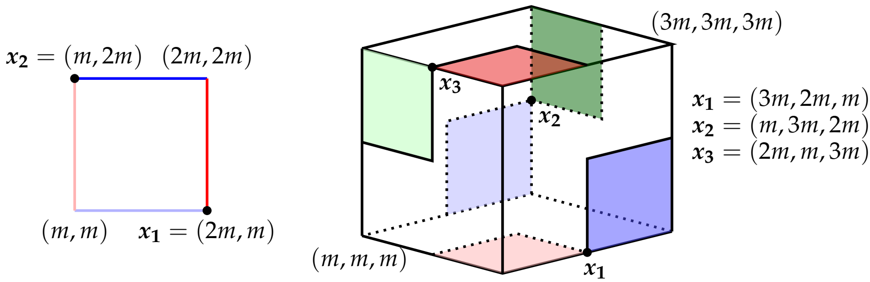


Figure 4.5: Lightly shaded $(d-1)$-dimensional regions in the figure denote $\boldsymbol{C_i}$'s with their maximum point being $\boldsymbol{x_{i+1}}$. Depending on which of the $\boldsymbol{x_i}$'s has largest value of f one of these $\boldsymbol{C_i}$'s has value of f on $\boldsymbol{x_i}$ smaller than that of f on the minimal point of a translate of $\boldsymbol{C_i}$ (strongly shaded region of the same colour) at $\boldsymbol{x_i}$.

This implies that there are $d-1$ sets $B_1, \ldots, B_{i-1}, B_{i+1}, \ldots, B_d$ and t "heights" $h_1, \ldots, h_t$ such that $\boldsymbol{A}_j$ can be written as $B_1 \times \ldots \times B_{i-1} \times \{h_j\} \times B_{i+1} \times \ldots \times B_d$ for every $1 \leq j \leq t$. Since $\boldsymbol{C}_i$ has size $(m+1) \times \ldots \times (m+1)$ each B_j has size $m+1$. Finally, since f is increasing, for $1 \leq j < k \leq t$ we have

$$\begin{aligned}\max f(\boldsymbol{A}_j) &= f(\boldsymbol{u}_j + \boldsymbol{x}_{i+1}) \\ &< f(\boldsymbol{u}_j + \boldsymbol{x}_i) \\ &< f(\boldsymbol{u}_k - (m-1)d\boldsymbol{e}_i + \boldsymbol{x}_i) \\ &= \min f(\boldsymbol{A}_k).\end{aligned}$$

The first equality follows since (4.20) implies $\boldsymbol{u}_j + \boldsymbol{x}_{i+1}$ is the largest point of $\boldsymbol{A}_j$ so since f is increasing we conclude that f is maximised over $\boldsymbol{A}_j$ at $\boldsymbol{u}_j + \boldsymbol{x}_{i+1}$. Similarly, we get the last equality as well. The first inequality follows since $\chi(\boldsymbol{u}_j) = i$. The second inequality follows since $j < k$ implies the i-th coordinate of $\boldsymbol{u}_j$ is smaller than that of $\boldsymbol{u}_k$ (since $\boldsymbol{u}_j$ and $\boldsymbol{u}_k$ belong to the same column along i-th dimension and since we named them according to their i-th coordinate) by at least dm (since $\boldsymbol{u}_j, \boldsymbol{u}_k \in \boldsymbol{T}$). This finishes our proof of Lemma 4.3.8. $\square$

This lemma provides us with a stack of subgrids $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_t$ on which f is increasing in dimension i. It is natural to try to iterate and now apply the lemma within each $\boldsymbol{A}_j$, but notice that we do not only want to find a lexicographic subgrid in some n different $\boldsymbol{A}_j$'s, but they also have to appear at the same positions in each of them. One can now apply a Ramsey result for $O(2^{F_{d-1}(n)})$ colours to ensure the found subgrids appear at the same locations. Consequently, this approach gives at best $F_d(n) \leq \mathrm{towr}_{d-1}(O_d(n))$. Fishburn and Graham [105](Section 4) follow a similar approach and their arguments hit the same barrier in terms of the bounds they can obtain.

We take a different approach in order to prove Theorem 4.1.2 (ii).

Theorem 4.3.9. *For $d \geq 3$, we have $F_d(n) \leq 2^{(c_d+o(1))n^{d-2}}$, where $c_d = \frac{1}{2}(d-1)!$.*

Proof. We prove the statement by induction on d.
The base case: $d = 3$.

Let $m = 2n^3$, $t = 2m\binom{m/n}{n}/\binom{m/(2n)}{n} = 2^{n+o(n)}$, and $N = d^2mt = 2^{n+o(n)}$. Consider an increasing array $f : [N]^3 \to \mathbb{R}$. By Lemma 4.3.8, we can assume w.l.o.g. that $[N]^3$ contains a stack of t subgrids $\boldsymbol{A}_1 = B_1 \times B_2 \times \{h_1\}, \ldots, \boldsymbol{A}_t = B_1 \times B_2 \times \{h_t\}$ of size $m \times m \times 1$ such that $h_1 < h_2 < \ldots < h_t$ and $f|_{\boldsymbol{A}_1} < f|_{\boldsymbol{A}_2} < \ldots < f|_{\boldsymbol{A}_t}$. We drop the third dimension from now on, and think of $\boldsymbol{A}_i$'s as 2-dimensional grids.

We split each $\boldsymbol{A}_i$ into $\left(\frac{m}{n}\right)^2$ smaller subgrids of size $n \times n$. Colour each such smaller subgrid red if its topmost leftmost corner is smaller than bottommost rightmost, and blue otherwise. As in the proof of Theorem 4.3.4, any n red subgrids in the same row of $\boldsymbol{A}_i$ give rise to an $n \times n$ subgrid of type $(1,2)$. If we further manage to find n layers of the stack, each having such a sequence of the same n red $n \times n$ subgrids we obtain an $n \times n \times n$ subgrid of type $(3,1,2)$ (using the property that $h_1 < h_2 < \ldots < h_t$ and $f|_{\boldsymbol{A}_1} < f|_{\boldsymbol{A}_2} < \ldots < f|_{\boldsymbol{A}_t}$). Similarly, if we find n layers each having the same sequence of n blue $n \times n$ subgrids in the same column we find an $n \times n \times n$ subgrid of type $(3,2,1)$.

By pigeonhole principle, each layer $\boldsymbol{A}_i$ of the stack has a row with $\frac{m}{2n}$ red subgrids or a column with at least $\frac{m}{2n}$ subgrids. This row or column can be chosen in $2 \cdot \frac{m}{n}$ ways, so there are $\frac{nt}{2m}$ layers having in the same row/column $\frac{m}{2n}$ red/blue subgrids. Let's say w.l.o.g. that it is the first row. Then, the first row of such a layer contains at least $\binom{m/(2n)}{n}$ tuples of n red subgrids. By pigeonhole principle, there are at least

$$\frac{\frac{nt}{2m}\binom{m/(2n)}{n}}{\binom{m/n}{n}} \geq n$$

layers having the same red n-tuples, as desired.
The induction step: suppose $d \geq 4$ and that the lemma holds for $d-1$.

It is easy to see that the desired estimate $F_d(n) \leq 2^{(c_d+o(1))n^{d-2}}$ follows from the induction hypothesis and the following recursive bound

$$F_d(n) \leq d^d F_{d-1}(n)^{(d-1)n+1} \text{ for every } d \geq 4 \text{ and } n \geq 2.$$

We now prove this inequality. Let $m = F_{d-1}(n)$, $t = (d-1)!n\binom{m}{n}^{d-1}$, and $N = d^2mt \leq d^d F_{d-1}(n)^{(d-1)n+1}$. Consider an increasing array $f : [N]^d \to \mathbb{R}$.

By Lemma 4.3.8, we can assume w.l.o.g. that $[N]^d$ contains a stack of t subgrids $\boldsymbol{A}_1 = B_1 \times \ldots \times B_{d-1} \times \{h_1\}, \ldots, \boldsymbol{A}_t = B_1 \times \ldots \times B_{d-1} \times \{h_t\}$ of size $m \times \ldots \times m \times 1$ such that $h_1 < h_2 < \ldots < h_t$ and $f|_{\boldsymbol{A}_1} < f|_{\boldsymbol{A}_2} < \ldots < f|_{\boldsymbol{A}_t}$.

Given $i \in [t]$, as $m = F_{d-1}(n)$, one can find a permutation $\sigma \in \mathfrak{S}_{d-1}$ and a subgrid $\boldsymbol{A}'_i \subset \boldsymbol{A}_i$ of size $n \times \ldots \times n \times 1$ such that $f|_{\boldsymbol{A}'_i}$ is of type σ. Since $\frac{t}{(d-1)!\binom{m}{n}^{d-1}} = n$, the pigeonhole principle implies the existence of a permutation $\sigma \in \mathfrak{S}_{d-1}$, an $n \times \ldots \times n$ subgrid $B'_1 \times \ldots \times B'_d$ of $B_1 \times \ldots \times B_d$, and n layers $1 \leq i_1 < \ldots < i_n \leq t$ such that for every $k \in [n]$, we have that $\boldsymbol{A}'_{i_k} = B'_1 \times \ldots \times B'_{d-1} \times \{h_{i_k}\}$, and that the restriction of f to $\boldsymbol{A}'_{i_k}$ is of type σ. As $f|_{\boldsymbol{A}_{i_1}} < \ldots < f|_{\boldsymbol{A}_{i_n}}$, the restriction of f to $B'_1 \times \ldots \times B'_{d-1} \times \{h_{i_1}, \ldots, h_{i_n}\}$ is an $n \times \ldots \times n$ array of type (d, σ). This shows $F_d(n) \leq N \leq d^d F_{d-1}(n)^{(d-1)n+1}$, as required. $\square$

4.4 CONCLUDING REMARKS

We obtain a major improvement on best known upper bounds for $M_d(n)$. However, our bounds are still off from the best known lower bound of $M_d(n) \geq n^{(1+o(1))n^{d-1}/d}$ due to Fishburn and Graham [105](Theorem 3). Perhaps the most interesting open question regarding $M_d(n)$ is to determine the behaviour in 2 dimensions.

Question 4.4.1. *What is the behaviour of $M_2(n)$? Is it closer to exponential or to double exponential in n?*

It is natural to ask whether our argument used to get a double exponential bound in the monotone case in 3 dimensions (Theorem 4.2.7) generalises to higher dimensions. Unfortunately, the natural generalisation of our approach to more dimensions gives a bound of the form $M_d(n) \leq \mathrm{towr}_{\lfloor d/2 \rfloor + 2}(O_d(n))$ which has a tower of height growing with d. However, this does still imply a better bound than Theorem 4.2.6 in 4 and 5 dimensions. The main issue preventing us from extending our argument to more dimensions is the fact that it seems hard to obtain asymmetric results which would allow us to find a monotone subarray with exponential size in at least 2 dimensions. For example, if we could find an $n \times 2^n \times 2^{n^2}$ monotone subarray within any array of size $2^{2^{O(n^2)}} \times 2^{2^{O(n^2)}} \times 2^{2^{O(n^2)}}$ we would obtain a double exponential bound $M_4(n) \leq 2^{2^{O(n^3)}}$. However, if we knew how to do this then by considering an array which is always increasing in the first dimension and has the same but arbitrary ordering for each 2-dimensional subarray with fixed value in the first dimension, we would also be able to get a better than double exponential bound in the 2-dimensional case, which leads us back to Question 4.4.1. Our better bounds in 3, 4 and 5 dimensions make it seem unlikely that a triple exponential is ever needed.

Question 4.4.2. *For $d \geq 4$ is $M_d(n)$ bounded from above by a double exponential in n^{d-1}?*

For the problem of determining $F_d(n)$, we completely settle the 2-dimensional case and give exponential upper bounds for $d \geq 3$. The best known lower bound $F_d(n) \geq (n-1)^d$, also due to Fishburn and Graham [105] is still only polynomial. We find the 3-dimensional case particularly interesting since via Lemma 4.3.8 it reduces to the following nice problem.

Question 4.4.3. *What is the smallest N such that given N increasing arrays of size $N \times N$ one can find an $n \times n$ lexicographic array of the same type appearing in the same positions in at least n of the arrays?*

In particular, is this N bounded by a polynomial in n or is it exponential in n.

The study of $M_d(n), F_d(n)$ and $L_d(n)$ while interesting in its own right is also closely related to various other interesting problems. We present just a few here.

4.4.1 *Long common monotone subsequence*

The problem of estimating $M_2(n)$ is closely related to the longest common monotone subsequence problem. A *common monotone subsequence* of two permutations $\pi, \sigma \in \mathfrak{S}_N$ is a set $I \subseteq [N]$ such that the restrictions of π and σ to I are either both increasing or both decreasing. A common monotone subsequence of more than two permutations is defined analogously. Given positive integers t, k and N, let $\mathrm{LMS}(t, k, N)$ denote the maximum ℓ such that any size-k multisubset $P \subseteq \mathfrak{S}_N$ contains a size-t multisubset P' such that the length of the longest common monotone subsequence of P' is at least ℓ.

We now describe the connection between $M_2(n)$ and $\mathrm{LMS}(t, k, N)$. Let $f\colon [N]^2 \to \mathbb{R}$ be a 2-dimensional array. Similarly to the first part of the proof of Theorem 4.2.3, we can show that $[N]^2$ contains a subgrid $R \times C$ of size $(\log N)^{1-o(1)} \times N^{1-o(1)}$ such that either $f|_{R\times C}$ is increasing in each column, or $f|_{R\times C}$ is decreasing in each column. For each $r \in R$, the restriction of f to the row $\{r\} \times C$ induces a permutation π_r of C, since f is assumed to be injective. (Note that π_r's are not necessarily distinct.) It is not hard to see that if among $(\log N)^{1-o(1)}$ permutations $\{\pi_r : r \in R\}$ of C there are n permutations whose longest common mononotone subsequence has length at least n, then $f|_{R\times C}$ contains an $n \times n$ monotone subarray. Therefore every $N \times N$ array contains a monotone subarray of size $n \times n$, where n is the maximum $t \in \mathbb{N}$ such that $\mathrm{LMS}(t, (\log N)^{1-o(1)}, N^{1-o(1)})$ is greater than or equal to t. By an iterative application of the Erdős-Szekeres theorem, one can take $n = (1/2 - o(1)) \log_2 \log_2 N$, or equivalently $N = 2^{2^{(2+o(1))n}}$.

The problem of determining the parameter $\mathrm{LMS}(t, k, N)$ for other ranges of t and k is also very appealing. For example, it would be interesting to have a good estimate for $\mathrm{LMS}(t, k, N)$ when t is fixed and k grows to infinity with N. We refer the reader to [28, 29, 46] for some related results in this direction.

4.4.2 *Ramsey type problems for vertex-ordered graphs*

One can place the problems we have studied in this chapter under the framework of (vertex-)ordered Ramsey numbers. For simplicity of presentation we choose to illustrate this through the 2-dimensional monotone subarray problem. Let $K^{(3)}_{N,N}$ be the 3-uniform hypergraph with vertex set $A \cup B$, where A and B are two copies of $[N]$, and edge set consisting of all those triples which intersect both A and B. Given an array $f : [N]^2 \to \mathbb{R}$, we can associate to f an edge-colouring χ of $K^{(3)}_{N,N}$ with two colours red and blue. For $i \in A$ and $j, j' \in B$ with $j < j'$, let $\chi(i, j, j') =$ red if and only if $f(i, j) < f(i, j')$. Similarly, for $j \in B$ and $i, i' \in A$ with $i < i'$, we assign colour red to (i, i', j) if and only if $f(i, j) < f(i', j)$. The following simple observation connects the multidimensional Erdős-Szekeres world with the ordered Ramsey world.

Observation 4.4.4. *Suppose there are two size-n subsets $\{a_1 < \ldots < a_n\} \subseteq A$, $\{b_1 < \ldots < b_n\} \subseteq B$ s.t.*

- $\{(a_i, b_j, b_{j+1}) : i \in [n], j \in [n-1]\}$ *is monochromatic under* χ,
- $\{(a_i, a_{i+1}, b_j) : j \in [n], i \in [n-1]\}$ *is monochromatic under* χ.

Then the restriction of f to $\{a_1, \ldots, a_n\} \times \{b_1, \ldots, b_n\}$ is monotone.

Now define $\mathrm{OR}(n)$ to be the smallest N such that in every red-blue colouring of the edges of $K^{(3)}_{N,N}$ we can always find two size-n subsets $\{a_1 < \ldots < a_n\} \subseteq A$ and $\{b_1 < \ldots < b_n\} \subseteq B$ with the aforementioned properties. From the observation, we know $M_2(n) \le \mathrm{OR}(n)$. A closer inspection of our proof of the inequality $M_2(n) \le 2^{2^{(2+o(1))n}}$ reveals that it actually gives $\mathrm{OR}(n) \le 2^{2^{(2+o(1))n}}$. Thus it is natural to ask whether $M_2(n)$ and $\mathrm{OR}(n)$ have the same order of magnitude.

4.4.3 *Canonical orderings of discrete structures*

An ordering of the edges of a (vertex-ordered) d-graph G with $V(G) \subset \mathbb{Z}$ is *lex-monotone* if one can find a permutation $\sigma \in \mathfrak{S}_d$ and a sign vector $\boldsymbol{s} \in \{-1, 1\}^d$ such that the edges $(a_1, \ldots, a_d)$ of G with $a_1 < \ldots < a_d$ are ordered according to the lexicographical order of the tuple $(s_{\sigma(1)} a_{\sigma(1)}, \ldots, s_{\sigma(d)} a_{\sigma(d)})$. An old result of Leeb and Prömel (see [192](Theorem 2.8)) says that for every $d, n \in \mathbb{N}$ there is a positive integer $\mathrm{LP}_d(n)$ such that every edge-ordering of a (vertex-ordered) complete d-graph on $\mathrm{LP}_d(n)$ vertices contains a copy of the complete d-graph on n vertices whose edges induce a lex-monotone ordering. Theorem 4.1.3 can be viewed naturally as a d-partite version (with a better bound) of this result. It would be interesting to know if our approach can lead to an improvement on the upper bound $\mathrm{LP}_d(n) \leq \mathrm{towr}_{2d}(O_d(n))$ for $d \geq 2$, due to Nešetřil and Rödl [193](Theorem 14). For other interesting results in edge-ordered Ramsey numbers, we refer the reader to [22, 109].

Theorem 4.1.3 is also related to the work of Nešetřil, Prömel, Rödl and B. Voigt [191] on linear orders of the combinatorial cube $[k]^n$ when k is fixed and n is large. For simplified presentations of this work, see [45, 205].

5

ROTA'S BASIS CONJECTURE FOR MATROIDS

5.1 INTRODUCTION

Given bases $B_1, \ldots, B_n$ in an n-dimensional vector space V, a *transversal basis* is a basis of V containing a single distinguished vector from each of $B_1, \ldots, B_n$. Two transversal bases are said to be *disjoint* if their distinguished vectors from B_i are distinct, for each i. In 1989, Rota conjectured (see [146](Conjecture 4)) that for any vector space V over a characteristic-zero field, and any choice of $B_1, \ldots, B_n$, one can always find n pairwise disjoint transversal bases.

Despite the apparent simplicity of this conjecture, it remains wide open, and has surprising connections to apparently unrelated subjects. Specifically, it was discovered by Huang and Rota [146] that there are implications between Rota's basis conjecture, the Alon–Tarsi conjecture [18] concerning enumeration of even and odd Latin squares, and a certain conjecture concerning the supersymmetric bracket algebra.

Rota also observed that an analogous conjecture could be made in the much more general setting of *matroids*, which are objects that abstract the combinatorial properties of linear independence in vector spaces. Specifically, a finite matroid $M = (E, \mathcal{I})$ consists of a finite ground set E (whose elements may be thought of as vectors in a vector space), and a collection $\mathcal{I}$ of subsets of E, called independent sets. The defining properties of a matroid are that:

- the empty set is independent (that is, $\emptyset \in \mathcal{I}$);
- subsets of independent sets are independent (that is, if $A' \subseteq A \subseteq E$ and $A \in \mathcal{I}$, then $A' \in \mathcal{I}$);
- if A and B are independent sets, and $|A| > |B|$, then an independent set can be constructed by adding an element of A to B (that is, there is $a \in A \backslash B$ such that $B \cup \{a\} \in \mathcal{I}$). This final property is called the *augmentation property*.

Observe that any finite set of elements in a vector space (over any field) naturally gives rise to a matroid, though not all matroids arise this way. A *basis* in a matroid M is a maximal independent set. By the augmentation property, all bases have the same size, and this common size is called the *rank* of M. The definition of a transversal basis generalises in the obvious way to matroids, and the natural matroid generalisation of Rota's basis conjecture is that for any rank-n matroid and any bases $B_1, \ldots, B_n$, there are n disjoint transversal bases.

Although Rota's basis conjecture remains open, various special cases have been proved. Several of these have come from the connection between Rota's basis conjecture and the Alon–Tarsi conjecture, which has since been simplified by Onn [196]. Specifically, due to work by Drisko [75] and Glynn [127] on the Alon–Tarsi conjecture, Rota's original conjecture for vector spaces over a characteristic-zero field is now known to be true whenever the dimension n is of the form $p \pm 1$, for p a prime. Wild [239] proved Rota's basis conjecture for so-called "strongly base-orderable" matroids, and used this to prove the conjecture for certain classes of matroids arising from graphs. Geelen and Humphries proved the conjecture for "paving" matroids [121], and Cheung [51] computationally proved that the conjecture holds for matroids of rank at most 4.

Various authors have also proposed variations and weakenings of Rota's basis conjecture. For example, Aharoni and Berger [8] showed that in any matroid one can cover the set of all the elements in $B_1, \ldots, B_n$ by at most $2n$ "partial" transversals, and Bollen and Draisma [37] considered an "online" version of Rota's basis conjecture, where the bases B_i are revealed one-by-one. In 2017, Rota's basis conjecture received renewed interest when it was chosen as the twelfth "Polymath" project, in which amateur and professional mathematicians from around the world collaborated on the problem. Some of the fruits of the project were a small improvement to Aharoni and Berger's theorem, and improved understanding of the online version of Rota's basis conjecture [202]. See [52] for Timothy Chow's proposal of the project, see [53, 54] for blog posts where much of the discussion took place, and see [55] for the Polymath wiki summarising most of what is known about Rota's basis conjecture.

One particularly natural direction to attack Rota's problem is to try to find lower bounds on the number of disjoint transversal bases. Rota's basis conjecture asks for n disjoint transversal bases, but it is not completely obvious that even two disjoint transversal bases must exist! Wild [239] proved some lower bounds for certain matroids arising from graphs, but the first nontrivial bound for general matroids was by Geelen and Webb [122], who used a generalisation of Hall's theorem due to Rado [207] to prove that there must be $\Omega\left(\sqrt{n}\right)$ disjoint transversal bases. Recently, this was improved by Dong and Geelen [74], who used a beautiful probabilistic argument to prove the existence of $\Omega\left(n/\log n\right)$ disjoint transversal bases. Here we improve this substantially and obtain the first linear bound.

Theorem 5.1.1. *For any $\varepsilon > 0$, the following holds for sufficiently large n. Given bases $B_1, \ldots, B_n$ of a rank-n matroid, there are at least $(1/2 - \varepsilon)\, n$ disjoint transversal bases.*

Of course, since matroids generalise vector spaces, this also implies the same result for bases in an n-dimensional vector space. We also remark that for the weaker fact that there exist $\Omega(n)$ disjoint transversal bases, our methods give a simpler proof.

In contrast to the previous work by Dong, Geelen and Webb, our approach is to show how to build a collection of transversal bases in an iterative fashion (reminiscent of augmenting path arguments in matching problems). It is tempting to imagine a future path to Rota's basis conjecture (at least in the case of vector spaces) using such an approach: by improving on our arguments, perhaps introducing some randomness, it might be possible to iteratively build a collection of $(1 - o(1))\, n$ transversal bases, and then it might be possible to use some sort of "template" or "absorber" structure to finish the job. This was precisely the approach taken in Keevash's celebrated proof of the existence of designs [198]. Actually, it has been observed by participants of the Polymath project (see [53]) that Rota's basis conjecture and the existence of designs conjecture both seem to fall into a common category of problems which are not quite "structured" enough for purely algebraic methods, but too structured for probabilistic methods.

Notation. We will frequently want to denote the result of adding and removing single elements from a set. For a set S and some $x \notin S$, $y \in S$, we write $S + x$ to mean $S \cup \{x\}$, and we write $S - y$ to mean $S \setminus \{y\}$.

5.2 FINDING MANY DISJOINT TRANSVERSAL BASES

In this section we prove theorem 5.1.1. It is convenient to think of $B_1, \ldots, B_n$ as "colour classes".

Definition 5.2.1. Let $U = \{(x, c) : x \in B_c, 1 \leq c \leq n\}$ be the set of all coloured elements that appear in one of $B_1, \ldots, B_n$. For $S \subseteq U$, let $\pi(S) = \{x : (x, c) \in S \text{ for some } c\}$ be its set of matroid elements. We say that a subset of elements of U is a *rainbow independent set* (RIS for short) if all its matroid elements are distinct and form an independent set, and all their colours are distinct.

Note that an RIS with size n corresponds to a transversal basis. We remark that RISs are sometimes also known as *partial transversals*. Note that two transversal bases are disjoint if and only if their corresponding RISs are disjoint as subsets of U.

Let $f = (1 - \varepsilon)\, n/2$. The basic idea is to start with a collection of f empty RISs (which are trivially disjoint), and iteratively enlarge the RISs in this collection, maintaining disjointness, until we have many disjoint transversal bases.

Let $\mathcal{S}$ be a collection of f disjoint RISs. We define the *volume* $\sum_{S \in \mathcal{S}} |S|$ of $\mathcal{S}$ to be the total number of elements in the RISs in $\mathcal{S}$. We will show how to modify $\mathcal{S}$ to increase its volume. We let $F = \bigcup_{S \in \mathcal{S}} S$ be the set of all currently used elements. One should think of F as being the set of all elements which we cannot add to any $S \in \mathcal{S}$ without violating the disjointness of RISs in $\mathcal{S}$.

We stress that in the following two subsections we fix a collection $\mathcal{S}$ and define F as above. All our definitions and claims are with respect to these F and $\mathcal{S}$. We will show that under certain conditions the size of $\mathcal{S}$ can be increased, at which point one needs to restart the argument from the beginning with a new $\mathcal{S}$ (and a new F). This is made precise in Section 5.2.3.

Remark. We remark that it is actually possible to reduce to the case where each B_c is disjoint, by making duplicate copies of all elements that appear in multiple B_c. So, instead of working with the universe U of element/colour pairs, one can alternatively think of U as being a collection of n^2 different matroid elements (each of which has a colour associated with it).

5.2.1 *Simple swaps*

Our objective is to increase the volume of $\mathcal{S}$. If an RIS $S \in \mathcal{S}$ is missing a colour c and there is $x \in B_c$ independent to the elements of S, such that $(x, c) \notin F$, then we can add (x, c) to S to create a larger RIS, increasing the volume of $\mathcal{S}$. We will want much more freedom than this: we also want to consider those elements that can be added to S after making a small change to S. This motivates the following definition.

Definition 5.2.2. Consider an RIS S and a colour b that does not appear in S. Say an element $(x, c) \in U$ (possibly $(x, c) \in F$) is (S, b)-*addable* if either

- $S + (x, c)$ is an RIS, or;
- There is $(x', c) \in S$ and $(y, b) \notin F$ such that $S - (x', c) + (y, b) + (x, c)$ is an RIS.

In the second case we say that y is a *witness* for the (S, b)-addability of (x, c). For $(x', c) \in S$ and $(y, b) \notin F$ when $S - (x', c) + (y, b)$ is an RIS we say it is the result of applying a *simple swap* to S.

If for some RIS $S \in \mathcal{S}$ missing a colour b there is an (S, b)-addable element $(x, c) \notin F$, then we can increase the volume of $\mathcal{S}$ by adding (x, c) to S, possibly after applying a simple swap to S. Note that we do not require $S \in \mathcal{S}$ for the definition of (S, b)-addability, though in practice we will only ever consider S that are either in $\mathcal{S}$ or slight modifications of RISs in $\mathcal{S}$.

Our next objective is to show that for any S missing a colour b, either there is an (S, b)-addable element that is not in F (which would allow us to increase the volume of $\mathcal{S}$, as above), or else there are *many* (S, b)-addable elements (which must therefore be in F). Although this will not allow us to immediately increase the volume of $\mathcal{S}$, it will allow us to transfer an element to S from some other $S' \in \mathcal{S}$, and this freedom to perform local modifications will be very useful.

Towards this end, we study which elements of S can be used in a simple swap.

Definition 5.2.3. Consider an RIS S and consider a colour b that does not appear on S. We say that a colour c appearing on S is (S, b)-*swappable* if there is a simple swap yielding an RIS $S + (y, b) - (x', c)$, with $(y, b) \notin F$ and $(x', c) \in S$. (For $S + (y, b) - (x', c)$ to be an RIS, we just need $\pi(S) + y - x'$ to be an independent set in our matroid.) We say that y is a witness for the (S, b)-swappability of c.

(Basically, a colour is (S, b)-swappable if we can replace it with a b-coloured element which is not in F). For a colour c we denote by $F_c = \{x \in B_c : (x, c) \in F\}$ the set of matroid elements which appear in $\mathcal{S}$ with colour c.

Claim 3. *For a nonempty RIS S and a colour b not appearing in S, either there is an (S, b)-addable element $(y, b) \notin F$ or there are at least $n - |F_b|$ colours which are (S, b)-swappable.*

Proof. For the purpose of contradiction, suppose that there is no (S, b)-addable element $(y, b) \notin F$, and that there are fewer than $n - |F_b|$ colours which are (S, b)-swappable. Let $S' \subseteq S$ be the set of all elements of S which have an (S, b)-swappable colour, so $|S'| < n - |F_b|$. Also $|S'| < |S|$ because otherwise we would have $|S| < n - |F_b|$, so by the augmentation property there would be $y \in B_b \setminus F_b$ such that $S + (y, b)$ is an RIS (meaning that $(y, b) \notin F$ would be (S, b)-addable). Repeating this argument for S' in place of S, there is $y \in B_b \setminus F_b$ such that $S' + (y, b)$ is an RIS. By repeatedly using the augmentation property, we can add $|S - S'| - 1$ elements of $S - S'$ to $S' + (y, b)$. This gives an RIS of size $|S|$ of the form $S + (y, b) - (x', c)$ for some $(x', c) \in S - S'$. But this means c is (S, b)-swappable, so $(x', c) \in S'$ by the definition of S'. This is a contradiction. $\square$

Now we show that all elements of an (S, b)-swappable colour which are independent to $\pi(S)$ are (S, b)-addable, unless there is an (S, b)-addable element not in F. (Recall that $\pi(S)$ is the set of matroid elements in S, without colour data.)

Claim 4. *Consider an RIS S with no element of a colour b and consider a colour c that is (S, b)-swappable with witness y. Either $S + (y, b)$ is an RIS (thus, $(y, b) \notin F$ is (S, b)-addable), or otherwise for any $x \in B_c$ independent of $\pi(S)$, (x, c) is (S, b)-addable with witness (y, b).*

Proof. Let (x', c) be the element with colour c in S. Consider some $x \in B_c$ independent to $\pi(S)$. Let $I = \pi(S) + x$ and $J = \pi(S) + y - x'$. By the augmentation property, there is an element of $I \backslash J$ that is independent of J; this element is either x' or x. In the former case $S + (y, b)$ is an RIS. In the latter case, $S + (y, b) - (x', c) + (x, c)$ is an RIS, showing that (x, c) is (S, b)-addable. $\square$

The following lemma gives a good illustration of how to use the ideas developed in this section to find many addable elements. It will be very useful later on.

Claim 5. *Let $S \in \mathcal{S}$ and let b be a colour which does not appear in S. Then either we can increase the volume of $\mathcal{S}$ or there are at least $(n - |S|)(n - f)$ elements that are (S, b)-addable.*

Proof. If there is an element $(y, b) \notin F$ which is (S, b)-addable, then we can directly add this element to S (making a simple swap if necessary), increasing the volume of $\mathcal{S}$. Otherwise, observe that $|F_b| \leq |\mathcal{S}| = f$, so by Claim 3 there are at least $n - f$ colours that are (S, b)-swappable. For each such colour c, by the augmentation property, there are at least $n - |S|$ elements $x \in B_c$ independent to all the elements of S, each of which is (S, b)-addable by Claim 4. That is to say, there are at least $(n - |S|)(n - f)$ elements which are (S, b)-addable, as claimed. $\square$

In our proof of theorem 5.1.1 we also make use of the following lemma. In the course of our arguments, when we need to find many addable elements with the colour, it will allow us to ensure that these elements are actually distinct.

Lemma 5.2.4. *Let S be an RIS. Then for each B_b, we can find an injection $\phi_b : S \to B_b$ such that for all $(x, c) \in S$, $\phi_b((x, c))$ is independent of $\pi(S - (x, c))$.*

Proof. Consider the bipartite graph G where the first part consists of the elements of S and the second part consists of the elements of B_b, with an edge between $(x, c) \in S$ and $y \in B_b$ if y is independent of $\pi(S - (x, c))$. We use Hall's theorem to show that there is a matching in this bipartite graph covering S. Indeed, consider some $W \subseteq S$. By the augmentation property, there are at least $|W|$ elements $y \in B_b$ such that $\pi(S - W) + y$ is an independent set, and again using the augmentation property, each of these can be extended to an independent set of the form $\pi(S) + y - x$ for some $(x, c) \in W$. That is to say, W has at least $|W|$ neighbours in G. $\square$

5.2.2 *Cascading swaps*

Informally speaking, for any $S_0 \in \mathcal{S}$ which is not a transversal basis, we have shown that either we can directly augment S_0, or there are many elements $(x_1, c_1) \in U$ with which we can augment S_0 after performing a simple swap. It's possible that each such (x_1, c_1) already appears in some other $S_1 \in \mathcal{S}$, but if this occurs we need not give up: we can transfer (x_1, c_1) from S_1 to S_0 and then continue to look for elements $(x_2, c_2) \in U$ with which we can augment $S_1 - (x_1, c_1)$ (again, possibly with a swap). We can iterate this idea, looking for sequences

$$S_1, \ldots, S_\ell \in \mathcal{S}; \qquad (x_1, c_1) \in S_1,\ (x_2, c_2) \in S_2, \ldots, (x_\ell, c_\ell) \in S_\ell,$$
$$(x_{\ell+1}, c_{\ell+1}) \notin \bigcup_{S \in \mathcal{S}} S$$

such that, after a sequence of simple swaps, each (x_i, c_i) is transferred from S_i to S_{i-1}, and then $(x_{\ell+1}, c_{\ell+1})$ can be added to S_ℓ. (We also need to ensure that the simple swaps we perform preserve disjointness of RISs in $\mathcal{S}$.) This transformation has the net effect of adding an element to S_0 and keeping the size of all other $S \in \mathcal{S}$ constant, thus increasing the volume of $\mathcal{S}$.

Crucially, because of the freedom afforded by simple swaps, each time we expand our search to consider longer cascades, our number of options for $(x_{\ell+1}, c_{\ell+1})$ increases. For sufficiently large ℓ, the number of options will be so great that there must be suitable $(x_{\ell+1}, c_{\ell+1})$ not appearing in any RIS in $\mathcal{S}$. In order to keep this analysis tractable, we will only consider transformations that cascade along a single sequence of RISs $S_0, \ldots, S_\ell$; we will iteratively construct this sequence of RISs in such a way that there are many possibilities $(x_i, c_i) \in S_i$ relative to the number of possibilities $(x_{i-1}, c_{i-1}) \in S_{i-1}$ in the previous step. The next definition makes precise the cascades that we consider.

Definition 5.2.5. Consider a sequence of distinct RISs $S_0, \ldots, S_{\ell-1} \in \mathcal{S}$. Say an element $(x_\ell, c_\ell) \notin S_0, \ldots, S_{\ell-1}$ is *cascade-addable with respect to* $S_0, \ldots, S_{\ell-1}$ if there is a colour c_0 and sequences

$$(x_1, c_1), \ldots, (x_{\ell-1}, c_{\ell-1}) \in U, \qquad y_0 \in B_{c_0}, \ldots, y_{\ell-1} \in B_{c_{\ell-1}},$$

such that the following hold.

- For each $1 \leq i \leq \ell - 1$, we have $(x_i, c_i) \in S_i$;
- c_0 does not appear in S_0, and (x_1, c_1) is (S_0, c_0)-addable with witness y_0;
- for each $0 \leq i \leq \ell - 1$, (x_{i+1}, c_{i+1}) is $(S_i - (x_i, c_i), c_i)$-addable with witness y_i;
- the colours $c_0, \ldots, c_\ell$ are distinct.

We call $c_0, c_1, \ldots, c_{\ell-1}$ *a sequence of colours freeing* (x_ℓ, c_ℓ).

We write $Q(S_0, \ldots, S_{\ell-1})$ for the set of all elements outside $S_0, \ldots, S_{\ell-1}$ which are cascade-addable with respect to $S_0, \ldots, S_{\ell-1}$.

We remark that if $\ell = 1$ then most of the conditions in the above definition become vacuous and an element being cascade-addable with respect to S_0 is equivalent to it being (S_0, c_0)-addable with a witness, for some colour c_0. Observe that if an element (x_ℓ, c_ℓ) is cascade-addable then we can transfer it into $S_{\ell-1}$, as the final step in a cascading sequence of simple swaps and transfers. The following lemma makes this precise.

Claim 6. *Suppose that (x_ℓ, c_ℓ) is cascade-addable with respect to $S_0, \ldots, S_{\ell-1}$ and $c_0, c_1, \ldots, c_{\ell-1}$ is a sequence of colours freeing (x_ℓ, c_ℓ). Then there are $S'_0 \ldots S'_{\ell-1} \subseteq S_0 \cup \cdots \cup S_{\ell-1} \cup B_{c_0} \cup \cdots \cup B_{c_{\ell-1}}$ such that replacing $S_0, \ldots, S_{\ell-1}$ with $S'_0, \ldots, S'_{\ell-1}$ in $\mathcal{S}$ results in a family $\mathcal{S}'$ of disjoint RISs of the same total volume as $\mathcal{S}$, in such a way that $S'_{\ell-1} + (x_\ell, c_\ell)$ is an RIS.*

Proof. Let $(x_1, c_1), \ldots, (x_{\ell-1}, c_{\ell-1}) \in U, y_0 \in B_{c_0}, \ldots, y_{\ell-1} \in B_{c_{\ell-1}}$ be as in the definition of cascade-addability. For each $i = 0, \ldots, \ell - 1$, let (x'_i, c_{i+1}) be the colour c_{i+1} element of S_i (which exists, because, from cascade-addability, (x_{i+1}, c_{i+1}) is $(S_i - (x_i, c_i), c_i)$-addable *with a witness*). For each $i = 1, \ldots, \ell - 2$, let $S'_i = S_i - (x_i, c_i) - (x'_i, c_{i+1}) + (y_i, c_i) + (x_{i+1}, c_{i+1})$. Let $S'_0 = S_0 - (x'_0, c_1) + (y_0, c_0) + (x_1, c_1)$ and $S'_{\ell-1} = S_{\ell-1} - (x_{\ell-1}, c_{\ell-1}) - (x'_{\ell-1}, c_\ell) + (y_{\ell-1}, c_{\ell-1})$. Let $\mathcal{S}'$ be the family formed by replacing $S_0, \ldots, S_{\ell-1}$ with $S'_0, \ldots, S'_{\ell-1}$ in $\mathcal{S}$. It is easy to check that $\mathcal{S}'$ has the same total volume as $\mathcal{S}$, so it remains to check that it is a family of disjoint RISs.

For $i = 1, \ldots, \ell - 2$, S'_i is an RIS because it comes from $S_i - (x_i, c_i)$ by making the change in the definition of (x_{i+1}, c_{i+1}) being $(S_i - (x_i, c_i), c_i)$-addable with witness y_i (and addability always produces an RIS by definition). Similarly $S'_{\ell-1} + (x_\ell, c_\ell)$ is an RIS. To see that S'_0 is an RIS we use that (x_1, c_1) is (S_0, c_0)-addable with witness y_0, and that c_0 does not appear in S_0, both of which come from the definition of cascade-addability.

It remains to show that the RISs $S'_0, \ldots, S'_{\ell-1}$ are disjoint from each other and the other RISs in $\mathcal{S}$. The elements (y_i, c_i) occur in only one RIS S'_i because they come from outside F (since they are addability witnesses), and because their colours $c_0, \ldots, c_{\ell-1}$ are distinct (from the definition of cascade-addability). The elements (x_i, c_i) occur in only one RIS because they get removed from S_i and added to S_{i-1}. □

The following lemma lets us build longer cascades.

Claim 7. *Suppose that $(x_\ell, c_\ell) \in S_\ell$ is cascade-addable with respect to $S_0, \ldots, S_{\ell-1}$ and $c_0, c_1, \ldots, c_{\ell-1}$ is a sequence of colours freeing (x_ℓ, c_ℓ). If (x, c) is $(S_\ell - (x_\ell, c_\ell), c_\ell)$-addable with a witness then either $(x, c) \in S_0 \cup \cdots \cup S_\ell \cup B_{c_0} \cup \cdots \cup B_{c_\ell}$ or (x, c) is cascade-addable with respect to $S_0, \ldots, S_\ell$.*

Proof. Suppose that $(x, c) \notin S_0, \ldots, S_\ell, B_{c_0}, \ldots, B_{c_\ell}$. For the definition of (x, c) being cascade-addable, all the conditions not involving (x, c) and (x_ℓ, c_ℓ) hold as a consequence of $(x_\ell, c_\ell) \in S_\ell$ being cascade-addable with respect to $S_0, \ldots, S_{\ell-1}$. It remains to check the conditions that $(x, c) \notin S_0, \ldots, S_\ell$ and that each of $c_0, \ldots, c_\ell, c$ are distinct, both of which hold as a consequence of our assumption $(x, c) \notin S_0, \ldots, S_\ell, B_{c_0}, \ldots, B_{c_\ell}$. □

In the next lemma, we essentially show that given $S_0, \ldots, S_{\ell-1}$, it is possible to choose S_ℓ in such a way that the number of cascade-addable elements increases.

Claim 8. *Consider a sequence of distinct RISs $S_0, \ldots, S_{\ell-1} \in \mathcal{S}$ with $1 \leq \ell < f = |\mathcal{S}|$. Then either we can modify $\mathcal{S}$ to increase its volume, or we can choose $S_\ell \neq S_0, \ldots, S_{\ell-1}$ from $\mathcal{S}$ such that*

$$|Q(S_0, \ldots, S_\ell)| \geq \frac{|Q(S_0, \ldots, S_{\ell-1})|}{f - \ell} \cdot (n - f - \ell) - (\ell + 1)n. \tag{5.1}$$

Proof. If $Q(S_0, \ldots, S_{\ell-1})$ contains an element (x, c) not in any $S \in \mathcal{S}$, then we can increase the volume of $\mathcal{S}$ with a cascading sequence of simple swaps and transfers (using Claim 6, noting that if $(x_\ell, c_\ell) \notin F$, then we can add (x_ℓ, c_ℓ) to $S'_{\ell-1}$ in that lemma to get a larger family of RISs).

Otherwise, all the elements of $Q(S_0, \ldots, S_{\ell-1})$ belong to some RIS $S \in \mathcal{S} \setminus \{S_0, \ldots S_{\ell-1}\}$ (since $Q(S_0, \ldots, S_{\ell-1})$ is defined to not contain any elements from $S_0, \ldots, S_{\ell-1}$). Choose $S_\ell \in \mathcal{S} \setminus \{S_0, \ldots S_{\ell-1}\}$ containing maximally many elements of $Q(S_0, \ldots, S_{\ell-1})$. Since the $f - \ell$ RISs $S \in \mathcal{S} \setminus \{S_0, \ldots S_{\ell-1}\}$ collectively contain all elements of $Q(S_0, \ldots, S_{\ell-1})$, our chosen RIS S_ℓ must contain a proportion of at least $1/(f-\ell)$ of the elements of $Q(S_0, \ldots, S_{\ell-1})$. In other words, if we let $Q = S_\ell \cap Q(S_0, \ldots, S_{\ell-1})$, we have

$$|Q| \geq \frac{|Q(S_0, \ldots, S_{\ell-1})|}{f - \ell}. \tag{5.2}$$

Apply lemma 5.2.4 to S_ℓ to obtain an injection ϕ_b, for every colour b. Fix some $(x_\ell, c_\ell) \in Q$ and a sequence of colours $c_0, \ldots, c_{\ell-1}$ freeing (x_ℓ, c_ℓ). We prove a sequence of claims about how many elements are swappable/addable with respect to $(S_\ell - (x_\ell, c_\ell), c_\ell)$, assuming we cannot increase the size of $\mathcal{S}$.

Claim. *There are at least $n - f$ colours which are $(S_\ell - (x_\ell, c_\ell), c_\ell)$-swappable.*

Proof. By Claim 3, either there is an $(S_\ell - (x_\ell, c_\ell), c_\ell)$-addable element $(y, c_\ell) \notin F$, or there are at least $n - |F_{c_\ell}| \geq n - f$ colours which are $(S_\ell - (x_\ell, c_\ell), c_\ell)$-swappable. In the former case, we can increase the volume of $\mathcal{S}$, by a cascading sequence of swaps and transfers (first consider $\mathcal{S}'$ from Claim 6, then move (x_ℓ, c_ℓ) from S_ℓ to $S'_{\ell-1}$, then add (y, c_ℓ) to $S_\ell - (x_\ell, c_\ell)$). $\square$

Claim. *There are at least $n - f$ colours c for which $(\phi_c((x_\ell, c_\ell)), c)$ is $(S_\ell - (x_\ell, c_\ell), c_\ell)$-addable.*

Proof. Let c be a colour which is $(S_\ell - (x_\ell, c_\ell), c_\ell)$-swappable with witness y, as in the previous claim. If y is independent to $\pi(S_\ell - (x_\ell, c_\ell))$, we can increase the volume of $\mathcal{S}$ by adding it to S_ℓ after a cascading sequence of swaps and transfers (first consider $\mathcal{S}'$ from Claim 6, then move (x_ℓ, c_ℓ) from S_ℓ to $S'_{\ell-1}$, then add (y, c_ℓ) to $S_\ell - (x_\ell, c_\ell)$). Otherwise, by Claim 4 applied with $b = c_\ell$, $S = S_\ell - (x_\ell, c_\ell)$, the element $(\phi_c((x_\ell, c_\ell)), c)$ is $(S_\ell - (x_\ell, c_\ell), c_\ell)$-addable. Here we are using that $(\phi_c((x_\ell, c_\ell)), c)$ is independent from $(S_\ell - (x_\ell, c_\ell), c_\ell)$ (which comes from the definition of ϕ_c in lemma 5.2.4). $\square$

Claim. *There are at least $n - f - \ell$ colours $c \notin \{c_0, \ldots, c_{\ell-1}\}$ for which $(\phi_c((x_\ell, c_\ell)), c)$ is $(S_\ell - (x_\ell, c_\ell), c_\ell)$-addable.*

Proof. This ensues from the previous claim and the fact that the only requirement on c, besides addability, is that it is different from the ℓ colours in $\{c_0, \ldots, c_{\ell-1}\}$. $\square$

We now prove the following:

$$|Q(S_0, \ldots, S_\ell)| \geq |Q|(n - \ell - f) - (\ell + 1)n. \tag{5.3}$$

From the last claim, we have $|Q|(n - \ell - f)$ elements of the form $(\phi_c((x_\ell, c_\ell)), c)$ which are all $(S_\ell - (x_\ell, c_\ell), c_\ell)$-addable, with c outside a sequence of colours freeing (x_ℓ, c_ℓ). Notice that these $(\phi_c((x_\ell, c_\ell)), c)$ are all distinct because ϕ_c is an injection. By Claim 7, each of these is cascade-addable with respect to $S_0, \ldots, S_\ell$, unless it appears in one of $S_0, \ldots, S_\ell$. The total number of elements in $S_0, \ldots, S_\ell$ is at most $(\ell + 1)n$, so we have found $|Q|(n - \ell - f) - (\ell + 1)n$ cascade-addable elements with respect to $S_0, \ldots, S_\ell$, as required by eq. (5.3).

The lemma immediately follows by combining eq. (5.2) and eq. (5.3). □

Now, we want to iteratively apply Claim 8 starting from some $S_0 \in \mathcal{S}$, to obtain a sequence $S_0, S_1, \ldots, S_h \in \mathcal{S}$. There are two ways this process can stop: either we find a way to increase the volume of $\mathcal{S}$, in which case we are done, or else we run out of RISs in $\mathcal{S}$ (that is, $h = f - 1$). We want to show that this latter possibility cannot occur by deducing from eq. (5.1) that the $|Q(S_0, \ldots, S_\ell)|$ increase in size at an exponential rate: after logarithmically many steps there will be so many cascade-addable elements that they cannot all be contained in the RISs in $\mathcal{S}$, and it must be possible to increase the volume of $\mathcal{S}$.

A slight snag with this plan is that eq. (5.1) only yields an exponentially growing recurrence if the "initial term" is rather large. To be precise, let C (depending on ε) be sufficiently large such that

$$C(1 + \varepsilon/2)^{\ell-1} \frac{1}{1 - \varepsilon} - \ell - 1 \geq C(1 + \varepsilon/2)^{\ell} \tag{5.4}$$

for all $\ell \geq 1$.

Claim 9. *For $S_0, \ldots, S_h$ as above, suppose that $|Q(S_0)| \geq Cn$ or $|Q(S_0, S_1)| \geq Cn$. Then, for $0 < \ell \leq \min\{h, \varepsilon n/2\}$, we have*

$$|Q(S_0, \ldots, S_\ell)| \geq C(1 + \varepsilon/2)^{\ell-1} n.$$

Proof. We first establish a technical inequality. Recall that $f = (1 - \varepsilon)n/2$, so

$$\frac{n - f - \ell}{f - \ell} \geq \frac{n - (1 - \varepsilon)n/2 - n\varepsilon/2}{(1 - \varepsilon)n/2} = \frac{1}{1 - \varepsilon}. \tag{5.5}$$

Now, let $Q_\ell = Q(S_0, \ldots, S_\ell)$. We proceed by induction. First observe that if $|Q_0| \geq Cn$ then eq. (5.1), eq. (5.5) and eq. (5.4) for $\ell = 1$ imply $|Q_1| \geq$

$Cn(n-f-1)/(f-1)-2n \geq (C/(1-\varepsilon)-2)n \geq Cn$, giving us the base case. If $|Q_\ell| \geq C\left(1+\varepsilon/2\right)^{\ell-1} n$, then once again using eq. (5.1), eq. (5.5) and eq. (5.4), we obtain

$$\begin{aligned}
|Q_{\ell+1}| &\geq \frac{C\left(1+\varepsilon/2\right)^{\ell-1} n}{f-\ell} \cdot (n-f-\ell) - (\ell+1)n \\
&= \left(C\left(1+\varepsilon/2\right)^{\ell-1} \frac{(n-f-\ell)}{f-\ell} - \ell - 1\right) n \\
&\geq \left(C\left(1+\varepsilon/2\right)^{\ell-1} \frac{1}{1-\varepsilon} - \ell - 1\right) n \\
&\geq C\left(1+\varepsilon/2\right)^{\ell} n.
\end{aligned}$$

□

If we could choose S_0, S_1 such that $|Q(S_0)| \geq Cn$ or $|Q(S_0, S_1)| \geq Cn$, then Claim 9 would imply that during the construction of $S_1, \ldots, S_h$ we never run out of RISs in $\mathcal{S}$ (that is, $h < f-1$). Indeed, otherwise $Q(S_0, \ldots, S_{\varepsilon n/2})$ would have size exponential in n, which is impossible. Therefore, the process must stop at some point when we find a way to increase the volume of $\mathcal{S}$. Provided we can again find suitable S_0, S_1 we can then repeat the arguments in this section, further increasing the volume of $\mathcal{S}$. After repeating these arguments enough times we will have obtained $f = (1-\varepsilon)\,n/2 \geq (1/2-\varepsilon)\,n$ disjoint transversal bases, completing the proof of theorem 5.1.1.

There may not exist suitable $S_0, S_1 \in \mathcal{S}$, but in the next section we will show that if at least $\varepsilon n/2$ of the RISs in S are not transversal bases, then it is possible to modify $\mathcal{S}$ without changing its volume, in such a way that suitable S_0, S_1 exist.

Remark. With the results we have proved so far, we can already find linearly many disjoint transversal bases. Indeed, if S_0 is not a transversal basis (missing a colour b, say), and the volume of $\mathcal{S}$ cannot be increased by adding an element to S_0 (possibly after a simple swap), then Claim 5 implies that there are at least $n-f$ elements which are (S_0, b)-addable, meaning that $|Q(S_0)| \geq n-f$. Take for example $\varepsilon = 4/5$, meaning that $f \leq n/10$ and $|Q(S_0)| \geq 9n/10$. We can check that eq. (5.4) holds for all $\ell \geq 1$ if $C = 9/10$. That is to say, as long as we have not yet completed $\mathcal{S}$ to a collection of disjoint transversal bases, we can keep increasing its volume without the considerations in the next section. This proves already that it is possible to find linearly many disjoint transversal bases.

Remark. It is not hard to add a term $(n-|S_\ell|)(n-f)$ to the right hand side of the inequality given by Claim 8 by considering also cascades along the sequence $S_0, \ldots, S_{\ell-1}$ of length strictly less than ℓ. However, since this increase is only significant when $|S_\ell|$ is not close to n, which may never be the case, we omit it from our argument for the sake of readability.

5.2.3 *Increasing the number of initial addable elements*

Consider a collection $\mathcal{S}$ of $f = (1-\varepsilon)\,n/2$ disjoint RISs, at least $\varepsilon n/2$ of which are not transversal bases. Recall the choice of C from the previous section, and let $D = 2C + 4$, so that $D\,(n - f - 1) - 2n \geq Cn$ for large n. We prove the following (for large n).

Claim 10. *We can modify $\mathcal{S}$ in such a way that at least one of the following holds.*

(a) The volume of $\mathcal{S}$ increases;

(b) the volume of $\mathcal{S}$ does not change, and there is now $S_0 \in \mathcal{S}$ missing at least D colours;

(c) the volume of $\mathcal{S}$ does not change, and there are now distinct $S_0, S_1 \in \mathcal{S}$ such that S_1 contains at least D elements that are (S_0, b)-addable, for some colour b.

This suffices for our proof of theorem 5.1.1; indeed, if S_0 is missing at least D colours, then by Claim 5, either we can increase the volume of $\mathcal{S}$ or there are at least $D\,(n-f) \geq Cn$ elements which are (S_0, b)-addable for every b not appearing in S_0, meaning that $|Q\,(S_0)| \geq Cn$. If S_1 contains at least D elements that are (S_0, b)-addable, then in the proof of Claim 8 with $\ell = 1$ we have $|Q| \geq D$ so either we can increase the volume of $\mathcal{S}$ or $|Q\,(S_0, S_1)| \geq D\,(n - f - 1) - 2n \geq Cn$ (recall eq. (5.3)).

Before proceeding to the proof of Claim 10, we first observe that using Claim 3 we can modify $\mathcal{S}$ to ensure that every $S \in \mathcal{S}$ that is not a transversal basis can be assigned a distinct missing colour $b\,(S)$. To see this, we iteratively apply the following lemma to $\mathcal{S}$.

Lemma 5.2.6. *Consider $f \leq n/2$ and let $\mathcal{S} = \left\{S_1, \ldots, S_f\right\}$ be a collection of disjoint RISs. We can either increase the size of $\mathcal{S}$ or we can modify $\mathcal{S}$ in such a way that the size of each S_i remains the same, and in such a way that that there is a choice of disjoint colours $\{b_1, \ldots, b_f\}$ for which any S_i that is not a transversal basis has no element of colour b_i.*

Proof. Suppose for some i that we found distinct colours $b_1, \ldots, b_{i-1}$ such that, for all S_j which are not transversal bases, no element of S_j is of colour

b_j. If S_i is a transversal basis we choose an arbitrary unused colour as b_i. Otherwise there is a colour, say c, not appearing in S_i. Then by Claim 3 either we can increase the size of $\mathcal{S}$ or there are at least $n - f \geq n/2$ colours which are (S_i, c)-swappable. At least one of these colours does not appear in $\{b_1, \ldots, b_{i-1}\}$, since $i - 1 < f \leq n/2$. Let b be such a colour and set $b_i = b$. By performing a simple swap, we transform S_i into a new RIS, still disjoint to all other $S_j \in \mathcal{S}$ and missing the colour b. □

Now we prove Claim 10.

Proof of Claim 10. Recall that we are assuming there are at least $\varepsilon n/2$ RISs in $\mathcal{S}$ that are not transversal bases. Let E be the largest integer such that there are at least $M_E = \left(\varepsilon / \left(4D^2\right)\right)^E n$ RISs in $\mathcal{S}$ missing at least E colours. We may assume $1 \leq E < D$. By lemma 5.2.6 we may assume that each $S \in \mathcal{S}$ which is not a transversal basis has a distinct missing colour $b(S)$. We describe a procedure that modifies $\mathcal{S}$ to increase E.

We create an auxiliary digraph G on the vertex set $\mathcal{S}$ as follows. For every $S_0 \in \mathcal{S}$ missing at least E colours, put an arc to S_0 from every $S_1 \in \mathcal{S}$ such that S_1 contains at least $E+1$ elements that are $(S_0, b(S_0))$-addable.

Say an $(E+1)$-*out-star* in a digraph is a set of $E+1$ arcs directed away from a single vertex. Our goal is to prove that there are M_{E+1} vertex-disjoint $(E+1)$-out-stars. To see why this suffices, consider an $(E+1)$-out-star (with centre S_1, say). We show how to transfer $E+1$ elements from S_1 to its out-neighbours, the end result of which is that S_1 is then missing $E+1$ colours. We will then be able to repeat this process for each of our out-stars.

For each of the $E+1$ out-neighbours S_0 of S_1 there are at least $E+1$ elements of S_1 which are $(S_0, b(S_0))$-addable. Therefore, for each such S_0 we can make a specific choice of such an $(S_0, b(S_0))$-addable element, in such a way that each of these $E+1$ choices are *distinct*. For each S_0 we can then transfer the chosen element from S_1 to S_0, possibly with a simple swap. These simple swaps will not create any conflicts, because any addability witness for any element in S_0 is in a colour unique to that S_0 (by the property from lemma 5.2.6). After this operation, S_i is now missing at least $E+1$ colours.

It will be a relatively straightforward matter to find our desired out-stars by studying the digraph G. First we show that G must have many edges.

Claim. *In the above auxiliary digraph, we may assume that every $S_0 \in \mathcal{S}$ missing at least E colours has in-degree at least $\varepsilon n/D$.*

Proof. By Claim 5 we can assume that there are at least $E(n - f)$ elements which are $(S_0, b(S_0))$-addable. All these elements appear in various $S \in \mathcal{S}$ (otherwise we can increase the volume of $\mathcal{S}$).

Let $N^{-}(S_0)$ be the set of all S_1 such that there is an arc from S_1 to S_0 in G (so $|N^{-}(S_0)|$ is the indegree of S_0). By definition, every $S \notin N^{-}(S_0)$ has at most E elements which are $(S_0, b(S_0))$-addable. Moreover, observe that every $S \in \mathcal{S}$ has fewer than D elements that are $(S_0, b(S_0))$-addable, or else (c) trivially occurs. It follows that

$$D|N^{-}(S_0)| + E(f - |N^{-}(S_0)|) \geq E(n-f),$$

so

$$|N^{-}(S_0)| \geq \frac{E\left((n-f)-f\right)}{D-E} \geq \frac{\varepsilon n}{D},$$

as desired. □

We have proved that G has at least $M_E \varepsilon n / D$ edges. Now we finish the proof by showing how to find our desired out-stars.

Claim. *G has at least M_{E+1} vertex-disjoint $(E+1)$-out-stars.*

Proof. We can find these out-stars in a greedy fashion. Suppose that we have already found t vertex-disjoint $(E+1)$-out-stars, for some $t < M_{E+1}$. We show that there must be an additional $(E+1)$-out-star disjoint to these. Let G' be obtained from G by deleting all vertices in the out-stars we have found so far. Each of these out-stars has $E+2$ vertices, so the number of arcs in G' is at least

$$\begin{aligned} M_E \frac{\varepsilon n}{D} - t(E+2) \cdot 2f &> M_E \frac{\varepsilon n}{D} - M_{E+1}(E+2) \cdot 2f \\ &= M_E \frac{\varepsilon n}{D} - \frac{M_E \varepsilon}{2D^2} \cdot (E+2) f \\ &\geq M_E \varepsilon \left(\frac{n}{D} - \frac{f}{D} \right) \\ &\geq M_E \varepsilon \cdot \frac{f}{D} \geq (E+1) f, \end{aligned}$$

where the last inequality holds for sufficiently large n, using the fact that M_E is linear in n. This means that G' (having at most f vertices) has a vertex with outdegree at least $E+1$, which means G' contains an $(E+1)$-out-star disjoint to the out-stars we have found so far. □

□

5.3 CONCLUDING REMARKS

In this chapter we proved that that given bases $B_1, \ldots, B_n$ in a matroid, we can find $(1/2 - o(1))\, n$ disjoint transversal bases. Although our methods do not extend past $n/2$, we do not think that there is a fundamental obstacle preventing related methods from going further. Indeed, by tracking the possible cascades of swaps more carefully, it might be possible to find $(1 - o(1))\, n$ disjoint transversal bases, or at least to find $(1 - o(1))\, n$ disjoint partial transversals each of size $(1 - o(1))\, n$. Although we cannot completely rule out the possibility that a full proof of Rota's basis conjecture could be obtained in this way, we imagine that more ingredients will be required. We are hopeful that ideas used to prove existence of designs (see [126, 198]) could be relevant, at least in the case of vector spaces.

Also, we remark that Rota's basis conjecture is reminiscent of some other problems concerning rainbow structures in graphs (actually, for a graphic matroid, Rota's basis conjecture can be interpreted as a conjecture about rainbow spanning forests in edge-coloured multigraphs). The closest one to Rota's basis conjecture seems to be the Brualdi–Hollingsworth conjecture [42], which posits that for every $(n-1)$-edge-colouring of the complete graph K_n, the edges can be decomposed into rainbow spanning trees. This conjecture has recently seen some exciting progress (see for example [24, 145, 186, 201]). We wonder if some of the ideas developed for the study of rainbow structures could be profitably applied to Rota's basis conjecture.

We also mention the following strengthening of Rota's basis conjecture due to Kahn (see [146]). This is simultaneously a strengthening of the Dinitz conjecture [101] on list-colouring of $K_{n,n}$, solved by Galvin [119].

Conjecture 5.3.1. *Given a rank-n matroid and bases $B_{i,j}$ for each $1 \le i, j \le n$, there exist representatives $b_{i,j} \in B_{i,j}$ such that each of the sets $\{b_{1,j}, \ldots, b_{n,j}\}$ and $\{b_{i,1}, \ldots, b_{i,n}\}$ are bases.*

The methods developed in this chapter are also suitable for studying conjecture 5.3.1. In particular, the argument used to prove theorem 5.1.1 can readily be modified to show the following natural partial result towards Kahn's conjecture.

Theorem 5.3.2. *For any $\varepsilon > 0$ the following holds for sufficiently large n. Given a rank-n matroid and bases $B_{i,j}$ for each $1 \le i \le n$ and $1 \le j \le f = (1-\varepsilon)n/2$, there exist representatives $b_{i,j} \in B_{i,j}$ and $L \subseteq \{1, \ldots, f\}$ such that each $\{b_{i,j} : i \in L\}$ is independent, and such that $\{b_{i,1}, \ldots, b_{i,n}\}$ is a basis for any $i \in L$ and $|L| \ge (1/2 - \varepsilon)n$.*

Note that if we are in the setting of conjecture 5.3.1 where bases are given for all $1 \leq i, j \leq n$ then the above theorem allows us to choose roughly which rows we would like to find our bases in.

Note also that if, for each fixed j, the bases $B_{1,j}, \ldots, B_{n,j}$ are all equal, then Kahn's conjecture reduces to Rota's basis conjecture. This observation also shows that theorem 5.3.2 implies theorem 5.1.1.

It is not hard to adapt the proof of theorem 5.1.1 to prove theorem 5.3.2. However, since it would require repeating most of the argument, we omit the details here. For interested readers we present the details in a companion note available on the arXiv [43].

6 TWO TOPICS IN EXTREMAL SET THEORY

6.1 Intersection spectrum of 3-chromatic intersecting hypergraphs

6.1.1 *Introduction*

A family $\mathcal{F}$ of sets is said to have *property B* if there exists a set X which properly intersects every set of the family, that is, $\emptyset \neq F \cap X \neq F$ for all $F \in \mathcal{F}$. The term was coined in the 1930s by Miller [184, 185] in honor of Felix Bernstein. In 1908, Bernstein [33] proved that for any transfinite cardinal number κ, any family $\mathcal{F}$ of cardinality at most κ, whose sets have cardinality at least κ, has property B. In the 60s, Erdős and Hajnal [89] revived the study of property B, and initiated its investigation for finite set systems, or hypergraphs. A hypergraph H consists of a *vertex set* $V(H)$ and an *edge set* $E(H)$, where every edge is a subset of the vertex set. As usual, H is *k-uniform* if every edge has size k. A hypergraph is *r-colourable* if its vertices can be coloured with r colours such that no edge is monochromatic. Note that a hypergraph has property B if and only if it is 2-colourable.

The famous problem of Erdős and Hajnal is to determine $m(k)$, the minimum number of edges in a k-uniform hypergraph which is not 2-colourable. This can be viewed as an analogue of Bernstein's result for finite cardinals. Clearly, one has $m(k) \leq \binom{2k-1}{k}$, since the family of all k-subsets of a given set of size $2k-1$ does not have property B. On the other hand, Erdős [79] soon observed that $m(k) \geq 2^{k-1}$. Indeed, if a hypergraph has less than 2^{k-1} edges, then the expected number of monochromatic edges in a random 2-colouring is less than 1, hence a proper 2-colouring exists. Thanks to the effort of many researchers [2, 30, 31, 50, 80, 120, 199, 206, 216, 224], the best known bounds are now

$$\Omega\left(\sqrt{k/\log k}\right) \leq m(k)/2^k \leq O(k^2), \tag{6.1}$$

proofs of which are now textbook examples of the probabilistic method [16]. Improving either of these bounds would be of immense interest.

The 2-colourability problem for hypergraphs has inspired a great amount of research over the last half century, with many deep results proved and methods developed. One outstanding example is the Lovász local lemma, originally employed by Erdős and Lovász [94] to show that a k-uniform hypergraph is 2-colourable if every edge intersects at most 2^{k-3} other edges. In addition to the Lovász local lemma, the seminal paper of Erdős and Lovász [94] from 1973 left behind a whole legacy of problems and results on the 2-colourability problem. Some of their problems were solved relatively soon [30, 31], others took decades [115, 150, 151] or are still the subject of ongoing research.

At the heart of some particularly notorious problems are intersecting hypergraphs. In an *intersecting* hypergraph (Erdős and Lovász called them *cliques*), any two edges intersect in at least one vertex. The study of intersecting families is a rich topic in itself, which has brought forth many important results such as the Erdős–Ko–Rado theorem. We refer the interested reader to [112]. With regards to colourability, the intersecting property imposes strong restrictions. For instance, it is easy to see that any intersecting hypergraph has chromatic number at most 3. Hence, the 3-chromatic ones are exactly those which do not have property B. On the other hand, every 3-chromatic intersecting hypergraph is "critical" in the sense that deleting just one edge makes it 2-colourable. These and other reasons (explained below) make the 2-colourability problem for intersecting hypergraphs very interesting. It motivated Erdős and Lovász to initiate the study of 3-chromatic intersecting hypergraphs, proving some fundamental results and raising tantalizing questions.

Analogously to $m(k)$, define $\tilde{m}(k)$ as the minimum number of edges in a k-uniform intersecting hypergraph which is not 2-colourable. The problem of estimating $\tilde{m}(k)$ seems much harder. While for non-intersecting hypergraphs, we know at least that $\lim_{k\to\infty} \sqrt[k]{m(k)} = 2$, no such result is in sight for $\tilde{m}(k)$. Clearly, the lower bound in (6.1) also holds for $\tilde{m}(k)$. However, the best known upper bound for $\tilde{m}(k)$ is exponentially worse. For any k which is a power of 3, an iterative construction based on the Fano plane yields a k-uniform 3-chromatic intersecting hypergraph with $7^{\frac{k-1}{2}}$ edges (see [2, 94]). Perhaps the main obstacle to improving this bound is that the probabilistic method does not seem applicable for intersecting hypergraphs. Erdős and Lovász also asked for the minimum number of edges in a k-uniform intersecting hypergraph with cover number k, which can be viewed as a relaxation of $\tilde{m}(k)$ since any k-uniform 3-chromatic intersecting hypergraph has cover number k. This problem was famously solved by Kahn [151].

In addition to the size of 3-chromatic intersecting hypergraphs, Erdős and Lovász also studied their "intersection spectrum". For a hypergraph H, define $I(H)$ as the set of all intersection sizes $|E \cap F|$ of distinct edges $E, F \in E(H)$. A folklore observation is that if a hypergraph is not 2-colourable, then there must be two edges which intersect in exactly one vertex, that is, $1 \in I(H)$. A very natural question is what else we can say about the intersection spectrum of a non-2-colourable hypergraph. In general, hypergraphs can be non-2-colourable even if their only intersection sizes are 0 and 1. There are various basic constructions for this, see e.g. [180]. For instance, consider K_N^{k-1}, the complete $(k-1)$-uniform hypergraph on N vertices. For N large enough, any 2-colouring of the edges will contain a monochromatic clique on k vertices by Ramsey's theorem. Let H be the hypergraph with $V(H) = E(K_N^{k-1})$ whose edges correspond to the k-cliques of K_N^{k-1}. Then H is a k-uniform non-2-colourable hypergraph with $I(H) = \{0, 1\}$.

Erdős and Lovász observed that the situation changes drastically for intersecting hypergraphs. In the aforementioned construction of the iterated Fano plane, the intersection spectrum consists of all odd numbers (between 1 and $k-1$). In particular, the *maximal intersection size* is $k-2$, and the *number of intersection sizes* is $(k-1)/2$. Astonishingly, not a single example (of a k-uniform 3-chromatic intersecting hypergraph) is known where these quantities are any smaller. Intrigued by this, Erdős and Lovász studied the corresponding lower bounds. Concerning the maximal intersection size, they (and also Shelah) could prove that $\max I(H) = \Omega(k/\log k)$ for any k-uniform 3-chromatic intersecting hypergraph H. This is in stark contrast to non-intersecting hypergraphs where we can have $\max I(H) = 1$ as discussed. In fact, Erdős and Lovász conjectured that a linear bound should hold, or perhaps even $k - O(1)$. Erdős [83] later offered \$100 for settling this question.

Finally, consider the number of intersection sizes. As already noted, we always have $1 \in I(H)$. Moreover, the above result on $\max I(H)$ adds another intersection size for sufficiently large k. Hence, 3-chromatic intersecting hypergraphs have a small intersection size, namely 1, and a relatively big intersection size. Recall that general non-2-colourable hypergraphs might only have two intersection sizes. However, Erdős and Lovász were able to show that intersecting hypergraphs must have at least one more. Using a theorem of Deza [72] on sunflowers, they proved that $i(k) \geq 3$ for sufficiently large k, where $i(k)$ is the minimum of $|I(H)|$ over all k-uniform 3-chromatic intersecting hypergraphs. They also remarked that they "cannot even prove" that $i(k)$ tends to infinity. This is particularly striking in view of the best known upper bound being $(k-1)/2$.

Conjecture 6.1.1 (Erdős and Lovász, 1973). *$i(k) \to \infty$ as $k \to \infty$.*

Despite the fact that over the years this problem has been reiterated many times by Erdős and other researchers [59, 78, 83, 99, 208], remarkably, the lower bound of three intersection sizes has withstood any improvement until now. Here, we prove Conjecture 6.1.1 in the following strong form.

Theorem 6.1.2. *The intersection spectrum of a k-uniform 3-chromatic intersecting hypergraph has size at least $\Omega(k^{1/2}/\log k)$.*

6.1.2 *Intersection spectrum*

In this section we will prove Theorem 6.1.2. We begin by gathering several properties of our hypergraphs that we are going to use.

6.1.2.1 *Preliminaries*

The following is a classical result of Erdős [79], already mentioned in the introduction.

Theorem 6.1.3. *Any non 2-colourable, k-uniform hypergraph has at least 2^{k-1} edges.*

This result provides us with a large number of edges among which we can look for different intersection sizes. The following simple result (inspired by Problem 13.16 in [180]) will give rise to certain restrictions on the distribution of intersection sizes across our hypergraph, which in turn will enable us to use a density increment argument.

Proposition 6.1.4. *Let $\mathcal{A}$ be a k-uniform and $\mathcal{B}$ a k'-uniform hypergraph on the same vertex set and with the same number of edges ℓ. Then*

$$\sum_{\{A,A'\}\subseteq E(\mathcal{A})} |A\cap A'| + \sum_{\{B,B'\}\subseteq E(\mathcal{B})} |B\cap B'| \geq \sum_{A\in E(\mathcal{A}),B\in E(\mathcal{B})} |A\cap B| - \ell(k+k')/2.$$

Proof. For any vertex x let us denote by a_x and b_x the number of edges in $\mathcal{A}$ and $\mathcal{B}$ that contain x, respectively. Each x will be a common vertex of exactly $\binom{a_x}{2}$, $\binom{b_x}{2}$ and $a_x b_x$ pairs of edges both in $\mathcal{A}$, both in $\mathcal{B}$ and one each in $\mathcal{A},\mathcal{B}$, respectively. Observe further that $\sum_x a_x = \ell k$ and $\sum_x b_x = \ell k'$. Putting these observations together, the desired inequality is equivalent to

$$\sum_x \binom{a_x}{2} + \sum_x \binom{b_x}{2} \geq \sum_x a_x b_x - \sum_x (a_x + b_x)/2.$$

Since $a_x^2 + b_x^2 \geq 2a_x b_x$, this completes the proof. $\square$

It will be convenient for us to introduce the following averaging functions. Given a hypergraph H and disjoint subsets $S, T \subseteq E(H)$ we define

$$\lambda_S := \frac{1}{\binom{|S|}{2}} \sum_{\{e,f\} \subseteq S} |e \cap f| \quad \text{and} \quad \lambda_{S,T} := \frac{1}{|S||T|} \sum_{e \in S, f \in T} |e \cap f|.$$

The following lemma encapsulates the aforementioned restrictions on the distribution of intersection sizes across a k-uniform hypergraph. Roughly speaking, it says that given two disjoint subsets of edges, on average, the intersection sizes inside the sets are at least as big as across. Moreover, a crucial ingredient of our density increment approach is that we can obtain a stronger inequality if we can find some vertices that exclusively belong to the edges of one set.

Lemma 6.1.5. *Let S, T be disjoint collections of ℓ edges of a k-uniform hypergraph H, with the property that there are x vertices of H which all belong to every edge in S and none of them belong to any edge in T. Then*

$$\frac{\lambda_S + \lambda_T}{2} \geq \lambda_{S,T} + \frac{x}{2} - \frac{k}{\ell - 1}.$$

Proof. Let $\mathcal{A}$ be the $(k-x)$-uniform hypergraph with edge set consisting of the edges in S with their x common vertices disjoint from any edge in T removed. Let $\mathcal{B}$ be the k-uniform hypergraph consisting of the edges in T. Since the intersection size drops by x for pairs of edges in S and it stays the same for pairs in T or one in S and one in T, Proposition 6.1.4 gives

$$\sum_{\{e,f\} \subseteq S} (|e \cap f| - x) + \sum_{\{e,f\} \subseteq T} |e \cap f| \geq \sum_{e \in S, f \in T} |e \cap f| - \ell(2k - x)/2.$$

Upon dividing by $\binom{\ell}{2}$ we obtain $\lambda_S + \lambda_T \geq x + \frac{2\ell}{\ell-1}\lambda_{S,T} - \frac{2k-x}{\ell-1} \geq x + 2\lambda_{S,T} - \frac{2k}{\ell-1}$. □

The following useful fact, which exploits the property of being 3-chromatic and intersecting, imposes further restrictions on the distribution of intersection sizes. In particular, it allows us to find many edges with large common intersection.

Proposition 6.1.6. *Let H be a k-uniform 3-chromatic and intersecting hypergraph, and $X \subseteq V(H)$. Then for any $0 \leq i \leq k - |X|$, there exists a set $X_i \subseteq V(H)$ of size $|X| + i$ such that at least a k^{-i} proportion of the edges containing X also contain X_i.*

Proof. We proceed by induction on i. The case $i = 0$ is trivial. Suppose now we have X_i with $|X_i| < k$. Since H is 3-chromatic it must contain an edge Y disjoint from X_i, as otherwise we can colour vertices in X_i red and all vertices outside blue to obtain a proper 2-colouring. Since H is intersecting every edge containing X_i intersects Y. By averaging, some vertex $x \in Y$ is contained in at least a $\frac{1}{k}$ proportion of these edges. Thus, adding x to X_i yields the desired X_{i+1}. □

Remark. The above proposition does not require the full strength of the 3-chromatic and intersecting assumptions. In fact we can replace them with the assumption that for any set $X \subseteq V(H)$ of size less than k, there exists a set $Y \subseteq V(H) \setminus X$ of size k such that every edge of H intersects Y. Since this proposition and Theorem 6.1.3 are the only places where we will use these assumptions in our proof of Theorem 6.1.2, we note that it actually holds if we only assume that this alternative property holds and that H has at least 2^{k-1} edges.

In our proof of Theorem 6.1.2, we will make use of a dependent random choice lemma. Dependent random choice is a powerful probabilistic technique which has recently led to a number of advances in Ramsey theory, extremal graph theory, additive combinatorics, and combinatorial geometry. Early variants of this technique were developed by Gowers [129], Kostochka and Rödl [164] and Sudakov [230]. In many applications, the technique is used to prove the useful fact that every dense graph contains a large subset U in which almost every set of t vertices has many common neighbours. For more information about dependent random choice and its applications, we refer the interested reader to the survey [110]. We are going to use the following variant of the dependent random choice lemma (Lemma 6.3 in [110]).

Lemma 6.1.7. *If $d > 0$, $t \leq n$ are positive integers, and G is a graph with $m > 4td^{-t}n$ vertices and at least $dm^2/2$ edges, then there is a vertex subset U with $|U| > 2n$ such that the proportion of t-subsets of U with less than n common neighbours in G is less than $(2t)^{-t}$.*

6.1.2.2 *Proof ideas*

In this subsection we will illustrate our proof ideas by sketching a slightly simpler argument which shows $|I(H)| \geq k^{1/3-o(1)}$. Our proof of Theorem 6.1.2, which will be presented in the next subsection, follows along very similar lines, with the exception of using dependent random choice in place of certain Ramsey arguments which we use here.

Let H be a k-uniform, 3-chromatic and intersecting hypergraph. Let $\lambda_1 < \ldots < \lambda_r$ denote the distinct intersection sizes in H, so $r = |I(H)|$. A natural way to approach our problem is to define a colouring of the complete graph with vertex set $E(H)$ where an edge is coloured according to the size of the intersection of its endpoints. We will refer to this colouring as the *intersection colouring*. Theorem 6.1.3 tells us that this graph is quite big, so if the number of colours r is small we could hope to use Ramsey's theorem to find a number of edges which make a monochromatic clique in the intersection colouring, i.e. all pairwise intersection sizes are the same. This would lead to a contradiction if there were about k^2 such edges (this is an easy consequence of [72]). Unfortunately, in an arbitrary colouring we cannot find this many, even if we assume $|I(H)| = 2$.

However, the well-known argument for bounding Ramsey numbers actually gives us more than just a monochromatic clique. If we repeatedly take out an arbitrary edge of H and only keep its majority colour neighbours, we keep at least a proportion of $1/r$ of the edges per iteration. If we repeat rt many times we can find a set X consisting of t edges that we took out which had the same majority colour, so in particular X is a monochromatic clique in the intersection colouring. Furthermore, we know that the size of the set of remaining edges Y has lost at most a factor of r^{rt} compared to the original number of edges. In addition, the complete bipartite graph between X and Y is also monochromatic in the same colour as X.

In particular, this provides us with a pair (X, Y) of disjoint subsets of edges of H with the property that any two edges in X as well as any pair of edges one in X and one in Y intersect in exactly λ_i vertices, for some λ_i. We call such a pair a λ_i-pair. Note that if we choose $|X| = t \approx k^{1/3}$ and assume $r \leq O(k^{1/3}/\log k)$ (otherwise we are done) then $|Y| \geq |E(H)|/r^{rt} \geq |E(H)|/k^{O(k^{2/3})}$.

Our strategy will be to show that given a λ_i-pair one can find a λ_j-pair, for some $j > i$, whose set X still has size t and the size of Y shrinks by at most a factor of $k^{O(k^{2/3})}$. Since by Theorem 6.1.3 we know $|E(H)| \geq 2^{k-1}$, we can repeat this procedure at least $\Omega(k^{1/3}/\log k)$ times to conclude there are at least this many different intersection sizes and complete the proof.

To do this, let (X, Y) be a λ_i-pair with $|X| = t \approx k^{1/3}$. We can apply Lemma 6.1.5 to X and any t-subset $Y' \subseteq Y$ to conclude that $\lambda_{Y'} \geq \lambda_i - 2k^{2/3}$. So in particular, the average intersection size in any subset of Y of size at least t cannot be much lower than λ_i. Next we take an arbitrary edge U in X and consider the intersections of edges in Y with U. We will separate between two cases depending on the structure of Y.

In the first case, many of the edges in Y have almost the same intersection with U. In this case we will find a collection of at least $|Y|/k^{O(k^{2/3})}$ edges in Y which contain the same set of vertices of size at least $\lambda_i - x$, where $x \approx 10k^{2/3}$. Then we apply Proposition 6.1.6 (with $i = x + 1$) to obtain a subset of edges of size at least $|Y|/k^{O(k^{2/3})}$ in which any pair of edges intersects in more than λ_i vertices. Applying once again the Ramsey argument, this time within this collection of edges, we find a λ_j-pair in which we only lost another factor of $k^{O(k^{2/3})}$ in terms of size of Y. Since all intersection sizes are larger than λ_i we know that $j > i$, so we found our desired new pair.

In the second case, the intersections of edges in Y with U are "spread out". Then we can find two disjoint subsets $A, B \subseteq Y$ both of size $|Y|/k^{O(k^{2/3})}$ with the property that there is a set of x vertices $W \subseteq U$ which belongs to every edge of A and is disjoint from all edges in B. By applying the Ramsey argument to the collection A and to the collection B we either find a desired λ_j-pair with $j > i$ or we find a t-subset $S \subseteq A$ and a t-subset $T \subseteq B$ such that all pairwise intersections inside S and T have size at most λ_i. In particular, $\lambda_S, \lambda_T \leq \lambda_i$. We now apply Lemma 6.1.5 to S and T, knowing that the $x = 10k^{2/3}$ vertices in W belong to every edge in S and none belong to any edge in T. This will give us $\lambda_{S,T} \leq \lambda_i - 4k^{2/3}$. Combining these three inequalities we obtain $\lambda_{S \cup T} < \lambda_i - 2k^{2/3}$, which contradicts our lower bound on the average intersection size among subsets of Y and completes the argument.

Remark. While one can develop the Ramsey type arguments of this section to prove Theorem 6.1.2, in the following subsection we choose to present the argument based on dependent random choice as it demonstrates a slightly different approach, is slightly shorter and we believe has greater potential for further improvement.

6.1.2.3 *Proof of Theorem 6.1.2*

Let H be a k-uniform, 3-chromatic and intersecting hypergraph where we assume throughout the proof that k is sufficiently large. Let $\lambda_1 < \ldots < \lambda_r$ denote the distinct intersection sizes in H, and set $\lambda_{r+1} = k$. Let us assume for the sake of contradiction that $r < \frac{\sqrt{k}}{51 \log_2 k}$. Let us set $t = 2\lceil \sqrt{k} \rceil$ and for all $1 \leq i \leq r + 1$

$$m_i := \frac{|E(H)|}{k^{25(i-1)t}} \geq 2^{k-1-25(i-1)t \log_2 k} \geq t, \tag{6.2}$$

where we used Theorem 6.1.3 in the first inequality and $i - 1 \leq r < \frac{\sqrt{k}}{51 \log_2 k}$ in the second. Our strategy is as follows. We will choose the largest i such that we can find a subset $A \subseteq E(H)$ of size m_i with the property that many pairs of edges in A intersect in at least λ_i vertices. We will then find such a subset for a larger i, reaching a contradiction.

Let us first specify what we mean by many pairs above. In order to make use of the dependent random choice lemma we will quantify it in terms of how many t-subsets of A consist of edges with all their pairwise intersections being of size smaller than λ_i. We call a set of edges λ_i-*small* if all their pairwise intersections have size strictly smaller than λ_i. Let i be the largest index such that there exists a subset $A \subseteq E(H)$ of size at least m_i with the property that at most half of the t-subsets of A are λ_i-small. Observe that since all intersections of edges in H are of size at least λ_1 there is no set of t edges of H which is λ_1-small. Hence, by taking $A = E(H)$ we get that $i = 1$ satisfies our condition, showing that i exists.

Our first goal is to show the following claim.

Claim. *There exists a pair (X, Y) of disjoint subsets of $E(H)$ such that*

1. $|X| = t$ *and* $|Y| \geq m_i/k^{3t}$,
2. *any two edges in X intersect in at most λ_i vertices,*
3. *any edge in X and any edge in Y intersect in at least λ_i vertices.*

Proof. Let $m := |A| \geq m_i$. By our choice of A, there are at least $\frac{1}{2}\binom{m}{t}$ t-subsets of A which contain at least one pair of edges with the intersection size at least λ_i. Since every pair of edges belongs to at most $\binom{m-2}{t-2}$ t-subsets of A, there are at least $\frac{1}{2}\binom{m}{t}/\binom{m-2}{t-2} = \frac{1}{t(t-1)}\binom{m}{2} \geq \frac{1}{8k}\frac{m^2}{2}$ pairs of edges in A with intersections of size at least λ_i.

We apply the dependent random choice lemma (Lemma 6.1.7) to the graph with vertex set A and edge set consisting of pairs of elements in A (so edges of H) which intersect in at least λ_i vertices. Choosing $d = \frac{1}{8k}$ we may take

$$n = \frac{m}{5td^{-t}} \geq \frac{m_i}{5t \cdot (8k)^t} \geq \frac{m_i}{k^{3t}} \geq m_{i+1} \geq t,$$

where in the last two inequalities we used (6.2).

This provides us with a subset $A' \subseteq A$ of size at least $n \geq m_{i+1}$ such that all but an $(2t)^{-t}$ proportion of t-subsets of A' have the property that there exist n edges of H each of which intersects every edge in the t-subset in at least λ_i vertices. Since $|A'| \geq m_{i+1}$, by our maximality assumption on i we know that more than half of the t-subsets of A' must be λ_{i+1}-small, i.e. have all pairs of edges intersecting in less than λ_{i+1}, so at most λ_i vertices. Therefore, there exists a t-subset $X \subseteq A'$ which is both λ_{i+1}-small and there is a set Y consisting of n edges of H with the property that any edge in X intersects any edge in Y in at least λ_i vertices. These X and Y satisfy the desired properties. $\square$

Our next step is to analyze average intersections among subsets of Y. The following claim tells us that any t-subset of Y must have average intersection size almost as big as λ_i.

Claim. *For any $Y' \subseteq Y$ of size t we have $\lambda_{Y'} \geq \lambda_i - 2\sqrt{k}$.*

Proof. Lemma 6.1.5 applied to X and Y' with $x = 0$ gives

$$\lambda_{Y'} \geq 2\lambda_{X,Y'} - \lambda_X - \frac{2k}{t-1} \geq \lambda_i - 2\sqrt{k},$$

since $\lambda_X \leq \lambda_i$, $\lambda_{X,Y'} \geq \lambda_i$ and by definition of t. $\square$

On the other hand our maximality assumption on i and the fact Y is large allows us to find a subset Y' as in the claim above which is λ_{i+1}-small, i.e. with $\lambda_{Y'} \leq \lambda_i$. This does not immediately lead to a contradiction since we lost a little bit in the application of Lemma 6.1.5 above. Our final step is to show one can find such Y' with an even smaller average intersection size which will give us a contradiction.

To this end let us fix an edge $U \in X$ and let $(A_1, B_1, X_1), \ldots, (A_m, B_m, X_m)$ be a collection of triples with the following properties:

- $A_1, \ldots, A_m, B_1, \ldots, B_m$ are distinct edges of H.
- For each $i \in [m]$, X_i is a subset of U of size $x = 10t$, and $X_i \subseteq A_i$ but $X_i \cap B_i = \emptyset$.
- m is maximal subject to the above conditions.

If $m < |Y|/4$ there is a subset $Y' \subseteq Y$ consisting of at least $|Y|/2$ edges with the property that no triple (A_i, B_i, X_i) as above exists inside Y'. Let us fix an $A \in Y'$. If we let $A' = A \cap U$ we know $|A'| \geq \lambda_i$. For any edge $B \in Y'$ we know that $|A' \setminus B| < x$, since otherwise we could have extended our family by using A, B and any x vertices in $A' \setminus B$. This condition can be rewritten as $|A' \cap B| > |A'| - x$. In particular, assuming $\lambda_i \geq x$, there exists a subset $A'' \subseteq A'$ of size $|A'| - x \geq \lambda_i - x$, which belongs to at least

$$\frac{|Y'|}{\binom{|A'|}{|A'|-x}} \geq \frac{|Y|/2}{\binom{k}{x}} \geq \frac{m_i}{k^{x+3t}} = \frac{m_i}{k^{13t}}$$

edges of H. If $\lambda_i < x$ we take $A'' = \emptyset$. Either way, using Proposition 6.1.6 (with $X = A''$ and $i = x+1$) we obtain a set of vertices of size larger than λ_i contained in at least $m_i/k^{x+1+13t} \geq m_i/k^{25t} \geq m_{i+1}$ edges of H. Since this means that any two edges in this collection intersect in more than λ_i vertices, we conclude that no set of t edges from this collection is λ_{i+1}-small. This contradicts the maximality of i.

Let us now assume $m \geq |Y|/4$. Since there are $\binom{|U|}{x}$ different choices for X_i, there are at least $m/\binom{|U|}{x} \geq \frac{|Y|}{4}/\binom{k}{x} \geq |Y|/k^x \geq m_{i+1}$ triples with the same X_i, which we denote by X'. This and the maximality of i allow us to find a set S of size t, consisting of edges A_i for which $X_i = X'$, which is λ_{i+1}-small, i.e. all pairs intersect in at most λ_i vertices. Similarly we find a set T of size t, consisting of edges B_i for which $X_i = X'$, which is λ_{i+1}-small. Let us now remove half of the edges from both S and T so that $|S \cup T| = t$. The above claim then tells us that $\lambda_{S\cup T} \geq \lambda_i - 2\sqrt{k}$. On the other hand Lemma 6.1.5 applied to S and T with $x = |X'|$ gives

$$\frac{\lambda_S + \lambda_T}{2} \geq \lambda_{S,T} + \frac{x}{2} - \frac{k}{t/2-1} \geq \lambda_{S,T} + 8\sqrt{k}.$$

Since $\lambda_S, \lambda_T \leq \lambda_i$ we obtain $\lambda_{S,T} \leq \lambda_i - 8\sqrt{k}$. Combining this with $\lambda_S, \lambda_T \leq \lambda_i$ and

$$\binom{t/2}{2}\lambda_S + \binom{t/2}{2}\lambda_T + (t/2)^2\lambda_{S,T} = \binom{t}{2}\lambda_{S\cup T}$$

we get $\lambda_{S\cup T} \leq \lambda_i - \frac{(t/2)^2}{\binom{t}{2}} \cdot 8\sqrt{k} < \lambda_i - 2\sqrt{k}$, which is a contradiction and completes the proof. □

6.1.3 *Concluding remarks*

We have proved the conjecture of Erdős and Lovász in a strong form, by showing that the intersection spectrum of any k-uniform 3-chromatic intersecting hypergraph has size polynomial in k. It would be very interesting to improve our result and obtain a linear lower bound on the number of intersection sizes. In particular, this would also improve the result of Erdős, Lovász and Shelah on the maximal intersection size. Some of the ideas behind our arguments could be useful in obtaining a better understanding of how the intersection spectrum can look like in general as well.

6.2 MINIMUM SATURATED FAMILIES OF SETS

6.2.1 *Introduction*

In extremal set theory, one studies how large, or how small, a family $\mathcal{F}$ can be, if $\mathcal{F}$ consists of subsets of some set and satisfies certain restrictions. A classical example in the area is the study of *intersecting families*. We say that a family $\mathcal{F}$ is *intersecting* if for every $A, B \in \mathcal{F}$ we have $A \cap B \neq \emptyset$. The following simple proposition, first noted by Erdős, Ko and Rado [93], gives an upper bound on the size of an intersecting family in $2^{[n]}$.

Proposition 6.2.1. *Let $\mathcal{F} \subseteq 2^{[n]}$ be intersecting, then $|\mathcal{F}| \leq 2^{n-1}$.*

This follows from the observation that for every set $A \subseteq [n]$ at most one of A and $\overline{A}$ (where $\overline{A} := [n] \setminus A$) is in $\mathcal{F}$. This bound is tight, which can be seen, e.g., by taking the family of all subsets of $[n]$ that contain the element 1. In fact, there are many more extremal examples (see [91]), partly due to the following proposition.

Proposition 6.2.2. *Let $\mathcal{F} \subseteq 2^{[n]}$ be intersecting, then there is an intersecting family in $2^{[n]}$ of size 2^{n-1} that contains $\mathcal{F}$. In other words, if $\mathcal{F} \subseteq 2^{[n]}$ is a maximal intersecting family, then it has size 2^{n-1}.*

Indeed, suppose that $\mathcal{F}$ is a maximal intersecting family of size less than 2^{n-1}. Then there is a set $A \subseteq [n]$ such that $A, \overline{A} \notin \mathcal{F}$. By maximality of $\mathcal{F}$, there exist sets $B, C \in \mathcal{F}$ such that $A \cap B = \emptyset$ and $\overline{A} \cap C = \emptyset$. In particular, $B \cap C = \emptyset$, a contradiction.

There have been numerous extensions and variations of Proposition 6.2.1. For example, the study of t-intersecting families [155] (where the intersection of every two sets has size at least t) and L-intersecting families [14] (where the size of the intersection of every two distinct sets lies in some set of integers L). Such problems were also studied for k-uniform families, i.e. families that are subsets of $[n]^{(k)}$ (see e.g. [10] and [210]). A famous example is the Erdős-Ko-Rado [93] theorem which states that if $\mathcal{F} \subseteq [n]^{(k)}$ is intersecting, and $n \geq 2k$, then $|\mathcal{F}| \leq \binom{n-1}{k-1}$, a bound which is again tight by taking the families of all sets containing 1. Another interesting generalisation of Proposition 6.2.2 looks for the maximum measure of an intersecting family under the p-biased product measure (see [9, 73, 104, 116]). A different direction, which was suggested by Simonovits and Sós [220], studies the size of intersecting families of structured families, such as graphs, permutations and sets of integers (see e.g. [41, 128]).

Here we are interested in a different extension of Propositions 6.2.1 and 6.2.2. Given $s \geq 2$, we say that a family $\mathcal{F} \subseteq 2^{[n]}$ is *s-saturated* if $\mathcal{F}$

contains no s pairwise disjoint sets, and furthermore $\mathcal{F}$ is maximal with respect to this property. An example for an s-saturated family is the set of all subsets of $[n]$ that have a non-empty intersection with $[s-1]$. In 1974 Erdős and Kleitman [92] made the following conjecture, which states that this example is the smallest s-saturated family in $2^{[n]}$.

Conjecture 6.2.3 (Erdős, Kleitman [92]). *Let $\mathcal{F} \subseteq 2^{[n]}$ be s-saturated. Then $|\mathcal{F}| \geq (1 - 2^{-(s-1)}) \cdot 2^n$.*

Note that by Proposition 6.2.2, Conjecture 6.2.3 holds for $s = 2$. Given a family $\mathcal{F} \subseteq 2^{[n]}$, define $\mathcal{F}^C = 2^{[n]} \setminus \mathcal{F}$, and $\overline{\mathcal{F}} = \{\overline{A} : A \in \mathcal{F}\}$. Then for every $s \geq 2$, if $\mathcal{F} \subseteq 2^{[n]}$ is s-saturated then $\overline{\mathcal{F}^C}$ is intersecting. Indeed, if $A \notin \mathcal{F}$ then $\overline{A}$ contains $s-1$ pairwise disjoint sets of $\mathcal{F}$, so if A and B are such that $\overline{A}$ and $\overline{B}$ are disjoint, then at least one of A and B is in $\mathcal{F}$, as otherwise $\mathcal{F}$ contains $2(s-1) \geq s$ pairwise disjoint sets, a contradiction. By Proposition 6.2.1, it follows that if $\mathcal{F}$ is s-saturated then $|\mathcal{F}| \geq 2^{n-1}$. Surprisingly, beyond this trivial lower bound, nothing was known. Moreover, Frankl and Tokushige [112] wrote in their recent survey that obtaining a lower bound of $(1/2 + \varepsilon)2^n$, i.e. a modest improvement over the trivial bound, is a challenging open problem. Here we prove such a result.

Theorem 6.2.4. *Let $\mathcal{F} \subseteq 2^{[n]}$ be s-saturated, where $s \geq 2$. Then $|\mathcal{F}| \geq (1 - 1/s)2^n$.*

In fact, Theorem 6.2.4 is a corollary of a multipartite version of the above problem. A sequence of s families $\mathcal{F}_1, \ldots, \mathcal{F}_s \subseteq 2^{[n]}$ is called *cross dependant* (see, e.g., [114]) if there is no choice of s sets $A_i \in \mathcal{F}_i$, for $i \in [s]$, such that $A_1, \ldots, A_s$ are pairwise disjoint. We call a sequence of s families $\mathcal{F}_1, \ldots, \mathcal{F}_s$ *cross saturated* if the sequence is cross dependant and is maximal with respect to this property, i.e. the addition of any set to any of the families results in a sequence which is not cross dependant. Our aim here is to obtain a lower bound on the $|\mathcal{F}_1| + \ldots + |\mathcal{F}_s|$. Note that if $\mathcal{F}$ is s-saturated then the sequence given by $\mathcal{F}_1 = \ldots = \mathcal{F}_s = \mathcal{F}$ is cross saturated. Hence, a lower bound on the sum of sizes of a cross saturated sequence of s families implies a lower bound on the size of an s-saturated family.

A simple example of a cross saturated sequence $\mathcal{F}_1, \ldots, \mathcal{F}_s$ can be obtained by taking $\mathcal{F}_1$ to be empty, and letting all other sets be $2^{[n]}$. This construction is a special case of a more general family of examples which we believe contains all extremal examples; we discuss this in Section 6.2.3. Our next result shows that this example is indeed a smallest example for a cross saturated sequence. Furthermore, it implies Theorem 6.2.4 by taking $\mathcal{F}_1 = \ldots = \mathcal{F}_s = \mathcal{F}$.

Theorem 6.2.5. *Let $\mathcal{F}_1, \ldots, \mathcal{F}_s \subseteq 2^{[n]}$ be cross saturated. Then $|\mathcal{F}_1| + \ldots + |\mathcal{F}_s| \geq (s-1)2^n$.*

We have two different approaches to this problem, each of which can be used to prove Theorem 6.2.5. As the proofs are short, and Conjecture 6.2.3 is still open, we feel that there is merit in presenting both proofs here in hope that they would give rise to further progress on Conjecture 6.2.3.

Our first approach makes use of an interesting connection to correlation inequalities. Let us start by defining the *disjoint occurrence* of two families. Given subsets $A, I \subseteq [n]$, let

$$\mathcal{C}(I, A) = \{S \subseteq [n] : S \cap I = A \cap I\}.$$

The *disjoint occurrence* of two families $\mathcal{A}, \mathcal{B} \subseteq 2^{[n]}$ is defined by

$$\mathcal{A} \square \mathcal{B} := \{A : \exists \textit{ disjoint } \text{sets } I, J \subseteq [n] \text{ s.t. } \mathcal{C}(I, A) \subset \mathcal{A} \text{ and } \mathcal{C}(J, A) \subset \mathcal{B}\}.$$

Note that when $\mathcal{A}$ and $\mathcal{B}$ are both increasing families (i.e. if $A \in \mathcal{A}$, and $A \subseteq B \subseteq [n]$ then $B \in \mathcal{A}$), $\mathcal{A} \square \mathcal{B}$ is the set of all subsets of $[n]$ which can be written as a disjoint union of a set from $\mathcal{A}$ and a set from $\mathcal{B}$. This notion of disjoint occurrence appears naturally in the study of percolation. Using it, one can express the probability that there are two edge-disjoint paths between two sets of vertices in a random subgraph, chosen uniformly at random, of a given graph.

Van den Berg and Kesten [237] proved that $|\mathcal{A} \square \mathcal{B}| \leq |\mathcal{A}||\mathcal{B}|/2^n$ for increasing families $\mathcal{A}, \mathcal{B} \subseteq 2^{[n]}$ and conjectured that this inequality should hold for general families. This was proved by Reimer [211] in a ground breaking paper and is currently known as the van den Berg-Kesten-Reimer inequality.

Disjoint occurrence is surprisingly suitable for the study of saturated families. For example, if $\mathcal{F}$ is 3-saturated then it is easy to see that $\mathcal{F}$ is increasing, so $\mathcal{F} \square \mathcal{F}$ is the family of sets that are disjoint unions of two sets from $\mathcal{F}$, which is exactly the family $\overline{\mathcal{F}^C}$. This observation alone implies an improved lower bound on $|\mathcal{F}|$ using the van den Berg-Kesten-Reimer inequality. We obtain a better bound using a variant of this inequality, which was first observed by Talagrand [232], and later played a major role in Reimer's proof of the van den Berg-Kesten-Reimer inequality in full generality.

Our second approach is algebraic: we define a polynomial for each set in a certain family related to $\mathcal{F}_1, \ldots, \mathcal{F}_s$, and show that these polynomials are linearly independent, thus implying that the family is not very large.

6.2.2 *The proof*

Before turning to the first proof of Theorem 6.2.5, we introduce the correlation inequality that we will need. We present its short proof for the sake of completeness.

Lemma 6.2.6 (Talagrand [232]). *Let $\mathcal{A}, \mathcal{B} \subseteq 2^{[n]}$ be increasing families. Then $|\mathcal{A} \,\square\, \mathcal{B}| \leq |\overline{\mathcal{A}} \cap \mathcal{B}|$.*

Remark. Before turning to the proof of Lemma 6.2.6, we remark that the statement of Lemma 6.2.6 holds even without the assumption that the families $\mathcal{A}$ and $\mathcal{B}$ are increasing. Furthermore, an equivalent version of this played a major role in Reimer's proof [211] of the van den Berg-Kesten-Reimer inequality.

Proof. We prove the statement by induction on n. It is easy to check it for $n = 1$. Let $n > 1$ and suppose that the statement holds for $n - 1$. Given a family $\mathcal{F} \subseteq 2^{[n]}$, denote by $\mathcal{F}_0$ the family of sets in $\mathcal{F}$ that do not contain the element n, and let $\mathcal{F}_1 = \{A \subseteq [n-1] : A \cup \{n\} \in \mathcal{F}\}$. In particular, $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq 2^{[n-1]}$ when $\mathcal{F}$ is an increasing family.

We have

$$\begin{aligned}
|\mathcal{A} \,\square\, \mathcal{B}| &= |(\mathcal{A} \,\square\, \mathcal{B})_0| + |(\mathcal{A} \,\square\, \mathcal{B})_1| \\
&= |\mathcal{A}_0 \,\square\, \mathcal{B}_0| + |\mathcal{A}_1 \,\square\, \mathcal{B}_0| + |\mathcal{A}_0 \,\square\, \mathcal{B}_1| - |(\mathcal{A}_1 \,\square\, \mathcal{B}_0) \cap (\mathcal{A}_0 \,\square\, \mathcal{B}_1)| \\
&\leq |\mathcal{A}_1 \,\square\, \mathcal{B}_0| + |\mathcal{A}_0 \,\square\, \mathcal{B}_1| \\
&\leq \left|\overline{\mathcal{A}_1} \cap \mathcal{B}_0\right| + \left|\overline{\mathcal{A}_0} \cap \mathcal{B}_1\right| \\
&= \left|(\overline{\mathcal{A}} \cap \mathcal{B})_0\right| + \left|(\overline{\mathcal{A}} \cap \mathcal{B})_1\right| \\
&= \left|\overline{\mathcal{A}} \cap \mathcal{B}\right|,
\end{aligned}$$

where the first inequality holds because $\mathcal{A}_0 \,\square\, \mathcal{B}_0 \subseteq (\mathcal{A}_1 \,\square\, \mathcal{B}_0) \cap (\mathcal{A}_0 \,\square\, \mathcal{B}_1)$, and the second one follows by induction. $\square$

We are now ready for the first proof of Theorem 6.2.5.

Proof of Theorem 6.2.5. Let $\mathcal{F}_1, \ldots, \mathcal{F}_s$ be cross saturated, where $s \geq 2$. Note that

$$\overline{\mathcal{F}_i{}^C} = \mathcal{F}_1 \,\square\, \ldots \,\square\, \mathcal{F}_{i-1} \,\square\, \mathcal{F}_{i+1} \,\square\, \ldots \,\square\, \mathcal{F}_s. \tag{6.3}$$

Indeed, $\forall A \notin \mathcal{F}_i$, $\overline{A}$ contains a disjoint union of sets from $\mathcal{F}_1, \ldots, \mathcal{F}_{i-1}, \mathcal{F}_{i+1}, \ldots, \mathcal{F}_s$ and, conversely, any $A \in \mathcal{F}_1 \,\square\, \ldots \,\square\, \mathcal{F}_{i-1} \,\square\, \mathcal{F}_{i+1} \,\square\, \ldots \,\square\, \mathcal{F}_s$ cannot be in $\mathcal{F}_i$ by cross dependence. By Lemma 6.2.6, the following holds for every $i \geq 2$.

$$\begin{aligned} \left|\mathcal{F}_i{}^C\right| &= \left|\overline{\mathcal{F}_i{}^C}\right| \\ &= \left|\mathcal{F}_1 \square \ldots \square \mathcal{F}_{i-1} \square \mathcal{F}_{i+1} \square \ldots \square \mathcal{F}_s\right| \\ &\leq \left|(\mathcal{F}_1 \square \ldots \square \mathcal{F}_{i-1}) \cap \overline{(\mathcal{F}_{i+1} \square \ldots \square \mathcal{F}_s)}\right|. \end{aligned} \tag{6.4}$$

Denote $\mathcal{G}_1 = \mathcal{F}_1{}^C$, and $\mathcal{G}_i = (\mathcal{F}_1 \square \ldots \square \mathcal{F}_{i-1}) \cap \overline{(\mathcal{F}_{i+1} \square \ldots \square \mathcal{F}_s)}$ for $i \geq 2$.

Claim 11. *$\mathcal{G}_i \cap \mathcal{G}_j = \emptyset$ for $1 \leq i < j \leq s$.*

Proof. Indeed, if $i = 1$ then $\mathcal{G}_1 \subseteq \mathcal{F}_1{}^C$ and $\mathcal{G}_j \subseteq \mathcal{F}_1$. Otherwise, if $A \in \mathcal{G}_i \cap \mathcal{G}_j$ with $2 \leq i < j$ then A is the disjoint union of elements from $\mathcal{F}_1, \ldots, \mathcal{F}_{j-1}$, so in particular (as the sets $\mathcal{F}_l$ are increasing) it is the disjoint union of elements from $\mathcal{F}_1, \ldots, \mathcal{F}_i$. Furthermore, since $i \geq 2$, A is also the complement (with respect to $[n]$) of a disjoint union of sets in $\mathcal{F}_{i+1}, \ldots, \mathcal{F}_s$, i.e. $\overline{A}$ is the disjoint union of sets in $\mathcal{F}_{i+1}, \ldots, \mathcal{F}_s$. But this means that $[n]$ is the disjoint union of sets from $\mathcal{F}_1, \ldots, \mathcal{F}_s$, a contradiction to the assumption that $\mathcal{F}_1, \ldots, \mathcal{F}_s$ form a cross saturated sequence. □

It follows from (6.3), (6.4) and Claim 11 that

$$\begin{aligned} |\mathcal{F}_1| + \ldots + |\mathcal{F}_s| &= s \cdot 2^n - (|\mathcal{F}_1{}^C| + \ldots + |\mathcal{F}_s{}^C|) \\ &\geq s \cdot 2^n - (|\mathcal{G}_1| + \ldots + |\mathcal{G}_s|) \\ &\geq s \cdot 2^n - 2^n = (s-1)2^n, \end{aligned} \tag{6.5}$$

thus completing the proof of Theorem 6.2.5. □

Our next approach is algebraic. Before presenting the proof, we introduce some definitions and an easy lemma. Let n be fixed and consider the vector space V (over $\mathbb{R}$) of functions from $\{0,1\}^n$ to $\mathbb{R}$. Note that this is a vector space of dimension 2^n. Given a subset $S \subseteq [n]$, let $P_S : \{0,1\}^n \to \mathbb{R}$ be defined by $P_S(x) = \prod_{i \in S} x_i$, where $x = (x_1, \ldots, x_n)^T \in \{0,1\}^n$, and let $x_S \in \{0,1\}^n$ be defined by $(x_S)_i = 1$ if and only if $i \in S$. The following lemma shows that $\{P_S : S \subseteq [n]\}$ is a linearly independent set in V (in fact, as V has dimension 2^n, it is a basis).

Lemma 6.2.7. *The set $\{P_S : S \subseteq [n]\}$ is linearly independent in V.*

Proof. Suppose that $\sum_{S \subseteq [n]} \alpha_S P_S = 0$, where $\alpha_S \in \mathbb{R}$, and not all α_S's are 0. Let T be a smallest set such that $\alpha_T \neq 0$. Note that $P_S(x_T) = 1$ if and only if $S \subseteq T$. Hence

$$0 = \sum_{S \subseteq [n]} \alpha_S P_S(x_T) = \sum_{S \subseteq [n], |S| \leq |T|} \alpha_S P_S(x_T) = \alpha_T,$$

a contradiction to the assumption that $\alpha_T \neq 0$. It follows that $\alpha_S = 0$ for every $S \subseteq [n]$, i.e. the polynomials $\{P_S(x) : S \subseteq [n]\}$ are linearly independent, as required. $\square$

We shall use the inner product on V which is defined by

$$\langle f, g \rangle = \sum_{x \in \{0,1\}^n} f(x) g(x). \tag{6.6}$$

It is easy to check that this is indeed an inner product; in fact, it is the standard inner product, if functions are viewed as vectors indexed by $\{0,1\}^n$.

We are now ready for the second proof of Theorem 6.2.5.

Alternative proof of Theorem 6.2.5. Let $\mathcal{F}_1, \ldots, \mathcal{F}_s$ be cross saturated, where $s \geq 2$. Given i and $A \in \overline{\mathcal{F}_i{}^C}$, recall that by (6.3), A can be written as the disjoint union of sets from $\mathcal{F}_1, \ldots, \mathcal{F}_{i-1}, \mathcal{F}_{i+1}, \ldots, \mathcal{F}_n$. For every such i and A, fix a representation

$$A = B \cup C, \tag{6.7}$$

where B is a disjoint union of sets from $\mathcal{F}_1, \ldots, \mathcal{F}_{i-1}$ and C is a disjoint union of sets from $\mathcal{F}_{i+1}, \ldots, \mathcal{F}_s$. Let

$$Q_{i,A}(x) = \prod_{j \in B} x_j \cdot \prod_{j \in C} (x_j - 1).$$

Let W_i be the family of polynomials $Q_{i,A}$, where $i \in [s]$ and $A \subseteq \overline{\mathcal{F}_i{}^C}$.

We shall show that the sets W_i are pairwise disjoint and that $W_1 \cup \ldots \cup W_s$ is linearly independent. This will follow from the following two claims, which state that each W_i is linearly independent and that W_i and W_j are orthogonal for distinct i and j.

Claim 12. *W_i is linearly independent for $i \in [s]$.*

Proof. Suppose that $\sum_{A \in \overline{\mathcal{F}_i{}^C}} \alpha_A Q_{i,A} = 0$, where $\alpha_A \in \mathbb{R}$ and not all α_A's are 0. Let A be a largest set such that $\alpha_A \neq 0$. Note that for every $A' \in \overline{\mathcal{F}_i{}^C}$, $Q_{i,A'}$ can be written as

$$Q_{i,A'} = P_{A'} + \sum_{S \subsetneq A'} \beta_{A',S} P_S,$$

where the values of $\beta_{A',S}$ depend on the representation of A' as in (6.7). Hence, by choice of A,

$$\begin{aligned}
0 &= \sum_{A' \in \overline{\mathcal{F}_i{}^C}, |A'| \leq |A|} \alpha_{A'} Q_{i,A'} \\
&= \sum_{A' \in \overline{\mathcal{F}_i{}^C}, |A'| \leq |A|} \alpha_{A'} (P_{A'} + \sum_{S \subsetneq A'} \beta_{A',S} P_S) \\
&= \alpha_A P_A + \sum_{|S| \leq |A|, S \neq A} \gamma_S P_S,
\end{aligned}$$

for some $\gamma_S \in \mathbb{R}$. However, since the P_S's are linearly independent (by Lemma 6.2.7), we have $\alpha_A = 0$, a contradiction. It follows that W_i is linearly independent, as required. $\square$

Claim 13. *W_i and W_j are orthogonal for $1 \leq i < j \leq s$.*

Proof. Let $A \in \overline{\mathcal{F}_i{}^C}$ and $A' \in \overline{\mathcal{F}_j{}^C}$, where $1 \leq i < j \leq s$. Write $A = B \cup C$ and $A' = B' \cup C'$ for the representations as in (6.7). Let $x \in \{0,1\}^n$. We claim that $Q_{i,A}(x) = 0$ or $Q_{j,A'}(x) = 0$. Indeed, if the former does not hold, then $x_i = 1$ for $i \in B$ and $x_i = 0$ for $i \in C$. Note that $B' \cap C \neq \emptyset$, because $\{\mathcal{F}_1, \ldots, \mathcal{F}_s\}$ is cross dependant. Hence, $x_i = 0$ for some $i \in B'$, which implies that $Q_{i,A'}(x) = 0$, as claimed. It easily follows that $\langle Q_{i,A}, Q_{j,A'} \rangle = 0$ (recall the definition of the inner product given in (6.6)), as required. $\square$

It follows from Claims 12 and 13 that $W_1 \cup \ldots \cup W_s$ is linearly independent, hence it has size at most the dimension of V, i.e. at most 2^n. But $|W_i| = |\mathcal{F}_i{}^C|$, thus, as in (6.5)

$$|\mathcal{F}_1| + \ldots + |\mathcal{F}_s| \geq (s-1)2^n,$$

as desired. $\square$

6.2.3 *Conclusion*

There are two main directions for further research that we would like to mention here.

The first is related to the tightness of Theorem 6.2.5. As mentioned in the introduction, the result is tight, which can be seen by taking $\mathcal{F}_1 = \emptyset$ and $\mathcal{F}_2 = \ldots = \mathcal{F}_s = 2^{[n]}$. In fact, this is a special case of the following class of examples: let $\mathcal{F}_1$ be any increasing family in $2^{[n]}$, let $\mathcal{F}_2 = \overline{\mathcal{F}_1{}^C}$ and let $\mathcal{F}_3 = \ldots = \mathcal{F}_s = 2^{[n]}$. Then $|\mathcal{F}_1| + |\mathcal{F}_2| = 2^n$ and it is easy to check that any set in $\mathcal{F}_1$ intersect every set in $\mathcal{F}_2$. Therefore, every such example yields a cross saturated set of smallest size. Furthermore, it is easy to see that these are the only examples for which $\mathcal{F}_3 = \ldots = \mathcal{F}_s$. It seems plausible that these are the only possible examples (up to permuting the order of the families). This problem of classifying all extremal examples, interesting in its own right, may give a hint on how to further improve the lower bound of the size of s-saturated families.

The second, and seemingly more challenging direction, is to improve on Theorem 6.2.4. We proved that if $\mathcal{F}$ is s-saturated then $|\mathcal{F}| \geq (1 - 1/s)2^n$, where the conjectured bound is $\left(1 - 2^{-(s-1)}\right) 2^n$. We note that it is possible to improve the lower bound slightly, to show that $|\mathcal{F}| \geq (1 - 1/s + \Omega(\log n/n))\, 2^n$, by running the argument of the first proof more carefully in the case where $\mathcal{F}_1 = \ldots = \mathcal{F}_s = \mathcal{F}$; we omit further details. It would be very interesting to obtain an improvement of error term $1/s$ to an expression exponential in s. We hope that our methods can be used to make further progress on this old conjecture.

Let us mention here a general class of examples of s-saturated families whose size is $\left(1 - 2^{-(s-1)}\right) 2^n$. We do not know of any other examples of s-saturated families, and feel that it is likely that if the conjecture holds, then these are the only extremal examples.

Example 6.2.8. *Given $s \geq 2$, let $\{I_1, \ldots, I_{s-1}\}$ be a partition of $[n]$. For each $i \in [s-1]$, pick a maximal intersecting family $\mathcal{F}_i$ of subsets of I_i; in particular, by Proposition 6.2.2, $|\mathcal{F}_i| = 2^{|I_i|-1}$. Define $\mathcal{F}$ as follows.*

$$\mathcal{F} = \{A \subseteq [n] : A \cap I_i \in \mathcal{F}_i \text{ for some } i \in [s-1].\}$$

It is easy to check that $\mathcal{F}$ is s-saturated as a family of subsets of $[n]$ and that it has size $\left(1 - 2^{-(s-1)}\right) 2^n$.

Note that this class of examples contains the example that was mentioned earlier, of the family of subsets of $[n]$ that intersect $[s-1]$.

Finally, we note the following interesting phenomenon.

Proposition 6.2.9. *If Conjecture 6.2.3 holds for $s+1$, then it holds for s.*

Indeed, suppose that Conjecture 6.2.3 holds for $s+1$, and let $\mathcal{F} \subseteq 2^{[n]}$ be s-saturated. Define $\mathcal{G} \subseteq 2^{[n+1]}$ as follows.

$$\mathcal{G} = \mathcal{F} \cup \{A \subseteq [n+1] : n+1 \in A\}.$$

Note that $\mathcal{G}$ is $(s+1)$-saturated (as a subset of $2^{[n+1]}$). Hence, by the assumption that the conjecture holds for $s+1$, we find that $|\mathcal{G}| \geq (1 - 2^{-s})\, 2^{n+1}$. Note also that $|\mathcal{G}| = |\mathcal{F}| + 2^n$. It follows that $|\mathcal{F}| \geq (1 - 2^{-s})\, 2^{n+1} - 2^n = \left(1 - 2^{-(s-1)}\right) 2^n$, as required.

BIBLIOGRAPHY

1. Abbott, H. L. A note on Ramsey's theorem. *Canad. Math. Bull.* **15**, 9–10 (1972).
2. Abbott, H. L. & Moser, L. On a combinatorial problem of Erdős and Hajnal. *Canad. Math. Bull.* **7**, 177–181 (1964).
3. Abu-Khazneh, A., Barát, J., Pokrovskiy, A. & Szabó, T. A family of extremal hypergraphs for Ryser's conjecture. *J. Combin. Theory Ser. A* **161**, 164–177 (2019).
4. Abu-Khazneh, A. & Pokrovskiy, A. Intersecting extremal constructions in Ryser's conjecture for r-partite hypergraphs. *J. Combin. Math. Combin. Comput.* **103**, 81–104 (2017).
5. Adler, I., Alon, N. & Ross, S. M. On the maximum number of Hamiltonian paths in tournaments. *Random Structures Algorithms* **18**, 291–296 (2001).
6. Aharoni, R. Ryser's conjecture for tripartite 3-graphs. *Combinatorica* **21**, 1–4 (2001).
7. Aharoni, R., Barát, J. & Wanless, I. M. Multipartite hypergraphs achieving equality in Ryser's conjecture. *Graphs Combin.* **32**, 1–15 (2016).
8. Aharoni, R. & Berger, E. The intersection of a matroid and a simplicial complex. *Trans. Amer. Math. Soc.* **358**, 4895–4917 (2006).
9. Ahlswede, R. & Katona, G. Contributions to the geometry of Hamming spaces. *Discr. Math.* **17**, 1–22 (1977).
10. Ahlswede, R. & Khachatrian, L. H. A pushing-pulling method: new proofs of intersection theorems. *Combinatorica* **19**, 1–15 (1999).
11. Ajtai, M., Komlós, J. & Szemerédi, E. A note on Ramsey numbers. *J. Combin. Theory Ser. A* **29**, 354–360 (1980).
12. Albertson, M., Bollobas, B. & Tucker, S. *The independence ratio and maximum degree of a graph* in *Proceedings of the Seventh Southeastern Conference on Combinatorics, Graph Theory, and Computing (Louisiana State Univ., Baton Rouge, La., 1976)* (1976), 43–50. Congressus Numerantium, No. XVII.
13. Alon, N. An extremal problem for sets with applications to graph theory. *J. Combin. Theory Ser. A* **40**, 82–89 (1985).

14. Alon, N., Babai, L. & Suzuki, H. Multilinear polynomials and Frankl-Ray-Chaudhuri-Wilson type intersection theorems. *J. Combin. Theory Ser. A.* **58**, 165–180 (1991).

15. Alon, N., Krivelevich, M. & Sudakov, B. Induced subgraphs of prescribed size. *J. Graph Theory* **43**, 239–251 (2003).

16. Alon, N. & Spencer, J. H. *The probabilistic method* Fourth Ed., xiv+375 (John Wiley & Sons, Inc., Hoboken, NJ, 2016).

17. Alon, N. & Sudakov, B. On graphs with subgraphs having large independence numbers. *J. Graph Theory* **56**, 149–157 (2007).

18. Alon, N. & Tarsi, M. Colorings and orientations of graphs. *Combinatorica* **12**, 125–134 (1992).

19. Alon, N., Gutin, G. & Krivelevich, M. Algorithms with large domination ratio. *J. Algorithms* **50**, 118–131 (2004).

20. Arora, S., Lovász, L., Newman, I., Rabani, Y., Rabinovich, Y. & Vempala, S. Local versus global properties of metric spaces. *SIAM J. Comput.* **41**, 250–271 (2012).

21. Bal, D. & DeBiasio, L. Partitioning random graphs into monochromatic components. *Electron. J. Combin.* **24**, P1.18 (2017).

22. Balko, M. & Vizer, M. Edge-ordered Ramsey numbers. *European J. Combin.* **87** (2020).

23. Balogh, J., Clemen, F. C., Lavrov, M., Lidickỳ, B. & Pfender, F. Making K_{r+1}-Free Graphs r-partite. *accepted in Combin. Probab. Comput.* (2020).

24. Balogh, J., Liu, H. & Montgomery, R. Rainbow spanning trees in properly coloured complete graphs. *Discrete Appl. Math.* **247**, 97–101 (2018).

25. Balogh, J. & Samotij, W. An efficient container lemma. *Discrete Anal.*, Paper No. 17, 56 (2020).

26. Barak, B., Rao, A., Shaltiel, R. & Wigderson, A. 2-source dispersers for $n^{o(1)}$ entropy, and Ramsey graphs beating the Frankl-Wilson construction. *Ann. of Math. (2)* **176**, 1483–1543 (2012).

27. Basu, S., Pollack, R. & Roy, M. F. *Algorithms in real algebraic geometry* viii+602 (Springer-Verlag, Berlin, 2003).

28. Beame, P. & Huynh-Ngoc, D. *On the Value of Multiple Read/Write Streams for Approximating Frequency Moments* in *2008 49th Annual IEEE Symposium on Foundations of Computer Science* (2008), 499–508.

29. Beame, P., Blais, E. & Huynh-Ngoc, D.-T. *Longest Common Subsequences in Sets of Permutations* 2009.
30. Beck, J. On 3-chromatic hypergraphs. *Discrete Math.* **24**, 127–137 (1978).
31. Beck, J. On a combinatorial problem of P. Erdős and L. Lovász. *Discrete Math.* **17**, 127–131 (1977).
32. Bennett, P., DeBiasio, L., Dudek, A. & English, S. Large monochromatic components and long monochromatic cycles in random hypergraphs. *European J. Combin.* **76**, 123–137 (2019).
33. Bernstein, F. Zur Theorie der trigonometrischen Reihe. *J. Reine Angew. Math.* **132**, 270–278 (1907).
34. Bodirsky, M. & Kára, J. The complexity of temporal constraint satisfaction problems. *J. ACM* **57**, Art. 9, 41 (2010).
35. Bohman, T. & Keevash, P. The early evolution of the H-free process. *Invent. Math.* **181**, 291–336 (2010).
36. Bohman, T. & Keevash, P. in *The Seventh European Conference on Combinatorics, Graph Theory and Applications* 489–495 (Ed. Norm., Pisa, 2013).
37. Bollen, G. P. & Draisma, J. An online version of Rota's basis conjecture. *J. Algebraic Combin.* **41**, 1001–1012 (2015).
38. Bollobás, B. *Combinatorics: set systems, hypergraphs, families of vectors, and combinatorial probability* (Cambridge University Press, 1986).
39. Bollobás, B. On generalized graphs. *Acta Math. Acad. Sci. Hungar.* **16**, 447–452 (1965).
40. Bollobás, B. & Hind, H. R. Graphs without large triangle free subgraphs. *Discrete Math.* **87**, 119–131 (1991).
41. Borg, P. Intersecting families of sets and permutations: a survey. *Int. J. Math. Game Theory Algebra* **21**, 543–559 (2012).
42. Brualdi, R. A. & Hollingsworth, S. Multicolored trees in complete graphs. *J. Combin. Theory Ser. B* **68**, 310–313 (1996).
43. Bucic, M., Kwan, M., Pokrovskiy, A. & Sudakov, B. *On Kahn's basis conjecture.* Companion Note. arXiv:1810.07464.
44. Bucic, M., Lidický, B., Long, J. & Wagner, A. Z. Partition problems in high dimensional boxes. *J. Combin. Theory Ser. A* **166**, 315–336 (2019).
45. Bukh, B. & Sevekari, A. Linear orderings of combinatorial cubes. *preprint arXiv:1906.11866* (2019).
46. Bukh, B. & Zhou, L. Twins in words and long common subsequences in permutations. *Israel J. Math.* **213**, 183–209 (2016).

47. Burkill, H. & Mirsky, L. Combinatorial problems on the existence of large submatrices. I. *Discrete Math.* **6**, 15–28 (1973).

48. Burkill, H. & Mirsky, L. Monotonicity. *J. Math. Anal. Appl.* **41**, 391–410 (1973).

49. Chattopadhyay, E. & Zuckerman, D. Explicit two-source extractors and resilient functions. *Ann. of Math. (2)* **189**, 653–705 (2019).

50. Cherkashin, D. D. & Kozik, J. A note on random greedy coloring of uniform hypergraphs. *Random Structures Algorithms* **47**, 407–413 (2015).

51. Cheung, M. *Computational proof of Rota's basis conjecture for matroids of rank 4* unpublished manuscript. `http://educ.jmu.edu/~duceyje/undergrad/2012/mike.pdf`. 2012.

52. Chow, T. *Proposals for polymath projects* `https://mathoverflow.net/q/231153`. 2017.

53. Chow, T. *Rota's basis conjecture: Polymath 12* `https://polymathprojects.org/2017/02/23/rotas-basis-conjecture-polymath-12/`. 2017.

54. Chow, T. *Rota's basis conjecture: Polymath 12* `https://polymathprojects.org/2017/03/06/rotas-basis-conjecture-polymath-12-2/`. 2017.

55. Chow, T. *Rota's conjecture* `http://michaelnielsen.org/polymath1/index.php?title=Rota%27s_conjecture`. 2017.

56. Chung, F. R. K. & Graham, R. L. Quasi-random set systems. *J. Amer. Math. Soc.* **4**, 151–196 (1991).

57. Chung, F. R. K. & Graham, R. L. Quasi-random tournaments. *J. Graph Theory* **15**, 173–198 (1991).

58. Chung, F. R. K., Graham, R. L. & Wilson, R. M. Quasi-random graphs. *Combinatorica* **9**, 345–362 (1989).

59. Chung, F. & Graham, R. *Erdős on graphs*, His legacy of unsolved problems, xiv+142 (A K Peters, Ltd., Wellesley, MA, 1998).

60. Chvátal, V. Monochromatic paths in edge-colored graphs. *J. Combinatorial Theory Ser. B* **13**, 69–70 (1972).

61. Cohen, G. *Two-source dispersers for polylogarithmic entropy and improved Ramsey graphs* in *STOC'16—Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing* (ACM, New York, 2016), 278–284.

62. Conlon, D. & Gowers, W. T. Combinatorial theorems in sparse random sets. *Ann. of Math. (2)*, 367–454 (2016).

63. Conlon, D., Fox, J., Lee, C. & Sudakov, B. Ordered Ramsey numbers. *J. Combin. Theory Ser. B* **122**, 353–383 (2017).

64. Conlon, D., Fox, J. & Sudakov, B. Hereditary quasirandomness without regularity. *Math. Proc. Cambridge Philos. Soc.* **164**, 385–399 (2018).

65. Conlon, D., Fox, J. & Sudakov, B. Ramsey numbers of sparse hypergraphs. *Random Structures Algorithms* **35**, 1–14 (2009).

66. Conlon, D., Hàn, H., Person, Y. & Schacht, M. Weak quasi-randomness for uniform hypergraphs. *Random Structures Algorithms* **40**, 1–38 (2012).

67. Cooper, J. N. Quasirandom permutations. *J. Combin. Theory Ser. A* **106**, 123–143 (2004).

68. Coregliano, L. N., Parente, R. F. & Sato, C. M. On the maximum density of fixed strongly connected subtournaments. *Electron. J. Combin.* **26**, Paper No. 1.44, 48 (2019).

69. Coregliano, L. N. & Razborov, A. A. On the density of transitive tournaments. *J. Graph Theory* **85**, 12–21 (2017).

70. Danzer, L., Grünbaum, B. & Klee, V. in *Proc. Sympos. Pure Math., Vol. VII* 101–180 (Amer. Math. Soc., Providence, R.I., 1963).

71. Dellamonica Jr., D. & Rödl, V. Hereditary quasirandom properties of hypergraphs. *Combinatorica* **31**, 165–182 (2011).

72. Deza, M. Solution d'un problème de Erdős-Lovász. *J. Combinatorial Theory Ser. B* **16**, 166–167 (1974).

73. Dinur, I. & Safra, S. On the hardness of approximating minimum vertex cover. *Ann. Math.* **162**, 439–485 (2005).

74. Dong, S. & Geelen, J. Improved bounds for Rota's basis conjecture. *Combinatorica* **39**, 265–272 (2019).

75. Drisko, A. A. On the number of even and odd Latin squares of order $p+1$. *Adv. Math.* **128**, 20–35 (1997).

76. Dudek, A., Retter, T. & Rödl, V. On generalized Ramsey numbers of Erdős and Rogers. *J. Combin. Theory Ser. B* **109**, 213–227 (2014).

77. Dudek, A. & Rödl, V. in *Ramsey theory* 63–76 (Birkhauser/Springer, New York, 2011).

78. Erdős, P. in *Second International Conference on Combinatorial Mathematics (New York, 1978)* 177–187 (New York Acad. Sci., New York, 1979).

79. Erdős, P. On a combinatorial problem. *Nordisk Mat. Tidskr.* **11**, 5–10, 40 (1963).

80. Erdős, P. On a combinatorial problem. II. *Acta Math. Acad. Sci. Hungar.* **15**, 445–447 (1964).

81. Erdős, P. On extremal problems of graphs and generalized graphs. *Israel J. Math.* **2**, 183–190 (1964).

82. Erdős, P. in *Graph theory, combinatorics, and applications, Vol. 1 (Kalamazoo, MI, 1988)* 397–406 (Wiley, New York, 1991).

83. Erdős, P. in *Extremal problems for finite sets (Visegrád, 1991)* 217–227 (János Bolyai Math. Soc., Budapest, 1994).

84. Erdős, P. Some remarks on the theory of graphs. *Bull. Amer. Math. Soc.* **53**, 292–294 (1947).

85. Erdős, P., Faudree, R. J., Rousseau, C. C. & Schelp, R. H. On cycle-complete graph Ramsey numbers. *J. Graph Theory* **2**, 53–64 (1978).

86. Erdös, P., Fon-Der-Flaass, D., Kostochka, A. V. & Tuza, Z. in. 1. Siberian Advances in Mathematics, 82–88 (1992).

87. Erdős, P., Gyárfás, A. & Pyber, L. Vertex coverings by monochromatic cycles and trees. *J. Combin. Theory Ser. B* **51**, 90–95 (1991).

88. Erdős, P., Győri, E. & Simonovits, M. in *Sets, graphs and numbers (Budapest, 1991)* 239–263 (North-Holland, Amsterdam, 1992).

89. Erdős, P. & Hajnal, A. On a property of families of sets. *Acta Math. Acad. Sci. Hungar.* **12**, 87–123 (1961).

90. Erdős, P., Hajnal, A. & Moon, J. W. A problem in graph theory. *Amer. Math. Monthly* **71**, 1107–1110 (1964).

91. Erdős, P. & Hindman, N. Enumeration of intersecting families. *Discr. Math.* **48**, 61–65 (1984).

92. Erdős, P. & Kleitman, D. J. Extremal problems among subsets of a set. *Discr. Math.* **8**, 281–194 (1974).

93. Erdős, P., Ko, C. & Rado, R. Intersection theorems for systems of finite sets. *Quart. J. Math. Oxford* **12**, 313–320 (1961).

94. Erdős, P. & Lovász, L. in *Infinite and finite sets (Colloq., Keszthely, 1973; dedicated to P. Erdős on his 60th birthday), Vol. II* 609–627. Colloq. Math. Soc. János Bolyai, Vol. 10 (1975).

95. Erdős, P. & Rado, R. Combinatorial theorems on classifications of subsets of a given set. *Proc. London Math. Soc. (3)* **2**, 417–439 (1952).

96. Erdős, P. & Rogers, C. A. The construction of certain graphs. *Canadian J. Math.* **14**, 702–707 (1962).

97. Erdös, P. & Szekeres, G. A combinatorial problem in geometry. *Compositio Math.* **2**, 463–470 (1935).

98. Erdős, P. & Szemerédi, A. On a Ramsey type theorem. *Period. Math. Hungar.* **2**, 295–299 (1972).

99. Erdős, P. in. 1-3. Trends in discrete mathematics, 53–73 (1994).

100. Erdős, P., Hajnal, A. & Tuza, Z. Local constraints ensuring small representing sets. *J. Combin. Theory Ser. A* **58**, 78–84 (1991).

101. Erdős, P., Rubin, A. L. & Taylor, H. *Choosability in graphs* in *Proceedings of the West Coast Conference on Combinatorics, Graph Theory and Computing (Humboldt State Univ., Arcata, Calif., 1979)* (Utilitas Math., Winnipeg, Man., 1980), 125–157.

102. Felsner, S., Fishburn, P. C. & Trotter, W. T. Finite three-dimensional partial orders which are not sphere orders. *Discrete Math.* **201**, 101–132 (1999).

103. Fernandez de la Vega, W. On the maximum cardinality of a consistent set of arcs in a random tournament. *J. Combin. Theory Ser. B* **35**, 328–332 (1983).

104. Filmus, Y. The weighted complete intersection theorem, *J. Combin. Theory A* **151**, 84–101 (2017).

105. Fishburn, P. C. & Graham, R. L. Lexicographic Ramsey theory. *J. Combin. Theory Ser. A* **62**, 280–298 (1993).

106. Fiz Pontiveros, G., Griffiths, S. & Morris, R. The Triangle-Free Process and the Ramsey Number $R(3, k)$. *Mem. Amer. Math. Soc.* **263**, v+125 (2020).

107. Fon-Der-Flaass, D. G., Kostochka, A. V. & Woodall, D. R. Transversals in uniform hypergraphs with property $(7, 2)$. *Discrete Math.* **207**, 277–284 (1999).

108. Fox, J. & Sudakov, B. Induced Ramsey-type theorems. *Adv. Math.* **219**, 1771–1800 (2008).

109. Fox, J. & Li, R. On edge-ordered Ramsey numbers. *Random Structures Algorithms* **57**, 1174–1204 (2020).

110. Fox, J. & Sudakov, B. Dependent random choice. *Random Structures Algorithms* **38**, 68–99 (2011).

111. Francetić, N., Herke, S., McKay, B. D. & Wanless, I. M. On Ryser's conjecture for linear intersecting multipartite hypergraphs. *European J. Combin.* **61**, 91–105 (2017).

112. Frankl, P. & Tokushige, N. Invitation to intersection problems for finite sets. *J. Combin. Theory, Ser. A* **144**, 157–211 (2016).

113. Frankl, P. & Wilson, R. M. Intersection theorems with geometric consequences. *Combinatorica* **1**, 357–368 (1981).

114. Frankl, P. & Kupavskii, A. Two problems on matchings in set families—in the footsteps of Erdős and Kleitman. *J. Combin. Theory Ser. B* **138**, 286–313 (2019).

115. Frankl, P., Ota, K. & Tokushige, N. Covers in uniform intersecting families and a counterexample to a conjecture of Lovász. *J. Combin. Theory Ser. A* **74**, 33–42 (1996).

116. Friedgut, E. On the measure of intersecting families, uniqueness and stability. *Combinatorica* **28**, 503–528 (2008).

117. Füredi, Z. Matchings and covers in hypergraphs. *Graphs Combin.* **4**, 115–206 (1988).

118. Füredi, Z. An upper bound on Zarankiewicz' problem. *Combin. Probab. Comput.* **5**, 29–33 (1996).

119. Galvin, F. The list chromatic index of a bipartite multigraph. *J. Combin. Theory Ser. B* **63**, 153–158 (1995).

120. Gebauer, H. On the construction of 3-chromatic hypergraphs with few edges. *J. Combin. Theory Ser. A* **120**, 1483–1490 (2013).

121. Geelen, J. & Humphries, P. J. Rota's basis conjecture for paving matroids. *SIAM J. Discrete Math.* **20**, 1042–1045 (2006).

122. Geelen, J. & Webb, K. On Rota's basis conjecture. *SIAM J. Discrete Math.* **21**, 802–804 (2007).

123. Geller, D. & Stahl, S. The chromatic number and other functions of the lexicographic product. *J. Combinatorial Theory Ser. B* **19**, 87–95 (1975).

124. Gerencsér, L. & Gyárfás, A. On Ramsey-type problems. *Ann. Univ. Sci. Budapest. Eötvös Sect. Math.* **10**, 167–170 (1967).

125. Girão, A., Letzter, S. & Sahasrabudhe, J. Partitioning a graph into monochromatic connected subgraphs. *J. Graph Theory, to appear*.

126. Glock, S., Kühn, D., Lo, A. & Osthus, D. The existence of designs via iterative absorption: hypergraph F-designs for arbitrary F. *accepted in Mem. Am. Mat. Soc.* (2020).

127. Glynn, D. G. The conjectures of Alon-Tarsi and Rota in dimension prime minus one. *SIAM J. Discrete Math.* **24**, 394–399 (2010).

128. Godsil, C. & Meagher, K. *Erdős-Ko-Rado Theorems: Algebraic Approaches* (Cambridge University Press, 2016).

129. Gowers, W. T. A new proof of Szemerédi's theorem for arithmetic progressions of length four. *Geom. Funct. Anal.* **8**, 529–551 (1998).

130. Gowers, W. T. Hypergraph regularity and the multidimensional Szemerédi theorem. *Ann. of Math. (2)* **166**, 897–946 (2007).

131. Gowers, W. T. Quasirandom groups. *Combin. Probab. Comput.* **17**, 363–387 (2008).

132. Griffiths, S. Quasi-random oriented graphs. *J. Graph Theory* **74**, 198–209 (2013).

133. Gromov, M. Local and global in geometry. *preprint* `www.ihes.fr/~gromov/wp-content/uploads/2018/08/1107.pdf` (2018).

134. Gustavsson, T. *Decompositions of large graphs and digraphs with high minimum degree* PhD thesis (Univ. of Stockholm, 1991).

135. Gyárfás, A. in *Irregularities of partitions (Fertőd, 1986)* 89–91 (Springer, Berlin, 1989).

136. Gyárfás, A. Vertex covers by monochromatic pieces—A survey of results and problems. *Discrete Mathematics* **339**, 1970–1977 (2016).

137. Gyárfás, A. & Lehel, J. A Ramsey-type problem in directed and bipartite graphs. *Period. Math. Hungar.* **3**, 299–304 (1973).

138. Gyárfás, A., Ruszinkó, M., Sárközy, G. N. & Szemerédi, E. An improved bound for the monochromatic cycle partition number. *J. Combin. Theory Ser. B* **96**, 855–873 (2006).

139. Hancock, R., Kabela, A., Král', D., Martins, T., Parente, R., Skerman, F. & J., V. No additional tournaments are quasirandom-forcing. *preprint arXiv:1912.04243*.

140. Hasse, H. Darstellbarkeit von Zahlen durch quadratische Formen in einem beliebigen algebraischen Zahlkörper. *J. Reine Angew. Math.* **153**, 113–130 (1924).

141. Haxell, P. E. & Scott, A. D. A note on intersecting hypergraphs with large cover number. *Electron. J. Combin.* **24**, P3.26 (2017).

142. Haxell, P. E. & Scott, A. D. On Ryser's conjecture. *Electron. J. Combin.* **19**, P23 (2012).

143. Helly, E. Über Mengen konvexer Körper mit gemeinschaftlichen Punkte. *Jahresbericht der Deutschen Mathematiker-Vereinigung* **32**, 175–176 (1923).

144. Henderson, J. R. *Permutation decomposition of (0, 1)-matrices and decomposition transversals* PhD thesis (California Institute of Technology, 1971).

145. Horn, P. Rainbow spanning trees in complete graphs colored by one-factorizations. *J. Graph Theory* **87**, 333–346 (2018).

146. Huang, R. & Rota, G.-C. On the relations of various conjectures on Latin squares and straightening coefficients. *Discrete Math.* **128**, 225–236 (1994).

147. Janzer, O. & Gowers, W. T. Improved bounds for the Erdős-Rogers function. *Advances in Combinatorics* **3** (2020).

148. Jenssen, M., Keevash, P., Long, E. & Yepremyan, L. Distinct degrees in induced subgraphs. *Proc. Amer. Math. Soc.* **148**, 3835–3846 (2020).

149. Jones, K. F. Independence in graphs with maximum degree four. *J. Combin. Theory Ser. B* **37**, 254–269 (1984).

150. Kahn, J. On a problem of Erdős and Lovász: random lines in a projective plane. *Combinatorica* **12**, 417–423 (1992).

151. Kahn, J. On a problem of Erdős and Lovász. II. $n(r) = O(r)$. *J. Amer. Math. Soc.* **7**, 125–143 (1994).

152. Kalmanson, K. On a theorem of Erdős and Szekeres. *J. Combinatorial Theory Ser. A* **15**, 343–346 (1973).

153. Kalyanasundaram, S. & Shapira, A. A note on even cycles and quasirandom tournaments. *J. Graph Theory* **73**, 260–266 (2013).

154. Kamčev, N., Krivelevich, M., Morrison, N. & Sudakov, B. The Kőnig graph process. *Random Structures Algorithms* **57**, 1272–1302 (2020).

155. Katona, G. Intersection theorems for systems of finite sets. *Acta Math. Acad. Sci. Hungar.* **15**, 329–337 (1964).

156. Kedlaya, K. *From Quadratic Reciprocity to Class Field Theory* (in Princeton Companion to Mathematics, Princeton University Press, 2008).

157. Kirkman, T. P. On a problem in combinations. *The Cambridge and Dublin Mathematical Journal*, 191–204 (1847).

158. Kohayakawa, Y., Mota, G. O. & Schacht, M. *Monochromatic trees in random graphs* in *Mathematical Proceedings of the Cambridge Philosophical Society* (2018), 1–18.

159. Kolaitis, P. G., Prömel, H. J. & Rothschild, B. L. K_{l+1}-free graphs: asymptotic structure and a 0-1 law. *Trans. Amer. Math. Soc.* **303**, 637–671 (1987).

160. Korándi, D., Roberts, A. & Scott, A. Exact stability for Turán's Theorem. *preprint arXiv:2004.10685* (2020).

161. Korándi, D., Lang, R., Letzter, S. & Pokrovskiy, A. Minimum degree conditions for monochromatic cycle partitioning. *J. Combin. Theory Ser. B* **146**, 96–123 (2021).

162. Korándi, D., Mousset, F., Nenadov, R., Škorić, N. & Sudakov, B. Monochromatic cycle covers in random graphs. *Random Structures & Algorithms* **53**, 667–691 (2018).

163. Kostochka, A. & Yancey, M. Ore's conjecture on color-critical graphs is almost true. *J. Combin. Theory Ser. B* **109**, 73–101 (2014).

164. Kostochka, A. V. & Rödl, V. On graphs with small Ramsey numbers. *J. Graph Theory* **37**, 198–204 (2001).

165. Kostochka, A. V. in. 2. Special issue: Paul Erdős and his mathematics, 275–285 (2002).

166. Král', D. & Pikhurko, O. Quasirandom permutations are characterized by 4-point densities. *Geom. Funct. Anal.* **23**, 570–579 (2013).

167. Krivelevich, M. K^s-Free Graphs Without Large K^r-Free Subgraphs. *Combin. Probab. Comput.* **3**, 349–354 (1994).

168. Krivelevich, M. On the minimal number of edges in color-critical graphs. *Combinatorica* **17**, 401–426 (1997).

169. Krivelevich, M., Kwan, M., Loh, P.-S. & Sudakov, B. The random k-matching-free process. *Random Structures Algorithms* **53**, 692–716 (2018).

170. Krivelevich, M. & Sudakov, B. in *More sets, graphs and numbers* 199–262 (Springer, Berlin, 2006).

171. Kruskal Jr., J. B. Monotonic subsequences. *Proc. Amer. Math. Soc.* **4**, 264–274 (1953).

172. Kwan, M. & Sudakov, B. Proof of a conjecture on induced subgraphs of Ramsey graphs. *Trans. Amer. Math. Soc.* **372**, 5571–5594 (2019).

173. Kwan, M. & Sudakov, B. Ramsey Graphs Induce Subgraphs of Quadratically Many Sizes. *Int. Math. Res. Not. IMRN, to appear* (2017).

174. Lang, R. & Lo, A. Monochromatic cycle partitions in random graphs. *Combin. Probab. Comput.* **30**, 136–152 (2021).

175. Lehel, J. in *Combinatorial mathematics (Marseille-Luminy, 1981)* 413–418 (North-Holland, Amsterdam, 1983).

176. Li, Y. & Zang, W. The independence number of graphs with a forbidden cycle and Ramsey numbers. *J. Comb. Optim.* **7**, 353–359 (2003).

177. Linial, N. Local-global phenomena in graphs. *Combin. Probab. Comput.* **2**, 491–503 (1993).

178. Linial, N. & Rabinovich, Y. Local and global clique numbers. *Journal of Combinatorial Theory, Series B* **61**, 5–15 (1994).
179. Linial, N. & Simkin, M. Monotone subsequences in high-dimensional permutations. *Combin. Probab. Comput.* **27**, 69–83 (2018).
180. Lovász, L. *Combinatorial problems and exercises* Second, 635 (North-Holland Publishing Co., Amsterdam, 1993).
181. Lovász, L. On minimax theorems of combinatorics. *Mat. Lapok* **26**, 209–264 (1978) (1975).
182. Lovász, L. On the ratio of optimal integral and fractional covers. *Discrete Math.* **13**, 383–390 (1975).
183. McKay, B. Database of Ramsey graphs. `http://users.cecs.anu.edu.au/bdm/data/ramsey.html`.
184. Miller, E. W. Concerning biconnected sets. *Fund. Math.* **29**, 123–133 (1937).
185. Miller, E. W. On a property of families of sets. *C. R. Soc. Sci. Lett. Varsovie, Cl. III* **30**, 31–38 (1937).
186. Montgomery, R., Pokrovskiy, A. & Sudakov, B. Decompositions into spanning rainbow structures. *Proc. Lond. Math. Soc. (3)* **119**, 899–959 (2019).
187. Morse, A. P. Subfunction structure. *Proc. Amer. Math. Soc.* **21**, 321–323 (1969).
188. Moshkovitz, G. & Shapira, A. Exact bounds for some hypergraph saturation problems. *J. Combin. Theory Ser. B* **111**, 242–248 (2015).
189. Nagle, B., Rödl, V. & Schacht, M. The counting lemma for regular k-uniform hypergraphs. *Random Structures Algorithms* **28**, 113–179 (2006).
190. Naor, M. *Constructing Ramsey graphs from small probability spaces* (Citeseer, 1992).
191. Nešetřil, J., Prömel, H. J., Rödl, V. & Voigt, B. Canonizing ordering theorems for Hales-Jewett structures. *J. Combin. Theory Ser. A* **40**, 394–408 (1985).
192. Nešetřil, J. Some nonstandard Ramsey-like applications. *Theoret. Comput. Sci.* **34**, 3–15 (1984).
193. Nešetřil, J. & Rödl, V. Statistics of orderings. *Abh. Math. Semin. Univ. Hambg.* **87**, 421–433 (2017).
194. Noga, A. & Shapira, A. Testing subgraphs of directed graphs. *J. Comput. System Sci.* **69**, 353–382 (2004).

195. O'Donnell, R. *Analysis of Boolean functions* xx+423 (Cambridge University Press, New York, 2014).

196. Onn, S. A colorful determinantal identity, a conjecture of Rota, and Latin squares. *Amer. Math. Monthly* **104**, 156–159 (1997).

197. Osthus, D. & Taraz, A. Random maximal H-free graphs. *Random Structures Algorithms* **18**, 61–82 (2001).

198. P. Keevash. The existence of designs. *preprint arXiv:1401.3665* (2014).

199. Pluhár, A. Greedy colorings of uniform hypergraphs. *Random Structures Algorithms* **35**, 216–221 (2009).

200. Pokrovskiy, A. Partitioning edge-coloured complete graphs into monochromatic cycles and paths. *J. Combin. Theory Ser. B* **106**, 70–97 (2014).

201. Pokrovskiy, A. & Sudakov, B. Linearly many rainbow trees in properly edge-coloured complete graphs. *J. Combin. Theory Ser. B* **132**, 134–156 (2018).

202. Polymath, D. H. J. *Rota's basis conjecture online for matroids* unpublished manuscript. `https://www.overleaf.com/8773999gccdbdmfdgkm`. 2017.

203. *Problèmes combinatoires et théorie des graphes* Colloque International CNRS held at the Université d'Orsay, Orsay, July 9–13, 1976, xiv+443 (Éditions du Centre National de la Recherche Scientifique (CNRS), Paris, 1978).

204. Prömel, H. J. & Rödl, V. Non-Ramsey graphs are $c \log n$-universal. *J. Combin. Theory Ser. A* **88**, 379–384 (1999).

205. Prömel, H. J. *Some remarks on natural orders for combinatorial cubes* in *Proceedings of the Oberwolfach Meeting "Kombinatorik" (1986)* **73** (1989), 189–198.

206. Radhakrishnan, J. & Srinivasan, A. Improved bounds and algorithms for hypergraph 2-coloring. *Random Structures Algorithms* **16**, 4–32 (2000).

207. Rado, R. A theorem on independence relations. *Quart. J. Math. Oxford Ser.* **13**, 83–89 (1942).

208. Raigorodskii, A. M. & Cherkashin, D. D. Extremal problems in hypergraph colouring. *Uspekhi Mat. Nauk* **75**, 95–154 (2020).

209. Ramsey, F. P. On a Problem of Formal Logic. *Proc. London Math. Soc. (2)* **30**, 264–286 (1929).

210. Ray-Chaudhuri, D. K. & Wilson, R. M. On *t*-designs. *Osaka J. Math.* **12**, 737–744 (1975).

211. Reimer, D. Proof of the van den Berg-Kesten conjecture. *Combin. Probab. Comput.* **9**, 27–32 (2000).

212. Roberts, A. & Scott, A. Stability results for graphs with a critical edge. *European J. Combin.* **74**, 27–38 (2018).

213. Rödl, V. & Skokan, J. Regularity lemma for *k*-uniform hypergraphs. *Random Structures Algorithms* **25**, 1–42 (2004).

214. Schacht, M. Extremal results for random discrete structures. *Ann. of Math. (2)*, 333–365 (2016).

215. Schacht, M. *Regularity lemma and its applications* Lecture Notes. `https://www.math.uni-hamburg.de/home/schacht/lnotes/GT/SzRL.pdf`.

216. Schmidt, W. M. Ein kombinatorisches Problem von P. Erdős und A. Hajnal. *Acta Math. Acad. Sci. Hungar.* **15**, 373–374 (1964).

217. Shearer, J. B. A note on the independence number of triangle-free graphs. II. *J. Combin. Theory Ser. B* **53**, 300–307 (1991).

218. Shelah, S. Erdős and Rényi conjecture. *J. Combin. Theory Ser. A* **82**, 179–185 (1998).

219. Siders, R. Monotone subsequences in any dimension. *J. Combin. Theory Ser. A* **85**, 243–253 (1999).

220. Simonovits, M. & Sós, V. Intersection theorems on structures. *Ann. Discr. Math.* **6**, 301–313 (1980).

221. Simonovits, M. & Sós, V. T. Hereditarily extended properties, quasi-random graphs and not necessarily induced subgraphs. *Combinatorica* **17**, 577–596 (1997).

222. Simonovits, M. & Sós, V. T. in. 3. Combinatorics, probability and computing (Oberwolfach, 2001), 319–344 (2003).

223. Spencer, J. Optimal ranking of tournaments. *Networks* **1**, 135–138 (1971).

224. Spencer, J. Coloring *n*-sets red and blue. *J. Combin. Theory Ser. A* **30**, 112–113 (1981).

225. Staton, W. Some Ramsey-type numbers and the independence ratio. *Trans. Amer. Math. Soc.* **256**, 353–370 (1979).

226. Steele, J. M. in *Discrete probability and algorithms (Minneapolis, MN, 1993)* 111–131 (Springer, New York, 1995).

227. Sudakov, B. A new lower bound for a Ramsey-type problem. *Combinatorica* **25**, 487–498 (2005).

228. Sudakov, B. A note on odd cycle-complete graph Ramsey numbers. *Electron. J. Combin.* **9**, Note 1, 4 (2002).

229. Sudakov, B. Large K_r-free subgraphs in K_s-free graphs and some other Ramsey-type problems. *Random Structures Algorithms* **26**, 253–265 (2005).

230. Sudakov, B. A few remarks on Ramsey-Turán-type problems. *J. Combin. Theory Ser. B* **88**, 99–106 (2003).

231. Szabó, T. & Tardos, G. A multidimensional generalization of the Erdős-Szekeres lemma on monotone subsequences. *Combin. Probab. Comput.* **10**, 557–565 (2001).

232. Talagrand, M. in *Probability in Banach Spaces, 9* 293–297 (Birkhäuser Boston, Boston, MA, 1994).

233. Tao, T. & Ziegler, T. Concatenation theorems for anti-Gowers-uniform functions and Host-Kra characteristic factors. *Discrete Anal.*, Paper No. 13, 60 (2016).

234. Thomason, A. in *Random graphs '85 (Poznań, 1985)* 307–331 (North-Holland, Amsterdam, 1987).

235. Thomason, A. in *Surveys in combinatorics 1987 (New Cross, 1987)* 173–195 (Cambridge Univ. Press, Cambridge, 1987).

236. Tuza, Z. *On the order of vertex sets meeting all edges of a* 3*-partite hypergraph* in *Proceedings of the First Catania International Combinatorial Conference on Graphs, Steiner Systems, and their Applications, Vol. 1 (Catania, 1986)* **24 A** (1987), 59–63.

237. Van den Berg, J. & Kesten, H. Inequalities with applications to percolation and reliability. *J. Appl. Probab.* **22**, 556–569 (1985).

238. Whitney, H. On the Abstract Properties of Linear Dependence. *Amer. J. Math.* **57**, 509–533 (1935).

239. Wild, M. On Rota's problem about n bases in a rank n matroid. *Adv. Math.* **108**, 336–345 (1994).

240. Wolfovitz, G. K_4-free graphs without large induced triangle-free subgraphs. *Combinatorica* **33**, 623–631 (2013).

CURRICULUM VITAE

PERSONAL DATA

Name	Matija Bucic
Date of Birth	October 04, 1993
Place of Birth	Zagreb, Croatia
Citizen of	Croatia

EDUCATION

2015 – 2016	University of Cambridge, Trinity College, Cambridge, United Kingdom *Final degree:* Master of Mathematics
2012 – 2015	University of Cambridge, Trinity College, Cambridge, United Kingdom *Final degree:* Bachelor of Arts

AWARDS AND SCHOLARSHIPS

Grants

- I was awarded an SNSF grant for the project: Local to Global Problems in Extremal and Probabilistic Combinatorics. Duration: 18 months, amount 83000USD.

Prizes and competitions

- I have received a Woods Prize 2014 for my mathematical essay on "Random Evolution of Graphs".
- 2 silver and 1 bronze medal at International Mathematical Olympiads (2010-2012).

Scholarships

- Senior Scholarship from Trinity College (2013-2016)
- Cambridge Summer Research Scholarship (2015)
- Trinity Overseas Bursary and Cambridge Overseas Trust (2012-2016) – full scholarship (26254£ p.a.)

Travel awards

- I was awarded a SIAM Student Travel Award 2020, Oberwolfach Leibniz Graduate Students Grant 2020 and a BCC 2019 travel award from Clay Mathematics Institute.

RESEARCH TALKS

Invited talks

- Joint Berlin-Tel Aviv workshop on Ramsey theory, *Tel-Aviv University*, Israel, 2020.
- Combinatorics Workshop in Oberwolfach Institute, *Oberwolfach Institute*, Germany, 2020.
- Workshop on Probabilistic and Extremal Combinatorics, *Banff Research Station*, Canada, 2019.
- British Combinatorial Conference, *University of Birmingham*, UK, 2019.
- SIAM Conference on Discrete Mathematics, *University of Colorado Denver*, USA, 2018.

Contributed talks

- Random Structures and Algorithms, *ETH Zürich*, Switzerland (July 2019).
- European Conference on Combinatorics, *TU Wien*, Austria (August 2017).
- Random Structures and Algorithms, *Adam Mickiewicz University in Poznań*, Poland (August 2017).

Seminar talks

- Mittagsseminar, *ETH Zürich* (March 2021).
- Tel Aviv Combinatorics Seminar, (November 2020).
- Stanford Online Combinatorics Seminar, (May 2020).
- Extremal and Probabilistic Combinatorics Webinar, (May 2020).
- Combinatorics seminar, *University of Warwick* (November 2019).
- Mittagsseminar, *ETH Zürich* (November 2019).
- Mittagsseminar, *ETH Zürich* (May 2019).
- Mittagsseminar, *ETH Zürich* (October 2018).
- Combinatorics Seminar, *University of Cambridge* (May 2018).
- Graduate Seminar in Probability, *ETH Zürich* (April 2018).
- Mittagsseminar, *ETH Zürich* (February 2018).
- Mittagsseminar, *ETH Zürich* (March 2017).
- Part III Seminar, *University of Cambridge* (March 2016).

TEACHING EXPERIENCE

- Course organiser and teaching assistant for Graph Theory at ETH (2018/2019,2019/2020,2020/2021)
- Course organiser for Probabilistic Methods in Combinatorics at ETH (2020/2021)
- Course organiser and teaching assistant for Algebraic Methods in Combinatorics at ETH (2019/2020)
- Teaching assistant for Graph Theory course at ETH (2017/2018)
- Supervising a masters thesis of Sven Heberle, resulting in the following publication: "Monochromatic trees in random tournaments", published in CPC.

- Supervising a Bachelor Thesis of Erik Jahn, resulting in the following publication: “2-factors with k cycles in Hamiltonian graphs”, accepted in JCTB.
- Supervised a Master Thesis by Michelle Sweering titled “Covering Problems in Hypergraphs”.
- Supervised a semester project by Michelle Sweering titled “Combinatorial Perspective on the Log-rank Conjecture”.
- Supervised a Master Thesis of Vedran Mihal titled “Generalised Ramsey Theories”.
- Leader of the Croatian International Mathematics Olympiad team at the IMO 2018
- Member of the Croatian International Competitions Team Selection Committee (2016-2020)

ORGANISATIONAL SKILLS AND SERVICE TO THE COMMUNITY

- Organising a full Minisymposium titled “Extremal Problems Involving Colouring”, accepted at SIAM DM 2020 (postponed due to Corona).
- I participated in the organisation of Random Structures and Algorithms conference in Zurich 2019 and Bennyfest, a one day conference celebrating Benny Sudakov’s 50th birthday.
- I have co-founded The European Mathematical Cup, and was a member of its Central Jury for five years. During this time the competition grew to over 200 high school students, from 14 countries.
- I have refereed papers for many major journals: Annals of Mathematics, Combinatorica, JCTB, JCTA, SIDMA, European Journal of Combinatorics, Electronic Journal of Combinatorics, Journal of Graph Theory, Discrete Mathematics, Utilitas Mathematica.

LANGUAGES

- Croatian (native speaker)
- English (full professional proficiency)

- German (limited working proficiency)
- Italian (elementary proficiency)

PUBLICATIONS

Preprints in peer-review:

8. Bucić, M., Gishboliner, L. & Sudakov, B. *Cycles of many lengths in Hamiltonian graphs* preprint arXiv:2104.07633. (2021).
9. Bucić, M., Glock, S. & Sudakov, B. *The intersection spectrum of 3-chromatic intersecting hypergraphs* preprint arXiv:2010.00495. (2020).
12. Bucić, M., Janzer, O. & Sudakov, B. *Counting H-free orientations of graphs* preprint arXiv:2106.08845. (2021).
24. Bucić, M. & Sudakov, B. *Large independent sets from local considerations* preprint arXiv:2007.03667. (2020).
25. Bucić, M., Sudakov, B. & Tran, T. *Erdős-Szekeres theorem for multidimensional arrays* preprint arXiv:1910.13318. (2019).

Articles in peer-reviewed journals:

1. Alon, N., Bucić, M., Kalvari, T., Kuperwasser, E. & Szabó, T. List Ramsey numbers. *J. Graph Theory* **96**, 109–128 (2021).
2. Alon, N., Bucić, M. & Sudakov, B. Large cliques and independent sets all over the place. *Proc. Amer. Math. Soc.* **149**, 3145–3157 (2021).
3. Bucić, M., Devlin, P., Hendon, M., Horne, D. & Lund, B. Perfect matchings and derangements on graphs. *J. Graph Theory* **97**, 340–354 (2021).
4. Bucić, M. An improved bound for disjoint directed cycles. *Discrete Math.* **341**, 2231–2236 (2018).
5. Bucić, M., Draganić, N. & Sudakov, B. Universal and unavoidable graphs. *Combin. Probab. Comput.* (to appear).
6. Bucić, M., Draganić, N., Sudakov, B. & Tran, T. Unavoidable hypergraphs. *J. Combin. Theory Ser. B* (to appear).
7. Bucić, M., Fox, J. & Sudakov, B. Clique minors in graphs with forbidden subgraphs and other restrictions. *Random Structures Algorithms* (to appear).
10. Bucić, M., Heberle, S., Letzter, S. & Sudakov, B. Monochromatic trees in random tournaments. *Combin. Probab. Comput.* **29**, 318–345 (2020).
11. Bucić, M., Jahn, E., Pokrovskiy, A. & Sudakov, B. 2-factors with k cycles in Hamiltonian graphs. *J. Combin. Theory Ser. B* **144**, 150–166 (2020).

13. Bucić, M., Korándi, D. & Sudakov, B. Covering random graphs by monochromatic trees and Helly-type results in hypergraphs. *Combinatorica* (to appear).

14. Bucić, M., Kwan, M., Pokrovskiy, A. & Sudakov, B. Halfway to Rota's basis conjecture. *Int. Math. Res. Not. IMRN*, 8007–8026 (2020).

15. Bucić, M., Kwan, M., Pokrovskiy, A., Sudakov, B., Tran, T. & Wagner, A. Z. Nearly-linear monotone paths in edge-ordered graphs. *Israel J. Math.* **238**, 663–685 (2020).

16. Bucić, M., Letzter, S. & Sudakov, B. 3-color bipartite Ramsey number of cycles and paths. *J. Graph Theory* **92**, 445–459 (2019).

17. Bucić, M., Letzter, S. & Sudakov, B. Directed Ramsey number for trees. *J. Combin. Theory Ser. B* **137**, 145–177 (2019).

18. Bucić, M., Letzter, S. & Sudakov, B. Monochromatic paths in random tournaments. *Random Structures Algorithms* **54**, 69–81 (2019).

19. Bucić, M., Letzter, S. & Sudakov, B. Multicolour bipartite Ramsey number of paths. *Electron. J. Combin.* **26**, Paper No. 3.60, 15 (2019).

20. Bucić, M., Letzter, S., Sudakov, B. & Tran, T. Minimum saturated families of sets. *Bull. Lond. Math. Soc.* **50**, 725–732 (2018).

21. Bucić, M., Lidický, B., Long, J. & Wagner, A. Z. Partition problems in high dimensional boxes. *J. Combin. Theory Ser. A* **166**, 315–336 (2019).

22. Bucić, M., Long, E., Shapira, A. & Sudakov, B. Tournament quasirandomness from local counting. *Combinatorica* **41**, 175–208 (2021).

Peer-reviewed conference contributions:

23. Bucić, M., Ornik, M. & Topcu, U. *Graph-Based Controller Synthesis for Safety-Constrained, Resilient Systems* in *56th Annual Allerton Conference on Communication, Control, and Computing* (2018).