On the structure of graphs with forbidden induced substructures

Zur Erlangung des akademischen Grades einer

Doktorin der Naturwissenschaften

von der KIT-Fakultät für Mathematik des
Karlsruher Instituts für Technologie (KIT)
genehmigte

Dissertation

von

Lea Weber

Tag der mündlichen Prüfung:
12.07.2023

Promotionsausschuss:

1. Referentin: Prof. Dr. Maria Axenovich
2. Referent: PD Dr. Torsten Ueckerdt

Abstract

One of the central goals in extremal combinatorics is to understand how the global structure of a combinatorial object, e.g. a graph, hypergraph or set system, is affected by local constraints. In this thesis we are concerned with structural properties of graphs and hypergraphs which locally do not look like some type of forbidden induced pattern. Patterns can be single subgraphs, families of subgraphs, or in the multicolour version colourings or families of colourings of subgraphs.

Erdős and Szekeres's quantitative version of Ramsey's theorem asserts that in every 2-edge-colouring of the complete graph on n vertices there is a monochromatic clique on at least $\frac{1}{2}\log n$ vertices. The famous Erdős-Hajnal conjecture asserts that forbidding fixed colourings on subgraphs ensures much larger monochromatic cliques. The conjecture is open in general, though a few partial results are known. The first part of this thesis will be concerned with different variants of this conjecture: A bipartite variant, a multicolour variant, and an order-size variant for hypergraphs.

In the second part of this thesis we focus more on order-size pairs; an order-size pair (n, e) is the family consisting of all graphs of order n and size e, i.e. on n vertices with e edges. We consider order-size pairs in different settings: The graph setting, the bipartite setting and the hypergraph setting. In all these settings we investigate the existence of *absolutely avoidable pairs*, i.e. fixed pairs that are avoided by all order-size pairs with sufficiently large order, and also *forcing densities* of order-size pairs (m, f), i.e. for n approaching infinity, the limit superior of the fraction of all possible sizes e, such that the order-size pair (n, e) does not avoid the pair (m, f).

Acknowledgements

This thesis is the result of my time in the Discrete Mathematics research group at KIT and I want to thank everyone who has shared some time in the group with me for always creating a fruitful environment for interesting (not only mathematical) discussions.

First and foremost I would like to thank my thesis advisor Maria Axenovich for sparking my interest in graph theory in the first place, for always throwing new interesting problems at me, and for her outstanding supervision and support. I am grateful for her always open door and ears, for the countless productive brainstorming sessions, for her great insight and invaluable feedback. Many thanks also to Torsten Ueckerdt for refereeing this thesis.

I would also like to thank Casey Tompkins, Richard Snyder, Jean-Sébastien Sereni, Alex Riasanovsky, Jószef Balogh, Felix Christian Clemen and Dhruv Mubayi for their valuable contributions and the productive collaboration.

A special thanks goes to my good friend Lucas for going through the entire thesis and providing valuable feedback on both language and content. Many thanks also to my mother for always supporting me and making tea a diolch i'r iaith Gymraeg am fod mor hyfryd.

Contents

List of Figures

List of Tables

Introduction

Outline of the thesis

In this thesis, we are concerned with the structure of graphs which do not contain some given induced *pattern*. Here a pattern can be a single graph, a family of graphs, a colouring of some graph or a family of colourings of graphs.

In Part I, we are concerned with different variants of the famous Erdős-Hajnal conjecture (EH-conjecture for short), which asserts that forbidding any graph H as an induced subgraph forces a large homogeneous set in the host graph.

Chapter 1: The bipartite version of the Erdős-Hajnal conjecture. It was shown by Erdős, Hajnal and Pach [60] that the EH-conjecture holds in the bipartite setting. Axenovich, Tompkins and the author [15] (see also Master thesis [125]) characterised for which forbidden induced subgraphs the size of a largest homogeneoues set is linear in the number of vertices – except for four open cases. We will show that for these cases it is also linear.

Chapter 2: Bipartite independence number for bounded maximum degree. Here we consider the following natural, yet seemingly not much studied, extremal problem in bipartite graphs: A *bihole* of size t in a bipartite graph G with a fixed bipartition is an independent set with exactly t vertices in each part; in other words, it is a copy of $K_{t,t}$ in the bipartite complement of G. Let $f(n, \Delta)$ be the largest k for which every $(n \times n)$ bipartite graph with maximum degree Δ in one of the parts has a bihole of size k. Thus, determining $f(n, \Delta)$ is the bipartite analogue of finding the largest independent set in graphs with a given number of vertices and bounded maximum degree. It has connections to the bipartite version of the Erdős-Hajnal conjecture, bipartite Ramsey numbers, and the Zarankiewicz problem. Our main result determines the asymptotic behaviour of $f(n, \Delta)$. More precisely, we show that for large but fixed Δ and n sufficiently large, $f(n,\Delta) = \Theta(\frac{\log \Delta}{\Delta} n)$. We further address more specific regimes of Δ, especially when Δ is a small fixed constant. In particular, we determine $f(n,2)$ exactly and obtain bounds for $f(n,3)$, though determining the precise value of $f(n,3)$ is still open. The results are joined work with Axenovich, Sereni and Snyder and are published in SIAM Journal on Discrete Mathematics, 35(2):1136-1148, 2021 [13].

Chapter 3: The multicolour Erdős-Hajnal conjecture. Here we will look at a multicolour version of the Erdős-Hajnal conjecture. Specifically, the most general multicolour version of the conjecture states that for any fixed integers k, s, s' and any s'-edge-colouring c of K_k, there exists $\varepsilon > 0$ such that in any s-edge-colouring of K_n that avoids c there is a clique on at least n^ε vertices, using at most $s-1$ colours. In particular, we reduce the multicolour EH-conjecture to the case where the number of colours is equal to or one more than the numbers of colours used in the forbidden pattern. Most of the results are joint work with Axenovich and Riasanovsky [12].

Chapter 4: The multicolour EH-conjecture for 3 colours and families of triangles. Here, we focus on quantitative aspects of the multicolour EH-conjecture in the case where the number of colours is $s = 3$, and the forbidden colourings are on triangles. More precisely, for a family $\mathcal{H}$ of triangles, each edge-coloured with colours from $\{r, b, y\}$, Forb$(n, \mathcal{H})$ denotes the family of edge-colourings of K_n using colours from $\{r, b, y\}$ and containing none of the colourings from $\mathcal{H}$. Let $h_2(n, \mathcal{H})$ be the maximum q such that any colouring from Forb$(n, \mathcal{H})$ has a clique on at least q vertices using at most two colours. We provide bounds on $h_2(n, \mathcal{H})$ for all families $\mathcal{H}$ consisting of at most three triangles. For most of them our bounds are asymptotically tight. This, in particular, extends a result of Fox, Grinshpun, and Pach, who determined $h_2(n, \mathcal{H})$ for $\mathcal{H}$ consisting of a rainbow triangle. In addition, we prove that for some $\mathcal{H}$, $h_2(n, \mathcal{H})$ corresponds to certain classical Ramsey numbers, smallest independence number in graphs of given odd girth, or some other natural graph theoretic parameters. The results are joined work with Axenovich and Snyder and are published in Discrete Mathematics, 345(5):112791, 2022 [14].

Chapter 5: The Erdős-Hajnal conjecture for order-size pairs. We consider a variant of the Erdős-Hajnal problem for r-graphs where we forbid a family of hypergraphs described by their orders and sizes. For graphs, we observe that if we forbid induced subgraphs on m vertices and f edges for any positive m and $0 \le f \le \binom{m}{2}$, then we obtain large homogeneous sets. For triple systems, in the first nontrivial case $m = 4$, for every $S \subseteq \{0, 1, 2, 3, 4\}$, we give bounds on the minimum size of a homogeneous set in a triple system where the number of edges spanned by every four vertices is not in S. For all S we determine if the growth rate is polylogarithmic. The results of this chapter are joined work with Axenovich and Mubayi and appear in the arXiv preprint https://arxiv.org/abs/2303.09578 [11].

In Part II, we are concerned with forbidding as induced subgraphs so called *order-size-pairs*: a class of graphs/hypergraphs defined by their orders and sizes. Here we do not focus on homogeneous sets, but rather on the question how many large graphs avoid a given small order-size pair. We focus on the existence of *absolutely avoidable* pairs (m, f), i.e. pairs which are not contained in any order size-pair (n, e) for n sufficiently large, and on the *forcing density* of a pair (m, f), i.e. for the order n approaching infinity, the limit superior of the fraction of all possible sizes e, such that the pair (n, e) forces the pair (m, f). We consider these problems for graphs, bipartite graphs, and hypergraphs.

Chapter 6: Order-size pairs in graphs. We call an order-size pair (m, f) of integers, $m \geq 1$, $0 \leq f \leq \binom{m}{2}$, *absolutely avoidable* if there is n_0 such that for any pair of integers (n, e) with $n > n_0$ and $0 \leq e \leq \binom{n}{2}$ there is a graph on n vertices and e edges that contains no induced subgraph on m vertices and f edges. Here we show that there are infinitely many absolutely avoidable pairs. We give a specific infinite set M such that for any $m \in M$ the pair $(m, \binom{m}{2}/2)$ is absolutely avoidable and show that for $m \geq 754$ either $(m, \lfloor \binom{m}{2}/2 \rfloor)$ or $(m, \lfloor \binom{m}{2}/2 \rfloor - 6m)$ is absolutely avoidable. In addition, we show that for any monotone integer function $q(m)$, $|q(m)| = O(m)$ there are infinitely many values of m such that the pair $(m, \binom{m}{2}/2 + q(m))$ is absolutely avoidable. Most of the results are joined work with Axenovich and have been accepted for publication by Journal of Combinatorics [16].

Chapter 7: Bipartite order-size pairs: We investigate the existence of absolutely avoidable pairs and forcing densities in the bipartite setting. The question whether there exist absolutely avoidable pairs in this setting remains open, but we show the existence of infinitely many pairs with forcing density 0 and also infinitely many pairs with forcing density 1.

Chapter 8: Order-size pairs in hypergraphs. We show that for any $r \geq 3$ and $m \geq m_0$, either the pair $(m, \lfloor \binom{m}{r}/2 \rfloor)$ or the pair $(m, \lfloor \binom{m}{r}/2 \rfloor - m - 1)$ is absolutely avoidable. We also show that for $r \geq 3$ most pairs (m, f) have forcing density 0. Further, we show that for $m > r$ there exists no non-trivial pair (m, f) of forcing density 1 and provide some general upper bounds on the forcing density. The results have been accepted for publication by Journal of Combinatorics [126].

Chapter 9: Positive forcing density of order-size pairs in hypergraphs. Answering a question from Chapter 8, we show that $(6, 10)$ is a pair of positive forcing density for $r = 3$ and conjecture that it is the unique such pair. Further, we find necessary conditions for a pair to have positive forcing density, supporting this conjecture. The results are joined work with Axenovich, Balogh and Clemen and have been submitted for publication to SIAM Journal on Discrete Mathematics [10].

Preliminaries

In this Section we will introduce some basic notation, concepts and previously known results that are used throughout the thesis. Less common notions will be introduced in the appropriate place within the chapters where they are needed. For a general introduction to graph theory we refer to the books of Diestel [52] and West [127].

General

For a finite set V and some $n \in \mathbb{N}$ let $\binom{V}{[n]}$ denote the set of all n-element subsets of V. For finite sets X and Y let $X \dot\cup Y$ denote the disjoint union of X and Y and let $X \times Y = \{(x, y) : x \in X, y \in Y\}$. For a positive real number x, let $[x] = \{0, 1, \dots, \lfloor x \rfloor\}$.

For two integers x, y, $x \leq y$, we denote by $[x, y]$ the set of all integers at least x and at most y. For two reals x, y, $x \leq y$, we use the standard notation (x, y), $[x, y)$, $(x, y]$, and $[x, y]$ for respective intervals of reals. For $x \in \mathbb{R}$ let $\{x\} = x - \lfloor x \rfloor$ denote the *fractional part* of x, i.e. $\{x\} \in [0, 1)$ and $\{x\} = x \pmod 1$.

Graphs, bipartite graphs, and hypergraphs

An r-uniform *hypergraph*, or *r-graph* G is a pair $G = (V, E)$ where V is the set of *vertices* and $E \subseteq \binom{V}{[r]}$ is the set of *edges* of G. The *uniformity* of G is r, and if $r = 2$, we refer to G as a *graph*. For an r-graph G, let $V(G)$ be the vertex set and $E(G)$ be the edge set of G. The *order* of G is $|V(G)|$ and the *size* of G is $|E(G)|$. For convenience we write $x_1 x_2 \cdots x_r$ for an edge $\{x_1, x_2, \cdots, x_r\}$ of an r-graph. An r-element subset $f \in \binom{V}{r}$ with $f \notin E$ is called a *non-edge* of G. The complement $\overline{G}$ of G is the r-graph with vertex set $V(\overline{G}) = V(G)$ and edge set $E(\overline{G}) = \binom{V}{[r]} \setminus E(G)$.

Two r-graphs G and G' are *isomorphic* if there is a bijection $f : V(G) \to V(G')$, such that $e \in \binom{V}{[r]}$ is an edge in G if and only if $f(e)$ is an edge in G'. All r-graphs G considered in this thesis are finite, i.e. $V(G)$ is finite and simple, i.e. they contain no multiple edges, since $E(G)$ is not a mulitset. Two vertices contained in a common edge are called *adjacent*. For a vertex $v \in V(G)$ the set $N(v) = \{u \in V(G) \setminus \{v\} : u \text{ is adjacent to } v\}$ is the *neighbourhood* of v, and the elements are the *neighbours* of v. The *degree* $d(u) = d_G(u)$ of a vertex u in an r-graph G is the total number of edges in G that contain u; we omit the subscript G if it is clear from the context. A vertex u of an r-graph is *isolated* if $d(u) = 0$ and a *leaf* if $d(u) = 1$. The *minimum degree* $\delta(G)$ is the smallest and the *maximum degree* $\Delta(G)$ is the largest degree of any vertex in $V(G)$.

For an r-graph G and $U \subseteq V(G)$, $F \subseteq E(G)$ let $G[U] = (U, E(G) \cap \binom{U}{[r]})$, $G - U = G[V(G) \setminus U]$, and $G - F = (V(G), E(G) \setminus F)$. A *subgraph* G' of an r-graph G is an r-graph with $V(G') \subseteq V(G)$ and $E(G') \subseteq E(G)$. G is called a *supergraph* of G'. We write $G' \subseteq G$. A subgraph G' of G is *induced* if $G' = G[V(G')]$. A *copy* of some r-graph H in an r-graph G is a subgraph $G' \subseteq G$ which is isomorphic to H. We call an r-graph G *H-free*, if it contains no induced copy of H, i.e. no induced subgraph of G is a copy of H. For a family of graphs $\mathcal{H}$ we say G is *$\mathcal{H}$-free* if G is H-free for all $H \in \mathcal{H}$. The vertex-disjoint union $G_1 \cup G_2$ of two r-graphs G_1 and G_2 is the r-graph F with vertex set $V = V_1 \dot\cup V_2$, such that $F[V_i] = G_i$ for $i = 1, 2$. For an r-graph G and an integer n let nG be the vertex-disjoint union of n copies of G.

Let $K_n^{(r)}$ denote the *complete* r-graph or *clique* on n vertices, i.e. the graph on n vertices in which all $\binom{n}{r}$ r-sets are edges; for $r = 2$, we simply write K_n. Note that for $r < n$, $K_n^{(r)}$ is a set of n isolated vertices. The *clique number* $\omega(G)$ of an r-graph G is $\omega(G) = \max\{n : K_n^r \subseteq G\}$. An *independent set* or *co-clique* in an r-graph G is a set $I \subseteq V(G)$ such that every r-element subset of I is a non-edge. The *independence number* $\alpha(G)$ of G is the size of a largest independent set in G. A *homogeneous set* is a clique or a co-clique. The size of largest homogeneous set in an r-graph G is denoted by $h(G) = \max\{\alpha(G), \omega(G)\}$.

In 2-graphs a *path* of length n, $n \geq 2$ consists of n vertices $v_1, \dots, v_n$ and $n-1$ edges $v_i v_{i+1}$, $i \in [n-1]$. We write P_n for the path of length n. For vertices u, v in some graph, a u-v-path is a path in G starting at u and ending in v. A graph G is *connected* if for any two vertices $u, v \in V(G)$ there exists an u-v-path.

A *cycle* of length n, $n \geq 3$ consists of n vertices $v_1, \dots, v_n$ and n edges $v_i v_{i+1}$, $i \in [n]$, indices taken modulo n. We write C_n for the cycle of length n. The *girth* of a graph G, denoted by $\mathrm{girth}(G)$ is the length of a shortest cycle in G. If G contains no cycle, G is called *acyclic* or *forest*, and we write $\mathrm{girth}(G) = \infty$. A *tree* is a connected forest. The *odd girth* of G, denoted by $\mathrm{girth}_{\mathrm{odd}}(G)$, is the length of a shortest cycle of odd length in G.

If a graph contains no odd cycle, it is *bipartite*.

Given a positive integer n and some p with $p \in [0,1]$, $\mathcal{G}(n,p)$ denotes the probability space on all n-vertex graphs that result from independently deciding whether to include each of the $\binom{n}{2}$ possible edges with probability p. This model is called the *Erdős–Rényi model of random graphs*. We call $G \in \mathcal{G}(n,p)$ a *random graph*.

A *vertex colouring* of an r-graph is a map $c : V(G) \to S$, where S is a set of colours. c is a *proper colouring* if each edge contains at least two vertices of distinct colours. The *chromatic number* $\chi(G)$ of an r-graph G ist the smallest number of colours used among all proper colourings of G. An r-graph G is *k-partite* if there is a partition $V(G) = V_1 \dot\cup \cdots \dot\cup V_k$, sucht that $|e \cap V_i| \leq 1$ for each $e \in E(G)$ and $i \in [k]$. A 2-partite r-graph is also called *bipartite*.

Bipartite 2-graphs

Let G be a bipartite graph with parts U and V of size m and n respectively, we write $G = ((U \dot\cup V), E)$, $E \subseteq U \times V$. We call such a graph an $(m \times n)$ bipartite graph. We shall often depict the sets U and V as sets of points on two horizontal lines in the plane and call U the top part and V the bottom part. We say that a graph is the *bipartite complement* of G if it has the same vertex set as G and its edge set is $(U \times V) \setminus E$. We denote the bipartite complement of a graph G by G^c. By $\tilde{\omega}(G)$ we denote the largest integer t such that there are $A \subseteq U$, $B \subseteq V$ with $|A| = |B| = t$ and $ab \in E$ for all $a \in A$, $b \in B$, i.e. A and B form a *biclique*. By $\tilde{\alpha}(G)$ we denote the largest integer t such that there are $A \subseteq U$, $B \subseteq V$ with $|A| = |B| = t$ and $ab \notin E$ for all $a \in A$, $b \in B$, i.e. A and B form a *co-biclique* or a *bihole*. A *homogeneous set* in a bipartite graph is a biclique or a bihole. Let $\tilde{h}(G) = \max\{\tilde{\alpha}(G), \tilde{\omega}(G)\}$ denote the size of a largest homogeneous set in G.

For bipartite graphs $H = ((U,V),E)$ and $G = ((A,B),E')$, we say that H is an *induced bipartite subgraph of G respecting sides* if $U \subseteq A$, $V \subseteq B$, and for any $u \in U$, $v \in V$, we have $uv \in E(H)$ if and only if $uv \in E(G)$. We say that a bipartite graph $H = ((U,V),E)$ is a *copy* of a bipartite graph $H^* = ((U^*,V^*),E^*)$ if H^* is isomorphic to H with isomorphism $\varphi : U^* \cup V^* \to U \cup V$ such that $\varphi(U^*) = U$ and $\varphi(V^*) = V$.

Let $K_{m,n}$ denote the *complete bipartite graph* with parts of sizes m and n and all possible edges.

Extremal graph theory

The extremal number $\mathrm{ex}_r(n, \mathcal{G})$ of a family $\mathcal{G}$ of r-graphs is defined as the maximum number of edges any r-graph on n vertices can have without containing any $G \in \mathcal{G}$ as a subgraph. If $\mathcal{G} = \{G\}$, we write $\mathrm{ex}_r(n, G) = \mathrm{ex}_r(n, \{G\})$.

By $T_r(n, l)$ we denote the complete balanced l-partite r-graph on n vertices, i.e. the l-partite r-graph in which each part has size $\lfloor \frac{n}{l} \rfloor$ or $\lceil \frac{n}{l} \rceil$ and any r vertices from r distinct parts form an edge. $T_2(n, l)$ is also called the *Turán graph*.

For $r = 2$ Turán's Theorem [122], proved in 1941, tells us that $\mathrm{ex}_2(n, K_t) = |E(T_2(n, t-1))| = \left(\frac{t-2}{t-1} + o(1)\right)\binom{n}{2}$. Erdős and Stone proved the following asymptotic generalisation: $\mathrm{ex}_2(n, H) = \left(\frac{t-2}{t-1} + o(1)\right)\binom{n}{2}$ for any H with $\chi(H) = t > 2$.

However, for bipartite graphs (i.e. graphs with $\chi = 2$) the Erdős-Stone theorem does not provide a tight bound; it is known that $\mathrm{ex}(n, G) = o(n^2)$ for general bipartite graphs. The *Zarankiewicz function* $z(m, n; s, t)$ denotes the maximum possible number of edges in a subgraph of $K_{m,n}$ which does not contain a copy of $K_{s,t}$. We write $z(n; t) = z(n, n; t, t)$ for the symmetric problem. It was proven by Kővári, Sós and Turán [98] that $z(m, n; s, t) < (s-1)^{1/t}(n-t+1)m^{1-1/t} + (t-1)m$. This was improved in the diagonal case by Znám [128] to $z(n; t) < (t-1)^{1/t}n^{2-1/t} + \frac{1}{2}(t-1)n$.

Much less is known for r-graphs with $r \geq 3$. For an r-graph H the *Turán density* is defined as $\pi_r(G) = \lim n \to \infty \frac{\mathrm{ex}_r(n,G)}{\binom{n}{r}}$, and the currently best known general bounds on the Turán density are

$$1 - \left(\frac{r-1}{m-1}\right)^{r-1} \leq \pi(K_m^r) \leq 1 - \binom{m-1}{r-1}^{-1},$$

due to Sidorenko [118] and de Caen [51].

Ramsey Theory

An *s-edge-colouring* of an r-graph G is a map $c : E(G) \to [s]$. A *monochromatic* subgraph H is a subgraph $H \subseteq G$, for which there exists some colour $i \in [s]$ such that $c(e) = i$ for all $e \in E(H)$.

Given r-graphs $H_1, \ldots, H_s$, the *Ramsey number* of $(H_1, \ldots, H_s)$ is $R_r(H_1, \ldots H_s)$, the minimum integer n, such that for any s-edge-colouring of K_n there exists $i \in [s]$, such that there is a monochromatic copy of H_i in colour i. Ramsey's theorem [114] states that this is well-defined. Note that we can also replace a single graph by a family of graphs in the definition: The Ramsey number $R_r(\{H_1, \dots, H_s\}, \{K_1, \dots H_t\})$ would then be the smallest integer n such that in any 2-edge-colouring of K_n there is either a monochromatic copy of some H_i in colour 1 or a monochromatic copy of some K_i in colour 2. If all H_is are complete, we write $R_r(k_1, \dots, k_s)$ for $R_r(K_{k_1}, \dots, K_{k_s})$.

For $r = 2$ we usually omit the index R_2 and simply write $R(H_1, H_2)$. Table 0.1 lists some upper and lower bounds on Ramsey numbers for 2-graphs which will be used throughout the thesis.

Ramsey number	Bound	Reference
$R(k,k)$	$\geq 2^{k/2}$	Erdős [53]
	$\leq 4^k$	Erdős and Szekeres [65]
$R(3,k)$	$\Omega(k^2/\log k)$	Kim [93]
	$O(k^2/\log k)$	Ajtai, Komlós, Szemeredi [1]
$R(4,k)$	$\Omega(k^{5/2}/\log^2 k)$	Bohman [23]
$R(C_5, K_k)$	$O(k^{3/2}/\sqrt{\log k})$	Caro et al. [39]
$R(\{C_3, C_4, C_5\}, K_k)$	$\Omega((k/\log k)^{4/3})$	Spencer [120]
$R(3,3,3)$	$= 17$	Greenwood and Gleason [81]

Table 0.1: Bounds on Classical Ramsey numbers

Note that the upper bound on the diagonal Ramsey number $R(k,k) \leq (1 + o(1))\frac{4^k}{4\sqrt{\pi k}}$ due to Erdős and Szekeres from 1935 has been improved to $R(k,k) \leq k^{-(c\log k)/(\log\log k)}4^k$ by Conlon [46] in 2009. Very recently, in March 2023, Campos, Griffiths, Morris and Sahasrabude [33] improved the bound to $R(k,k) \leq (4 - 2^{-7})^k$. However, in this thesis, we will only use the old bound listed in the table.

For $r = 3$, the best known bounds in the diagonal case are due to Erdős, Hajnal and Rado [61]. They showed that there exist positive constants c, c' such that

$$2cn^2 < R_3(k,k) < 2^{2^{c'n}}.$$

Part I

The Erdős-Hajnal conjecture

Introduction and basic notions

A homogeneous set in an r-graph is a clique or an independent set. We write $h(G)$ for the size of largest homogeneous set in an r-graph G. In 1935 Erdős and Szekeres [65] proved that for any 2-graph G of size n we have $h(G) \geq \frac{1}{2}\log n$. On the other hand, a well-known theorem by Erdős [55] shows that for any n there exists a 2-graph G on n vertices with $h(G) \leq 2\log n$. Erdős and Hajnal [59] conjectured that this behaviour changes if one only considers H-free graphs G for any fixed graph H.

We say that an r-graph H has the *Erdős Hajnal-property* or simply EH-*property* if there is a constant $\epsilon = \epsilon_H > 0$ such that every n-vertex H-free r-graph G satisfies $h(G) \geq n^{\epsilon}$. Erdős and Hajnal [59] conjectured the following in 1989:

Conjecture 0.1 (Erdős, Hajnal [59])**.** *Any 2-graph has the Erdős-Hajnal property.*

In the same paper they investigated *perfect graphs*, i.e. graphs G for which neither G nor its complement $\overline{G}$ contains an induced odd cycle of length at least 5. They were able to prove that for a perfect graph G, we have $h(G) \geq \sqrt{|G|}$. Thus, one could also formulate Conjecture 0.1 by asking for a large induced perfect subgraph in any H-free graph.

The conjecture remains open, see for example a survey by Chudnovsky [43], as well as [6, 27, 72], to name a few. However, some partial results are known. When H is a fixed graph and G is an H-free n-vertex graph, Erdős and Hajnal [59] proved that $h(G) \geq 2^{c\sqrt{\log n}}$. This was recently improved to $h(G) \geq 2^{c\sqrt{\log n \log\log n}}$ by Bucić, Nguyen, Scott, and Seymour [31].

Only very few graphs are known to have the EH-property and the only operation known to preserve the EH-property, i.e. with which one can build larger graphs with the EH-property from smaller ones, is the so called *blow-up*: For a 2-graph H with vertex set $V(H) = \{v_1, \ldots, v_k\}$ and any other pairwise vertex-disjoint r-graphs $F_1, \ldots, F_k$, we define the *blow-up* $H(F_1, \ldots, F_k)$ as the 2-graph obtained by taking pairwise vertex-disjoint copies of $F_1, \ldots, F_k$ with an edge between vertices from F_i and F_j if and only if $v_i v_j \in E(H)$.

Lemma 0.2 (Alon, Pach, Solymosi [6])**.** *If $H, F_1, \ldots, F_k$ have the Erdős–Hajnal property, then so does $H(F_1, \ldots, F_k)$.*

All graphs on up to four vertices are known to have the EH-property and the constants in the quantitative version are known. With Lemma 0.2 one can show that all but four graphs on 5 vertices also have the EH-property, the remaining ones being the bull, C_5, P_5 and $\overline{P_5}$. For the bull the conjecture was proven to hold by Chudnovsky and Safra [44], and for C_5 it was proven by Chudnovsky, Scott, Seymour and Spirkl [45], so the only graph on up to 5 vertices for which the conjecture remains open, is P_5.

One can also consider a bipartite version of the Erdős-Hajnal conjecture. Here one asks for a largest homogeneous set, i.e. a biclique or a bihole in a bipartite graph not containing some forbidden induced bipartite subgraph. For a bipartite graph G, the size of a largest homogeneous set is denoted by $\tilde{h}(G) = \max\{\tilde{\alpha}(G), \tilde{\omega}(G)\}$. It is implicit from a result of Erdős, Hajnal and Pach [60] that for any bipartite H with the smaller part of size k, any H-free bipartite graph G satisfies $\tilde{h}(G) = \Omega(|G|^{1/k})$. It was shown in [15] that for bipartite graphs H containing a cycle, any H-free bipartite graph G satisfies $\tilde{h}(G) \in o(n)$, and for all but at most four bipartite graphs H with acyclic H and bipartite complement H^C, any H-free bipartite graph satisfies $\tilde{h}(G) = cn$ for some constant $c(H)$. In Chapter 1 we will look at these open cases and in Chapter 2 we determine the size of a largest bihole in a bipartite graph with bounded maximum degree, which will lead to improved EH-coefficients for bipartite graphs H with only one vertex in one part.

Erdős and Hajnal [59] further stated a multicoloured version of Conjecture 0.1 asserting that for any fixed integer $k \geq 3$ and for any fixed s'-edge-coloured clique K on k vertices, for $s \geq s' \geq 2$, there is a positive constant $a = a(K)$ such that any s-edge-colouring of a complete graph on n vertices with no copy of K contains a clique on $\Omega(n^a)$ vertices using at most $s - 1$ colours.

In Chapter 3 we will make this precise and reduce the multicolour EH-conjecture to the case where the number of colours s is equal to or one more than the number of colours used on the forbidden colouring. We further state a size variant of the multicolour EH-conjecture and show that it holds for $s \in \{2,3\}$ colours. We also generalise Lemma 0.2 to an arbitrary number of colours.

In Chapter 4 we will investigate the multicolour EH-conjecture for 3 colours and forbidden patterns on triangles, i.e. we are looking for large 2-coloured cliques in 3-coloured complete graphs which do not contain some given forbidden colourings on triangles. In particular, there might be more than one forbidden triangle-colouring. We show that the Erdős-Hajnal conjecture holds true in this setting. We focus on the quantitative version of the conjecture and provide asymptotic bounds on the sizes of the largest 2-edge-coloured cliques for all families of forbidden patterns containing at most 3 triangle-colourings.

The Erdős-Hajnal conjecture fails for r-graphs, $r \geq 3$, already when H is a complete graph of size $r+1$. Indeed, well-known results on off-diagonal hypergraph Ramsey numbers show that there are n-vertex r-graphs that do not have a clique on $r+1$ vertices and do not have co-cliques on $f_r(n)$ vertices, where f_r is an iterated logarithmic function (see [110] for the best known results).

Moreover, the following result (Claim 1.3. in [80]) tells us exactly which r-graphs, $r \geq 3$, have the EH-property. Here D_2 is the unique 3-graph on 4 vertices with exactly 2 edges.

Theorem 0.3 (Gishboliner, Tomon [80]). *Let $r \geq 3$. If F is an r-graph on at least $r+1$ vertices and $F \neq D_2$, then there is an F-free r-graph H on n vertices such that $h(H) = (\log n)^{O(1)}$.*

It is natural to consider the EH-property for families of r-graphs instead of a single r-graph. We call an r- graph F on m vertices and f edges an *(m,f)-graph*, we call the pair (m,f) the *order-size pair* for F and we say that an r-graph H is (m,f)-free if it contains no induced copy of an (m,f)-graph. Similarly, we say a pair (m,f) (or a family of pairs $Q = \{(m_1,f_1),\ldots,(m_k,f_k)\}$) has the EH-property if any (m,f)-free (or Q-free) r-graph H satisfies $h(H) \geq |H|^\epsilon$ for some ϵ only depending on (m,f). In Chapter 5 we show that for $r=2$ any order-size pair has the EH-property. We then fix $r=3$ and $m=4$ and consider all possible families Q of order-size pairs with these parameters. For each such Q we give bounds on $h(H)$ for any Q-free H and determine which families Q have the EH-property.

Chapter 1 The bipartite variant of the Erdős-Hajnal conjecture – quantitative version

1.1 Introduction

Let $\mathrm{Forb}(n, H)$ denote the set of all bipartite graphs with parts of size n which do not contain a copy of H as an induced bipartite subgraph respecting sides. Recall that we call a bipartite graph H-free if it does not contain an induced copy of H. Recall that $\tilde{h}(G) = \min\{\tilde{\alpha}(G), \tilde{\omega}(G)\}$ denotes the size of a largest balanced homogeneous set, i.e. a biclique or a bihole/co-biclique in G. Let

$$\tilde{h}(n, H) = \tilde{h}(\mathrm{Forb}(n, H)) = \min\{\tilde{h}(G) : \; G \in \mathrm{Forb}(n, H)\}.$$

Lemma 1.1. *If H is an induced bipartite subgraph of K (respecting sides), we have $\tilde{h}(n, H) \geq \tilde{h}(n, K)$.*

Proof. Any H-free bipartite graph is also K-free by assumption. Thus, we have $\mathrm{Forb}(n, H) \subseteq \mathrm{Forb}(n, K)$, and thus, $\tilde{h}(n, H) = \min\{\tilde{h}(G) : \; G \in \mathrm{Forb}(n, H)\} \geq \min\{\tilde{h}(G) : \; G \in \mathrm{Forb}(n, K)\} = \tilde{h}(n, K)$. □

It is implicit from a result of Erdős, Hajnal and Pach [60] that for any bipartite H with the smaller part of size k, we have $\tilde{h}(n, H) = \Omega(n^{1/k})$. If either H or its bipartite complement H^c contain a cycle, then it must contain either C_4, C_6 or C_8. A standard probabilistic argument, see for example [15], shows that in this case we have $\tilde{h}(n, H) = O(n^{1-\epsilon})$ for some positive ϵ.

Axenovich, Tompkins and the author [15] addressed the question of when $\tilde{h}(n, H)$ is linear in n. We say that a bipartite graph H is *strongly acyclic* if neither H nor its bipartite complement H^c contain a cycle. They showed that for all but at most four strongly acyclic graphs H, $\tilde{h}(n, H)$ is linear in n. Let

$$\mathcal{H} = \{\tilde{P}_5, P_6, S_{1,2,3}, P_7\},$$

the set of these four strongly acyclic graphs, given in Figure 1.1.

Theorem 1.2 (Axenovich, Tompkins, Weber [15])**.** *Let H be a strongly acyclic bipartite graph. If neither H nor H' is in $\mathcal{H}$, there is a positive constant $c = c(H)$ such that $\tilde{h}(n, H) \geq cn$.*

Moreover, for several graphs H the value of $\tilde{h}(n, H)$ was determined exactly; the

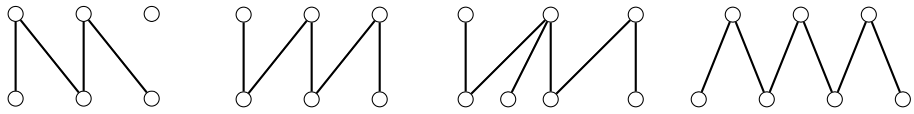


Figure 1.1: The set $\mathcal{H} = \{\tilde{P}_5, P_6, S_{1,2,3}, P_7\}$.

results are summarised in Table 1.1. Here $M_{k,k}$ is the bipartite graph with vertex set $\{v_1, v_2\}\dot\cup\{s_1, \ldots, s_{2k+2}\}$, such that v_1 is incident to $\{s_1, \ldots, s_k, s_{2k+1}\}$, v_2 is incident to $\{s_{k+1}, \ldots, s_{2k+1}\}$, and $H_{k,k} = 2K_{1,k}$. Then any strongly acyclic bipartite graph with at most two vertices in its smaller part is an induced subgraph of $M_{k,k}$ for some k; one can easily show that any strongly acyclic bipartite graph H with $H \notin \mathcal{H}$ has at most 2 vertices in its smaller part. $\tilde{P}_3$ denotes the (2×2) bipartite graph which is a P_3 and an isolated vertex. Using Lemma 1.1 one obtains the given bounds.

forbidden H	$\tilde{h}(n, H)$
not strongly acyclic	$o(n)$
$H \subseteq M_{k,k}$	$\geq \frac{n}{30\lvert V(H)\rvert}$
$H \in \tilde{\mathcal{H}}$	?
$H \subseteq H_{k,k}$	$\geq \frac{n}{2k}$
P_4	$= \lceil \frac{n}{3} \rceil$
$2K_2$	$= \lceil \frac{n}{2} \rceil$
$\tilde{P}_3$	$\geq \lceil \frac{2n}{5} \rceil$

Table 1.1: Bipartite EH-coefficients from [15]

In the remaining part of this chapter we will deal with the four open cases. In Chapter 2, we will look at large biholes in bipartite graphs with bounded maximum degree in one part, which will yield the upper bound on $\tilde{h}(n, H)$ for H with one vertex in the smaller part. The results are summarised in Table 1.2.

forbidden H	$\tilde{h}(n, H)$	Source
$H \subseteq K_{1,k}$	$\geq \frac{1}{2}\frac{\log \Delta(H)}{\Delta(H)} n$	Chapter 2, [13]
P_7	$\geq \frac{n}{10^{13}}$	Proposition 1.5
P_6	$\geq \frac{n}{10^{13}}$	Corollary 1.9
$S_{1,2,3}$	$\geq \frac{n}{6}$	[2], Proposition 1.10
$\tilde{P}_5$	$\geq \frac{n}{6}$	Corollary 1.13

Table 1.2: Improved bipartite EH-coefficients

In particular, the results from Table 1.2 together with Theorem 1.2 imply the following:

Theorem 1.3. *Let H be a bipartite graph. Then H is strongly acyclic if and only if there exists some $c = c(H) > 0$, s.t. $\tilde{h}(n, H) \geq cn$.*

Scott, Seymour and Spirkl [117] proved the following theorem:

Theorem 1.4 (Scott, Seymour, Spirkl [117])**.** *For every bipartite graph H which is a forest and every τ with $0 < \tau \leq 1$ there exists $\epsilon = \epsilon(H) > 0$, such that any H-free $(n \times n)$ bipartite graph G with at most $(1-\tau)n^2$ edges satisfies $\tilde{\alpha}(G) \geq \epsilon n$.*

Note that one can obtain Theorem 1.3 from Theorem 1.4 in the following way: Let H be a strongly acyclic bipartite graph and let G be any $(n \times n)$ bipartite H-free graph. If $|E(G)| \leq \frac{1}{2}n^2$, by Theorem 1.4 we have $\tilde{\alpha}(G) \geq \epsilon(H)n$. Otherwise, G^c is H^c-free and has at most $\frac{1}{2}n^2$ edges, so by Theorem 1.4 we have that $\tilde{\omega}(G) = \tilde{\alpha}(G^c) \geq \epsilon(H^c)n$.

However, the proof of Theorem 1.4 does not provide any specific constants $\epsilon(H)$. It uses, amongst others, the existence of hypergraph Ramsey numbers, for which we have the lower bound $R_k(n, n) \geq t_{k-1}(cn^2)$, where $t_k(x)$ is the tower function defined by $t_1(x) = x$ and $t_{i+1}(x) = 2^{t_i(x)}$. At one step in the proof one iterates over all forests on a fixed number of vertices, where k is the number of vertices in the forest, and in each iteration, n is the Ramsey number from the previous step. Even when only considering strongly acyclic bipartite graphs in $\tilde{\mathcal{H}}$ on 6 or 7 vertices, following the proof of Theorem 1.4 will only give $\epsilon^{-1} > R_k(n, n)$ for some huge n, which is considerably weaker than the bounds obtained in the next section.

Note that the notion of large bicliques and co-bicliques in ordered bipartite graphs with forbidden induced subgraphs corresponds to the notion of submatrices of all 0's or of all 1's in binary matrices with forbidden submatrices. Here *ordered* means that the vertices in the two parts are ordered according to the rows and columns of the matrices, thus, one forbidden submatrix only forbids one specific ordering of the corresponding bipartite graph. A paper by Korándi, Pach and Tomon [96] addresses a similar question for matrices. In addition, one could interpret bipartite graphs as set systems consisting of all the neighbourhoods of vertices from one part. Structural properties of these graphs in terms of VC-dimension of the respective set system in connection to the Erdős-Hajnal conjecture are addressed for example by Fox, Pach and Suk [72].

1.2 Specific bounds in the linear regime

The goal of this section is to deal with the four remaining cases $\tilde{P}_5, S_{1,2,3}, P_6$ and P_7. In particular, we will show that in each of those cases we obtain $\tilde{h}(n, H) \geq c(H)n$, for some positive constants $c(H)$, which will then prove Theorem 1.3. Note that $\tilde{P}_5$ and P_6 are induced subgraphs of $S_{1,2,3}$ and P_7 respectively, so by Lemma 1.1 it suffices to find positive constants $c(S_{1,2,3})$ and $c(P_7)$, even though these might not be best possible for $c(\tilde{P}_5)$ and $c(P_6)$.

1.2.1 Paths on 6 and 7 vertices

Proposition 1.5. *We have* $\tilde{h}(n, P_7) \geq \frac{n}{544^4 \cdot 3 \cdot 2^5}$.

The proof is partly inspired by a result by Bousquet, Lagoutte and Thomassé [27] on the Erdős-Hajnal conjecture for paths and antipaths. Before we prove the statement, we need some auxiliary lemmata:

Lemma 1.6. *Let* $G = (A\dot{\cup}B, E)$ *be an* $(n_1 \times n_2)$ *bipartite graph with* $n_1, n_2 \geq 2$. *Then* G *contains either a co-biclique or a connected component with parts of sizes at least* $\frac{n_1}{3}$ *and* $\frac{n_2}{3}$ *in* A *and* B *respectively.*

Proof. Let, for some index set I, the connected components of G have parts A_i and B_i of sizes a_i, b_i, respectively, $A_i \subseteq A$, $B_i \subseteq B$, $i \in I$. Let $G_i = G[A_i\dot{\cup}B_i]$ be the i^{th} connected component of G.

We can assume that for each i we either have $a_i < \frac{n_1}{3}$ or $b_i < \frac{n_2}{3}$, since otherwise there is a connected component with parts of sizes $\frac{n_1}{3}$, $\frac{n_2}{3}$.

Then we have $a_i < \frac{n_1}{3}$ and $b_i < \frac{n_2}{3}$ for each $i \in I$: Assume there is $i \in I$ s.t. $a_i \geq \frac{n_1}{3}$ or $b_i \geq \frac{n_2}{3}$, say $a_i \geq \frac{n_1}{3}$. Then by our assumption above we must have $b_i < \frac{n_2}{3}$. Hence, we find a co-biclique with parts A_i, $B \setminus B_i$ of sizes at least $\frac{n_1}{3}$ and $\frac{2n_2}{3}$ respectively.

Let I_1 be the set of indices for which we have $\frac{a_i}{n_1} \leq \frac{b_i}{n_2}$. Let $I_2 = I \setminus I_1$. Let $X_1 = \bigcup_{i \in I_1} A_i$, $Y_1 = \bigcup_{i \in I_1} B_i$, $X_2 = A \setminus A_1$, $Y_2 = B \setminus B_2$, $x_1 = |X_1|$, $x_2 = |X_2|$, $y_1 = |Y_1|$, $y_2 = |Y_2|$. Consider the co-biclique with parts Y_1, X_2. We can assume that either $y_1 < \frac{n_2}{3}$ or $x_2 < \frac{n_1}{3}$. If $x_1 < \frac{n_1}{3}$, then clearly $x_2 > \frac{2n_1}{3}$, so assume $y_1 < \frac{n_2}{3}$. Then $x_1 < \frac{n_1}{3}$ since for each $i \in I_1$, $\frac{a_i}{n_1} \leq \frac{b_i}{n_2}$. Thus, in either case we have $x_2 > \frac{2n_1}{3}$.

Consider a minimal subset $I_3 \subseteq I_2$ such that $X_3 = \bigcup_{i \in I_3} A_i \subseteq X_2$ has size $x_3 > \frac{n_1}{3}$.

Then $x_3 < \frac{2n_1}{3}$, otherwise for any $i \in I_3$, $|X_3 \setminus A_i| > \frac{2n_1}{3} - \frac{n_1}{3} = \frac{n_1}{3}$. In particular, we could have taken $I_3 \setminus \{i\}$ instead of I_3, contradicting its minimality. Thus, $Y_3 = \bigcup_{i \in I_3} B_i$ has size less than $\frac{2n_2}{3}$, since $\frac{b_i}{n_2} < \frac{a_i}{n_1}$ for $i \in I_3 \subseteq I_2$. This implies that X_3 and $B \setminus Y_3$ form a co-biclique with parts of sizes at least $\frac{n_1}{3}$ and $\frac{n_2}{3}$.

Note that $\frac{1}{3}$ is best possible: For $n_1, n_2 \geq 3$ take the pairwise disjoint union of 3 bicliques with part sizes in $\left\{ \left\lfloor \frac{n_1}{3} \right\rfloor, \left\lceil \frac{n_1}{3} \right\rceil, \left\lfloor \frac{n_2}{3} \right\rfloor, \left\lceil \frac{n_2}{3} \right\rceil \right\}$. $\square$

Lemma 1.7. *For every $k \geq 2$ there exists $\epsilon_k > 0$ and c_k with $0 < c_k \leq 1$, such that every connected $(n_1 \times n_2)$ bipartite graph $G = (U \dot\cup V, E)$ with $n_1, n_2 \geq 2$ satisfies one of the following:*

- *There exists a vertex $v \in A$ of degree more than $\epsilon_k n_2$ or a vertex $v \in B$ of degree more than $\epsilon_k n_1$; or*
- *for every vertex v, G contains an induced P_k starting at v; or*
- *G contains a co-biclique with part sizes $c_k n_1$ and $c_k n_2$.*

In particular, we can set $\epsilon_2 = c_2 = 1$ and $\epsilon_k = \frac{\epsilon_{k-1}}{3+\epsilon_{k-1}}$ and $c_k = c_{k-1}\frac{(1-\epsilon_k)}{3}$ for $k \geq 3$.

Proof by induction on k. For $k = 2$, the second item trivially holds, since G is connected and thus, every vertex is the endpoint of an edge. Set $\epsilon_2 = c_2 = 1$.

If $k > 2$, let $\epsilon_k = \frac{\epsilon_{k-1}}{3+\epsilon_{k-1}}$ and $c_k = c_{k-1}\frac{(1-\epsilon_k)}{3}$. Then we have $\epsilon_k = \epsilon_{k-1}\frac{(1-\epsilon_k)}{3}$, and so we have $c_k = \epsilon_k$.

Assume the first item does not hold, i.e. the maximum degree in A is at most $\epsilon_k n_2$ and the maximum degree in B is at most $\epsilon_k n_1$. We will show that the 2nd or 3rd item must hold then.

Let v be any vertex in $V(G)$, w.l.o.g. $v \in A$ (for $v \in B$ simply swap the roles of A and B in the proof) and set $A' = A \setminus \{v\}$ and $B' = B \setminus N(v)$. Then we have that $a = |A'| = n_1 - 1$ and $b = |B'| \geq (1 - \epsilon_k)n_2$. Then by Lemma 1.6, $G[A' \dot\cup B']$ contains either a co-biclique or a connected component S with parts of sizes $n_1' \geq a/3 \geq \frac{n_1 - 1}{3}$ and $n_2' \geq b/3 = \frac{(1-\epsilon_k)n_2}{3}$. If S is a co-biclique, the third item holds and we are done, so S is a connected component.

Let $w \in N(v)$ be adjacent to S (which exists, by connectivity of G). Consider the graph $G' = G[S \dot\cup \{w\}]$. Then the maximum degree in $A \cap V(G')$ is still at most $\epsilon_k n_2 = \epsilon_{k-1}(\frac{(1-\epsilon_k)n_2}{3}) \leq \epsilon_{k-1} n_2' \leq \epsilon_{k-1}(n_2' + 1)$ and the maximum degree in $B' \cap V(G')$

is still at most $\epsilon_k n_1 = \epsilon_{k-1}(\frac{(1-\epsilon_k)n_1}{3}) \leq \epsilon_{k-1} n_1'$. Thus, by the induction hypothesis, either the second or third item holds in G', which either gives a P_k in G starting at v (vw and a P_{k-1} starting at w) or to a co-biclique with part sizes $c_{k-1}\frac{a}{3} \geq c_{k-1}\frac{n_1-1}{3} \geq c_{k-1}(1-\epsilon_k)\frac{n_1}{3} = c_k n_1$ and $c_{k-1}(\frac{b}{3}+1) \geq c_{k-1}(1-\epsilon_k)\frac{n_2}{3} = c_k n_2$. □

This proves the existence of either a large co-biclique or a vertex of high degree in P_7-free bipartite graphs.

The following lemma is due to Erdős, Hajnal and Pach [60], rephrased and proven in a slightly more general form by Scott, Seymour and Spirkl [117]. It provides the existence of either a sparse or a dense subgraph of any H-free bipartite graph.

Lemma 1.8 (Scott, Seymour, Spirkl [117]). *Let H be a $(k \times l)$ bipartite graph and let $\epsilon > 0$. Then there exists $\gamma > 0$ with the following property: Let G be an H-free $(n \times n)$ bipartite graph with $n > 0$ and parts A, B; then there exist an $(\gamma n \times \gamma n)$ bipartite subgraph G' of G, such that either*

$$\Delta(G') < \epsilon\gamma n \qquad \text{or} \qquad \delta(G') > (1-\epsilon)\gamma n.$$

In particular, $\gamma = \min\{\frac{1}{2}, (k+l)^{-1}, (\epsilon/2)^k/l\}$ is sufficient.

Now we can finally prove Proposition 1.5.

Proof of Proposition 1.5. Let $\epsilon = \epsilon_7$ from Lemma 1.7. By the recursion given ($\epsilon_2 = 1$, $\epsilon_k = \frac{\epsilon_{k-1}}{3+\epsilon_{k-1}}$ for $k \geq 3$) we have $\epsilon_2 = 1$, $\epsilon_3 = \frac{1}{4}$, $\epsilon_4 = \frac{1}{15}$, $\epsilon_5 = \frac{1}{48}$, $\epsilon_6 = \frac{1}{147}$ and $\epsilon_7 = \frac{1}{544} = c_7$. Let $\gamma = \min\{\frac{1}{2}, \frac{1}{7}, (\epsilon/2)^3/4\} = \frac{(\epsilon/2)^3}{4}$.

Let G be an $(n \times n)$ bipartite graph which is P_7-free. We want to show that $\tilde{h}(G) \geq \frac{\epsilon_7^4}{3 \cdot 2^5} n$. By Lemma 1.8, G contains a $(\gamma n \times \gamma n)$ bipartite subgraph G' with either $\Delta(G) < \epsilon\gamma n$ or $\delta(G) > (1-\epsilon)\gamma n$.

Assume we have $\Delta(G') < \epsilon\gamma n$ (in the other case, consider the bipartite complement G'^c. Then since $P_7^c = P_7$, G'^c is P_7-free and has $\Delta(G'^c) = \gamma n - \delta(G'^c) < \epsilon\gamma n$). By Lemma 1.6, G' contains either a connected component S or a co-biclique with parts of size $\gamma n/3$. In the latter case we are done, so assume the former. Let $G'' = G'[S]$.

Then according to Lemma 1.7, G'' contains a co-biclique with parts of sizes at least $c_7\gamma\frac{1}{3}n$. Thus, in either case we find a large homogeneous set, and in particular, we have

$$\tilde{h}(G) \geq \min\left\{\frac{1}{3}\gamma n, \frac{1}{3}c_7\gamma n\right\} = \frac{1}{3}\epsilon_7\frac{(\epsilon_7/2)^3}{4} = \frac{\epsilon_7^4}{3 \cdot 2^5}n = \frac{n}{544^4 \cdot 3 \cdot 2^5} < \frac{n}{10^{13}}. \qquad \square$$

Corollary 1.9. *We have* $\tilde{h}(n, P_6) \geq \frac{n}{544^4 \cdot 3 \cdot 2^5}$.

Proof. Since P_6 is an induced subgraph of P_7, by Lemma 1.1 we have $\tilde{h}(n, P_7) \leq \tilde{h}(n, P_6)$. □

1.2.2 REMAINING BIPARTITE GRAPHS

Let $H = S_{1,2,3}$ denote the following bipartite graph on 7 vertices:

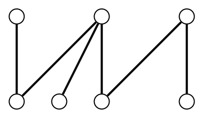


Figure 1.2: The graph $S_{1,2,3}$

Proposition 1.10 (Alecu, Atminas, Lozin, Zamaraev [2])**.** *We have* $\tilde{h}(n, S_{1,2,3}) \geq \frac{n}{6}$.

A partial proof of this proposition appears, with some gaps, in [2]. We give a complete proof here.

We use a decomposition scheme, using the language from [2, 69], called *canonical decomposition*:

We define a *good split* as the decomposition of a bipartite graph $G = (U \dot\cup V, E)$ into two non-empty bipartite graphs $G_1 = G[U_1 \dot\cup V_1]$ and $G_2 = G[U_2 \dot\cup V_2]$, where $U = U_1 \dot\cup U_2$ and $V = V_1 \dot\cup V_2$, such that each of the graphs $G[U_1 \dot\cup V_2]$ and $G[U_2 \dot\cup V_1]$ is either complete or empty.

If we want to canonically decompose a bipartite graph G, we find a good split into G_1 and G_2, and then recursively find good splits in G_1 and G_2. If we cannot find a good split in a bipartite graph, it is called *canonically indecomposable*. If we can recursively decompose a bipartite graph by good splits until all components consist of a single vertex, the graph is called *totally decomposable*.

We need the following two characterisations:

Lemma 1.11 (Fouquet et al. [69])**.** *A bipartite graph H is totally decomposable by canonical decomposition if and only if it is*$\{P_7, S_{1,2,3}\}$*-free.*

Lemma 1.12 (Alecu et al. [2])**.** *Any canonically indecomposable $S_{1,2,3}$-free $(n \times n)$ bipartite graph G containing a P_7 satisfies* $\tilde{h}(G) \geq \frac{n}{4}$.

In the proof of Lemma 1.12 a lemma from Lozin [104] is used, characterising $S_{1,2,3}$-free graphs containing a P_7. Now we can prove Proposition 1.10.

Proof of Proposition 1.10: Let G be an $S_{1,2,3}$-free $(n \times n)$ bipartite graph. If G is canonically indecomposable, by Lemma 1.11, G contains a P_7, so by Lemma 1.12, we have $\tilde{h}(G) \geq \frac{n}{4}$.

So assume G is not indecomposable, so we can find a good split and decompose $G = (U \dot{\cup} V, E)$ into two graphs $G_1 = G[U_1 \dot{\cup} V_1]$ and $G'_1 = G[U'_1 \dot{\cup} V'_1]$. W.l.o.g. $|U_1| \geq |U'_1|$. We will recursively decompose G_k into two graphs $G_{k+1} = G_k[U_{k+1} \dot{\cup} V_{k+1}]$ and $G'_{k+1} = G_k[U'_{k+1} \dot{\cup} V'_{k+1}]$ with $|U_{k+1}| \geq |U'_{k+1}|$ until G_k does not contain a good split, i.e. until we are left with a canonically indecomposable bipartite graph G_k. Note that we only further decompose the graphs G_k, never G'_k. In particular, for any $j \in [k]$ we have $U = U'_1 \dot{\cup} U'_2 \dot{\cup} \cdots \dot{\cup} U'_k \dot{\cup} U_k$.

Case 1: $|U_k| \geq \frac{2n}{3}$. If $|V_k| \geq \frac{2n}{3}$, then by Lemma 1.12 $\tilde{h}(G'_k) \geq \frac{1}{4}\frac{2n}{3} = \frac{n}{6}$, so assume $|V_k| < \frac{2n}{3}$. Then we have $|V'_1 \dot{\cup} V'_2 \dot{\cup} \cdots \dot{\cup} V'_k| > \frac{n}{3}$, and $G[U_k \dot{\cup} V'_i]$ is either complete or empty for each $i \leq k$. Thus, taking U_k and some of the sets V'_i, we have $\tilde{h}(G) \geq \frac{1}{2}\frac{n}{3} = \frac{n}{6}$.

Case 2: $|U_k| < \frac{2n}{3}$. Since $U = U'_1 \dot{\cup} U'_2 \dot{\cup} \cdots \dot{\cup} U'_k \dot{\cup} U_k$, we have $|U'_1 \dot{\cup} \cdots \dot{\cup} U'_k| > \frac{n}{3}$. Pick the smallest j s.t. $|U'_1| + |U'_2| + \cdots + |U'_j| > \frac{n}{3}$. Since $|U'_1 \dot{\cup} U'_2 \dot{\cup} \cdots \dot{\cup} U'_j \dot{\cup} U_j| = n$ and by definition of j, we obtain $|U'_j| + |U_j| \geq \frac{2n}{3}$ and thus, since $|U_j| \geq |U'_j|$, we have $|U_j| \geq \frac{n}{3}$.

Case 2.1: $|V_j| \geq \frac{n}{6}$. Then for $i \leq j$, each $G[U'_i \dot{\cup} V_j]$ is either complete or empty. Since $\sum_{i=1}^{j} |U'_i| > \frac{n}{3}$, we find a collection of sets U'_i that span $\frac{1}{2}\frac{n}{3}$ vertices that form a biclique or a co-biclique with V'_j, i.e. $\tilde{h}(G) \geq \frac{n}{6}$.

Case 2.2: $|V_j| < \frac{n}{6}$. Then for $i \leq j$, each $G[U_j, V'_i]$ is either complete or empty. Since $|U_j| \geq \frac{n}{3}$, and $\left|\dot{\bigcup}_{i=1}^{j} V'_i\right| \geq \frac{5n}{6}$, we have $\tilde{h}(G) \geq \tilde{h}\left(G\left[U'_j \dot{\cup} \left(\dot{\bigcup}_{i \leq j} V_i\right)\right]\right) \geq \frac{n}{6}$.

Thus, in any case, we have $\tilde{h}(G) \geq \frac{n}{6}$. □

Corollary 1.13. *We have $\tilde{h}(n, \tilde{P}_5) \geq \frac{n}{6}$.*

Proof. Since $\tilde{P}_5$ is an induced subgraph of $S_{1,2,3}$, by Lemma 1.1 and Proposition 1.10 we obtain $\frac{n}{6} \leq \tilde{h}(n, S_{1,2,3}) \leq \tilde{h}(n, \tilde{P}_5)$. □

1.3 Concluding remarks

While Theorem 1.3 gives a full characterisation for which forbidden bipartite graphs H the function $\tilde{h}(n, H)$ is linear in n, we only know the asymptotic behaviour. The constants $c(H)$ obtained in Propositions 1.5 and 1.10 are probably far from optimal; an easy to see upper bound on $c(H)$ for any graph in $\tilde{\mathcal{H}}$ is $\frac{1}{3}$, which can be obtained by taking three pairwise disjoint copies of $K_{n/3,n/3}$.

For $H \in \{S_{1,2,3}, \tilde{P}_5\}$ we have $c(H) \leq \frac{n}{4}$, as can be seen by considering the *blow-up* of C_8, i.e. the $(4m \times 4m)$ bipartite graph G with parts $U = U_1 \dot\cup U_2 \dot\cup U_3 \dot\cup U_4$ and $V = V_1 \dot\cup V_2 \dot\cup V_3 \dot\cup V_4$, where $|U_i| = |V_i| = m$ for $i \in [4]$, and uv with $u \in U_i$, $v \in V_j$ is an edge if and only if $j \in \{i, i+1\} \pmod 4$. This graph is H-free and satisfies $\tilde{\alpha}(G) = \tilde{\omega}(G) = \frac{n}{4}$. Thus, $\frac{n}{6} \leq c(H) \leq \frac{n}{4}$.

Also, there is no other non-trivial upper bound on any of those constants $c(H)$, so it would be interesting to exactly determine those constants.

CHAPTER 2 BIPARTITE INDEPENDENCE NUMBER IN GRAPHS WITH BOUNDED MAXIMUM DEGREE

2.1 INTRODUCTION

The problem of finding $g(n, \Delta)$, the smallest possible size of a largest independent set in an n-vertex graph with given maximum degree Δ is not very difficult. Indeed, one can consider the graph that is the disjoint union of $\lfloor n/(\Delta+1) \rfloor$ complete graphs on $\Delta + 1$ vertices each and a complete graph on the remaining vertices. This shows that $g(n, \Delta) \leq \lceil n/(\Delta+1) \rceil$. On the other hand, every n-vertex graph of maximum degree Δ contains an independent set of size $\lceil n/(\Delta+1) \rceil$, obtained for example by the greedy algorithm. Consequently, $g(n, \Delta) = \lceil n/(\Delta+1) \rceil$. The situation is more interesting for regular graphs, see Rosenfeld [116] for a more detailed analysis. The analogous problem in the bipartite setting is more complex: determining the smallest possible *bipartite independence number* of a bipartite graph with maximum degree Δ is still unresolved, even for $\Delta = 3$.

Recall that a *bihole* of size k in a bipartite graph $G = (A \dot\cup B, E)$ with a given bipartition A, B, is a pair (X, Y) with $X \subseteq A$, $Y \subseteq B$ such that $|X| = |Y| = k$, and such that there are no edges of G with one endpoint in X and the other endpoint in Y. Thus, the size of the largest bihole can be viewed as a bipartite version of the usual independence number. This chapter is devoted to studying the behaviour of this function. We write $\log$ for the natural logarithm.

Definition 2.1. *Let $f(n, \Delta)$ be the largest integer k such that any $(n \times n)$ bipartite graph $G = (A \dot\cup B, E)$ with $d(a) \leq \Delta$ for all $a \in A$ contains a bihole of size k. Let $f^*(n, \Delta)$ be the largest integer k such that any $(n \times n)$ bipartite graph G with $\Delta(G) \leq \Delta$ contains a bihole of size k.*

While $f(n, \Delta)$ is defined by restricting the maximum degree in one part of the graph, $f^*(n, \Delta)$ is its 'symmetric' version. Observe that $f(n, \Delta) \leq f^*(n, \Delta)$ for any natural numbers n and Δ for which these functions are defined.

Theorem 2.2. *There exists an integer Δ_0 such that if $\Delta \geq \Delta_0$ and $n > 5\Delta \log \Delta$, then*

$$f^*(n, \Delta) \geq f(n, \Delta) \geq \frac{1}{2} \cdot \frac{\log \Delta}{\Delta} n.$$

The proof of Theorem 2.2 follows by first determining the value of $f(n, 2)$ (see Theorem 2.6) and then reducing the general problem of bounding $f(n, \Delta)$ by applying the bound on $f(n,2)$ in conjunction with a probabilistic argument. A slightly weaker bound was obtained by Feige and Kogan, [66]. They proved that $f(n,\Delta) \geq cn\log\Delta/\Delta$, for any constant $c < 1/2$. The proof is a very similar probabilistic approach to the one used here, but instead of looking for a subset of vertices inducing a graph with maximum degree at most 2 in A, the authors immediately look for a bihole. Thus, our Theorem 2.2 is a modest improvement of this result giving a constant term equal to $1/2$; this has been improved recently by Chakraborti [41] to $(1-\varepsilon)$, see Theorem 2.17 in Section 2.4.

Theorem 2.3. *For any $\varepsilon \in (0,1)$ there is an integer $\Delta_0 = \Delta(\epsilon)$ such that if $\Delta \geq \Delta_0$ and $n \geq \Delta$, then*

$$f(n,\Delta) \leq (2+\varepsilon)\cdot\frac{\log\Delta}{\Delta}n.$$

In addition, there is a constant c such that $f^*(n,\Delta) \leq c\frac{\log\Delta}{\Delta}n$.

Theorems 2.3 and 2.2 thus determine $f(n,\Delta)$ asymptotically for sufficiently large, but fixed Δ and growing n. We state this concisely in the following corollary.

Corollary 2.4. *For every $\varepsilon \in (0,1)$ there exists an integer $\Delta_0 = \Delta_0(\varepsilon)$ such that the following holds. For any $\Delta \geq \Delta_0$ there is $N_0 = N_0(\Delta)$ such that for any $n \geq N_0$,*

$$\frac{1}{2}\cdot\frac{\log\Delta}{\Delta}n \leq f(n,\Delta) \leq (2+\varepsilon)\cdot\frac{\log\Delta}{\Delta}n.$$

Corollary 2.5. *There is a Δ_0 such that if $\Delta \geq \Delta_0$, then $f^*(n,\Delta) = \Theta\left(\frac{\log\Delta}{\Delta}n\right)$.*

There is a gap between the lower bound in Theorem 2.2 (and also the improvement, see Theorem 2.17) and the upper bound in Theorem 2.3. We leave the closing of this gap as an open problem; see Section 2.4.

Given Corollary 2.4, it is natural to consider the behaviour of $f(n,\Delta)$ when Δ is small. In this case, we have only the following modest results. The bounds are obtained as corollaries of general bounds that become less and less precise as Δ grows; see Table 2.1 and Section 2.3.2 for more explicit values.

Theorem 2.6. *For any $\Delta \geq 2$ and any $n \in \mathbb{N}$ we have $f(n,\Delta) \geq \lfloor\frac{n-2}{\Delta}\rfloor$. Moreover, for any $n \in \mathbb{N}$ we have $f(n,1) = \lfloor n/2\rfloor$, $f(n,2) = \lceil n/2\rceil - 1$, and there exists n_0 such that if $n > n_0$, then $0.3411n < f(n,3) \leq f^*(n,3) < 0.4591n$.*

We also consider the other end of the regime for Δ, when Δ is close to n. In particular, when Δ is linear in n, say $\Delta = n - cn$, it follows from Theorem 2.3 that $f(n,n-cn) = O(\log n)$. Furthermore, it is not too difficult to show that $f(n,n-cn) =$

Δ	Lower bound	Upper bound
3	0.34116	0.4591
4	0.24716	0.4212
5	0.18657	0.3887
6	0.14516	0.3621
7	0.11562	0.3395
8	0.09384	0.3201
9	0.07735	0.3031
10	0.06459	0.2882

Table 2.1: Explicit asymptotic lower and upper bounds on $f(n,\Delta)$, divided by n, obtained for small values of Δ, for n large enough.

$\Omega(\log n)$ (see Proposition 2.15 for details). When Δ is much larger (i.e. $\Delta = n - o(n)$), bounding $f(n,\Delta)$ bears a strong connection to the Zarankiewicz problem, which will be discussed in the next section, and we are able to obtain the following result. We formulate it in terms of a bound on the degrees guaranteeing a bihole of constant size t. Let

$$\Delta_n(t) := \max\{q \,:\, f(n,q) = t\}.$$

Theorem 2.7. *Let $t \geq 4$ be an integer. There is a positive constant C and an integer N_0 such that if $n > N_0$, then $n - Cn^{1-1/t} \leq \Delta_n(t) \leq n - Cn^{1-\frac{2}{t+1}}$. In addition there is an integer N_0, such that if $n > N_0$, then $\Delta_n(2) = n - n^{1/2}(1+o(1))$ and $\Delta_n(3) = n - n^{2/3}(1+o(1))$.*

The main results of this chapter are joined work with Axenovich, Sereni and Snyder [13].

This chapter is structured as follows. We describe connections between the function $f(n,\Delta)$, classical bipartite Ramsey numbers, and the Erdős-Hajnal conjecture in Section 2.2. We prove Theorems 2.2 and 2.3 in Section 2.3.1 and prove Theorem 2.6 and establish the values for Table 2.1 in Section 2.3.2. We prove Theorem 2.7 in Section 2.3.3. Section 2.4 provides concluding remarks and open questions.

2.2 Related problems

The function $f(n,\Delta)$ is closely related to the bipartite version of the Erdős-Hajnal conjecture, bipartite Ramsey numbers, and the Zarankiewicz function.

Note that $f(n,\Delta)$ for Δ sublinear in n corresponds to $\tilde{h}(n,H)$ for H a bipartite graph with one vertex in one part of degree $\Delta+1$. Indeed, a bipartite graph not having a star with $\Delta + 1$ leaves is a graph with vertices in one part having degrees at most Δ. Note that in our bipartite version of the Erdős-Hajnal conjecture we respect sides, so H can also have an arbitrary number of isolated vertices in the larger part. In such an H-free graph G with degree at most Δ in one part there are clearly no complete bipartite graphs with $\Delta + 1$ vertices in each part, so the largest homogeneous set is a bihole, as Δ is sublinear in n. Its size is thus determined by $f(n, \Delta)$.

Furthermore, the parameter $f(n, \Delta)$ bears a connection to bipartite Ramsey numbers. If H_1 and H_2 are bipartite graphs, then the *bipartite Ramsey number* $\mathrm{br}(H_1, H_2)$ is the smallest integer n such that any red-blue colouring of the edges of $K_{n,n}$ produces a red copy of H_1 or a blue copy of H_2 respecting sides. Thus, if $f(n, \Delta) = k$, then $\mathrm{br}(K_{1,\Delta+1}, K_{k,k}) = n$. For results on bipartite Ramsey numbers, see Caro and Rousseau [40], Thomason [121], Hattingh and Henning [87], Irving [89], and Beineke and Schwenk [21].

Finally, considering the bipartite complement, determining $f(n, \Delta)$ is related to the Zarankiewicz problem in bipartite graphs. Recall that $z(n; t)$ denotes the maximum number of edges in a subgraph of $K_{n,n}$ with no copy of $K_{t,t}$. Finding a large bihole in a bipartite graph is the same as finding a large copy of $K_{t,t}$ in the bipartite complement, where the bipartite complement has large minimum degree on one side (this is spelled out more carefully in Section 2.3.3). There is some literature on the Zarankiewicz problem for t large (see, for example, Balbuena et al. [19, 20], Čulík [49], Füredi and Simonovits [77], Griggs and Ouyang [82], and Griggs, Simonovits and Thomas [83]). However, most of these results address the case when t is close to $n/2$, or when the results do not lead to improvements on our bounds.

2.3 Bounds on $f(n, \Delta)$

In this section we establish upper and lower bounds on $f(n, \Delta)$ for various ranges of Δ. First, we establish the exact value of $f(n, 2)$ that is used in other results. Then, we treat the case when Δ is fixed but large. We then move on to the case when Δ is a small fixed constant, and finally, when Δ is large, i.e. close to n.

Lemma 2.8. *For every positive integer n, we have $f(n, 2) = \lceil n/2 \rceil - 1$.*

Proof. To see that $f(n, 2) < n/2$, simply consider an even cycle C_{2n} on $2n$ vertices. It remains to establish the lower bound. Let $H = (A \dot\cup B, E)$ be an $(n \times n)$ bipartite graph where the degree of each vertex in A is at most 2. Note that we may assume without loss of generality that the degree of each vertex in A is exactly 2. Consider the auxiliary multi-graph G with vertex set B, in which two vertices are adjacent if and only if they have a common neighbour in H. Consequently, there is a natural bijection between the edges of G and A, and thus, G has n vertices and n edges. We assert that G contains a set E' of edges and a set V' of vertices each of size $\lceil n/2 \rceil - 1$, and such that no edge in E' has a vertex in V'. Note then that this pair of sets corresponds to parts of a bihole in H of size $\lceil n/2 \rceil - 1$, thus proving that $f(n, 2) \geq \lceil n/2 \rceil - 1$. The rest of the proof is devoted to proving the above assertion.

To this end, we consider the (connected) components of G: a component C is *dense* if $|E(G[C])| \geq |C|$. Let $S_1, \ldots, S_k$ be the components of G, enumerated such that $S_1, \ldots, S_m$ are dense and the others are not. Note that we must have at least one dense component, so $m \geq 1$, and it could be that all components are dense. Let x be the number of components of G that are not dense, that is, $x := k - m \in \{0, \ldots, k-1\}$. Let v and e be the number of vertices and edges, respectively, in the union of all dense components of G. Then the total number of edges in non-dense components of G is $n-e$ and the total number of vertices in these components is $n - v$. In addition, the number of vertices in non-dense components is at least the number of edges plus the number of components. Thus, $n - v \geq n - e + x$, so $x \leq e - v$.

Let G' be a subgraph of G with precisely $\lceil n/2 \rceil - 1$ edges and consisting of $S_1, \ldots, S_q$ and a connected subgraph of S_{q+1}, for some $q \in \{0, \ldots, k-1\}$. In particular, if S_1 has at least $\lceil n/2 \rceil - 1$ edges, then G' is a connected subgraph of S_1. It suffices to show that G' has at most $\lfloor n/2 \rfloor + 1$ vertices, since we can then choose a set V' of $\lceil n/2 \rceil - 1$ vertices in $V(G) \setminus V(G')$, which along with $E' := E(G')$ will form the sought pair (V', E'). To this end, first notice that if G' has at most one non-dense component, then the number of vertices of G' is at most $|E'|+1$, which is at most $\lceil n/2 \rceil \leq \lfloor n/2 \rfloor + 1$, as desired. Suppose now that G' has more than one non-dense component. It follows that G' contains all dense components of G. Let x' be the number of non-dense components of G'. Then $x' \leq x$. The number of edges in dense components of G' is e, and thus, the number of edges in non-dense components of G' is $\lceil n/2 \rceil - 1 - e$. This implies that the number of vertices in non-dense components of G' is at most $(\lceil n/2 \rceil - 1 - e) + x' \leq (\lceil n/2 \rceil - 1 - e) + x$. Adding the number v of vertices in dense components of G and the number of vertices in non-dense components of G', we see that the total number of vertices in G' is at most $v + ((\lceil n/2 \rceil - 1 - e) + x) \leq \lceil n/2 \rceil - 1 \leq \lfloor n/2 \rfloor + 1$. This concludes the proof. $\square$

2.3.1 Proof of Theorems 2.2 and 2.3

The upper bound given in Theorem 2.3 comes from suitably modifying a random bipartite graph $G \in \mathcal{G}\left(n, n, \frac{\Delta}{n}\right)$. The idea of the proof of the lower bound given in Theorem 2.2 is as follows. Let $G = (A \dot{\cup} B, E)$ be an $(n \times n)$ bipartite graph with $d(x) \leq \Delta$ for every $x \in A$. We choose an appropriate parameter s and choose a subset S of B uniformly at random from the set of all s-element subsets of B and consider the set T of vertices in A that have at least $\Delta - 2$ neighbours in S. Lemma 2.8 can then be applied to the bipartite graph induced on parts $(T, B \setminus S)$, as in this bipartite graph every vertex in T has degree at most 2. Intuitively, the set T should be "large enough" to guarantee a large bihole in G. Floors and ceilings, when not relevant, are ignored in what follows. We start by establishing the lower bound, that is Theorem 2.2.

Proof of Theorem 2.2. Consider an arbitrary bipartite graph with parts A and B each of size n so that the degrees of vertices in A are at most Δ. Up to adding edges arbitrarily, we may assume without loss of generality that each vertex in A has degree exactly Δ. Choose a subset S of B of size $(1-2x)n - 2$ randomly and uniformly among all such subsets, where $x := \frac{1}{2}\frac{\log \Delta}{\Delta}$. We assume that $n > 5\Delta \log \Delta$ and $\Delta \geq \Delta_0$ is chosen large enough to satisfy the last inequality in the proof. Let X be the random variable counting the number of vertices in A with at least $\Delta - 2$ neighbours in S. Then

$$\mathbb{E}[X] \geq n \cdot h(x, n, \Delta),$$

where $h(x, n, \Delta)$ denotes the probability that an arbitrary vertex in A has exactly $\Delta - 2$ neighbours in S. Since we may assume that every vertex in A has degree exactly Δ, we have

$$h(x, n, \Delta) = \binom{\Delta}{\Delta - 2}\binom{n - \Delta}{(1-2x)n - \Delta}\binom{n}{(1-2x)n - 2}^{-1}.$$

Observe that if $\mathbb{E}[X] \geq 2xn + 2$, then there is a set A' of at least $2xn + 2$ vertices in A, each sending at most 2 edges to $B \setminus S$. Since $|B \setminus S| = 2xn + 2$, Lemma 2.8 implies that there is a bihole between A' and $B \setminus S$ of size at least xn. Thus, it is sufficient to prove that $h(x, n, \Delta) \geq 2x + 2/n$. Let us now verify this inequality.

Recall that $x = \frac{1}{2}\frac{\log \Delta}{\Delta}$. Let $\alpha = 1 - 2x$, so $\alpha = 1 - \frac{\log \Delta}{\Delta} = \frac{\Delta - \log \Delta}{\Delta} \in (0, 1)$. Note that $\alpha n \geq \Delta$ since $n \geq 5\Delta \log \Delta$. Let $\beta = \frac{1}{\alpha} - 1$. Then $\beta = \frac{\log \Delta}{\Delta - \log \Delta}$. We have

$$h(x, n, \Delta) = \frac{\binom{\Delta}{\Delta-2}\binom{n-\Delta}{(1-2x)n-\Delta}}{\binom{n}{(1-2x)n-2}}, \qquad \text{i.e.}$$

$$
\begin{aligned}
h(x,n,\Delta) &= \frac{\binom{\Delta}{\Delta-2}\binom{n-\Delta}{(1-2x)n-\Delta}}{\binom{n}{(1-2x)n-2}} \\
&= \binom{\Delta}{2}\prod_{j=2}^{\Delta-1}\left(\frac{\alpha n - j}{n-j}\right)\cdot\left[\frac{(2xn+2)(2xn+1)}{n(n-1)}\right] \\
&> \binom{\Delta}{2}\prod_{j=2}^{\Delta-1}\left(\frac{\alpha n - j}{n-j}\right)\cdot (2x)^2 \\
&= \binom{\Delta}{2}(2x)^2\alpha^{\Delta-2}\prod_{j=2}^{\Delta-1}\left(1-\frac{\beta j}{n-j}\right) \\
&\geq \binom{\Delta}{2}(2x)^2\alpha^{\Delta-2}\left(1-\frac{\beta\Delta}{n-\Delta}\right)^{\Delta-2} \qquad (2.1)\\
&\geq \binom{\Delta}{2}(2x)^2\alpha^{\Delta}\left(1-\frac{\beta\Delta}{n-\Delta}\right)^{\Delta}, \qquad (2.2)
\end{aligned}
$$

where (2.1) holds because the function $j \mapsto \frac{\beta j}{n-j}$ is increasing, as $\beta > 0$. Now, expressing β in terms of Δ, we note that

$$
\frac{\beta\Delta}{n-\Delta} \leq \frac{\Delta\log\Delta/(\Delta-\log\Delta)}{5\Delta\log\Delta-\Delta} = \frac{\log\Delta}{(5\log\Delta-1)(\Delta-\log\Delta)} \leq 1,
$$

and therefore Bernoulli's inequality can be applied to (2.2). It follows that

$$
\begin{aligned}
h(x,n,\Delta) &> \binom{\Delta}{2}(2x)^2(1-2x)^{\Delta}\left(1-\frac{\Delta^2\log\Delta}{(\Delta-\log\Delta)(n-\Delta)}\right) \\
&\geq \binom{\Delta}{2}(2x)^2(1-2x)^{\Delta}\left(1-\frac{4\Delta\log\Delta}{n}\right) \qquad (2.3)\\
&\geq \binom{\Delta}{2}(2x)^2(1-2x)^{\Delta}\frac{1}{5}, \qquad (2.4)
\end{aligned}
$$

where (2.3) follows since $\frac{1}{n-\Delta} < \frac{2}{n}$ and $\log\Delta < \Delta/2$, and (2.4) holds since $n > 5\Delta\log\Delta$. Now, note that $(1-2x)^{\Delta} = \left(1-\frac{\log\Delta}{\Delta}\right)^{\Delta} \geq \frac{1}{2}e^{-\frac{\log\Delta}{\Delta}\cdot\Delta} = \frac{1}{2\Delta}$. Thus, from (2.4) we obtain

$$
h(x,n,\Delta) > \binom{\Delta}{2}(2x)^2\frac{1}{10\Delta} = (2x)\frac{(\Delta-1)\log\Delta}{20\Delta}.
$$

Finally, to bound the right-hand side of the above inequality from below, observe that

$$
(2x)\frac{(\Delta-1)\log\Delta}{20\Delta} \geq (2x)\left(1+\frac{1}{40}\log\Delta\right) = 2x+\frac{\log^2\Delta}{40\Delta} \geq 2x+\frac{2}{5\Delta\log\Delta} \geq 2x+\frac{2}{n},
$$

where these inequalities hold for sufficiently large Δ. Accordingly, $h(x,n,\Delta) > 2x+\frac{2}{n}$, which concludes the proof of Theorem 2.2. $\square$

To prove Theorem 2.3 we shall need to use Chernoff's bound. Specifically, we use the following version (see [90], Corollary 21.7, p.401, for example).

Lemma 2.9 (Chernoff's bounds). *Let X be a random variable with distribution* $\mathrm{Bin}(N, p)$ *and $\delta \in (0,1)$. Then*

$$\mathbb{P}[X \geq (1+\delta)\,\mathbb{E}\,X] \leq \exp\left(-\frac{\delta^2}{3}\mathbb{E}X\right) \quad \textit{and} \tag{2.5}$$

$$\mathbb{P}[X \leq (1-\delta)\,\mathbb{E}\,X] \leq \exp\left(-\frac{\delta^2}{2}\,\mathbb{E}\,X\right). \tag{2.6}$$

Proof of Theorem 2.3. Let $\varepsilon' \in (0,1)$ be arbitrary and let $\varepsilon := \varepsilon'/8$. We shall assume that $\Delta \geq \Delta_0(\varepsilon)$ is sufficiently large such that our inequalities hold. In particular, we assume that $\Delta \geq 27$. Suppose that $n \geq \frac{3\Delta}{5\log(\Delta/2)}$. Set $N := (1+\varepsilon)n$ and $\Delta' := (1-\varepsilon)\Delta$, so in particular $\Delta' \geq 13.5$. We consider first $H \in \mathcal{G}\left(N, N, \frac{\Delta'}{N}\right)$, that is, H is a random $(N \times N)$ bipartite graph with parts A and B, where each edge ab with $a \in A$ and $b \in B$ is chosen independently with probability Δ'/N. We first establish that the random graph H contains no "large" biholes with fairly large probability. In the following, for subsets $X \subseteq A$ and $Y \subseteq B$, let $e(X, Y)$ denote the number of edges with one endpoint in X and the other in Y.

(A). *With probability at least 0.75, any two subsets $X \subset A$ and $Y \subset B$ with $|X| = |Y| = \frac{2N\log\Delta'}{\Delta'}$ satisfy $e(X,Y) > 0$.*

Proof. Set $m := \frac{2N\log\Delta'}{\Delta'}$ and note that m is therefore at least $\frac{6}{5}$. Suppose that $X \subset A$ and $Y \subset B$ both have size m. Then

$$\mathbb{P}[e(X,Y) = 0] = \left(1 - \frac{\Delta'}{N}\right)^{m^2}.$$

Let p be the probability that there is a pair (X, Y), with $X \subset A$ and $Y \subset B$, $|X| = |Y| = m$, such that $e(X,Y) = 0$. Forming a union bound over all possible pairs of sets of size m, we have

$$p \leq \binom{N}{m}^2 \left(1 - \frac{\Delta'}{N}\right)^{m^2} \leq \left(\frac{Ne}{m}\right)^{2m} e^{-\frac{\Delta' m^2}{N}} = \left(\frac{\Delta' e}{2\log\Delta'} e^{-\log\Delta'}\right)^{2m} = \left(\frac{e}{2\log\Delta'}\right)^{2m}$$
$$\leq 0.25.$$

Here, we used the standard estimates $\binom{t}{k} \leq \left(\frac{t\cdot e}{k}\right)^k$, $1-x \leq e^{-x}$, the fact that $(e/2\log\Delta') < 0.53$ because $\Delta' \geq 13.5$, as well as the inequality $2m \geq \frac{12}{5}$. This establishes (A). ■

Now, let $g(\varepsilon) = 2\log\left(\frac{4(1+\varepsilon)}{3\varepsilon}\right)$. We shall show that, with probability sufficiently large for our purposes, at least n of the vertices of A have degree at most $\Delta' + \sqrt{3g(\varepsilon)\Delta'}$.

(B). *With probability greater than 0.25, the number of vertices $v \in A$ with more than $\Delta' + \sqrt{3g(\varepsilon)\Delta'}$ neighbours in B is at most $\frac{\varepsilon}{1+\varepsilon}N$.*

Proof. We use standard concentration inequalities to show that the degree of every vertex in A is approximately Δ'. For each vertex $v \in A$, let X_v be the degree of v in H. Noting that $\mathbb{E}[X_v] = \Delta'$, we apply (2.5) from Lemma 2.9 with $\delta := \sqrt{3g(\varepsilon)/\Delta'} < 1$ to obtain

$$\mathbb{P}[X_v \geq (1+\delta)\Delta'] \leq \exp\left(-\frac{\left(\sqrt{3g(\varepsilon)/\Delta'}\right)^2}{3}\Delta'\right) = e^{-g(\varepsilon)} = (0.75)^2\left(\frac{\varepsilon}{1+\varepsilon}\right)^2$$
$$< 0.75 \cdot \frac{\varepsilon}{1+\varepsilon}.$$

Letting X be the random variable counting those vertices $v \in A$ with $X_v \geq (1+\delta)\Delta'$, by Markov's inequality, we deduce that $\mathbb{P}[X \geq \frac{\varepsilon}{1+\varepsilon}N] < 0.75$, thereby establishing (B). ■

It follows from (A) and (B) that with positive probability, H has no large biholes, and at least $N - \frac{\varepsilon}{1+\varepsilon}N = n$ of the vertices in A have degree at most $\Delta' + \sqrt{3g(\varepsilon)\Delta'} \leq \frac{1}{1-\varepsilon}\Delta' = \Delta$, which holds for sufficiently large Δ depending on ε. We now fix such a graph H. We can thus choose a subset A' of A of size n such that every vertex in A' has degree at most Δ in H. Now, arbitrarily choosing a subset B' of B of size n, we know that the subgraph H' of H induced by $A'\dot{\cup}B'$ is an $(n \times n)$ bipartite graph with maximum degree Δ on one side and without a bihole of size larger than

$$2(1+\varepsilon)n\left(\frac{\log(\Delta')}{(1-\varepsilon)\Delta}\right) < (2+8\varepsilon)\left(\frac{\log\Delta}{\Delta}\right)n = (2+\varepsilon')\left(\frac{\log\Delta}{\Delta}\right)n.$$

In order to obtain an upper bound on $f^*(n,\Delta)$, all that is required is to make the example obtained above have bounded maximum degree in both parts. Thus, it suffices to apply Chernoff's inequality to all vertices (instead of just the vertices in A). We may have to remove more vertices after doing this, but the loss will only be reflected in the constant. This completes the proof of Theorem 2.3. □

2.3.2 Bounding $f(n,\Delta)$ for small Δ

We have already established a part of Theorem 2.6 via Lemma 2.8. Namely, we showed that $f(n,2) = \lceil n/2 \rceil - 1$. It is not hard to see that $f(n,1) = \lfloor n/2 \rfloor$. Thus, our aim in this section is to investigate the behaviour of $f(n,3)$ more closely, and to complete the proof of Theorem 2.6. First, let us note the following lower bound on $f(n, \Delta)$, valid for all integers n and Δ greater than 1. In the following, for a vertex subset X of a graph let $N(X)$ be the neighbourhood of X, i.e. $N(X) = \bigcup_{x \in X} N(x)$.

Proposition 2.10. *If n and Δ are two integers greater than 1, then $f(n, \Delta) \geq \lfloor \frac{n-2}{\Delta} \rfloor$.*

Proof. We shall prove this by induction on Δ with the base case $\Delta = 2$ following from Lemma 2.8. Let $H = (A \dot\cup B, E)$ be an $(n \times n)$ bipartite graph, such that the degree of each vertex in part A is equal to Δ, $\Delta \geq 3$.

Consider a set X of $\lfloor (n-2)/\Delta \rfloor$ vertices in B. If $|N(X)| \leq n - \lfloor (n-2)/\Delta \rfloor$, then X and $A \setminus N(X)$ form a bihole with at least $\lfloor (n-2)/\Delta \rfloor$ vertices in each part. Otherwise, $|N(X)| > n - \lfloor (n-2)/\Delta \rfloor$. Let $G' := G[N(X) \dot\cup (B \setminus X)]$. Then each of the parts of G' has size at least $n - \lfloor (n-2)/\Delta \rfloor \geq n - (n-2)/\Delta$ and the maximum degree of vertices of $N(X)$ in G' is at most $\Delta - 1$. Thus, by induction G' has a bihole of size at least $\left\lfloor \frac{1}{\Delta - 1}(n - (n-2)/\Delta - 2) \right\rfloor = \lfloor \frac{n-2}{\Delta} \rfloor$. $\square$

It follows from the above proposition that $f(n, 3) \geq \lfloor (n-2)/3 \rfloor$. However, this lower bound can be improved slightly by choosing a random subset of B and considering the neighbourhood of this set in A, similarly as in the proof of the lower bound in Theorem 2.2.

Lemma 2.11. *If n and Δ are two integers greater than 1, then $f(n, \Delta) \geq f(\lfloor \xi n \rfloor, \Delta - 1)$, where $\xi = \xi(\Delta)$ is a solution to the inequality $1 - \xi^{\Delta} \geq \xi$.*

Proof. For simplicity we omit floors in the following. Let G be a bipartite graph with parts A and B each of size n such that the vertices in A have degrees at most Δ. We shall show that there is a set $S \subset B$, such that $|S| = (1-\xi)n$ and such that $|N(S)| \geq \xi n$. To do this, we shall choose S randomly and uniformly out of all subsets of B of size $(1-\xi)n$ and show that the expected number X of vertices from A with at least one neighbour in S is at least ξn. Indeed, if p is the probability for a fixed vertex in A not to have a neighbour in S, then

$$p = \frac{\binom{n-\Delta}{(1-\xi)n}}{\binom{n}{(1-\xi)n}}.$$

Using the identity $\binom{n-l}{k} / \binom{n}{k} = \binom{n-k}{l} / \binom{n}{l}$, we see that $p = \binom{\xi n}{\Delta} / \binom{n}{\Delta}$. Now, using the inequality $\binom{\delta n}{r} \leq \delta^r \binom{n}{r}$, which is valid for every $\delta \in (0,1)$, we find that

$$p = \frac{\binom{\xi n}{\Delta}}{\binom{n}{\Delta}} \leq \xi^{\Delta}.$$

Thus, $\mathbb{E}[X] = n(1-p) \geq n(1-\xi^{\Delta}) \geq n\xi$ by our choice of ξ. Consequently, with positive probability $|N(S)| \geq \xi n$. We also have $|B \setminus S| = \xi n$. Since each vertex of $N(S)$ sends at most $\Delta - 1$ edges to $B \setminus S$, it follows that there is a bihole between $N(S)$ and $B \setminus S$ of size $f(\xi n, \Delta - 1)$. This completes the proof. □

We make explicit some lower bounds obtained using Lemma 2.11 (and Lemma 2.8).

Corollary 2.12. *There exists N_0 such that if $n \geq N_0$, then $f(n,3) > 0.34116n$, $f(n,4) > 0.24716n$, $f(n,5) > 0.18657n$, and $f(n,6) > 0.14516n$.*

The next natural step regarding small values of Δ is to evaluate how good the bounds written in Corollary 2.12 are. The following upper bounds are obtained by analysing the pairing model (also known as the *configuration model*) to build random regular graphs, tailored to the bipartite setting.

Lemma 2.13. *Let Δ be an integer greater than 2, and assume that $\beta \in (0,1/2)$ is such that*

$$\frac{(1-\beta)^{2\Delta(1-\beta)}}{\beta^{2\beta}(1-\beta)^{2(1-\beta)}(1-2\beta)^{\Delta(1-2\beta)}} < 1.$$

Then there exists $N_0 = N_0(\beta)$ such that for every $n > N_0$ we have $f(n,\Delta) \leq f^*(n,\Delta) < \beta n$. In particular, for n sufficiently large there exists a Δ-regular $(n \times n)$ bipartite graph with no bihole of size at least βn.

Proof. We shall work with the configuration model of Bollobás [25] suitably altered to produce a bipartite graph. Fix an integer n and consider two sets of Δn (labelled) vertices each: $X = \{x_1^1, \ldots, x_1^{\Delta}, \ldots, x_n^1, \ldots, x_n^{\Delta}\}$ and $Y = \{y_1^1, \ldots, y_1^{\Delta}, \ldots, y_n^1, \ldots, y_n^{\Delta}\}$. Choose a perfect matching F between X and Y uniformly at random. We call F a *pairing*.

Given a pairing F, for each $i \in \{1, \ldots, n\}$ the vertices $x_i^1, \ldots, x_i^{\Delta}$ are identified with a new vertex x_i, and similarly the vertices $y_i^1, \ldots, y_i^{\Delta}$ are identified with a new vertex y_i. This yields a multi-graph G'. We prove that with positive probability G' is a simple graph. To see why this holds, first notice that the total number of different pairings is $(\Delta \cdot n)!$. Second, each fixed (labelled) Δ-regular $(n \times n)$ bipartite graph arises from precisely $(\Delta!)^{2n}$ different pairings (because for each vertex x_i we can freely permute the vertices $\{x_i^1, \ldots, x_i^{\Delta}\}$). Third, McKay, Wormald and Wysocka [106] proved that the number of different labelled Δ-regular $(n \times n)$ bipartite graphs is

$$(1+o(1)) \exp\left(-\frac{(\Delta-1)^2}{2} \cdot ((\Delta-1)^2+1)\right) \frac{(\Delta \cdot n)!}{(\Delta!)^{2n}},$$

which is at least $c\frac{(\Delta\cdot n)!}{(\Delta!)^{2n}}$ for some $c > 0$. Combining these three facts, we find that $\mathbb{P}[G'$ is a graph$] \geq c > 0$, as announced.

Now, fix $k = k(n) = \beta \cdot n$ for some $\beta \in (0,1)$. For each $i \in \{1,\ldots,n\}$, set $X_i := \{x_i^1,\ldots,x_i^\Delta\}$ and $Y_i := \{y_i^1,\ldots,y_i^\Delta\}$. Fixing a family of k sets $\mathcal{X} = \{X_{i_1},\ldots,X_{i_k}\}$ and also $\mathcal{Y} = \{Y_{i_1},\ldots,Y_{i_k}\}$, let $W(\mathcal{X},\mathcal{Y})$ be the event that the union U of the sets in $\mathcal{X} \cup \mathcal{Y}$ spans no edge in F (and thus, corresponds to a bihole in G'). Let us now find the probability of $W(\mathcal{X},\mathcal{Y})$. There are $2\Delta k$ edges incident to vertices in U. The number of ways to choose these $2\Delta k$ edges is as follows: if the edge is incident to a vertex in $\bigcup_{j=1}^k X_{i_j}$, then its other vertex must belong to $Y \setminus \bigcup_{j=1}^k Y_{i_j}$, and hence there are

$$(\Delta n - \Delta k)\cdots(\Delta n - 2\Delta k + 1)$$

different ways of choosing the edges incident to a vertex in $\bigcup_{j=1}^k X_{i_j}$. The situation is analogous for edges incident to $\bigcup_{j=1}^k Y_{i_j}$, yielding a total of $(\Delta n - \Delta k)^2 \cdots (\Delta n - 2\Delta k + 1)^2$ ways to choose the $2\Delta k$ edges incident to a vertex in U. For each such choice there are $(\Delta n - 2\Delta k)!$ ways to choose the remaining edges, for a total of $(\Delta n - 2\Delta k)! \cdot (\Delta n - \Delta k)^2 \cdots (\Delta n - 2\Delta k + 1)^2$ different pairings in which U spans no edge. It follows that

$$\mathbb{P}(W(\mathcal{X},\mathcal{Y})) = \frac{(\Delta n - 2\Delta k)! \cdot (\Delta n - \Delta k)^2 \cdots (\Delta n - 2\Delta k + 1)^2}{(\Delta n)!} = \frac{((\Delta n - \Delta k)!)^2}{(\Delta n)!(\Delta n - 2\Delta k)!}.$$

Let $W = \bigcup_{\mathcal{X},\mathcal{Y}} W(\mathcal{X},\mathcal{Y})$ be the event that F contains a bihole of size k. Taking the union bound over all $\binom{n}{k}^2$ choices of $(\mathcal{X},\mathcal{Y})$, we find that

$$\mathbb{P}(W) \leq \binom{n}{k}^2 \frac{((\Delta n - \Delta k)!)^2}{(\Delta n)!(\Delta n - 2\Delta k)!}. \tag{2.7}$$

Using Stirling's approximation,

$$\sqrt{2\pi n}\left(\frac{n}{e}\right)^n \exp\left(-\frac{1}{12n+1}\right) \leq n! \leq \sqrt{2\pi n}\left(\frac{n}{e}\right)^n \exp\left(-\frac{1}{12n}\right),$$

in (2.7), and ignoring the exponential factors (they can be bounded from above by $\exp(1/(cn))$ for some positive constant c, and hence be made arbitrarily close to 1), we thus obtain

$$\binom{n}{k}^2 \approx \frac{1}{2\pi\beta(1-\beta)n}\left(\frac{1}{\beta^{2\beta}(1-\beta)^{2(1-\beta)}}\right)^n$$

and

$$\frac{(\Delta n - \Delta k)!^2}{(\Delta n)!(\Delta n - 2\Delta k)!} \approx \frac{1-\beta}{\sqrt{1-2\beta}}\left(\frac{(1-\beta)^{2\Delta(1-\beta)}}{(1-2\beta)^{\Delta(1-2\beta)}}\right)^n.$$

Hence, if

$$\frac{(1-\beta)^{2\Delta(1-\beta)}}{\beta^{2\beta}(1-\beta)^{2(1-\beta)}(1-2\beta)^{\Delta(1-2\beta)}} < 1,$$

then $\mathbb{P}(W) \to 0$ as $n \to \infty$. Thus, with positive probability G' is a Δ-regular graph with no bihole of size at least $k = \beta n$ (for n sufficiently large), as stated. □

Performing explicit computations in Lemma 2.13 for specific values of Δ yields the following bounds (see also Table 2.1).

Corollary 2.14. *There exists N_0 such that if $n \geq N_0$, then $f^*(n, 3) < 0.4591n$, $f^*(n, 4) < 0.4212n$, $f^*(n, 5) < 0.3887n$, and $f^*(n, 6) < 0.3621n$.*

In particular, one sees that $f(n, 3) \leq f^*(n, 3) < 0.4591n$ for sufficiently large n. Thus, combined with our earlier work, it follows that $0.3411n < f(n, 3) < 0.4591n$. It would be very interesting to improve either the lower or upper bound.

2.3.3 BOUNDING $f(n, \Delta)$ WHEN Δ IS LARGE

In this section we address the behaviour of $f(n, \Delta)$ for large Δ and prove Theorem 2.7. Before doing so, let us note the following simple result, which shows that Theorem 2.3 is tight (up to constants) when Δ is linear in n.

Proposition 2.15. *For any $\varepsilon \in (0, 1)$ there is a constant $c = c(\varepsilon)$ such that for n sufficiently large $f(n, (1-\varepsilon)n) \geq c \log n$.*

Proof. Let $\varepsilon \in (0, 1)$. To show the lower bound on $f(n, (1-\varepsilon)n)$, consider an $(n \times n)$ bipartite graph G with parts A, B, such that $d(x) \leq (1-\varepsilon)n$ for every $x \in A$. Letting G^c be the bipartite complement of G, we see that G^c has at least εn^2 edges. The result then follows from the fact that for any $\varepsilon \in (0, 1)$ and sufficiently large n, any $(n \times n)$ bipartite graph with εn^2 edges contains a $K_{t,t}$ where $t = c \log n$ for some constant $c = c(\varepsilon)$. This can be proved using the standard Kővári-Sós-Turán [98] double counting argument. □

Therefore, the behaviour of $f(n, \Delta)$ is clear whenever Δ is linear, aside from more precise estimates of the constants involved. What happens when Δ is very large, more precisely, when $\Delta = n - o(n)$? This is partly addressed in Theorem 2.7, which we now prove.

Proof of Theorem 2.7. Recall that the classical Zarankiewicz number, $z(n;t)$, is the largest number of edges in an $(n \times n)$ bipartite graph that contains no copy of $K_{t,t}$. Assume first that $t \geq 4$.

The lower bound on $\Delta_n(t)$ follows from standard bounds on Zarankiewicz numbers. Indeed, $z(n;t) \leq Cn^{2-1/t}$ for some constant $C = C(t)$, see for example [98]. Thus, any $(n \times n)$ bipartite graph on at least $Cn^{2-1/t}$ edges contains a copy of $K_{t,t}$, and so any $(n \times n)$ bipartite graph on at most $n^2 - Cn^{2-1/t}$ edges contains a bihole of size t. In particular, any $(n \times n)$ bipartite graph with maximum degree at most $n - Cn^{1-1/t}$ contains a bihole of size t. So the announced lower bound in Theorem 2.7 holds.

To determine the stated upper bound on $\Delta_n(t)$, we shall prove the existence of a $K_{t,t}$-free bipartite $(n \times n)$ graph with the additional constraint that the minimum degree of vertices (on one side) is large. For that, we shall alter the standard random construction used to prove lower bounds on Zarankiewicz numbers. For a graph F, we shall carefully control $X = X_F$, the total number of copies of $K_{t,t}$ in F, as well as $X(v) = X_F(v)$, the number of copies of $K_{t,t}$ containing a vertex v in F.

Let $N := 2n$ and $p := cN^{-2/(t+1)}$, for a constant c to be determined later. Consider a bipartite binomial random graph $G' \in \mathcal{G}(N, n, p)$ with parts A and B of sizes N and n, respectively. By Markov's inequality, $\mathbb{P}[X \geq 2\,\mathbb{E}[X]] \leq 1/2$. Since $d(v)$, for $v \in A$, is distributed as $\mathrm{Bin}(n,p)$, Chernoff's inequality (2.6) from Lemma 2.9 with $\varepsilon := 1/2$, implies that with high probability, every vertex $v \in A$ has degree at least $pn/2$. So with positive probability we have $X \leq 2\,\mathbb{E}[X] = 2\binom{N}{t}\binom{n}{t}p^{t^2} \leq 2\binom{N}{t}^2 p^{t^2}$ and $d(v) \geq pn/2$ for every $v \in A$.

Fix a bipartite graph G with these properties, i.e. G is a bipartite graph with parts A and B, the number $X = X_G$ of copies of $K_{t,t}$ satisfies $X \leq 2\binom{N}{t}^2 p^{t^2}$ and $d(v) \geq pn/2$ for every $v \in A$. Observe that there are fewer than n vertices v in A with $X(v) > \frac{2t}{n} 2\binom{N}{t}^2 p^{t^2}$. Indeed, otherwise $X \geq n\frac{4t}{n}\binom{N}{t}^2 p^{t^2}/t$, a contradiction.

Let $A' \subset A$ be a set of n vertices such that $X(v) \leq \frac{2t}{n} 2\binom{N}{t}^2 p^{t^2}$ for all $v \in A'$. Let H' be the subgraph of G induced by $A' \dot\cup B$. Finally, let H be obtained from H' by removing an edge from each copy of $K_{t,t}$. Thus, H has no copies of $K_{t,t}$. It remains to check that the degrees of vertices in A are sufficiently large. Indeed, for any $v \in A$, we have $d_H(v) \geq d_{H'}(v) - X(v)$, i.e.

$$\begin{aligned} d_H(v) &\geq d_{H'}(v) - X(v) \\ &\geq np/2 - \frac{4t}{n}\binom{N}{t}^2 p^{t^2} \\ &\geq Np/4 - \frac{4t}{n}N^{2t}p^{t^2} \\ &\geq \frac{c}{4}N^{1-\frac{2}{t+1}} - \frac{4t}{n}N^{2t}p^{t^2} \\ &> (c/4 - 4tc^{t^2})n^{1-\frac{2}{t+1}} \\ &\geq \frac{1}{16}n^{1-\frac{2}{t+1}}, \end{aligned}$$

where the last inequality holds for $c = 1/2$. This concludes the proof of the general upper bound for $t \geq 4$.

Now, let $t = 2$ or $t = 3$. We have $z(n;2) = (1 + o(1))\, n^{3/2}$ and $z(n;3) = (1 + o(1))\, n^{5/3}$. For the former, the upper bound on $z(n;2)$ is due to Reiman [115] and the lower bound is due to Erdős, Renyi and Sós [63], and independently Brown [30]. For the latter, the lower bound is due to Brown [30] (with an improvement by Alon, Rónyai and Szabó [8]), and the upper bound is due to Füredi [76]. In fact, the constructions giving lower bounds on $z(n;2)$ and $z(n;3)$ are almost regular and therefore show that there are $(n \times n)$ bipartite graphs with no $K_{2,2}$ (no $K_{3,3}$) with minimum degree at least $(1 + o(1))n^{1/2}$ (at least $(1 + o(1))n^{2/3}$), respectively. This completes the proof. □

2.4 Concluding remarks

We have made progress in determining the asymptotic behaviour of $f(n, \Delta)$. However, we could not obtain better bounds for small Δ. The most glaring open problem is the case $\Delta = 3$.

Open Problem 2.16. *Determine the value of $f(n, 3)$ for n sufficiently large.*

We were able to show that for any $\varepsilon \in (0, 1)$ and fixed (but large) Δ, if n is sufficiently large, then

$$\frac{1}{2} \cdot \frac{\log \Delta}{\Delta} n \leq f(n, \Delta) \leq (2 + \varepsilon) \cdot \frac{\log \Delta}{\Delta} n.$$

It would be interesting to close the gap between these two bounds.

Chakraborti [41] considered a similar problem, replacing maximum by average degree, and proved the following:

Theorem 2.17 (Chakraborti [41]). *For every $\epsilon \in (0,1)$ there exists $\Delta_0 = \Delta_0(\epsilon)$, such that the following holds: For each $\Delta \geq \Delta_0$, there exists N_0, such that for any $n \geq N_0$, if G is an $(n \times n)$ bipartite graph with average degree $\Delta \geq \Delta_0$, then G contains a bihole of size $(1-\epsilon)\frac{\log \Delta}{\Delta}n$.*

Clearly any $(n \times n)$ bipartite graph with maximum degree Δ in one part also has average degree at most Δ, which means that for any $\varepsilon \in (0,1)$ and fixed (but large) Δ, if n is sufficiently large, then

$$(1-\varepsilon) \cdot \frac{\log \Delta}{\Delta} n \leq f(n, \Delta) \leq (2+\varepsilon) \cdot \frac{\log \Delta}{\Delta} n.$$

Open Problem 2.18. *Close the gap between the lower and upper bounds from Theorem 2.17 and Theorem 2.3, respectively.*

CHAPTER 3 THE MULTICOLOUR VERSION OF THE ERDŐS-HAJNAL CONJECTURE

3.1 INTRODUCTION

In this and the following chapter, we will consider the multicolour version of the Erdős-Hajnal conjecture, which asserts that for any fixed integer $k \geq 3$ and for any fixed s'-edge-coloured clique K on k vertices, for $s \geq s' \geq 2$, there is a positive constant $a = a(s, K)$ such that any s-edge-colouring of a clique on n vertices with no copy of K contains a clique on $\Omega(n^a)$ vertices using at most $s-1$ colours. We will make this more precise with the following definitions.

We shall consider edge-colourings of complete graphs using colours in $[s]$ for some integer s. An *s-edge-colouring* c of the complete graph K_n on vertex set $[n]$ is a map $c : \binom{[n]}{2} \to [s]$. We denote by $|c|$ the number of colours from $[s]$ for which c^{-1} is not empty. Note that an s-edge-colouring c of K_n can be seen as an edge-partition of G into s colour classes, i.e. $K_n = G_1 \cup \cdots \cup G_s$, where G_i corresponds to a maximal subgraph of K_n whose edges are assigned colour i under c. Here G_i can be an empty graph if $|c| < s$.

For an s-edge-colouring c of K_n and an s'-edge-colouring c' of K_k, we say that c is c'-free if c does not contain a copy of c', i.e. for any $V \subseteq [n]$ such that $|V| = k$ and for any bijection $\phi : V \to [k]$, there are two vertices $x, y \in V$ such that $c(xy) \neq c'(\phi(x)\phi(y))$. Typically we assume that k is fixed and n is large, i.e. the c'-free property is a local condition on the colouring. One can think of the colouring c' as a forbidden colour pattern. One of the key questions considered is how the local restrictions impact global properties, in particular how large the homogeneous number must be:

A *homogeneous set* in an s-edge-colouring c of K_n is a set $X \subseteq [n]$ that has a colour "missing", i.e. $|\{c(xy) : x, y \in X\}| < s$. The size of a largest homogeneous set of c is denoted by $h_{s-1}(c)$ (since at most $s-1$ colours are used on the homogeneous set), or if the number of colours is clear from the context, simply $h(c)$. Note that any homogeneous set is an independent set in some colour class G_i, $i \in [s]$. Thus, we have $h(c) = \max_{i \in [s]} \alpha(G_i)$. For an s-edge-colouring c' of K_k, $k \leq n$, we also define $h_{s-1}(n, c') = \min\{h(c) \mid c \text{ is a } c'\text{-free } s\text{-edge-colouring of } K_n\}$.

Note that for $s = 2$ one colour of c corresponds to the edges of some n-vertex graph G and the other colour corresponds to the edges of $\overline{G}$. Then in particular, we have

$h(c) = \max\{\alpha(G), \omega(G)\}$, which coincides with the definition of a homogeneous set in 2-graphs.

Definition 3.1. *Let c' be an s'-edge-colouring of K_k and let $s \geq s'$. If there is a positive constant $\epsilon = \epsilon(c', s)$ and a constant $C > 0$, such that any c'-free s-edge-colouring c of K_n satisfies $h(c) \geq Cn^{\epsilon}$, we say that c' has the EH-property for s colours. We call $\epsilon(c', s)$ the EH-exponent for c' and s colours if it exists.*

Conjecture 3.2 (Erdős, Hajnal [59]). *Let k, s' be integers with $k, s' \geq 2$. Then for any $s \geq s'$, any edge-colouring c' of K_k with $|c'| = s'$ has the EH-property for s colours.*

On could extend the arguments in [59] from two colours to multiple colours to show that in the above setting $h(n, c') = \Omega(n^{\sqrt{\log n}})$.

Note that c' might not use all colours in $[s]$ and it is not immediately obvious whether a larger number of colours in the edge-colouring of the host clique forces larger homogeneous sets. We show that we can reduce the problem to the case when the number of colours in the edge-colouring of a large clique is the same or one larger than the number of colours in the forbidden pattern c'. The following is joint work with Axenovich and Riasanovsky [12].

Theorem 3.3. *Let c' be an edge-colouring of a clique and s be an integer with $s > |c'|$. Then c' has the EH-property for s colours if and only if c' has the EH-property for $s + 1$ colours.*

Corollary 3.4. *Let c' be an s'-edge-colouring of a clique. Then the EH-conjecture holds for c' if and only if c' has the EH-property for s' and $s' + 1$ colours.*

This chapter is structured as follows. We extend Lemma 0.2 by Alon, Pach, and Solymosi [6] on graph blow-ups to a multicoloured version in Section 3.2. We prove Theorem 3.3 and Corollary 3.4 in Section 3.3. In Section 3.4 we consider two special cases of edge-colourings c of small cliques and forbid them as sub-colourings. We show that for those two cases allowing an additional colour in the c-free edge-colouring does not necessarily yield a larger homogeneous set. We state some concluding remarks and open problems in Section 3.5.

3.2 THE MULTICOLOUR EH-PROPERTY UNDER BLOW-UPS

For an edge-coloured clique H with vertex set $V(H) = \{v_1, \ldots, v_k\}$ and edge-coloured cliques $F_1, \ldots, F_k$, we define the *blow-up* $H(F_1, \ldots, F_k)$ as the edge-coloured clique obtained by taking pairwise vertex-disjoint copies of $F_1, \ldots, F_k$ and colouring the edge between a vertex of the copy of F_i and a vertex of the copy of F_j according to the colour of v_iv_j in H for any edge v_iv_j of H.

The following theorem is a straightforward generalisation of Lemma 0.2 by Alon, Pach, and Solymosi [6] for 2-edge-coloured graphs.

Theorem 3.5. *If edge-coloured cliques $H, F_1, \dots, F_k$ have the EH-property for s colours, so does the blow-up $H(F_1, \dots, F_k)$.*

Proof. Let H, F be edge-coloured cliques having the EH-property for s colours with exponents $\epsilon(H)$ and $\epsilon(F)$, respectively. Let $V(H) = \{v_1, \dots, v_k\}$ and for simplicity write $G = H(F) = H(F, K_1, \dots, K_1)$ for the blow-up of H where we only replaced the vertex v_1. Let the resulting colouring be c. Note that it suffices to prove that $H(F)$ has the EH-property for s colours, since we can just replace one vertex of H by a clique F_i at a time, and each blow-up will have the desired property.

Let s' be the number of colours used in $H(F)$, and for $s \geq s'$ let c be an s-edge-colouring of K_n that is $H(F)$-free for n sufficiently large. Let

$$\delta = \delta(H, F) = \frac{\epsilon(F)\epsilon(H)}{\epsilon(H) + k\epsilon(F)}.$$

We shall show that $h(c) \geq n^\delta$.

Let $m := n^{\delta/\epsilon(H)}$. Assume first that some subset $U \subset V(G)$ of size m contains no copy of H under c. Then, since H has the EH-property for s colours, we know that there is a homogeneous set of size at least $m^{\epsilon(H)} \geq n^\delta$ in U, so we are done. Thus, we can assume that each subset of $V(G)$ of size m contains a copy of H. It implies that the number $\#H$ of copies of H in G satisfies

$$\#H \geq \binom{n}{m}\binom{n - |H|}{m - |H|}^{-1} = \frac{n!}{(n-m)!m!}\frac{(n-m)!(m-k)!}{(n-k)!} = \frac{n!(m-k)!}{m!(n-k)!}. \tag{1}$$

Here, we use the fact that there are $\binom{n}{m}$ ways to choose subsets of size m in $V(G)$, each of them containing a copy of H. If we fix a copy of H, there are $\binom{n-|H|}{m-|H|}$ sets of size m containing that copy.

Consider the set X_H of ordered k-tuples $(v'_1, \dots, v'_k)$ of vertices in G such that they induce a copy of H with vertices $v'_1, \dots, v'_k$ playing the roles of $v_1, \dots, v_k$, respectively. Consider the set X_{H_0} of ordered $(k-1)$-tuples $(v'_2, \dots, v'_k)$ of vertices in G such that they induce a copy of $H_0 = H[v_2, \dots, v_k]$ with vertices $v'_2, \dots, v'_k$ playing the role of $v_2, \dots, v_k$, respectively, in some copy of H on vertex set $\{v'_1, \dots, v'_k\}$. We have that

$|X_{H_0}| \leq n(n-1)\ldots(n-k+2) = n!/(n-k+1)!$ since we can embed the first vertex in at most n ways, the second in at most $n-1$ ways, and so on. On the other hand, $|X_H| \geq \#H$. Thus, there is an ordered tuple $(v_2', \ldots, v_k')$ of vertices in G and a set W of at least $|X_H|/|X_{H_0}|$ vertices in $V(G) - \{v_2', \ldots, v_k'\}$ such that for any $w \in W$ the vertices $w, v_2', \ldots, v_k'$ induce a copy of H and play the roles of $v_1, v_2, \ldots, v_k$, respectively. In addition, since G is $H(F)$-free, $G[W]$ is F-free. Since F has the EH-property,

$$
\begin{aligned}
h(G[W], c) &\geq |W|^{\epsilon(F)} \\
&\geq \left(\frac{|X_H|}{|X_{H_0}|}\right)^{\epsilon(F)} \\
&\geq \left(\frac{\#H}{n!/(n-k+1)!}\right)^{\epsilon(F)} \\
&\geq \left(\frac{n!(m-k)!}{m!(n-k)!}\frac{(n-k+1)!}{n!}\right)^{\epsilon(F)} \\
&\geq \left(\frac{n-k+1}{m(m-1)\cdots(m-k+1)}\right)^{\epsilon(F)} \\
&\geq \left(\frac{n}{m^k}\right)^{\epsilon(F)} \\
&\geq n^{\epsilon(F)(1-k\delta/\epsilon(H))} \\
&\geq n^{\delta}.
\end{aligned}
$$

This implies that $h(c) \geq h(G[W]) \geq n^{\delta}$. $\square$

We also state a slight variant of the blow-up lemma applied to families of graphs. Let H be an edge-colouring of a clique on vertices $v_1, \ldots, v_k$ and $\mathcal{F}$ be a family of edge-coloured cliques. Then we define the blow-up $H(\mathcal{F})$ to be the family $\{H(F, K_1, \ldots, K_1) : F \in \mathcal{F}\}$. Running almost identical to the previous lemma proof, we have

Theorem 3.6. *If an edge-coloured clique H and a family of edge-coloured cliques $\mathcal{F}$ have the EH-property for s colours, so does the blow-up $H(\mathcal{F})$.*

3.3 Allowing more colours than used in the forbidden pattern

Proof of Theorem 3.3. Let c' be an edge-colouring of a clique, let s be an integer with $s > |c'|$ and let n be a sufficiently large integer.

"⇒": Assume that for any c'-free s-edge-colouring c of K_n we have $h(c) \geq n^{\epsilon}$ for some $\epsilon > 0$. Now let c'' be a c'-free $(s+1)$-edge-colouring of K_n. We want to show that $h(c'') \geq n^{\epsilon}$ as well.

Since $s > |c'|$, there exists a colour $a \in [s]$ which is not used in c'. The colour $s+1$ is also not used in c'. Now recolour all edges of colour $s+1$ in c'' with colour a and call the resulting colouring c'''. Then c''' is an s-edge-colouring of K_n which is c'-free, since the edges having colours from c' are the same in c'' and c'''. Thus, by our assumption there is a homogeneous set X in K_n under c''' of size at least n^ϵ.

In particular, X avoids some colour $a' \in [s]$ under c'''. If $a' \neq a$, then X avoids a' under c''', then X also avoids a' under c''. If $a = a'$, then X avoids a and $s+1$ under c''. Thus, in any case, X is a homogeneous set under c''. In particular, $h(c'') \geq n^\epsilon$.

"$\Leftarrow$": Assume that for any c'-free $(s+1)$-edge-colouring c of K_n we have $h(c) \geq n^{3\epsilon}$ for some $\epsilon > 0$. Now let c'' be a c'-free s-colouring of K_n. We want to show that $h(c'') \geq n^\epsilon$.

We shall construct an $(s+1)$-edge-colouring c''' of K_n as follows: Recolour each edge from c'' with colour $s+1$ with probability $\frac{1}{2}$, and leave the colour from c'' with probability $\frac{1}{2}$. Since the colour $s+1$ is not used in c', the new colouring c''' is c'-free, and thus, by assumption, $h(c''') \geq n^{3\epsilon}$. Now assume $h(c'') < n^\epsilon$. Then under c'' every set $Y \subseteq V(K_n)$ of size $|Y| = n^\epsilon$ contains at least one edge of each colour in $[s]$.

Using the properties of a random graph $G \in \mathcal{G}(n, \frac{1}{2})$, for any $\delta > 0$ and n sufficiently large, any subset of $n^{2\delta}$ vertices in K_n has an edge of colour $s+1$ under c''' with probability approaching 1 as n grows. On the other hand, we know that in c'' each subset $Y \subseteq V(K_n)$ of size $|Y| = n^\epsilon$ induces an edge of colour i for each $i \in [s]$. Thus, using Turán's theorem [122], a given subset $X \subseteq V(K_n)$ of size $|X| = n^{2\epsilon}$ induces at least $x = \Omega\left(\binom{n^{2\epsilon}}{2}\frac{1}{n^\epsilon}\right) = \Omega(n^{3\epsilon})$ edges of colour i, for any $i \in [s]$, under the colouring c''.

The probability that all these edges of colour i are recoloured with $s+1$ is at most $(1/2)^x$. Thus, the probability that some subset of $n^{2\epsilon}$ vertices misses some colour from $[s]$ under c''' is at most

$$p = \binom{n}{n^{2\epsilon}} s(1/2)^x \leq 2^{n^{2\epsilon}\log n - n^{3\epsilon}}.$$

We see that p approaches zero as n grows. Therefore, with high probability, all subsets of $n^{2\epsilon}$ vertices induce edges of all colours under c'''. Thus, with high probability $h(c''') < n^{2\epsilon}$, a contradiction to our assumption that $h(c''') \geq n^{3\epsilon}$. □

3.4 Special cases

Corollary 3.4 shows that for an integer s and an edge-colouring c of a clique using all colours from $[s]$, one only needs to check whether c has the EH-property for s and for $s+1$ colours. Intuitively one might think that having an extra colour in the host graph allows for larger homogeneous sets. In this section we will consider two special colourings, for $s=2$ and $s=3$ colours, which each have the EH-property for s colours. In both cases we do not know if they still have the EH-property for $s+1$ colours, but in both cases we can show, that the size of a largest homogeneous set does not grow with an extra colour, in the second case it is strictly smaller.

3.4.1 Rainbow triangle and an extra colour

Consider the rainbow triangle, i.e. the 3-edge-colouring c of K_3 in which the edges have colours $1, 2$ and 3. The structure of c-free 3-colourings of cliques is known and is called a *Gallai colouring* [79, 86]. It is known that c has the EH-property for 3 colours, see for example [70], see also Theorem 4.7. In particular, we have

$$h_2(n,c) \in \Theta\left(n^{1/3}\log^2 n\right).$$

Next, we shall give a construction, providing an upper bound on the size of largest homogeneous set in any c-free 4-colouring of K_n. We will show that $h_3(n,c) \in O(h_2(n,c))$.

The lexicographic product $c' \times c''$ of two edge-colourings c' of K' and c'' of K'' is the edge-colouring of the blow-up $K'(K'', K'', \ldots, K'')$ as defined in Section 3.2. For a given colouring c' of some clique K and some colours i_1, i_2, i_3, we denote by $S^{c'}_{ijk}$ the size of largest clique in K that only uses colours from $\{i,j,k\}$ under c'. It is easy to see that $S^{c'\times c''}_{ijk} = S^{c'}_{ijk} \cdot S^{c''}_{ijk}$.

Lemma 3.7. *We have* $h_3(n,c) \leq O(n^{1/3}\log^2 n)$.

Proof. Let c_i, $i \in [3]$ be a 3-edge-colouring of $K_{n^{1/3}}$ using colours $[4]\setminus\{i\}$, satisfying $h_2(c_i) \in O(\log n)$. Note that such colourings exist and could be chosen by randomly assigning one of the three colours to each edge uniformly. Also note that c_i is c-free for $i \in [3]$. Let $c_4 = c_1 \times c_2 \times c_3$ be the lexicographic product of c_1, c_2 and c_3, i.e. a 4-edge-colouring of K_n. This is a construction very similar to one used in [70].

Claim: c_4 is c-free and $h_3(c_4) \leq O(n^{1/3}\log^2 n)$.

Consider first $c_2 \times c_3$. Since none of the c_i's contains c, we only need to consider triangles T in c' that have two vertices in some c_3-coloured clique K, and another vertex in a different c_3-coloured clique K'. But since the edges between two different c_3-coloured cliqued all have the same colour, T is not rainbow. Thus, $c_2 \times c_3$ is c-free. Similarly, we conclude that $c_4 = c_1 \times (c_2 \times c_3)$ is c-free.

For the size of a largest homogenous set in c_4 consider the following:

$$\begin{aligned} S^{c_4}_{123} &= S^{c_1}_{123} \cdot S^{c_2}_{123} \cdot S^{c_3}_{123} = O(\log n)O(\log n)O(\log n), \\ S^{c_4}_{124} &= S^{c_1}_{124} \cdot S^{c_2}_{124} \cdot S^{c_3}_{124} = O(\log n)O(\log n)n^{1/3}, \\ S^{c_4}_{134} &= S^{c_1}_{134} \cdot S^{c_2}_{134} \cdot S^{c_3}_{134} = O(\log n)n^{1/3}O(\log n) \quad \text{and} \\ S^{c_4}_{234} &= S^{c_1}_{234} \cdot S^{c_2}_{234} \cdot S^{c_3}_{234} = n^{1/3}O(\log n)O(\log n). \end{aligned}$$

Since $h_3(c_4) = \max\{S^{c_4}_{ijk} : \{i,j,k\} \subseteq [4], |\{i,j,k\}| = 3\}$, we have that $h_3(c_4) \leq O(n^{1/3} \log^2 n)$. □

3.4.2 2-EDGE-COLOURED K_4 AND AN EXTRA COLOUR

Let c be the 2-edge-colouring of K_4 in which each colour class induces P_4. Note that c having the EH-property for 2 colours is equivalent to P_4 having the EH-property. Any P_4-free graph is a *co-graph* (see for example [22, 48] for properties of co-graphs), which is in particular a perfect graph, and thus, by a Theorem of Erdős and Hajnal [59], any P_4-free graph G contains a homogeneous set of size $\sqrt{|V(G)|}$. In particular, we have $h_1(n, c) = n^{1/2}$.

We will show that $h_2(n, c) < h_1(n, c)$.

Lemma 3.8. *There exists a 3-edge-colouring c' of K_n which is c-free which satisfies $h_2(c') \leq O(n^{2/5} \log^{9/5} n)$.*

Proof. By Bohman's [23] upper bound on the Ramsey number $R(4, t) = \Omega(t^{5/2}/\log^2 t)$, we know that for n sufficiently large there exists a graph H on n vertices with $\omega(H) < 4$ and $\alpha(H) < Cn^{2/5} \log^{4/5} n$, for some positive constant C.

Define a 3-edge-colouring of K_n with colours in [3] on the vertex set of H as follows: the edges not in H are coloured 3, and each edge of H is coloured 1 with probability 1/2 and 2 with probability 1/2. Note that in this colouring each K_4 has an edge of colour 3, and therefore there is no copy of c.

Claim: With positive probability $h_2(c') = O(n^{2/5}\log^{9/5} n)$.

Letting $q(n) = 8\alpha(H)\log(n)$, we shall show that any set of $q(n)$ vertices induces edges of all three colours under c'. Let X be a fixed set of $q(n)$ vertices. Then by Turán's theorem [122], any graph G on n vertices with $\alpha(G) < r$ has at least $\frac{1}{r}\binom{n}{2}$ edges. Thus, the number of edges induced by X in H is at least

$$e_X = \frac{1}{\alpha(H)}\binom{q(n)}{2} \geq \frac{q^2(n)}{4\alpha(H)}.$$

Then the probability that X induces only edges of colours 2 and 3 in c' or that X induces only edges of colours 1 and 3 in c' is at most

$$p_X \leq 2\cdot 2^{-e_X} \leq 2\cdot 2^{-q^2(n)/4\alpha(H)}.$$

Using the union bound over all $q(n)$-element subsets of $[n]$, we have that the probability that c' contains a $q(n)$-vertex set inducing edges of only two colours is at most

$$\begin{aligned}\binom{n}{q(n)}p_X &\leq n^{q(n)}2^{1-q^2(n)/4\alpha(H)} = 2^{\left(q(n)\log n + 1 - q^2(n)/4\alpha(H)\right)}\\ &= 2^{\left(8\alpha(H)\log^2(n)+1-16\alpha(H)\log^2(n)\right)}\\ &< 1,\end{aligned}$$

using the definition of $q(n)$. Thus, with positive probability there is a desired colouring. $\square$

We remark that we did not attempt to optimise any of the constants involved.

3.5 CONCLUDING REMARKS

The multicolour Erdős-Hajnal conjecture is concerned with the existence of large homogeneous sets in edge-coloured cliques that do not contain a copy of a given colouring on small subcliques. It could be that the number of colours used in a large clique is strictly larger than the number of colours used in a forbidden subclique-colouring. Although intuitively it seems that having more colours on a large clique allows for larger homogeneous sets, no formal proof of this is known and it is actually not clear whether it is true.

We showed that the multicolour EH-conjecture could be reduced to the situation when the large clique uses the same set of colours as the forbidden colouring or maybe one more. This brings us to the following special cases, in a sense smallest, for which the EH-conjecture is known to be true for the number of colours used in the forbidden colouring, but not any more once additional colours are allowed. In Section 3.4 we provided upper bounds on $h_s(n, c)$ for those two colourings, which show, that the size of a largest homogeneous set in large c-free edge-colourings does not grow when allowing an extra colour, in the second case it even decreases. We still do not know if those two colourings have the EH-property for $s + 1$ colours at all:

Open Problem 3.9. *Does the* 2*-edge colouring of* K_4 *in which each colour class is isomorphic to* P_4 *have the EH-property for* 3 *colours?*

Open Problem 3.10. *Does the rainbow triangle have the EH-property for* 4 *colours?*

Chapter 4 The Erdős-Hajnal conjecture for three colours and families of triangles

4.1 Introduction

In this chapter, we will consider the multicolour version of the Erdős-Hajnal conjecture for $s = 3$ colours and forbidden families of colourings on triangles. We use the definitions and notations from Section 3, but reproduce them here for this special case.

We say that a clique K' edge-coloured with a colouring c contains a *copy of* (or simply contains) a clique K on vertex set $\{1, \ldots, k\}$ with an edge-colouring c', if there is a set of k vertices $\{v_1, \ldots, v_k\}$ in K' and a bijection $\phi : \{1, \ldots, k\} \to \{v_1, \ldots, v_k\}$ such that $c'(ij) = c(\phi(i)\phi(j))$, for all i, j, $1 \leq i < j \leq k$.

In this chapter, to avoid confusion, we will use the term *pattern* for a forbidden colouring of a subgraph, so the term *colouring* will refer to the colouring of a larger graph in which we forbid certain patterns. We consider the case when the number of colours is three and the forbidden patterns are imposed on triangles, but there could be more than one forbidden pattern. Specifically, we investigate all sets of at most three patterns. We provide all our results in the Tables 4.1, 4.2, 4.3. One can immediately see from Table 4.1 that the Erdős-Hajnal conjecture holds true in this setting. We focus on the quantitative version of the conjecture and provide asymptotic bounds on the sizes of the largest 2-edge-coloured cliques.

All of the colourings considered here use colours r, b, and y, corresponding to 'red', 'blue', and 'yellow'. The complete graph on n vertices is denoted by K_n. We call an edge-coloured complete graph K_k using at most two colours on its edges a *two-coloured k-clique*. We also call the set of vertices of a two-coloured clique a *two-coloured set*. For a family $\mathcal{H}$ of patterns using colours r, b, y, we define the class of *$\mathcal{H}$-free edge-colourings* of K_n as a family using colours r, b, y, and containing none of the patterns from $\mathcal{H}$. We denote the family of all $\mathcal{H}$-free edge-colourings by $\mathrm{Forb}(n, \mathcal{H})$.

For an edge-colouring c, let

$$h_2(c) = \max\{k \in \mathbb{N} \mid c \text{ contains a two-coloured } k\text{-clique}\}, \text{ and}$$

$$h_2(n, \mathcal{H}) = \min\{h_2(c) \mid c \in \mathrm{Forb}(n, \mathcal{H})\}.$$

In particular, each edge-colouring of K_n using three colours and not containing patterns from $\mathcal{H}$ contains a clique on $h_2(n, \mathcal{H})$ vertices using at most two colours. In addition, there is an $\mathcal{H}$-free colouring with every clique on more than $h_2(n, \mathcal{H})$ vertices using all three colours.

We consider all sets $\mathcal{H}$ of at most three patterns on triangles using colours from $\{r, b, y\}$. We also write, for example, rrb to represent a colouring of a triangle with one blue and two red edges. These patterns are $rrr, bbb, yyy, rrb, rry, bbr, bby, yyr, yyb, rby$. Note that if for two families $\mathcal{H}, \mathcal{H}'$ there is a permutation π of colours, such that $\mathcal{H}'$ is obtained by applying π to each pattern in $\mathcal{H}$, we have $h_2(n, \mathcal{H}) = h_2(n, \mathcal{H}')$. If for two sets of avoiding patterns, $\mathcal{H}$ and $\mathcal{H}'$, we have that $\mathcal{H} \subseteq \mathcal{H}'$, then $h_2(n, \mathcal{H}) \leq h_2(n, \mathcal{H}')$. Indeed, this holds since any $\mathcal{H}'$-free colouring is also an $\mathcal{H}$-free colouring.

Two of the entries of our tables are expressed in terms of functions $f(n)$ and $g(n)$ that are interesting in their own right. For a graph G, let G^2 be the square of G, i.e. the graph on the same vertex set as G with two vertices adjacent if and only if they are at distance at most 2 in G. Let

$$f(G) = \max\{\alpha(G), \omega(G^2)\} \text{ and } f(n) = \min\{f(G) : |G| = n, \omega(G) = 2\}.$$

Further, recall that the *odd girth*, $\text{girth}_{\text{odd}}(G)$, of a graph G is the length of a shortest odd cycle in G. Let

$$g(n) = \min\{\alpha(G) : |G| = n, \text{girth}_{\text{odd}}(G) \geq 7\}.$$

$\mathcal{H}$	$h_2(n, \mathcal{H})$	Results
$\{\triangle\}$	$\Theta(n^{1/3} \log^2 n)$	[70]
$\{\triangle\}$	$\Theta(\sqrt{n \log n})$	Lemma 4.1.1
$\{\triangle\}$	$\lceil \sqrt{n} \rceil$	Lemma 4.1.2

Table 4.1: Bounds on $h_2(n, \mathcal{H})$ for families $\mathcal{H}$ of one pattern on a triangle

The main results of this chapter are joined work with Axenovich and Snyder [13].

This chapter is structured as follows. We give classical and preliminary results and more definitions in Section 4.2. Section 4.3 contains most of the constructions we use, and hence yields the upper bounds on $h_2(n, \mathcal{H})$ listed in Tables 4.1, 4.2, 4.3. The remaining part of this chapter provides lemmata and their proofs for the corresponding lower bounds on $h_2(n, \mathcal{H})$. Section 4.7 provides final remarks and open questions.

We remark that 'log' for us always denotes the base 2 logarithm.

$\mathcal{H}$	$h_2(n,\mathcal{H})$	Results
$\{\triangle,\triangle\},\{\triangle,\triangle\},\{\triangle,\triangle\},$ $\{\triangle,\triangle\}$	$\lceil\sqrt{n}\rceil$	Lemma 4.2.1
$\{\triangle,\triangle\}$	$\Theta(\sqrt{n})$	Lemma 4.2.2
$\{\triangle,\triangle\},\{\triangle,\triangle\},\{\triangle,\triangle\}$	$\Theta(\sqrt{n\log n})$	Lemmata 4.2.3, 4.2.6
$\{\triangle,\triangle\}$	$\Omega(\sqrt{n\log n}), O(\sqrt{n}\log^{3/2} n)$ $f(n)\le h_2(n,\mathcal{H})\le 2f(n)$	Lemma 4.2.5 Lemma 4.2.4
$\{\triangle,\triangle\}$	$\Omega(n^{2/3}\log^{-3/2} n), O(n^{2/3}\sqrt{\log n})$	Lemma 4.2.7

Table 4.2: Bounds on $h_2(n,\mathcal{H})$ for families $\mathcal{H}$ of two patterns on triangles

4.2 CONNECTIONS TO OTHER RESULTS, PRELIMINARY RESULTS, AND MORE DEFINITIONS

The conclusion of the multicolour Erdős-Hajnal conjecture could be restated as: there is a positive constant $a = a(K)$ such that in any s-edge-colouring of a clique on n vertices with no copy of K there is a colour class with independence number $\Omega(n^a)$. Thus, the multicolour Erdős-Hajnal conjecture not only extends the respective conjecture for graphs, but puts the problem in the framework of Ramsey problems defined through some parameter p, where one seeks a largest clique coloured with a fixed number of colours, such that the parameter p of each colour class is bounded by a given number. For example, the classical Ramsey theorems are stated for the parameter p being equal to the clique number, while the Erdős-Hajnal conjecture has a formulation as a Ramsey number with parameter p, for p equal to the independence number; see other papers [7,42,62,75,78,91,94,105,107], where Ramsey problems with parameter p have been considered for p equal to the diameter, the minimum degree, the connectivity, and the chromatic number. Fox, Grinshpun and Pach [70] addressed the multicoloured Erdős-Hajnal conjecture when K is a rainbow triangle, i.e. a complete graph on 3 vertices edge-coloured using three distinct colours. Among other results, they proved that any such colouring with $s = 3$ colours contains a clique using at most 2 colours that has order at least $\Omega(n^{1/3}\log^2 n)$. Moreover, they showed that this bound is tight.

For a given colouring c in r, b, y, we denote by S^c_{rb}, S^c_{ry}, and S^c_{by} the size of a largest clique using only colours from $\{r, b\}$, only from $\{r, y\}$, and only from $\{b, y\}$, respectively. A pattern is *monochromatic* if only one colour occurs, i.e. rrr, bbb or yyy. For a subset of colours, e.g., $\{r, b\}$, we say that a graph with all edges coloured r or b is a *red/blue graph* (or a *blue/red graph*). If this graph is a clique, we refer to it as a *red/blue clique*.

$\mathcal{H}$	$h_2(n,\mathcal{H})$	Results
$\{\triangle,\triangle,\triangle\},\{\triangle,\triangle,\triangle\}$	$\lceil\sqrt{n}\rceil$	Lemma 4.3.1
$\{\triangle,\triangle,\triangle\},\{\triangle,\triangle,\triangle\}$	$\lceil n/2\rceil$	Lemmata 4.3.2, 4.3.3
$\{\triangle,\triangle,\triangle\}$	$\lceil n/2\rceil+1$	Lemma 4.3.5
$\{\triangle,\triangle,\triangle\},\{\triangle,\triangle,\triangle\}$	$\Omega(\sqrt{n}), O(\sqrt{n\log n})$	Lemmata 4.3.4, 4.3.6
$\{\triangle,\triangle,\triangle\}$	$\Theta(n^{2/3})$	Lemma 4.3.7
$\{\triangle,\triangle,\triangle\}$	no $\mathcal{H}$-free col. for $n\geq 17$	[81]
$\{\triangle,\triangle,\triangle\}$	$\Omega(\sqrt{n\log n})$, $O(\sqrt{n}\log^{3/2} n)$	Lemma 4.3.9
$\{\triangle,\triangle,\triangle\}$	$2\lfloor n/5\rfloor+\epsilon(n)$	Lemma 4.3.10
$\{\triangle,\triangle,\triangle\}$	$\lceil(n-1)/3\rceil\leq h_2(n,\mathcal{H})\leq 2\lceil n/5\rceil$	Lemma 4.3.12
$\{\triangle,\triangle,\triangle\},\{\triangle,\triangle,\triangle\}$, $\{\triangle,\triangle,\triangle\}$	$\lceil n/2\rceil$	Lemmata 4.3.13 4.3.21, 4.3.23
$\{\triangle,\triangle,\triangle\}$	$\Omega(\sqrt{n\log n})$, $O(\sqrt{n}\log^{3/2} n)$	Lemma 4.3.14
$\{\triangle,\triangle,\triangle\}$	$\Omega(\sqrt{n\log n})$, $O(n^{2/3}\sqrt{\log n})$	Lemma 4.3.15
$\{\triangle,\triangle,\triangle\}$	$\Omega(n^{2/3}\log^{1/3} n)$, $O(n^{3/4}\log n)$ $g(n)\leq h_2(n,\mathcal{H})\leq 2g(n)$	Lemma 4.3.18 Lemma 4.3.17
$\{\triangle,\triangle,\triangle\}$	$\Omega(n^{2/3}\log^{-1/3} n)$, $O(n^{2/3}\sqrt{\log n})$	Lemma 4.3.19
$\{\triangle,\triangle,\triangle\}$	$\Theta(\sqrt{n\log n})$	Lemma 4.3.20
$\{\triangle,\triangle,\triangle\}$	$\lceil 3n/7\rceil+\epsilon_1(n)$	Lemma 4.3.22
$\{\triangle,\triangle,\triangle\}$	$\Omega(n^{3/4}\log^{-3/2} n), O(n^{3/4}\sqrt{\log n})$	Lemma 4.3.24
$\{\triangle,\triangle,\triangle\},\{\triangle,\triangle,\triangle\}$, $\{\triangle,\triangle,\triangle\}$	$\Omega(n^{2/3}\log^{-2} n), O(n^{2/3}\sqrt{\log n})$	Lemma 4.3.25

Table 4.3: Bounds on $h_2(n,\mathcal{H})$ for families $\mathcal{H}$ of three patterns on triangles
$\epsilon(n)=0$ if $n\equiv 0 \pmod 5$, $\epsilon(n)=1$ if $n\equiv 1$, and $\epsilon(n)=2$ otherwise;
$\epsilon_1(n)=1$ if $n\equiv 2 \pmod 7$ and $\epsilon_1(n)=0$, otherwise.

If a pattern is not rainbow, the colour used on more than one edge is called the *majority colour*. For any given colour, e.g., red, we say that a vertex v has *red degree* k if it is incident to exactly k red edges. We denote by $N_r(v)$ the *red neighbourhood* of v, the set of all vertices joined to v by edges coloured red.

For positive integers k and ℓ, we let $R(k,\ell)$ denote the usual Ramsey number for cliques, i.e. the smallest N such that no matter how the edge set of K_N is coloured with red and blue, there is a red K_k or a blue K_ℓ in this colouring. For graphs F and F', we let $R(F,F')$ denote the graph Ramsey number of F and F', i.e. the smallest N such that no matter how the edge set of K_N is coloured with red and blue, there is a red copy of F or a blue copy of F' in this colouring. Here a copy of F is a graph isomorphic to F. This definition easily extends to the case when either F or F' is replaced by a finite family of graphs. As usual, we let $V(G), E(G), \chi(G)$, $\alpha(G)$, $\omega(G)$, and $\Delta(G)$ denote the vertex set, edge set, chromatic number, independence number, clique number, and the maximum degree of a graph G, respectively. For a clique with vertex set V, we say that $|V|$ is the *size of the clique*, and we often write "a clique V" instead of "a clique on the vertex set V". Most of the colourings we define in Section 4.3 are obtained by 'blowing up' other colourings. We make this precise as follows. Suppose c is an edge-colouring of K_k on vertex set $[k]$ and let $V_1, \ldots, V_k$ be non-empty, pairwise disjoint sets. The $(V_1, \ldots, V_k)$*-blow-up* of c is the edge-colouring of the complete k-partite graph with parts V_i for $i \in [k]$, such that all edges between V_i and V_j have colour $c(ij)$ for all $1 \leq i < j \leq k$. See also Section 3 for more on blow-ups of colourings.

Observation 4.1. *Let c_1, c_2 be 3-edge-colourings of K_{n_1}, K_{n_2}. Now let $V_1, \ldots, V_{n_1}$ be vertex-disjoint sets of size $|V_i| = n_2$ each and consider the $(V_1, \ldots, V_{n_1})$-blow-up of c_1, where each V_i is coloured according to c_2. Let the resulting colouring be c. Then we have $S_k^c = S_k^{c_1} \cdot S_k^{c_2}$ for any $k \in \{rb, ry, by\}$.*

A graph is *triangle-free* if it does not contain K_3 as a subgraph. We shall need some results about triangle-free graphs and certain Ramsey numbers.

Theorem 4.2 (Kim [93])**.** *For every sufficiently large $n \in \mathbb{N}$ there exists a triangle-free graph G on n vertices with $\alpha(G) \leq 9\sqrt{n \log n}$.*

The following result gives a corresponding lower bound for the independence number of triangle-free graphs:

Theorem 4.3 (Ajtai et al. [1])**.** *For any integer $t \geq 3$, we have $R(3,t) = O(t^2/\log t)$. That is, there is a constant C, such that in any red/blue edge-colouring of K_n with $n = Ct^2/\log t$ there is either a red K_3 or a blue K_t.*

Translating the above theorem into the language of independent sets in triangle-free graphs, we have:

Corollary 4.4. *For any triangle-free graph G on n vertices, $\alpha(G) = \Omega(\sqrt{n \log n})$.*

The following result provides an upper bound on the chromatic number of any n-vertex triangle-free graph.

Theorem 4.5 (Poljak, Tuza [113])**.** *For any triangle-free graph G on n vertices, $\chi(G) \leq 4\sqrt{n/\log n}$.*

We shall also need some known bounds on Ramsey numbers listed in Table 0.1 in the Preliminaries.

Observation 4.6. *A lower bound on some Ramsey number $R(s,t)$ corresponds to the existence of a 2-edge-colouring of K_n with no red K_s and no blue K_t. For example, $R(3,k) = \Omega(k^2/\log k)$ gives the existence of a red/blue edge-colouring of K_n with no red K_3 and no blue K_t for any $t > C\sqrt{n \log n}$ for some constant C.*

Recall that a pattern containing all three colours r, b, and y is called *rainbow*. Colourings not containing rainbow triangles are called *Gallai colourings*. We will need the following fundamental theorem, which asserts that Gallai colourings have a specific structure.

Theorem 4.7 (Gallai [79])**.** *In any Gallai colouring of the complete graph on at least two vertices, the vertex set can be partitioned into at least two non-empty parts such that*

- *for any two distinct parts, all edges between them are of the same colour;*
- *the total number of colours used between parts is at most two.*

4.3 Constructions

Here we will list some explicit or probabilistic 3-edge-colourings of K_n which are $\mathcal{H}$-free for certain families, which we will then refer to in order to prove upper bounds on $h_2(n, \mathcal{H})$. We describe the colourings and bound the size of a largest 2-coloured set for each of them. The constructions are ordered by increasing order of magnitude of h_2. For each constructed colouring c an upper bound on $h_2(c)$ is established. We remark that these upper bounds are asymptotically tight, but since these facts are not needed for our results, we omit proofs. Some of the constructions are explicit and give exact bounds, while others rely on probabilistic/Ramsey results, and therefore only give asymptotic bounds. Lastly, in each construction where we obtain asymptotic upper bounds, divisibility conditions are ignored, and floors and ceilings are omitted for simplicity of presentation.

Construction 2.1 ($\lceil \sqrt{n} \rceil$, none of $\triangle$, $\triangle$, $\triangle$, $\triangle$)

Let $n \geq 3$ be an integer, and assume first that $n = m^2$ for some integer m. Let the vertex set be $\{v_{ij} : i, j \in [m]\}$, a set of m^2 vertices. Define the colouring c of $E(V)$ as follows:

$$c(v_{ik}v_{jl}) = \begin{cases} \text{red} & \text{if } i = j, k \neq l, \\ \text{blue} & \text{if } i \neq j, k = l, \\ \text{yellow} & \text{if } i \neq j, k \neq l. \end{cases}$$

The red graph is the disjoint union of m cliques of size m, where $m \geq 2$. The same holds for the blue graph. It is not difficult to see that this colouring contains none of the stated patterns.

Claim: $h_2(c) \leq \lceil \sqrt{n} \rceil$.

Proof. Let $V_i = \{v_{ik} \mid k \in [m]\}, U_i = \{v_{ji} \mid j \in [m]\}, i \in [m]$. Any blue/yellow clique can use at most one vertex of any given V_i, any red/yellow clique can use at most one vertex of any given U_j, so we have $S^c_{by}, S^c_{ry} \leq m$. In fact, it is not hard to find blue/yellow and red/yellow cliques of size m, so $S^c_{by} = S^c_{ry} = m$. Assume there is a red/blue clique of size $m+1$. Then it uses two vertices $x, y \in V_i$ and one vertex $z \in V_j$ for some $i \neq j$ (here we are using $m \geq 2$). But then z can have a blue edge to at most one of $\{x, y\}$, a contradiction. Thus, we have $S^c_{rb}, S^c_{ry}, S^c_{by} \leq m = \sqrt{n}$. □

Now let $n \in \mathbb{N}$ be arbitrary and take the smallest $m \in \mathbb{N}$, such that $n \leq m^2$. Note that $m = \lceil \sqrt{n} \rceil$. Take the construction described above with m^2 vertices and arbitrarily remove $m^2 - n$ vertices. Then the size of a largest two-coloured clique is still at most $m = \lceil \sqrt{n} \rceil$.

Construction 2.2 ($O(\sqrt{n})$, none of $\triangle$, $\triangle$)

Let $k = \sqrt{n}$. By Bohman's [23] result $R(4, t) = \Omega(t^{5/2} / \log^2 t)$ and observation 4.6, we may consider a blue/yellow edge-colouring of K_k (for sufficiently large n) with no yellow clique of size 4 and no blue clique of size larger than $Ck^{2/5} \log^{4/5} k$, for some constant C. Call this colouring c'. Now, let $V_1, \ldots, V_k$ be pairwise vertex-disjoint sets of size $\sqrt{n}$ each and consider the $(V_1, \ldots, V_k)$-blow-up of c'. Inside each V_i colour the edges in red/yellow with no yellow clique of size 4 and no red clique on more than $Ck^{2/5} \log^{4/5} k$ vertices. Call the resulting colouring c.

Observe that any triangle in this colouring contains edges from either one V_i (in which case it is coloured in red and yellow), from two distinct V_i's (in which case it is a bbr, bby, yyr or yyy triangle), or from three distinct V_i's (then it uses only colours blue and yellow). Thus, there are no rby, rrb triangles.

Claim: $h_2(c) = O(\sqrt{n})$.

Proof. Note that any red/blue clique must be the blow-up of a blue clique in c' with red cliques inside each V_i. By construction, we then have $S^c_{rb} \le (C(\sqrt{n})^{2/5} \log^{4/5}(\sqrt{n}))^2 \le \sqrt{n}$, for sufficiently large n. For red/yellow and blue/yellow cliques we have $S^c_{by}, S^c_{ry} \le 3 \cdot \sqrt{n}$. $\square$

Construction 2.3 ($O(\sqrt{n \log n})$, none of $\triangle, \triangle, \triangle$)

By Theorem 4.2, for n large enough there is a triangle-free graph G on n vertices with $\alpha(G) \le 9\sqrt{n \log n}$. By Theorem 4.5, we have that $\chi(G) \le 4\sqrt{n/\log n}$. Consider a partition of $V(G)$ into $\chi(G)$ independent sets $V_1, \dots, V_{\chi(G)}$ (each of size at most $\alpha(G)$). Now consider the following 3-edge-colouring c of K_n. Fix a copy of G in K_n and colour it red, then colour all edges with both endpoints in the same V_i blue and all remaining edges (between two different V_i's that are not in G) yellow.
Observe that the blue graph is a disjoint union of cliques and the red graph is triangle-free, so there are no rrr, *bbr* or *bby* triangles.

Claim: $h_2(c) = O(\sqrt{n \log n})$.

Proof. A blue/yellow clique in this colouring corresponds to an independent set in G, i.e. we have $S^c_{by} = \alpha(G) \le 9\sqrt{n \log n}$. Since a red/yellow clique contains at most one vertex from any V_i, we have $S^c_{ry} \le \chi(G) \le 4\sqrt{n/\log n}$. A red/blue clique contains vertices from at most two V_i's, since otherwise there is a red triangle in G, i.e. we have $S^c_{rb} \le 2\alpha(G) \le 18\sqrt{n \log n}$. $\square$

Construction 2.4 ($O(\sqrt{n \log n})$, none of $\triangle, \triangle, \triangle$)

Take $k = \sqrt{n \log n}$ and consider a blue/yellow edge-colouring c' of K_k without a monochromatic clique of size more than $2 \log k$, which exists by the bound on the Ramsey number $R(t,t) \ge 2^{t/2}$. Let $V_1, \dots, V_k$ be pairwise vertex-disjoint sets each of size $n/k = \sqrt{n/\log n}$, and consider the $(V_1, \dots, V_k)$-blow-up of c'. Colour every edge with both endpoints in V_i red for each $i \in [k]$. Let us denote by c the resulting colouring.

Observe that each triangle in c is either monochromatic red with all vertices in V_i, $i \in [k]$, a *bbr* or *yyr* triangle (one edge in a red clique and two edges of the same colour to another clique), or one of *bbb*, *bby*, *yyy*, *yyb* (vertices from three different red cliques).

Claim: $h_2(c) = O(\sqrt{n \log n})$.

Proof. Any blue/yellow clique contains at most one vertex from each V_i, i.e. we have $S^c_{by} = k = \sqrt{n \log n}$. Any red/blue (red/yellow) clique in c corresponds to a blue (yellow) clique in c', so we have $S^c_{rb} = S^c_{ry} \leq (n/k) \cdot 2 \log k = \sqrt{n/\log n} \cdot 2 \log(\sqrt{n \log n}) \leq 2\sqrt{n \log n}$. $\square$

CONSTRUCTION 2.5 ($O(\sqrt{n \log n})$, NONE OF $\triangle, \triangle, \triangle$)

Let $k = \sqrt{n \log n}$. Consider a red/blue edge-colouring c' of K_k without a monochromatic clique of size larger than $2 \log k$. Such a colouring exists by the bound on the Ramsey number $R(t,t) \geq 2^{t/2}$. Take $m = n/k$ pairwise vertex-disjoint copies of K_k each coloured according to c' with vertex sets $V_1, \ldots, V_m$. Colour every edge between V_i and V_j yellow for all distinct $i, j \in [m]$. Call the resulting colouring c.
Observe that each triangle in c is either coloured in red and blue (all vertices are in the same V_i), or contains at least two yellow edges (vertices in at least two different V_i's).

Claim: $h_2(c) = O(\sqrt{n \log n})$.

Proof. Any red/blue clique contains vertices from at most one V_i, so $S^c_{rb} = k = \sqrt{n \log n}$. Any red/yellow (blue/yellow) clique contains at most $2 \log k$ vertices from each V_i, so we have $S^c_{ry}, S^c_{by} \leq (n/k) \cdot 2 \log k = \sqrt{n/\log n} \cdot \log(\sqrt{n \log n}) \leq 2\sqrt{n \log n}$. $\square$

CONSTRUCTION 2.6 ($O(\sqrt{n} \log^{3/2} n)$, NONE OF $\triangle, \triangle, \triangle, \triangle$)

From the lower bound on $R(3,t)$, consider a triangle-free graph H with vertex set $[n]$ and independence number at most $9\sqrt{n \log n}$. Define c, a colouring of the edges of a complete graph on vertex set $[n]$ as follows: the edges not in H are coloured yellow, and each edge of H is coloured red with probability $1/2$ or blue with probability $1/2$. Note that in this colouring each triangle has a yellow edge, and therefore contains none of the stated patterns.

Claim: With positive probability $h_2(c) = O(\sqrt{n} \log^{3/2} n)$.

Proof. Letting $q(n) = 80\sqrt{n} \log^{3/2} n$, we shall show that any set of $q(n)$ vertices induces edges of all three colours. Let X be a fixed set of $q(n)$ vertices.

Now we use Turán's theorem [122], which tells us that a graph G on n vertices with $\alpha(G) < r$ has at least $\frac{1}{r}\binom{n}{2}$ edges. Since $\alpha(G[X]) < 10\sqrt{n \log n}$, the number of edges induced by X in H is at least

$$e_X = \frac{1}{10\sqrt{n \log n}}\binom{q(n)}{2} \geq \frac{q^2(n)}{4 \cdot 10\sqrt{n \log n}}.$$

Then the probability that X induces only yellow and blue edges in c or that X induces only yellow and red edges in c is at most

$$p_X \leq 2 \cdot 2^{-e_X} \leq 2 \cdot 2^{-q^2(n)/40\sqrt{n \log n}}.$$

Using the union bound over all $q(n)$-element subsets of $[n]$, we have that the probability that c contains a $q(n)$-vertex set inducing edges of only two colours is at most

$$\binom{n}{q(n)} p_X \leq n^{q(n)} 2^{1-q^2(n)/40\sqrt{n \log n}} = 2^{\left(q(n)\log n + 1 - q^2(n)/40\sqrt{n \log n}\right)} < 1,$$

using the definition of $q(n)$. Thus, there is a desired colouring with positive probability (we remark that we did not attempt to optimize the constants here). $\square$

Construction 2.7 ($O(n^{2/3})$, none of $\triangle, \triangle, \triangle, \triangle$)

Consider $n^{1/3}$ pairwise vertex-disjoint red cliques of size $n^{1/3}$ each and colour all edges in-between blue. This is a red/blue edge-colouring of $K_{n^{2/3}}$. Consider $n^{1/3}$ pairwise vertex-disjoint copies of $K_{n^{2/3}}$ each coloured as above, and colour all remaining edges yellow. Call the resulting colouring c.
Observe that the red graph is a disjoint union of cliques, so any triangle contains either 3, 1 or 0 red edges, if it has its vertices in 1, 2 or 3 different red cliques respectively, so there is no triangle containing exactly two red edges. Assume there are vertices u, v, w, such that $c(uv) = b$ and $c(vw) = y$. Then u and v are in the same red/blue clique and w is in a different red/blue clique, so $c(uw) = y$, and there are no rby, bby triangles.

Claim: $h_2(c) = O(n^{2/3})$.

Proof. It is not difficult to see that $S^c_{rb} = O(n^{2/3})$. Note that any red/yellow clique contains vertices from at most one red clique inside each copy of $K_{n^{2/3}}$. Since each red clique has size $n^{1/3}$ and there are $n^{1/3}$ copies of $K_{n^{2/3}}$, we obtain $S^c_{ry} = O(n^{2/3})$ (see Observation 4.1). Finally, any blue/yellow clique contains at most one vertex from each red clique inside every copy of $K_{n^{2/3}}$. Since there are $n^{1/3}$ red cliques inside each $K_{n^{2/3}}$, we have $S^c_{by} = O(n^{2/3})$. $\square$

Construction 2.8 ($O(n^{2/3}\sqrt{\log n})$, **none of** △, △, △, △)

Let $k = n^{2/3}\sqrt{\log n}$. From the lower bound on $R(3,t)$/Observation 4.6 consider a red/blue edge-colouring c' of K_k with no red K_3 and no blue K_t for any $t > 9\sqrt{k \log k}$. Let $V_1, \dots, V_k$ be pairwise disjoint sets each of size n/k, and consider the $(V_1, \dots, V_k)$-blow-up of c'. Colour all edges with both endpoints in V_i yellow for each $i \in [k]$ and call the resulting colouring c.
Observe that each triangle is either yyy (all vertices in the same V_i), yrr or ybb (one edge within a V_i and two edges to a different V_i) or one of rrb, bbr, bbb (all vertices in different V_i's).

Claim: $h_2(c) = O(n^{2/3}\sqrt{\log n})$.

Proof. Since any red/blue clique can contain at most one vertex from each V_i, we have $S^c_{rb} = k$. Further, as c' contains no red triangle, we obtain $S^c_{ry} \leq 2 \cdot (n/k)$. Lastly, as c' contains no blue K_t for any $t > 9\sqrt{k \log k}$, we have that $S_{by} \leq 9\sqrt{k \log k} \cdot (n/k)$. Thus, by our choice of k it follows that $\max\{S^c_{rb}, S^c_{ry}, S^c_{by}\} \leq Cn^{2/3}\sqrt{\log n}$, for some constant C. $\square$

Construction 2.9 ($O(n^{2/3}\sqrt{\log n})$, **none of** △, △, △, △)

Let $k = n^{1/3}/\sqrt{\log n}$ and consider the trivial edge-colouring c' of K_k where every edge is coloured blue. Let $V_1, \dots, V_k$ be pairwise vertex-disjoint sets of size n/k each and consider the $(V_1, \dots, V_k)$-blow-up of c'. From the lower bound on $R(3,t)$, there is a red/yellow edge-colouring c'' of $K_{n/k}$ with no red K_3 and no yellow K_t for any $t > 9\sqrt{(n/k)\log(n/k)}$ and sufficiently large n. Colour the edges inside V_i according to c'' for all $i \in [k]$, and call the resulting colouring c.
Observe that each triangle is either one of yyy, yyr or rry (all vertices in the same V_i), one of bbr, bby (vertices in two different V_i's) or bbb (vertices in three different V_i's).

Claim: $h_2(c) = O(n^{2/3}\sqrt{\log n})$.

Proof. First, note that any red/yellow clique can contain vertices from at most one of the V_i's. It easily follows that $S^c_{ry} = n/k$. Furthermore, using the bounds on the sizes of red and yellow cliques inside any given V_i, we obtain $S^c_{by} \leq k \cdot 9\sqrt{(n/k)\log(n/k)}$, and $S^c_{rb} \leq 2k$. Thus, our choice of k implies that $\max\{S^c_{rb}, S^c_{ry}, S^c_{by}\} \leq Cn^{2/3}\sqrt{\log n}$, for some constant C. □

Construction 2.10 ($O(n^{2/3}\sqrt{\log n})$, none of $\triangle, \triangle, \triangle, \triangle$)

Let $k = n^{2/3}\sqrt{\log n}$ and consider $m = n/k$ pairwise vertex-disjoint copies of K_k with vertex sets $V_1, \dots, V_m$ where $V_i = \{v_{i1}, \dots, v_{ik}\}$ for each $i = 1, \dots, m$. Let c' be an edge-colouring of K_k in red/yellow with no red K_3 and no yellow clique larger than $9\sqrt{k \log k}$. Colour the cliques induced by V_i according to c' for each $i = 1, \dots, m$. For $i \neq j$ colour the edge $v_{is}v_{jt}$ blue if $s = t$ and yellow otherwise. Call this final colouring c.

Observe that there is no rrr and no rrb triangle since c' contains no red K_3, and any two incident red edges are completely contained in some V_i, which contains no blue edges. Similarly, any two incident blue edges are contained in a blue clique, so there are no *bbr* and no *bby* triangles.

Claim: $h_2(c) = O(n^{2/3}\sqrt{\log n})$.

Proof. Consider the sets $U_t = \{v_{it} : i = 1, \dots, m\}$ for $t = 1, \dots, k$. Observe that any red/yellow clique can contain at most one vertex from each U_t. Therefore, we obtain $S^c_{ry} \leq k$. From our bound on the sizes of largest yellow cliques in each V_i, it follows that $S^c_{by} \leq (n/k) \cdot 9\sqrt{k \log k}$. Lastly, since there is no clique using both red and blue edges, there is no red triangle and each blue clique has size n/k, so we have that $S^c_{rb} = n/k$. Thus, our choice of k implies $\max\{S^c_{rb}, S^c_{ry}, S^c_{by}\} \leq Cn^{2/3}\sqrt{\log n}$ for a constant C. □

Construction 2.11 ($O(n^{3/4}\sqrt{\log n})$, none of $\triangle, \triangle, \triangle, \triangle$)

Let $k = \sqrt{n}$. By the lower bound on $R(3, t)$, take a blue/yellow edge-colouring c' of K_k with no blue K_3 and no yellow clique of size greater than $9\sqrt{k \log k}$. Let $V_1, \dots, V_k$ be pairwise vertex-disjoint sets each of size n/k and consider the $(V_1, \dots, V_k)$-blow-up of c'. Colour each V_i in red/yellow with no red K_3 and no yellow clique of size greater than $9\sqrt{(n/k)\log(n/k)}$. Let the resulting colouring be c.

Note that c contains neither a monochromatic red triangle nor a monochromatic blue triangles, since c' contains no blue K_3 and the colouring inside each V_i contains no red

K_3. Since the V_i's contain no blue edges, there is no rrb triangle, and it is easy to see that there is no rainbow triangle in c.

Claim: $h_2(c) = O(n^{3/4}\sqrt{\log n})$.

Proof. As there is no blue K_3 in c', any red/blue clique contains vertices from at most two of the V_i's, and since there is no red K_3 inside the $V_i's$, we see that $S^c_{rb} \leq 4$. Moreover, since yellow cliques in each V_i have size at most $9\sqrt{(n/k)\log(n/k)}$ and the edges between any two distinct V_i and V_j are coloured in blue/yellow, we obtain $S^c_{by} \leq k \cdot 9\sqrt{(n/k)\log(n/k)}$. Finally, since any yellow clique in c' has size at most $9\sqrt{k \log k}$, it follows that $S^c_{ry} \leq 9\sqrt{k\log k} \cdot (n/k)$. By our choice of k we have $\max\{S^c_{rb}, S^c_{by}, S^c_{ry}\} \leq Cn^{3/4}\sqrt{\log n}$, for some constant C. □

CONSTRUCTION 2.12 ($2\left\lfloor \frac{n}{5} \right\rfloor + \epsilon$, **NONE OF** $\triangle, \triangle, \triangle, \triangle$)

Consider the red/blue colouring c' of K_5 with no monochromatic triangle. Let $n \geq 5$ be an integer, take $\left\lceil \frac{n}{5} \right\rceil$ pairwise vertex-disjoint copies of K_5 coloured according to c', and delete some vertices from one of these copies to make sure that the total number of vertices is n. Finally, colour all remaining edges between these copies yellow. Denote by c the resulting colouring.
Observe that there are no monochromatic red or blue triangles and that each triangle contains either no yellow edges (if it is contained in a red/blue K_5) or at least two yellow edges (if it contains vertices of at least two distinct red/blue K_5's.)

Claim: $h_2(c) \leq 2\left\lfloor \frac{n}{5} \right\rfloor + \epsilon$, where $\epsilon = 0$ if $n \equiv 0 \pmod 5$, $\epsilon = 1$ if $n \equiv 1$, and 2 otherwise.

Proof. The largest red/blue clique has size 5, and any red/yellow or blue/yellow clique contains at most two vertices from each of the $\lfloor n/5 \rfloor$ copies of K_5 and at most ϵ vertices from the remaining vertices, so we obtain $S^c_{by}, S^c_{ry} \leq 2\left\lfloor \frac{n}{5} \right\rfloor + \epsilon$. □

CONSTRUCTION 2.13 ($\leq 2\left\lceil \frac{n}{5} \right\rceil$, **NONE OF** $\triangle, \triangle, \triangle, \triangle$)

Consider the red/blue colouring c' of K_5 with no monochromatic triangle. Let $n \geq 5$ be an integer and let $V_1, \ldots, V_5$ be pairwise disjoint sets of sizes $\left\lceil \frac{n}{5} \right\rceil$ or $\left\lfloor \frac{n}{5} \right\rfloor$ such that $\sum_{i=1}^5 |V_i| = n$. Consider the $(V_1, \ldots, V_5)$-blow-up of c' and colour every edge within V_i yellow for $i = 1, \ldots, 5$. Denote by c the resulting colouring.

Observe that there is no monochromatic red or blue triangle in c and that the yellow graph forms a disjoint union of cliques, so there is no triangle with exactly two yellow edges.

Claim: $h_2(c) \leq 2\left\lceil \frac{n}{5} \right\rceil$.

Proof. The largest red/blue clique has size 5, and any red/yellow or blue/yellow clique contains at most two of the parts V_i's, so we obtain $S^c_{by}, S^c_{ry} \leq 2\left\lceil \frac{n}{5} \right\rceil$. □

CONSTRUCTION 2.14 ($\leq \left\lceil \frac{3n}{7} \right\rceil + 1$, NONE OF △, △, △, △)

Consider K_7 with vertex set $\{v_0, \ldots, v_6\}$. Define a red/blue/yellow edge-colouring c' of K_7 as follows. For distinct $i, j \in \{0, \ldots, 6\}$ set

$$c'(v_iv_j) = \begin{cases} b, & \text{if } i-j = \pm 1 \pmod 7, \\ y, & \text{if } i-j = \pm 2 \pmod 7, \\ r, & \text{if } i-j = \pm 3 \pmod 7, \end{cases}$$

see Figure 4.1 for an illustration.

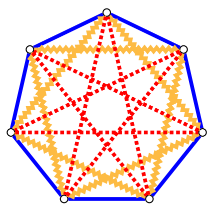


Figure 4.1: $\mathcal{H}$-good colouring of K_7

Note that c' contains no monochromatic blue or red triangles, since the blue and red graph form 7-cycles. Since vertices at distance 2 along the cycle are coloured yellow, c' contains no bbr triangles. Finally, consider triangles containing two incident yellow edges. By symmetry, we may assume that the vertices of this triangle are $0, 2$, and 5. Then since $5-2=3$, the third edge must be red. Thus, there are no yyb triangles in c'.

Now, let $n \geq 7$ be an integer and let $V_0, \ldots, V_6$ be pairwise disjoint vertex sets $V_0, \ldots, V_6$ of sizes $x = \left\lfloor \frac{n}{7} \right\rfloor$ or $w = \left\lceil \frac{n}{7} \right\rceil$ such that $\sum_{i=0}^{6} |V_i| = n$. The parts of sizes x and w are arranged cyclically according to the following orders depending on when n is $0, 1, 2, 3, 4, 5, 6$ modulo 7, respectively: $xxxxxxx$, $wxxxxxx$, $wxwxxxx$, $wwxxwxx$, $wwxxwwx$, $wwwwwxx$, $wwwwwwx$. Consider the $(V_0, \ldots, V_6)$-blow-up of c' and colour every edge within V_i yellow for $i = 0, \ldots, 6$. Call the resulting colouring c, and note that c still does not contain any of the stated patterns.

Claim: $h_2(c) \leq \lceil \frac{3n}{7} \rceil + \epsilon_1(n)$, where $\epsilon_1(n) = 1$ if $n \equiv 2 \pmod 7$ and $\epsilon_1(n) = 0$, otherwise.

Proof. If some clique contains vertices from four different V_i's, then it induces all three colours. Thus, any two-coloured clique contains vertices from at most three different V_i's. To prove the upper bound we may assume n is not divisible by 7. Let us call a partition set V_i *big* if it has size w, and otherwise *small*. Write $n = 7x + r$ for non-negative integers x, r with $1 \leq r \leq 6$. Given our distribution of sizes, it is not difficult to check that any 2-coloured clique contains vertices from at most one big set if $r = 1$, at most two big sets if $2 \leq r \leq 4$, and at most three big sets if $r = 5, 6$. Thus, if $r = 1$ the largest 2-coloured clique has size $2x + w = 3x + 1 = \lceil 3n/7 \rceil$. If $r = 3, 4$, then the largest 2-coloured clique has size $2w + x = 3x + 2 = \lceil 3n/7 \rceil$. If $r = 5, 6$, the largest 2-coloured clique has size $3w = 3x + 3 = \lceil 3n/7 \rceil$. Lastly, suppose $r = 2$. The largest 2-coloured clique has size $2w + x = 3x + 2$. On the other hand, $\lceil 3n/7 \rceil = \lceil 3x + 6/7 \rceil = 3x + 1$. Hence, the largest 2-coloured clique has size $\lceil 3n/7 \rceil + 1$. This completes the proof of the claim. $\square$

CONSTRUCTION 2.15 ($\lceil \frac{n}{2} \rceil$, NONE OF $\triangle, \triangle, \triangle, \triangle, \triangle, \triangle, \triangle$)

Consider the following 3-edge-colouring c of K_n: take disjoint blue cliques V_1, V_2 of sizes $\lfloor \frac{n}{2} \rfloor$ and $\lceil \frac{n}{2} \rceil$, put a maximum red matching in between and colour all other edges yellow.
Observe that each triangle is either monochromatic blue (if it is contained in V_1 or V_2) or one of rby, yyb (if it has w.l.o.g. one vertex in V_1 and two in V_2).

Claim: $h_2(c) \leq \lceil \frac{n}{2} \rceil$.

Proof. The vertex set of any blue/red clique of size at least 3 is contained in either V_1 or V_2, and thus, $S^c_{rb} \leq \lceil \frac{n}{2} \rceil$. Any red/yellow clique contains at most one vertex from each of V_i's, $i = 1, 2$, so we have $S^c_{ry} = 2$. Consider a largest blue/yellow clique X. Then it has x_1 vertices in V_1, each one has a red neighbour in V_2, so X can contain at most $|V_2| - x_1$ vertices in V_2, i.e. we have $S^c_{by} \leq |V_2| = \lceil \frac{n}{2} \rceil$. $\square$

Construction 2.16 ($\lceil \frac{n}{2} \rceil + 1$, **none of** $\triangle, \triangle, \triangle, \triangle, \triangle, \triangle$)

Consider the following 3-edge-colouring c of K_n for $n \geq 3$: Take a red clique V_1 of size $\lceil \frac{n}{2} \rceil$ and vertex-disjoint from it a blue clique V_2 of size $\lfloor \frac{n}{2} \rfloor$ and colour all edges between V_1 and V_2 yellow.
Observe that each triangle is either monochromatic red or blue (if it is contained in V_1 or V_2) or one of yyr, yyb (if it has vertices in both V_1 and V_2).

Claim: $h_2(c) \leq \lceil \frac{n}{2} \rceil + 1$.

Proof. Any red/blue clique contains only vertices from either V_1 or V_2. Any red/yellow clique contains at most one vertex from V_2 and any blue/yellow clique contains at most one vertex from V_1, so we have $\max\{S^c_{rb}, S^c_{ry}, S^c_{by}\} \leq |V_1| + 1$. □

4.4 Forbidding $\mathcal{H}$ with $|\mathcal{H}| = 1$

We start with forbidding only one pattern, i.e. up to swapping colours we only need to consider 3 families $\mathcal{H} \in \{\{\triangle\}, \{\triangle\}, \{\triangle\}\}$. According to Fox et al. [70], if $\mathcal{H} = \{\triangle\}$, then $h_2(n, \mathcal{H}) = \Theta(n^{1/3} \log^2 n)$.

Lemma 4.1.1. *Let* $\mathcal{H} = \{\triangle\}$*. Then we have* $h_2(n, \mathcal{H}) = \Theta(\sqrt{n \log n})$.

Proof. Consider an arbitrary red/blue/yellow edge-colouring c of K_n that has no red triangle. By the upper bound on the Ramsey number $R(3, k)$ (Theorem 4.3) we see that c contains a blue/yellow K_k with $k = \Omega(\sqrt{n \log n})$, so $h_2(n, \{rrr\}) = \Omega(\sqrt{n \log n})$. Moreover, Construction 2.3 shows that $h_2(n, \{rrr\}) = O(\sqrt{n \log n})$. □

Lemma 4.1.2. *Let* $\mathcal{H} = \{\triangle\}$*. Then we have* $h_2(n, \mathcal{H}) = \lceil \sqrt{n} \rceil$.

Proof. For the lower bound, let c be an $\mathcal{H}$-free red/blue/yellow edge-colouring of K_n. Let Δ_r be the maximum red degree in c. If $\Delta_r \geq \lfloor \sqrt{n} \rfloor$, there exists a vertex v which has a red neighbourhood $N_r(v)$ of size at least $\lfloor \sqrt{n} \rfloor$. Note that this neighbourhood does not contain a blue edge, since together with v that would create an rrb triangle. Thus, $N_r(v) \cup \{v\}$ spans a red/yellow clique of size at least $\lfloor \sqrt{n} \rfloor + 1 \geq \lceil \sqrt{n} \rceil$. Otherwise, we have $\Delta_r \leq \lfloor \sqrt{n} \rfloor - 1$ and so the graph induced on red edges may be vertex-coloured with $\lfloor \sqrt{n} \rfloor$ colours. One colour class has size at least $\lceil n / \lfloor \sqrt{n} \rfloor \rceil \geq \lceil \sqrt{n} \rceil$. This forms a blue/yellow clique of size at least $\lceil \sqrt{n} \rceil$ as required. Finally, the upper bound on $h_2(n, \mathcal{H})$ follows from Construction 2.1. □

4.5 Forbidding $\mathcal{H}$ with $|\mathcal{H}| = 2$

Proposition 4.2. *Any family $\mathcal{H}$ consisting of two distinct patterns can be obtained by applying a colour permutation to all patterns in one of the following families:*

- $\{\triangle, \triangle\}, \{\triangle, \triangle\}, \{\triangle, \triangle\}, \{\triangle, \triangle\}$,
- $\{\triangle, \triangle\}$,
- $\{\triangle, \triangle\}, \{\triangle, \triangle\}, \{\triangle, \triangle\}, \{\triangle, \triangle\}$,
- $\{\triangle, \triangle\}$.

Proof. We split the cases according to rainbow and monochromatic patterns:

- $\mathcal{H}$ contains no rainbow and no monochromatic pattern.
  Case 1: the majority colour is the same, i.e. w.l.o.g. $\triangle\ \triangle$.
  Case 2: the majority colour is different, say red and blue.

  Case 2.1: non-majority colours are both not yellow, i.e. $\triangle\ \triangle$.

  Case 2.2: yellow is a non-majority colour in one pattern, i.e. $\triangle\ \triangle$.

  Case 2.3: yellow is a non-majority colour in both patterns, i.e. $\triangle\ \triangle$.
  This gives us $\triangle\ \triangle$, $\triangle\ \triangle$, $\triangle\ \triangle$, $\triangle\ \triangle$.

- $\mathcal{H}$ contains a rainbow but no monochromatic triangle, w.l.o.g. $\triangle\ \triangle$.

- $\mathcal{H}$ contains a monochromatic triangle and no rainbow triangle, w.l.o.g. $\triangle$. Then the second pattern is $\triangle$, $\triangle$, $\triangle$ or $\triangle$.

- $\mathcal{H}$ contains a monochromatic triangle and a rainbow triangle w.l.o.g. $\triangle$, $\triangle$.

This completes the proof. □

4.5.1 $\mathcal{H}$ contains no rainbow and no monochromatic pattern

Lemma 4.2.1. *Let $\mathcal{H} \in \{\{\triangle, \triangle\}, \{\triangle, \triangle\}, \{\triangle, \triangle\}, \{\triangle, \triangle\}\}$. Then we have $h_2(n, \mathcal{H}) = \lceil \sqrt{n} \rceil$.*

Proof. For the lower bound note that we have $\{rrb\} \subseteq \mathcal{H}$ or $\{rry\} \subseteq \mathcal{H}$ for all of these families. Thus, by $h_2(n, \{rrb\}) = h_2(n, \{rry\})$ and Lemma 4.1.2, we obtain $h_2(n, \mathcal{H}) \geq h_2(n, \{rrb\}) = \lceil \sqrt{n} \rceil$. The upper bound follows from Construction 2.1. □

4.5.2 $\mathcal{H}$ contains a rainbow but no monochromatic pattern

Lemma 4.2.2. *Let $\mathcal{H} = \{\triangle, \triangle\}$. Then we have $h_2(n, \mathcal{H}) = \Theta(\sqrt{n})$.*

Proof. The lower bound follows from Lemma 4.1.2: $h_2(n, \mathcal{H}) \geq h_2(n, \{rrb\}) \geq \lceil \sqrt{n} \rceil$.

The upper bound follows from Construction 2.2. $\square$

4.5.3 $\mathcal{H}$ contains a monochromatic but no rainbow pattern

Since our forbidden family contains a monochromatic triangle, by Lemma 4.1.1 we have the lower bound $h_2(n, \mathcal{H}) \geq h_2(n, \{rrr\}) = \Omega(\sqrt{n \log n})$.

Lemma 4.2.3. *Let $\mathcal{H} \in \{\{\triangle, \triangle\}, \{\triangle, \triangle\}\}$. Then we have $h_2(n, \mathcal{H}) = \Theta(\sqrt{n \log n})$.*

Proof. Since $\{rrr\} \subseteq \mathcal{H}$, Lemma 4.1.1 implies that $h_2(n, \mathcal{H}) \geq h_2(n, \{rrr\}) = \Omega(\sqrt{n \log n})$. The upper bound follows from Construction 2.3. $\square$

Recall, that for a graph G, $f(G) = \max\{\alpha(G), \omega(G^2)\}$ and $f(n) = \min\{f(G) : |G| = n,\ \omega(G) = 2\}$. The following lemma shows that determining the value of $f(n)$ is closely linked to determining the value of $h_2(n, \mathcal{H})$, where $\mathcal{H} = \{\triangle, \triangle\}$.

Lemma 4.2.4. *Let $\mathcal{H} = \{\triangle, \triangle\}$. Then $f(n) \leq h_2(n, \mathcal{H}) \leq 2f(n)$.*

Proof. For the upper bound consider a triangle-free graph G on n vertices such that $f(G) = f(n)$, i.e. $\alpha(G) \leq f(n)$ and $\omega(G^2) \leq f(n)$. Colour the edges of G red, colour each edge from $E(G^2) \setminus E(G)$ yellow, and colour all remaining edges blue. We see that there are no red triangles and any two adjacent red edges form an rry triangle. Note that $S_{by} = \alpha(G) \leq f(n)$, $S_{ry}^c = \omega(G^2) \leq f(n)$ and $S_{rb}^c \leq 2\alpha(G) \leq 2f(n)$. Here, the statement on S_{rb}^c holds since in any red/blue clique, the red graph forms a matching.

For the lower bound, consider an arbitrary $\mathcal{H}$-free colouring c of K_n. Let G be the red graph. Then $S_{by} = \alpha(G)$. Since there is no rrb triangle, each triangle containing two red edges is an rry triangle, so $S_{ry}^c \geq \omega(G^2)$. Thus, $h_2(c) \geq \max\{\alpha(G), \omega(G^2)\} \geq f(n)$. $\square$

Lemma 4.2.5. *Let $\mathcal{H} = \{\triangle, \triangle\}$. Then we have*

$$h_2(n, \mathcal{H}) = \Omega(\sqrt{n \log n}) \text{ and } h_2(n, \mathcal{H}) = O(\sqrt{n} \log^{3/2} n).$$

Proof. The lower bound follows from the lower bound on $h_2(n, \{rrr\})$, Lemma 4.1.1. The upper bound follows from Construction 2.6. $\square$

Lemma 4.2.6. *Let* $\mathcal{H} = \{\triangle, \triangle\}$. *Then we have* $h_2(n, \mathcal{H}) = \Theta(\sqrt{n \log n})$.

Proof. The lower bound follows from the lower bound on $h_2(n, \{rrr\})$, Lemma 4.1.1.

The upper bound follows from a result by Guo and Warnke [84], that implies that there are two edge-disjoint triangle free subgraphs G and G' of K_n, each with independence number at most $c\sqrt{n \log n}$. We colour the edges of G, the edges of G' blue and the rest of the edges yellow. Then $S^c_{yr} = \alpha(G') \leq c\sqrt{n \log n}$, $S^c_{yb} = \alpha(G) \leq c\sqrt{n \log n}$, and $S^c_{rb} \leq 6$ since red and blue graphs are triangle-free.

Here, we include a proof of a slightly weaker bound ($h_2(n, \mathcal{H}) = O(\sqrt{n} \log n)$), which is easily obtained by random packing:
We shall find two edge-disjoint triangle-free graphs each with sufficiently small independence number. We shall colour the edges of one of them red, the other one blue, and the rest of the edges yellow. Then the largest bicoloured clique will have size at most the size of the largest independence set of each of these triangle-free graphs.

As we cannot directly guarantee that the desired packing exists, we shall deal with a small overlap. Consider a graph G on a vertex set $[N]$ that is triangle-free, such that $\alpha(G) = O(\sqrt{N \log N})$. In particular, we have that $\Delta(G) = O(\sqrt{N \log N})$ and $|E(G)| = O(N^{3/2}\sqrt{\log N})$. The following claim asserts that we can find a copy G' of G on vertex set $[N]$ such that $|E(G) \cap E(G')|$ is small. A similar proof appears in Konarski and Żak [95], but we include the short proof for convenience of the reader.

Claim 1: *There is a copy G' of G on vertex set $[N]$ such that* $|E(G) \cap E(G')| \leq |E(G)|^2/\binom{N}{2}$.

Proof of Claim 1: Consider a permutation $\pi : [N] \to [N]$ chosen uniformly at random and apply it to G. Let $E_\pi = \{\pi(u)\pi(v) : uv \in E(G)\}$. For each edge $e \in E(G)$ we say that e is *bad* if $e \in E_\pi$, and we let X be the random variable counting bad edges. For each edge $e = uv \in E(G)$ there are $2|E(G)|(N-2)!$ permutations that can make e bad. Hence

$$\mathbb{P}(e \text{ is bad}) = \frac{2|E(G)|(N-2)!}{N!} = \frac{|E(G)|}{\binom{N}{2}}.$$

Thus, $\mathbb{E}[X] = |E(G)|^2/\binom{N}{2}$, and so there is a permutation σ such that

$$|E(G) \cap E_\sigma| \leq \frac{|E(G)|^2}{\binom{N}{2}}. \qquad \blacksquare$$

Consider the union of G with its isomorphic image G' on $[N]$ granted by the above Claim. Since $|E(G)| = O(N^{3/2}\sqrt{\log N})$, we obtain $|E(G) \cap E(G')| \leq |E(G)|^2/\binom{N}{2} = O(N \log N)$. Let G'' be the graph on $[N]$ with edge set $E(G) \cap E(G')$. Then it has at least $N/2$ vertices with degree $O(\log N)$. These vertices induce a graph with independence number at least $\Omega(N/\log N)$, so $\alpha(G'') = \Omega(N/\log N)$. Let X be a largest independent set in G'', and let N be selected such that $n = |X|$. In particular, $n = \Omega(N/\log N)$, and $N = O(n \log n)$. Now, colour the edges of $G[X]$ red, edges of $G'[X]$ blue and the rest yellow. We see that $S_{rb} \leq 5$ since there are no red and no blue triangles. We have that $S_{by} \leq \alpha(G)$ since any blue/yellow clique corresponds to an independent set in G. Thus,

$$S_{by} = O(\sqrt{N \log N}) = O(\sqrt{n \log n \log(n \log n)}) = O(\sqrt{n} \log n),$$

and the lemma follows. $\square$

4.5.4 $\mathcal{H}$ contains a rainbow and a monochromatic pattern

Lemma 4.2.7. *Let $H = \{\triangle, \triangle\}$. Then $h_2(n, \mathcal{H}) = \Omega(n^{2/3}/\log^{3/2} n)$ and $h_2(n, \mathcal{H}) = O(n^{2/3}\sqrt{\log n})$.*

Proof. The upper bound follows from Construction 2.8.

For the lower bound, consider an arbitrary $\mathcal{H}$-free colouring c of K_n. Consider the vertex sets of red components, which we refer to as *blobs*. Note that all edges between any two blobs are of the same colour, either blue or yellow, otherwise there is a rainbow triangle in c. Since the colouring induced by each blob is Gallai, by Theorem 4.7 we have that each blob is a disjoint union of sets which we call *sub-blobs*, so that all edges between any two sub-blobs are of the same colour and the total number of colours between sub-blobs is at most 2. Note that since each blob is a red connected component, one of the colours between sub-blobs must be red and another is blue or yellow. Note also that each sub-blob spans a blue/yellow clique. Otherwise, there is a red triangle in c.

We shall delete some vertices of the graph such that c restricted to the remaining part is easier to analyse. Specifically, we will end up with a colouring c'' of a complete graph on at least $C''n/\log^2 n$ vertices (for some constant C'') in which all blobs contain the same number of sub-blobs and all sub-blobs overall have the same size. In addition, this colouring will have only red and blue edges between sub-blobs of any given blob.

We can assume that each blob has at least two vertices because if there are at least $n/2$ blobs of size 1, they correspond to a blue/yellow clique on at least $n/2$ vertices.

It could be assumed, without loss of generality, that non-red edges between sub-blobs of any given blob are blue. Indeed, either at least $n/4$ vertices are spanned by blobs with red/blue between the sub-blobs or at least $n/4$ vertices are spanned by blobs with red/yellow between sub-blobs. Let us assume the former.

We shall split the sub-blobs according to sizes. Let X_i be the union of all sub-blobs of sizes from 2^i to $2^{i+1} - 1$, $i = 0, \ldots, \log n$. Consider i for which X_i is largest, i.e. $|X_i| \geq \frac{1}{4}n/\log n$. Delete at most half of the vertices from sub-blobs in X_i so that all of them are of the same size, and call the resulting set X_i'. Now, consider a colouring c' that is the restriction of c to X_i'. We see that c' has the same structure as c but with all sub-blobs of the same size and total number of vertices $n' \geq \frac{1}{8}n/\log n$. Let k be the size of each sub-blob. If $k > n^{2/3}$ we are done since each sub-blob spans a blue/yellow clique, and then $S_{by} > n^{2/3}$. Thus, $k < n^{2/3}$. Let Y_j be the union of blobs each having sizes from 2^j to $2^{j+1} - 1$, $j = 0, \ldots, \log n'$. Consider Y_j of largest size so that $|Y_j| \geq n'/\log n'$. By deleting at most half of the vertices in Y_j we can assume that all blobs in Y_j have the same number of vertices, and hence exactly the same number of sub-blobs. Denote the number of sub-blobs by ℓ. Again, by restricting c' to the resulting set, we have an $\mathcal{H}$-free colouring c'' on $n'' \geq C''n/\log^2 n$ vertices (for some constant C'') with each blob having ℓ sub-blobs and each sub-blob having k vertices. Recall that $k < n^{2/3}$.

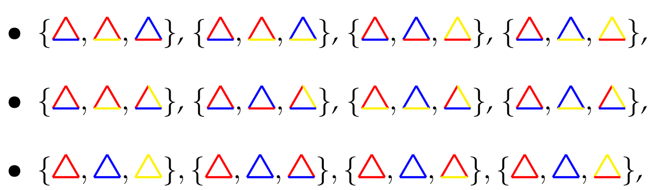
Now we shall analyse c''. Since the red graph is triangle-free, each blob has a blue/yellow clique of order at least $C\sqrt{\ell \log \ell} \cdot k$ for some constant $C > 0$, by Corollary 4.4. Taking a union of these cliques over all blobs, we see that

$$S_{by} \geq C\sqrt{\ell \log \ell} \cdot k \cdot n''/k\ell = C\sqrt{\log \ell} \cdot n''/\sqrt{\ell} \geq CC''\sqrt{\log \ell/\ell} \cdot n/(\log^2 n).$$

Thus, if $\ell < n^{2/3}$, we are done as in this case $S_{by} \geq C'n^{2/3}/\log^{3/2} n$, for some constant C'. Thus, $\ell \geq n^{2/3}$. However, in this case pick a blob and pick a vertex from each sub-blob of this blob. This gives a red/blue clique on $\ell \geq n^{2/3}$ vertices. $\square$

4.6 Forbidding $\mathcal{H}$ with $|\mathcal{H}| = 3$

Proposition 4.3. *Any family $\mathcal{H}$ consisting of 3 distinct patterns can be obtained by applying a colour permutation to all patterns in one of the following families:*



- $\{\triangle, \triangle, \triangle\}, \{\triangle, \triangle, \triangle\}, \{\triangle, \triangle, \triangle\}, \{\triangle, \triangle, \triangle\},$
$\{\triangle, \triangle, \triangle\}, \{\triangle, \triangle, \triangle\}, \{\triangle, \triangle, \triangle\}, \{\triangle, \triangle, \triangle\}, \{\triangle, \triangle, \triangle\},$
- $\{\triangle, \triangle, \triangle\}, \{\triangle, \triangle, \triangle\}, \{\triangle, \triangle, \triangle\}, \{\triangle, \triangle, \triangle\}.$

Proof. We split the cases according to rainbow and monochromatic patterns:

- $\mathcal{H}$ contains no rainbow and no monochromatic triangle.
Case 1: exactly two patterns have the same majority colour, w.l.o.g. $\triangle$ $\triangle$, third pattern has different majority colour, say blue. Then the third pattern is either $\triangle$ or $\triangle$.
Case 2: patterns have different majority colour, w.l.o.g, $rr*, bb*, yy*$.
Then either all non-majority colours are distinct, w.l.o.g. $\triangle$, $\triangle$, $\triangle$, or there are only two different non-majority colours, w.l.o.g $\triangle$, $\triangle$, $\triangle$.

- $\mathcal{H}$ contains a rainbow and a no monochromatic pattern. Then, the other two patterns are listed in the first item of the proof of Proposition 4.2.

- $\mathcal{H}$ contains a monochromatic and no rainbow pattern.
Case 1: There are at least two monochromatic patterns, say we have $\triangle$ $\triangle$. Then all the options for the third pattern up to permutation of patterns are $\triangle$, $\triangle$, $\triangle$, $\triangle$.
Case 2: There is only one monochromatic pattern, say $\triangle$ and two non-monochromatic patterns. We have the following cases:

  Case 2.1: Both non-monochromatic triangles have majority colour red: $\triangle$ $\triangle$ $\triangle$.

  Case 2.2: Exactly one non-monochromatic triangle has majority colour red, w.l.o.g. $\triangle$. Then all the options for the 3rd pattern are $\triangle$, $\triangle$, $\triangle$, $\triangle$.

  Case 2.3: None of the non-monochromatic triangles has majority colour red. Then they either have the same majority colour, w.l.o.g. $\triangle$ $\triangle$, or we have $bb * yy*$.
Then either both non-majority colours are red ($\triangle$ $\triangle$), exactly one non-majority colour is red (w.l.o.g. $\triangle$ $\triangle$) or no non-majority colour is red ($\triangle$ $\triangle$).

- $\mathcal{H}$ contains a rainbow and a monochromatic pattern, w.l.o.g., $\triangle$ and $\triangle$. Then the third pattern either has a red majority colour, or other majority colour, say blue. This gives the following options for the third pattern: $\triangle$, $\triangle$, $\triangle$, $\triangle$.

This completes the proof. □

4.6.1 $\mathcal{H}$ contains no rainbow and no monochromatic pattern

Lemma 4.3.1. *Let* $\mathcal{H} \in \{\{\triangle, \triangle, \triangle\}, \{\triangle, \triangle, \triangle\}\}$. *Then we have* $h_2(n, \mathcal{H}) = \lceil \sqrt{n} \rceil$.

Proof. The lower bound follows from Lemma 4.1.2 since $h_2(n, \mathcal{H}) \geq h_2(n, \{rrb\}) = \lceil \sqrt{n} \rceil$. The upper bound follows from Construction 2.1. $\square$

Lemma 4.3.2. *Let $\mathcal{H} = \{\triangle, \triangle, \triangle\}$. Then we have $h_2(n, \mathcal{H}) = \lceil \frac{n}{2} \rceil$.*

Proof. The upper bound follows from Construction 2.15.

For the lower bound, let c be an $\mathcal{H}$-free edge-colouring of K_n.

Case 1: There is a red triangle in c.
Let the vertex set of a red triangle be $\{u, v, w\}$. Then there cannot be a blue edge adjacent to the triangle. Assume the contrary, i.e. there is a blue edge au. Then av cannot be red or blue, since then uva would induce an rrb or a bbr triangle respectively, i.e. av has to be yellow. The same holds for aw, but then vwa forms a yyr triangle, a contradiction. Assume there is a blue edge xz in the graph. Since each of x and z send only red and yellow edges to $\{u, v, w\}$, and each of x and z send at most one yellow edge to $\{u, v, w\}$, x and z have a common red neighbour in $\{u, v, w\}$, say u. But then uxz is a rrb triangle, a contradiction. Thus, if the graph contains a red triangle, it contains no blue edge and hence, we have a red/yellow clique of size n.

Case 2: c contains no red triangle.
We show that in this case the red graph is bipartite. We need to show that there is no red odd cycle. Assume the contrary, and let $v_1 v_2 \cdots v_k v_1$ ($k \geq 5$) be a shortest red odd cycle. Then we cannot have any red chord of the cycle, since that would create a shorter red odd cycle. Assume there is an index i such that $v_1 v_i$ and $v_1 v_{i+1}$ have the same colour. But then $v_1 v_i v_{i+1}$ create a bbr or a yyr triangle. Also, the edge $v_1 v_3$ has to be yellow, since otherwise $v_1 v_2 v_3$ creates a bbr triangle. Similarly, $v_1 v_{k-1}$ is yellow. But then, combining these two facts we obtain that all edges of the form $v_1 v_i$ with i odd are yellow, including $v_1 v_{k-2}$. Then $v_{k-2} v_{k-1} v_1$ forms a yyr triangle, a contradiction.

Thus, we have no odd red cycle, so the red graph is bipartite. But then in any bipartition there is a bipartite class of size at least $\lceil \frac{n}{2} \rceil$ in which only colours blue and yellow appear. Hence, we have a 2-coloured set of size $\lceil \frac{n}{2} \rceil$. $\square$

Lemma 4.3.3. *Let $\mathcal{H} = \{\triangle, \triangle, \triangle\}$. Then we have $h_2(n, \mathcal{H}) = \lceil \frac{n}{2} \rceil$ for $n \neq 7$ and $h_2(7, \mathcal{H}) =$ 3.*

Proof. The upper bound follows from Construction 2.15.

For the lower bound, let c be an $\mathcal{H}$-free colouring of K_n on a vertex set V. Assume first that there is a monochromatic triangle. Because of symmetry on forbidden patterns, we may assume that there is a red triangle. Let R be a largest red clique in c. Then $|R| \geq 3$. Note that every vertex not in R sends a yellow or a blue edge to R. Then every vertex outside of R sends at most one yellow edge to R (otherwise we have a yyr triangle). In addition, no vertex outside of R sends both red and blue edges to R, otherwise we get an rrb triangle.

Thus, $V - R = O \cup P$, where O is the set of vertices in $V - R$ such that each edge between O and R is yellow or red and P is the set of vertices in $V - R$ such that each edge between P and R is yellow or blue. Note that O and P are disjoint.

Every vertex from O sends $|R| - 1$ red edges to R. Then any two vertices in O have a common red neighbour in R, and hence there cannot be a blue edge induced by O. Similarly, any two vertices in P have a common blue neighbour in R and hence, there cannot be a yellow edge induced by P.

Thus, we have either $|P| \geq \lceil \frac{n}{2} \rceil$ which yields a red/blue clique of size $\lceil \frac{n}{2} \rceil$ or $|R \cup O| \geq \lceil \frac{n}{2} \rceil$, which is a red/yellow clique of desired size.

It remains to deal with the case where we have no monochromatic triangle. In this case, the red neighbourhood of any vertex induces a yellow clique, the blue neighbourhood induces a red clique, and the yellow neighbourhood induces a blue clique, so the maximum degree at each vertex must be at most 6, i.e. we only need to consider colourings of K_n with $n \leq 7$.

For $n = 7$ every vertex must have degree 2 in any colour, so each colour class is a 2-factor. Since there is no monochromatic triangle, each colour class must be a C_7 and up to isomorphism there is a unique such colouring (see also Construction 2.14). One can create such a colouring by ordering the vertices cyclically and colouring an edge with vertices at distance 1, 2, 3 along the cycle yellow, red, and blue respectively. In this colouring the largest 2-coloured clique has size 3.

For $n \leq 6$, observe that there is no red C_5, otherwise all other edges induced by the vertex set of this C_5 are yellow since there are no rrr and no rrb triangles. But then there is a yyr triangle. Thus, the red graph has no odd cycles and so is bipartite. Therefore, the blue/yellow graph contains a clique of size $\lceil \frac{n}{2} \rceil$. $\square$

4.6.2 $\mathcal{H}$ contains a rainbow but no monochromatic pattern

Lemma 4.3.4. *Let $\mathcal{H} = \{\triangle, \triangle, \triangle\}$. Then we have $h_2(n, \mathcal{H}) = \Omega(\sqrt{n})$ and $h_2(n, \mathcal{H}) = O(\sqrt{n \log n})$.*

Proof. By Lemma 4.2.2 we obtain $h_2(n, \mathcal{H}) \geq h_2(n, \{rrb, rby\}) = \Omega(\sqrt{n})$. The upper bound follows from Construction 2.4. $\square$

Lemma 4.3.5. *Let $\mathcal{H} = \{\triangle, \triangle, \triangle\}$. Then we have $h_2(n, \mathcal{H}) = \lceil \frac{n}{2} \rceil + 1$.*

Proof. For the lower bound, let c be an $\mathcal{H}$-free colouring of K_n. Assume there is a vertex v incident to a red edge vx and a blue edge vy. But then the edge xy cannot be coloured. Thus, w.l.o.g at least $\lceil \frac{n}{2} \rceil$ vertices are not incident to a red edge. Taking a maximum set of vertices not incident to a red edge and an arbitrary additional vertex (if it exists) creates a blue/yellow clique, i.e. we have $h_2(n, \mathcal{H}) \geq \lceil \frac{n}{2} \rceil + 1$.
The upper bound follows from Construction 2.16. $\square$

Lemma 4.3.6. *Let $\mathcal{H} = \{\triangle, \triangle, \triangle\}$. Then we have $h_2(n, \mathcal{H}) = \Omega(\sqrt{n})$ and $h_2(n, \mathcal{H}) = O(\sqrt{n \log n})$.*

Proof. By Lemma 4.2.2 we obtain $h_2(n, \mathcal{H}) \geq h_2(n, \{rrb, rby\}) = \Omega(\sqrt{n})$. The upper bound follows from Construction 2.5. $\square$

Lemma 4.3.7. *Let $\mathcal{H} = \{\triangle, \triangle, \triangle\}$. Then we have $h_2(n, \mathcal{H}) = \Theta(n^{2/3})$.*

Proof. For the lower bound consider an $\mathcal{H}$-free edge-colouring c of K_n on vertex set V. Consider a partition of V into sets $A_1, \ldots, A_m$ such that A_1 is maximum sized red clique in c, and for each $i \geq 2$, A_i is a maximum sized red clique in c contained in $V - (A_1 \cup \cdots \cup A_{i-1})$. Note that $|A_i| = 1$ is allowed here. Moreover, for each $i \neq j$ there is at least one non-red edge between A_i and A_j. We shall show that either there is a 2-coloured clique of a desired size or there are at least $n/2$ vertices such that the colouring restricted to these vertices is formed by pairwise vertex-disjoint red cliques such that between any two such cliques all edges are blue or all edges are yellow.

First, assume there is a blue edge uv with $u \in A_i$ and $v \in A_j$, for some $i \neq j$. Then every edge between A_i and A_j incident to u or v must be blue, since otherwise a rainbow or rrb triangle is formed. Assume there is an edge wz with $w \in A_i$ and $z \in A_j$ that is not incident to uv. If wz is red, then uz cannot be coloured without forming a forbidden pattern. Similarly, wz cannot be yellow. It follows that if there is a blue edge between any A_i and A_j, then all edges between A_i and A_j must be blue.

We claim that for each A_i, either A_i sends red/yellow edges to every other A_j, or A_i sends only blue/yellow edges to every other A_j. Suppose otherwise, so that there is $i, j, k \in [m]$ such that all edges between A_i and A_j are blue, and all edges between A_i and A_k are red/yellow and there is at least one red edge. We assume first that $k < i$.

In this case, A_k was chosen as a largest clique before A_i. There must be at least one red and at least one yellow edge between A_k and A_i, so there exist vertices u, v, w such that $v \in A_i$, $u, w \in A_k$, and $c(uv) = r$ and $c(vw) = y$. Pick a vertex $z \in A_j$. Then zv is blue, so we cannot use blue between A_j and A_k: otherwise vzw is a *bby* triangle. Similarly, we cannot colour zw red, because then vzw is a rainbow triangle. It follows that zw is yellow. But then we cannot colour zu without forming an rrb or rainbow triangle. The argument is similar if $i < k$. Indeed, in this case we find vertices u, v, w such that $v \in A_k$, $u, w \in A_i$ and $c(vw) = r$ and $c(vu) = y$. Pick a vertex $z \in A_j$ and note that both uz and uw are blue. In this case, note that if vz is red, then vzw is an rrb triangle. If it is blue, then uvz is a *bby* triangle, and if it is yellow, then vzw is a rainbow triangle. This is a contradiction. Therefore, we say that A_i is of *Type I* if it sends only blue/yellow edges to all other A_j's. Otherwise, we say that A_i is of *Type II*.

Given the above, we can now break the proof into two cases:

Case 1: There are at least $n/2$ vertices in cliques of Type II. In this case, we have a red/yellow clique of size at least $n/2$.

Case 2: At least $n/2$ vertices are in cliques of Type I. Relabel and denote by $V_1, \ldots, V_k$ the red cliques of Type I. Recall that all edges between V_i and V_j, $i \neq j$ must be blue or all of them must be yellow. Then we denote by $c(V_i, V_j)$ the colour of the edges between V_i and V_j.

Note that $k < n^{2/3}$, otherwise we have a blue/yellow clique of that size.
Let $\mathcal{I} \subseteq [k]$ be the subset of indices such that $|V_i| \geq \frac{1}{4}n^{1/3}$ iff $i \in \mathcal{I}$. Split each V_i, $i \in \mathcal{I}$ into disjoint sets $B_{i,j}$ and C_i, $j = 1, \ldots, m_i$, with $|B_{i,j}| = \frac{1}{4}n^{1/3}$ and $|C_i| < \frac{1}{4}n^{1/3}$. Then we have

$$\sum_{i \in \mathcal{I}} |C_i| + \sum_{i \notin \mathcal{I}} |V_i| < \frac{1}{4}n^{1/3}n^{2/3} = \frac{n}{4}.$$

Thus, there are $s > \frac{n/2 - n/4}{n^{1/3}/4} = n^{2/3}$ sets $B_{i,j}$. Consider a blue/yellow edge-colouring c' of K_s with vertex set $\{B_{i,j} : i \in \mathcal{I}, j \in [m_i]\}$ where

$$c'(B_{i,j}, B_{i',j'}) = \begin{cases} c(V_i, V_{i'}), & i \neq i' \\ blue, & i = i'. \end{cases}$$

Then in c', there is no *bby* triangle, so in c' there is a monochromatic set of size at least $\sqrt{s}$. Indeed, the blue graph in c' is a pairwise vertex-disjoint union of cliques, so either one of these cliques has at least $\sqrt{s}$ vertices, or there are at least $\sqrt{s}$ such cliques and thus, there is a yellow clique on $\sqrt{s}$ vertices. Such a blue clique in c' corresponds to a blue/red clique in c, such a yellow clique in c' corresponds to a red/yellow clique in c

of size $\sqrt{s}|B_{i,j}| > \sqrt{n^{2/3}}n^{1/3}/4 = n^{2/3}/4$ in c.

The upper bound follows from Construction 2.7. $\square$

4.6.3 $\mathcal{H}$ contains a monochromatic but no rainbow pattern

Lemma 4.3.8. *Let $\mathcal{H} = \{\textcolor{red}{\triangle}, \textcolor{blue}{\triangle}, \textcolor{yellow}{\triangle}\}$. Then there is no $\mathcal{H}$-free colouring for $n \geq 17$.*

Proof. We know that the Ramsey number $R(3, 3, 3) = 17$, so there is no edge-3-colouring of K_n without a monochromatic K_3 for $n \geq 17$. $\square$

Lemma 4.3.9. *Let $\mathcal{H} = \{\textcolor{red}{\triangle}, \textcolor{blue}{\triangle}, \triangle\}$. Then $h_2(n, \mathcal{H}) = O(\sqrt{n} \log^{3/2} n)$ and $h_2(n, \mathcal{H}) = \Omega(\sqrt{n \log n})$.*

Proof. The lower bound holds since $h_2(n, \mathcal{H}) \geq h_2(n, \{rrr\}) = \Omega(\sqrt{n \log n})$. The upper bound follows from Construction 2.6. $\square$

Lemma 4.3.10. *Let $\mathcal{H} = \{\textcolor{red}{\triangle}, \textcolor{blue}{\triangle}, \triangle\}$. Then $h_2(n, \mathcal{H}) = 2\left\lfloor \frac{n}{5} \right\rfloor + \epsilon$, where $\epsilon = 0$ if $n \equiv 0 \pmod 5$, $\epsilon = 1$ if $n \equiv 1 \pmod 5$, and $\epsilon = 2$ otherwise.*

Proof. To see the lower bound, consider an $\mathcal{H}$-free colouring of $E(K_n)$. Observe that the red degree of every vertex is at most 2. Indeed, since there are no rry or rrr triangles, the entire red neighbourhood of a given vertex must induce a blue clique. But as there is no blue triangle, each red neighbourhood has at most 2 vertices. Thus each red component is either a path or a cycle of length at least 4. Among all such red graphs, the one with the smallest independent set is a union of pairwise vertex-disjoint C_5's, and if n is not divisible by 5, a component on at most 4 vertices. This matches exactly the Construction 2.12 and gives a blue/yellow clique of size at least $2\lfloor n/5 \rfloor + \epsilon$.

The upper bound follows from Construction 2.12. $\square$

Lemma 4.3.11. *If a family $\mathcal{H}$ contains three patterns with different majority colours, then $h_2(n, \mathcal{H}) \geq \left\lceil \frac{n-1}{3} \right\rceil$.*

Proof. Consider an $\mathcal{H}$-free colouring c and an arbitrary vertex v. Let N_r, N_b, and N_y be the red, blue, and yellow neighbourhoods of v, respectively. Then we see that each of these sets induces a 2-coloured clique. Since at least one of the sets N_r, N_b, N_y has size at least $\left\lceil \frac{n-1}{3} \right\rceil$, the result follows. $\square$

Lemma 4.3.12. *Let $\mathcal{H} = \{\textcolor{red}{\triangle}, \textcolor{blue}{\triangle}, \triangle\}$. Then $\left\lceil \frac{n-1}{3} \right\rceil \leq h_2(n, \mathcal{H}) \leq 2\left\lceil \frac{n}{5} \right\rceil$.*

Proof. The lower bound follows from Lemma 4.3.11. The upper bound follows from Construction 2.13. $\square$

Lemma 4.3.13. *Let $\mathcal{H} = \{\triangle, \triangle, \triangle\}$. Then we have $h_2(n, \mathcal{H}) = \lceil \frac{n}{2} \rceil$.*

Proof. For the lower bound, let c be an $\mathcal{H}$-free colouring of K_n. Then there are no two adjacent red edges, so the red graph forms a matching, i.e. there is a blue/yellow clique of size $\lceil \frac{n}{2} \rceil$. The upper bound follows from Construction 2.15. $\square$

Lemma 4.3.14. *Let $\mathcal{H} = \{\triangle, \triangle, \triangle\}$. Then $h_2(n, \mathcal{H}) = \Omega(\sqrt{n \log n})$ and $h_2(n, \mathcal{H}) = O(\sqrt{n} \log^{3/2} n)$.*

Proof. For the lower bound, by Lemma 4.1.1 we have : $h_2(n, \mathcal{H}) \geq h_2(n, \{rrr\}) = \Omega(\sqrt{n \log n})$. The upper bound follows from Construction 2.6. $\square$

Lemma 4.3.15. *Let $\mathcal{H} = \{\triangle, \triangle, \triangle\}$. Then $h_2(n, \mathcal{H}) = \Omega(\sqrt{n \log n})$ and $h_2(n, \mathcal{H}) = O(n^{2/3} \sqrt{\log n})$.*

Proof. For the lower bound, by Lemma 4.1.1 we have : $h_2(n, \mathcal{H}) \geq h_2(n, \{rrr\}) = \Omega(\sqrt{n \log n})$. The upper bound follows from Construction 2.10. $\square$

Although there is a gap between the lower and the upper bound in Lemma 4.3.15, we are able to prove the following lemma concerning the structure of colourings with no patterns in $\{\triangle, \triangle, \triangle\}$.

Lemma 4.3.16. *Let $\mathcal{H} = \{\triangle, \triangle, \triangle\}$ and let c be an $\mathcal{H}$-free colouring of K_n. Then either $h_2(c) = \Omega(n^{2/3} \log^{1/3} n)$ or at least $n/4$ vertices span pairwise vertex-disjoint blue cliques with only red/yellow edges between them, with the red graph forming a matching between any two distinct blue cliques.*

Proof. Let c be an $\mathcal{H}$-free colouring of K_n. Start by partitioning the vertex set into blue/red cliques by greedily picking a maximal red/blue clique at each step. Let $\mathcal{C}$ be the set of these cliques. Note that there is a yellow edge between any two cliques from $\mathcal{C}$. Within each clique the red graph is a matching and the red edges between any two components also form a matching (otherwise we have rrb or rrr triangles). Let $\mathcal{C}'$ be the set of cliques from $\mathcal{C}$ on at least 4 vertices each.

Assume first that at least $n/2$ vertices are spanned by cliques from $\mathcal{C}'$. We shall show first that there is no blue edge between any two cliques from $\mathcal{C}'$. Assume $U, V \in \mathcal{C}'$ with $|U| \geq |V|$ and there is a blue edge $u'v$, $u' \in U$, $v \in V$. If there is a yellow edge $u''v$, for some $u'' \in U$, then $u'u''$ must be red (since otherwise $u'u''v$ is a *bby* triangle). Let $u \in U - \{u', u''\}$. Since red forms a matching within U, we have $c(u''u) = c(u'u) = b$. Then uv cannot be blue ($u''uv$ would induce a *bby* triangle), and it cannot be yellow ($u'vu$ would induce a *bby* triangle), so it must be red. Since this holds for each $u \in U - \{u', u''\}$, we must have $|U| \leq 3$, for v cannot have two red neighbours in U, a contradiction to our assumption that $|U| \geq 4$. Thus, the assumption that there is a yellow edge is wrong, so all edges from v to U are blue or red. But then $U \cup \{v\}$ would form a larger red/blue clique and contradict how we greedily chose $\mathcal{C}$. Thus, between any two cliques from $\mathcal{C}'$ there is no blue edge and red forms a matching between any two cliques from $\mathcal{C}'$. Thus our structural result follows by choosing at least half the vertices from each clique of $\mathcal{C}'$ so that these vertices induce a blue clique.

Now assume that at least $n/2$ vertices are spanned by cliques from $\mathcal{C} - \mathcal{C}'$, i.e. cliques of size at most 3 each. We have that $|\mathcal{C} - \mathcal{C}'| \geq n/6$. Each of the cliques from $\mathcal{C} - \mathcal{C}'$ either spans a blue triangle or not. Let $\mathcal{C}'' \subseteq \mathcal{C} - \mathcal{C}'$ be the set of cliques forming blue triangles. We distinguish the following cases:

Case 1: $|\mathcal{C}''| \geq n/12$. Since there is a yellow edge between any two cliques from $\mathcal{C}$, there can't be a blue edge between any two members of $\mathcal{C}''$, otherwise we create a *bby* triangle. Thus, we can pick one vertex from each member of $\mathcal{C}''$ and have a red/yellow clique of size at least $n/12 \in \Omega(n^{2/3})$.

Case 2: $|\mathcal{C}''| < n/12$, i.e. at least $n/12$ cliques from $\mathcal{C} - \mathcal{C}'$ do not span a blue triangle. Let G be the subgraph spanned by vertices of $\mathcal{C} - (\mathcal{C}' \cup \mathcal{C}'')$ with the inherited colouring. Assume that G contains a blue C_5. Then all edges within this cycle must be red (no blue K_3, no *bby* triangle), so we have a red/blue K_5, which contains an rrb triangle, a contradiction. Thus, the blue subgraph of G is C_5-free, and since $R(C_5, K_k) = O(k^{3/2}/\sqrt{\log k})$ we have a red/yellow clique of a desired size. $\square$

Recall that $g(n)$ is a smallest possible independence number of an n-vertex graph that has no cycles of length 3 and no cycles of length 5, i.e. that has an odd girth at least 7.

Lemma 4.3.17. *We have that* $g(n) \leq h_2(n, \{\triangle, \triangle, \triangle\}) \leq 2g(n)$.

Proof. For the upper bound, let G be an n-vertex graph with odd girth at least 7 and independence number $g(n)$. Colour its edges red, the edges corresponding to pairs of vertices at distance two in G yellow, and all remaining edges blue. Clearly, we have no rrr and no rrb triangles. Assume that there is a yyr triangle. Since vertices of any yellow edge are endpoints of a red path of length 2, we see that a yyr triangle implies the existence of an rrr triangle, or a red cycle of length 5. Note that $S_{yb} = \alpha(G) = g(n)$. Consider a largest yellow/red clique X. We see that since the yellow colour class is induced P_3-free in X, the yellow edges form disjoint cliques in X, and there are at most two of them since there are no red triangles. So, X contains a yellow clique on at least $|X|/2$ vertices, i.e. $|X|/2 \leq \alpha(G) = g(n)$, thus, $S_{ry} = |X| \leq 2g(n)$. Similarly, let Y be a largest red/blue clique. The red edges in it must form a matching, so there is a blue clique of size at least $|Y|/2$, in particular $|Y|/2 \leq \alpha(G) = g(n)$. Thus, $S_{rb} = |Y| \leq 2g(n)$.

For the lower bound, consider an $\{rrr, rrb, yyr\}$-free colouring of a complete graph on n vertices. The red graph G does not have 5-cycles because otherwise all other edges induced by the vertices of that cycle must be yellow, forcing a yyr triangle. Thus, the red graph has odd girth at least 7. We have that $S_{by} = \alpha(G) \geq g(n)$. $\square$

Lemma 4.3.18. *Let $\mathcal{H} = \{\triangle, \triangle, \triangle\}$. Then $h_2(n, \mathcal{H}) = \Omega(n^{2/3} \log^{1/3} n)$ and $h_2(n, \mathcal{H}) = O(n^{3/4} \log n)$.*

Proof. By Lemma 4.3.17 it is sufficient to bound $g(n)$.

By Caro et al. [39], we have $R(C_5, K_t) \leq C\frac{t^{3/2}}{\sqrt{\log t}}$, i.e. any C_5-free graph on n vertices has independence number at least $C'n^{2/3} \log^{1/3} n$. Thus, $g(n) \geq C'n^{2/3} \log^{1/3} n$.

By a result by Spencer [120] we have $R(\{C_3, C_4, C_5\}, K_t) \geq C\,(t/\log t)^{4/3}$ for some positive constant C, i.e. for n sufficiently large there exists a graph G on n vertices with no C_3, C_4, C_5 and $\alpha(G) \leq C'n^{3/4} \log n$. Thus, $g(n) \leq C''n^{3/4} \log n$. $\square$

Lemma 4.3.19. *Let $\mathcal{H} = \{\triangle, \triangle, \triangle\}$. Then we have*

$$h_2(n, \mathcal{H}) = \Omega(n^{2/3} / \log^{1/3} n) \text{ and } \; h_2(n, \mathcal{H}) = O(n^{2/3} \sqrt{\log n}).$$

Proof. For the lower bound, let c be an $\mathcal{H}$-free colouring of K_n, and let the vertex set be V.

We shall argue that either our lower bound holds or there is a subset of at least $n/4$ vertices that is a pairwise disjoint union of red/yellow cliques with only blue edges in between. We shall conclude by showing that such a colouring has a large 2-coloured clique of a desired size.

Consider a maximal blue clique B. Then each vertex in $V - B$ sends either a red or a yellow edge to B. Moreover, each vertex in $V - B$ sends at most one red and at most yellow edge to B. Let V_1 be the set of vertices in $V - B$ that send red edges to B. Then these edges form a family Q of pairwise vertex-disjoint stars with centres in B.

Note that there is no red odd cycle in V_1. Indeed, otherwise this cycle contains vertices from three distinct stars from Q. In particular, there is a vertex v in one star from Q in V_1 that sends red edges in V_1 to two distinct stars from Q, say with centres $w_1, w_2 \in B$. Then w_1v and w_2v are yellow, so vw_1w_2 is a yyb triangle, a contradiction. So, if V_1 has Cn vertices then the red graph induced by V_1 has an independent set of size at least $Cn/2$. This gives $S_{yb} \geq Cn/2$. So, we can assume that $|V_1| < Cn$. We can also assume that $|B| < Cn$. So, let $V_2 = V - (B \cup V_1)$. Thus, $|V_2| > n/2$ (by taking $C = 1/4$).

Now, there are only yellow and blue edges between V_2 and B and moreover the yellow edges among those form pairwise vertex-disjoint stars with centres in B. Let $R_1, \dots, R_m$ be the intersections of the vertex sets of these stars and V_2. In particular, V_2 is the union of the R_is. Note that there are no yellow edges between R_i's and there are no blue edges within R_i's, otherwise we obtain a byy triangle.

There is no vertex $v \in V_2$ that sends a red edge to two different R_i's: assume the contrary, i.e. we have red edges w_1v and w_2v with w_1, w_2 belonging to different R_i's. But then w_1w_2 connects two different R_i's, but can be neither red nor blue without creating a forbidden pattern, a contradiction. Thus, the red graph whose edges have endpoints in different R_i's is bipartite. Let $V_3 \subseteq V_2$ be a larger part of such a bipartition (i.e. $|V_3| \geq n/4$) and let $T_i = R_i \cap V_3$. Then the T_i's are red/yellow and all edges in between are blue.

The remaining part of the proof shows that in any colouring c' of K_n that is formed by pairwise vertex-disjoint red/yellow cliques $T_1, \dots, T_m$ with all edges between them blue has a large bi-coloured clique. This will imply the lower bound $h_2(c) = \Omega(n^{2/3}/(\log^{1/3} n))$. The logarithmic factors here could probably be improved.

We shall split the T_i's according to sizes. Let X_i be the union of all T_j of sizes from 2^i to $2^{i+1} - 1$, where $i = 0, \dots, \log n$. Consider a largest X_i, i.e. $|X_i| \geq |V_3|/\log n \geq n/(4\log n)$. Delete at most half the vertices from each T_j in X_i such that all members T_j of X_i have the same size, say k. Let the resulting set be X'. Let c' be the colouring that results from restricting c to X'. Then c' consists of red/yellow cliques of the same size k and only blue edges in between. In addition, $|X'| \geq n/(8\log n)$. Then we have $S_{ry} \geq k$. Moreover, since the red graph is triangle-free, Corollary 4.4 implies that we may find a yellow clique of size $C\sqrt{k \log k}$ inside each of the red/yellow cliques. Hence,

$$S_{by} \geq C\sqrt{k \log k}\frac{|X'|}{k} \geq C'\frac{n}{\log n}\sqrt{\frac{\log k}{k}}.$$

If $k \leq n^{2/3}/\log^{1/3} n$, then $S_{by} \geq C''n^{2/3}/\log^{1/3} n$. Otherwise, we get a large red/yellow clique, which concludes the proof.

The upper bound follows from Construction 2.9. $\square$

Lemma 4.3.20. *Let $\mathcal{H} = \{\triangle, \triangle, \triangle\}$. Then we have $h_2(n, \mathcal{H}) = \Theta(\sqrt{n \log n})$.*

Proof. For the lower bound by Lemma 4.1.1 we have $h_2(n, \mathcal{H}) \geq h_2(n, \{rrr\}) = \Theta(\sqrt{n \log n})$. The upper bound follows from Construction 2.3. $\square$

Lemma 4.3.21. *Let $\mathcal{H} = \{\triangle, \triangle, \triangle\}$. Then we have $h_2(n, \mathcal{H}) = \lceil \frac{n}{2} \rceil$.*

Proof. For the lower bound, let c be an $\mathcal{H}$-free colouring of K_n. Assume that c contains a red odd cycle. Let $v_1 \cdots v_k v_1$, $k \geq 5$ be a shortest odd red cycle. First let $k \geq 7$. Without loss of generality, we have $c(v_1v_3) = b$. Then $c(v_1v_4) = y$, since otherwise v_1, v_3, v_4 induce a bbr triangle. Thus, we have $c(v_1v_i) = b$ for i odd and $c(v_1v_i) = y$ for i even, and thus, $c(v_1v_{k-1}) = y$. Since $c(v_1v_3) = b$, we also have $c(v_3v_k) = y$, and hence $c(v_3v_{k-1}) = b$. Thus, we must have $c(v_{k-1}v_2) = y$, since otherwise v_2, v_3, v_{k-1} induce a bbr triangle, but then v_1, v_2, v_{k-1} induce a yyr triangle, a contradiction.
If $k = 5$, w.l.o.g. we have $c(v_1v_3) = b$. Then we must have $c(v_1v_4) = y = c(v_3v_5)$, since otherwise v_1, v_3, v_4 or v_1, v_3, v_5 induce a bbr triangle. But then we must have $c(v_2v_4) = b = c(v_2v_5)$, since otherwise v_1, v_2, v_4 or v_2, v_3, v_5 induce a yyr triangle. But then v_2, v_4, v_5 induce a bbr triangle, a contradiction.
Thus, the red graph is bipartite, so c contains a blue/yellow clique of size at least $\lceil \frac{n}{2} \rceil$.

The upper bound follows from Construction 2.15. $\square$

Lemma 4.3.22. *Let $\mathcal{H} = \{\triangle, \triangle, \triangle\}$. Then we have $h_2(n, \mathcal{H}) = \lceil \frac{3n}{7} \rceil + \epsilon_1(n)$, where $\epsilon_1(n) = 1$ if $n \equiv 2 \pmod 7$, and $\epsilon_1(n) = 0$ otherwise.*

Proof. The upper bound follows from Construction 2.14.

For the lower bound, let c be an $\mathcal{H}$-free colouring of a n-vertex graph on a vertex set V. Assume first that there is a blue clique B of size $|B| \geq 3$. Then each vertex not in B sends at most one yellow edge to B (otherwise we have a yyb triangle) and no vertex not in B sends both blue and red edges to B (otherwise we have a bbr triangle). Let A be the set of vertices from $V - B$ that send only blue and yellow edges to B and let $O = V - B - A$, i.e. each vertex from O sends a red edge to B. Then, in particular, each vertex from O sends no blue edges to B, thus, it must send at least two red edges and at most one yellow edge to B. Thus for any two vertices of A there is a vertex in B joined to both of them with blue edges. Similarly, for any two vertices of O there is a vertex in B joined to both of them with red edges. Then A induces no red edges and neither does O. Note that B, A and O span the whole graph; $B \cup A$ induces no red edge and O contains no red edge. Then $S^c_{by} \geq \max\{|B \cup A|, |O|\} \geq \lceil \frac{n}{2} \rceil$.

Thus, we can assume that both the red and blue graph are triangle-free. We shall show that the blue graph has a special structure, in particular it must be a blow-up of a cycle on 7 vertices.

If the blue graph contains no odd cycle, then it is bipartite and hence $S_{ry} \geq \lceil \frac{n}{2} \rceil$. Thus, we can assume that there is a blue odd cycle. The blue graph also does not contain a cycle length 5 since otherwise all other edges spanned by the vertices of this cycle are yellow, producing a yyb triangle. Assume the shortest blue odd cycle has length $k \geq 9$. Let $C = C_k = v_1 \dots v_k v_1$ be a shortest blue odd cycle with $k \geq 9$. Then C has no blue chord, otherwise there is a shorter blue odd cycle. Fix a vertex, v_1, and order chords incident to v_1 as they appear on the cycle, i.e. $v_1v_3, v_1v_4, \dots, v_1v_{k-2}$. Each chord is red or yellow. There are no two consecutive yellow chords, otherwise we have a yyb triangle. We have that v_1v_3 and v_1v_{k-2} are yellow, otherwise there is a *bbr* triangle. In addition v_1v_4 is red, otherwise there are two consecutive yellow chords. Assume that there are two consecutive red chords incident to v_1. Let these chords be, without loss of generality, v_1v_i and v_1v_{i+1}, for $i > 4$. Then v_4v_i and v_4v_{i+1} must be yellow, resulting in a yyb triangle. Thus, the chords incident to v_1 must have alternating colours $yry \dots ry$. However, this is impossible since the number of such chords is even.

Thus, a shortest blue odd cycle has length 7. Let the vertex set of such a 7-cycle be $Y = \{v_1, \dots, v_7\}$. Then $c(v_iv_j) = y$ if v_i and v_j are at distance two on C and $c(v_iv_j) = r$ if v_i and v_j are at distance three on C.

Let $x \in V - Y$. Note that x can send at most 3 yellow edges to Y, otherwise we have a yyb triangle. Assume x sends no blue edge to Y. Then it sends at least four red edges to Y, whose endpoints contain two vertices at distance 3 on C, which creates an rrr triangle, a contradiction. Thus, x sends at least one blue edge to Y. Assume x sends exactly one blue edge to Y, say to v_1. Then xv_2 and xv_7 must be yellow. Then xv_3 and xv_6 must be red. But this implies that xv_3v_6 forms an rrr triangle, a contradiction. Note that x sends at most two blue edges to C, otherwise there is a shorter blue odd cycle. Since x sends at least one, at most two, and not exactly one blue edge to Y, x sends exactly two blue edges to Y. The endpoints of these two edges must be at distance 2 on the cycle, otherwise there is a shorter odd cycle. Without loss of generality, let these endpoints be v_2, v_7. Consider a blue cycle $xv_2v_3 \cdots v_7x$. It must have the same colour structure as C, i.e. in particular x "mimics" v_1, i.e. $c(xv_i) = c(v_1v_i)$ for all $v_i \in Y - \{v_1\}$. Since x was chosen arbitrarily outside of any blue cycle, we have that each vertex in $V - Y$ "mimics" some vertex on C and thus, the colouring contains a spanning blow-up of a colouring c restricted to Y.

More specifically, we have that V is a disjoint union of parts $V_0, \dots, V_6$ such that for any $i, j \in \{0, 1, \dots, 6\}$, with i, j, and any $x \in V_i, z \in V_j$, $c(xz) = b$ if $|i - j| = 1 \pmod 7$, $c(xz) = y$ if $|i - j| = 2 \pmod 7$, $c(xz) = r$ if $|i - j| = 3 \pmod 7$. All the edges induced by each V_i are yellow. That is, we have a colouring with a structure as in Construction 2.14 (where the V_i's have variable sizes).

For each $i = 0, \dots, 6$, let $U_i = V_i \cup V_{i+1} \cup V_{i+2}$ and $W_i = V_i \cup V_{i+2} \cup V_{i+4}$ (indices computed modulo 7). Then each U_i spans a blue/yellow clique and each W_i spans a red/yellow clique. We have that $|W_0| + \cdots + |W_6| = 3n$, thus, there is W_i of size at least $\lceil 3n/7 \rceil$. This proves the lower bound, except when $n \equiv 2 \pmod 7$.

So let $n \equiv 2 \pmod 7$, i.e. $n = 7k + 2$ for an integer k. We shall show that there is a 2-coloured clique of size at least $\lceil 3n/7 \rceil + 1$. In order to do so, we first shall show that the sets V_i's pairwise differ in size by at most 1.

If there is an index i', such that $|W_{i'}| \geq \lceil 3n/7 \rceil + 1$, we are done, so assume $|W_i| \leq \lceil 3n/7 \rceil$ for $i = 0, \dots, 6$. Now assume there is an index i' such that $|W_{i'}| \leq \lfloor 3n/7 \rfloor - 1$. But then we have $\sum_{i=0}^{6} |W_i| \leq \lfloor 3n/7 \rfloor - 1 + 6\lceil 3n/7 \rceil = \lfloor 3(7k+2)/7 \rfloor - 1 + 6\lceil 3(7k+2)/7 \rceil = 21k + 5 < 21k + 6 = 3n$, a contradiction. Thus, $|W_i|$ and $|W_j|$ differ by at most 1 for $0 \leq i < j \leq 6$. Similarly, all U_i's differ in size by at most 1.

Note that $W_i \cap W_{i+2} = V_{i+2} \cup V_{i+4}$ for $i = 0, \dots, 6$ (indices computed modulo 7), so the symmetric difference $W_i \triangle W_{i+2} = V_i \cup V_{i+6}$. By the above observation, that means in particular that $|V_{i-1}|$ and $|V_i|$ differ by at most 1, for each $i = 0, 1, \dots, 6$. Similarly, by considering two consecutive U_i's, we see that $|V_i|$ and $|V_{i+3}|$ differ in size by at most 1. Then it is clear that $||V_0| - |V_2|| \leq 2$. Assume that $|V_0| = t$ and $|V_2| = t + 2$, for some t. Then $|V_1| = t + 1$ and $|V_3| \geq t + 1$. Thus, $|U_1| \geq 3t + 4$ that implies that $3t + 3 \leq |U_6| = |V_6| + t + t + 1$, that in turn implies that $|V_6| \geq t + 2$. This contradicts the fact that $|U_6|$ and $|U_0|$ differ by at most 1. It shows that $||V_i| - |V_{i+2}|| \leq 1$. Together with the fact that $||V_i| - |V_{i+1}|| \leq 1$ and $||V_i| - |V_{i+3}|| \leq 1$, we see that any two set V_i, V_j, $i, j \in \{0, 1, \dots, 6\}$ differ in size by at most 1. Thus, V_i's have sizes either $\lceil n/7 \rceil$ or $\lfloor n/7 \rfloor$. Since $n = 7k + 2$, there are exactly two parts V_i, V_j of sizes $\lceil n/7 \rceil$. No matter how they are located, there is a third part V_g such that $V_i \cup V_j \cup V_g$ is either W_m or U_m for some m. This gives a two-coloured clique on $2\lceil n/7 \rceil + \lfloor n/7 \rfloor = \lceil 3n/7 \rceil + 1$ vertices.

This concludes the proof. $\square$

Lemma 4.3.23. *Let $\mathcal{H} = \{\triangle, \triangle, \triangle\}$. Then we have $h_2(n, \mathcal{H}) = \lceil \frac{n}{2} \rceil$.*

Proof. For the lower bound let c be an $\mathcal{H}$-free colouring. Let v be an arbitrary vertex and denote by N_r, N_b, N_y its red, blue and yellow neighbourhoods, respectively.

Then N_r does not contain a red edge, so it must induce a monochromatic clique that is either blue or yellow. Without loss of generality, we assume it is blue (a symmetric argument deals in the case when it is yellow).

Between N_b and N_y there cannot be a single blue or yellow edge, so between them we have a complete bipartite red graph. Thus, N_b and N_y cannot induce red edges, so N_b induces a monochromatic blue and N_y a monochromatic yellow clique. Now consider the bipartite graph between N_r and N_y. There can be no incident blue and yellow edges and each vertex in N_y sends at most one yellow edge to N_r and each vertex in N_r sends at most one blue edge to N_y. Likewise in the bipartite graph between N_b and N_r, the yellow edges form a matching and no vertex is incident to both blue and yellow.

If one of the sets N_r, N_b, N_y contains at most 1 vertex, then the larger of the other two sets together with v is a 2-coloured a clique of size at least $\lceil n/2 \rceil$, as required.

So we may assume that each of the sets N_r, N_b, N_y contains at least 2 vertices. We consider two cases:

Case 1: There is a vertex $w^* \in N_r$ that sends only yellow edges to N_y.
Every vertex in $N_r^* := N_r - \{w^*\}$ sends only red edges to N_y (otherwise, we obtain either a yyb or bby triangle). This implies that there are no red edges between N_r^* and N_b: if otherwise, we obtain a monochromatic red triangle with vertices in N_r^*, N_b, and N_y, using the fact that the bipartite graph with parts N_b and N_y is entirely red. Moreover, since the yellow edges between N_r^* and N_b form a matching, we have that all edges between N_r^* and N_b must be blue. Now, consider the sets $V_1 = N_r^* \cup N_b \cup \{v\}$ and $V_2 = \{w^*\} \cup N_y \cup \{v\}$. Note that V_1 is a red/blue clique and V_2 is a red/yellow clique. One of them must have size at least $\lceil n/2 \rceil$, completing the proof in this case.

Case 2: No vertex in N_r sends only yellow edges to N_y.
Since no blue and yellow edges are incident in the bipartite graph between N_r and N_y, every vertex in N_r sends at least one red edge to N_y. Now the proof is similar to the previous case: we have that no edges between N_r and N_b are red (otherwise, we obtain a monochromatic red triangle). Hence, as before all edges between N_r and N_b are blue. So $N_r \cup N_b \cup \{v\}$ is a red/blue clique and $\{v\} \cup N_y$ is a yellow clique. One of these two sets has size at least $\lceil n/2 \rceil$, and this completes

the proof of the lower bound.

The upper bound follows from Construction 2.15 with red and yellow swapped. $\square$

4.6.4 $\mathcal{H}$ contains a monochromatic and a rainbow pattern

Lemma 4.3.24. *Let* $\mathcal{H} = \{\triangle, \triangle, \triangle\}$. *Then* $\Omega(n^{3/4}/\log^{3/2} n) = h_2(n, \mathcal{H}) = O(n^{3/4}\sqrt{\log n})$.

Proof. For the lower bound, consider an $\mathcal{H}$-free colouring c of K_n. The structure of c is very similar to the structure of $\{rrr, rby\}$-free colourings. Consider red components and call their vertex sets *blobs*. Assume that each blob has at least three vertices. This assumption can be done since if there are at least $n/2$ vertices spanned by red components on at most two vertices, these vertices contain a blue/yellow clique on at least $n/4$ vertices. Since the colouring is Gallai, Theorem 4.7 implies that each blob is a union of sets (which we refer to as *sub-blobs*) with all edges between any two sub-blobs of the same colour and such that the set of colours between all sub-blobs is either $\{r, b\}$ or $\{r, y\}$. Moreover, each sub-blob sends only red edges to some other sub-blob of its blob. All edges between any two blobs are of the same colour, either blue or yellow, otherwise there is a rainbow triangle. Lastly, there are no red edges contained in any sub-blob, because otherwise we obtain a monochromatic red triangle.

As in the lemma on $\{rrr, rby\}$-free colourings, (Lemma 4.2.7) we can assume that for some constant C'' there is a subset of $n' \geq C''n/\log^2 n$ vertices such that all sub-blobs have the same size, k, and all blobs contain the same number, ℓ, of sub-blobs. Assume first that at least half the vertices are spanned by blobs with blue/red between sub-blobs. Then, since the red and blue graph are both triangle-free, there are at most 5 sub-blobs in each blob, and each sub-blob is blue/yellow. By taking a largest sub-blob in each such blob, we see that $S_{by} \geq n'/10 = \Omega(n/\log^2 n)$.

Now, assume that at least half the vertices are spanned by blobs with red/yellow between sub-blob. All edges between blobs are yellow or blue and all sub-blobs are blue/yellow. As the red graph is triangle-free, applying Corollary 4.4 yields that there are at least $C\sqrt{\ell \log \ell}$ sub-blobs in each blob, such that there are only yellow edges between them. By taking the union of these sets over all blobs we have that

$$S_{by} \geq C\frac{n'}{k\ell}\sqrt{\ell \log \ell} \cdot k = C\frac{n'\sqrt{\log \ell}}{\sqrt{\ell}}.$$

If $\ell < \sqrt{n}$ then $S_{by} = \Omega(n^{3/4}/\log^{3/2} n)$. Otherwise, $\ell \geq \sqrt{n}$. By picking a yellow clique from each sub-blob and selecting a set of blobs with only yellow edges between them (using Corollary 4.4 again, and the fact that the red and blue graphs are both triangle-free), we have

$$S_{ry} \geq C' \cdot \ell\sqrt{k \log k}\sqrt{n'/(k\ell)\log(n'/(k\ell))} = \Omega(\sqrt{\ell n'}) = \Omega(n^{3/4}/\log n).$$

The upper bound follows from Construction 2.11. □

Lemma 4.3.25. *Let $\mathcal{H} \in \{\{\triangle, \triangle, \triangle\}, \{\triangle, \triangle, \triangle\}, \{\triangle, \triangle, \triangle\}\}$. Then $h_2(n, \mathcal{H}) = \Omega(n^{2/3}/\log^{3/2} n)$ and $h_2(n, \mathcal{H}) = O(n^{2/3}\sqrt{\log n})$.*

Proof. The lower bound follows from Lemma 4.2.7: $h_2(n, \mathcal{H}) \geq h_2(n, \{rrr, rby\}) = \Omega(n^{2/3}/\log^{3/2} n)$.

For the upper bound in case $\mathcal{H} \in \{\{rby, rrr, bbr\}, \{rby, rrr, bby\}\}$, we use Construction 2.8 with blue and yellow swapped. For the upper bound in case $\mathcal{H} = \{rby, rrr, rrb\}$ we use Construction 2.9. □

4.7 Concluding remarks

We have determined $h_2(n, \mathcal{H})$ asymptotically up to logarithmic factors for nearly all families $\mathcal{H}$ of at most three patterns. Aside from improving logarithmic factors, there are two major gaps left. First, for the family $\mathcal{H}_0 := \{\triangle, \triangle, \triangle\}$ we were able to show that (see Lemma 4.3.15)

$$\Omega(\sqrt{n \log n}) = h_2(n, \mathcal{H}_0) = O(n^{2/3}\sqrt{\log n}).$$

We believe that the upper bound is the correct answer (up to logarithmic terms).

Our second gap comes from the family $\mathcal{H}_1 := \{\triangle, \triangle, \triangle\}$. We showed that $h_2(n, \mathcal{H}_1)$ is related to the function $g(n)$ defined as the smallest independence number of an n-vertex graph of odd girth at least 7:

$$g(n) = \min\{\alpha(G) : |G| = n \text{ and } G \text{ is } \{C_3, C_5\}\text{-free}\}.$$

In particular, we showed that $g(n) \leq h_2(n, \mathcal{H}_1) \leq 2g(n)$. It follows that good bounds on the Ramsey number $R(\{C_3, C_5\}, K_n)$ translate to good bounds on $h_2(n, \mathcal{H}_1)$. Using known results on the Ramsey numbers $R(C_5, K_n)$ and $R(\{C_3, C_4, C_5\}, K_n)$, by

Lemma 4.3.18 we have

$$\Omega(n^{2/3}\log^{1/3} n) = h_2(n, \mathcal{H}_1) = O(n^{3/4}\log n).$$

A consequence of the work of Bohman and Keevash [24] and Warnke [124] is that the C_5-free process with high probability terminates in a graph whose independence number is $\Theta(n^{3/4}\log^{3/4} n)$. We suspect that the behaviour of the independence number does not change much if one forbids triangles in addition to C_5's. Thus, we conjecture that our upper bound on $h_2(n, \mathcal{H}_1)$ is close to the truth:

Conjecture 4.8.

$$g(n) = \Omega(n^{3/4})$$

and thus,

$$h_2(n, \{\triangle, \triangle, \triangle\}) = \Omega(n^{3/4}).$$

From our reduction to the function $g(n)$, this would follow from the corresponding upper bound on $R(\{C_3, C_5\}, K_n)$. This is likely to be challenging, however, as cycle-complete Ramsey numbers are widely open when the cycle lengths are small and fixed.

Lastly, we decided to stop our investigation at three forbidden patterns. Forbidding more patterns in many cases makes the problem of finding large 2-coloured cliques simple. For this reason (and also for the sake of brevity) we did not pursue this line further. Still, one of course may consider families of forbidden patterns of size four and larger.

Chapter 5 The Erdős-Hajnal conjecture for order-size pairs

5.1 Introduction

In this chapter, we consider families determined by a given set of orders and sizes. Several special cases of this have been extensively studied over the years (see for example Erdős and Hajnal [58]). For $0 \le f \le \binom{m}{r}$, we call an r-graph F on m vertices and f edges an (m, f)-*graph* and we call the pair (m, f) the *order-size pair* for F. Say that H is (m, f)-free if it contains no induced copy of an (m, f)-graph. If $Q = \{(m_1, f_1), \dots, (m_t, f_t)\}$, say that H is Q-free if H is (m_i, f_i)-free for all $i = 1, \dots, t$.

Definition 5.1. *Given $r \ge 2$ and $Q = \{(m, f_1), \dots, (m, f_t)\}$, let $h(n, Q) = h_r(n, Q)$ be the minimum of $h(H)$, taken over all n-vertex Q-free r-graphs H. Say that Q has the* EH*-property if there exists $\epsilon = \epsilon_Q > 0$ such that $h(n, Q) > n^{\epsilon}$.*

For example $h_3(n, \{(4, 0), (4, 2)\}) = k$ means that any n-vertex 3-graph in which any 4 vertices induce 1, 3, or 4 edges has a homogenous set of size k, and there is an r-graph H as above with $h(H) = k$. We may omit the subscript r in the notation $h_r(n, Q)$ if it is obvious from the context. When $Q = \{(m, f)\}$ we use the simpler notation $h(n, m, f)$ instead of $h(n, \{(m, f)\})$. Let us make two simple observations:

$$h_r(n, Q) \le h_r(n, Q') \qquad \text{if} \qquad Q \subseteq Q', \tag{5.1}$$

$$h_r(n, Q) = h_r(n, \overline{Q}) \qquad \text{where} \qquad \overline{Q} = \left\{ \left(m, \binom{m}{r} - f \right) : (m, f) \in Q \right\}. \tag{5.2}$$

Our first result concerns 2-graphs, where we show that forbidding a single order-size pair already guarantees large homogeneous sets.

Proposition 5.2. *For any integers m, f with $m \ge 2$ and $0 \le f \le \binom{m}{2}$ there exists $c > 0$ such that $h_2(n, m, f) > c\, n^{1/(m-1)}$.*

It seems a challenging problem to give good upper bounds on $h_2(n, m, f)$. For example, determining $h_2(n, m, \binom{m}{2})$ is equivalent to determining off-diagonal Ramsey numbers.

Our remaining results are in the hypergraph case $r = 3$ and $m = 4$. We shall be considering sets Q of pairs $(4, i)$ for $i \in \{0, 1, 2, 3, 4\}$. We do not need to consider sets Q that contain both $(4,0)$ and $(4,4)$ because Ramsey's theorem guarantees that for sufficiently large n we cannot avoid both of them. Using complementation (5.2), this leaves us with the following sets:

- $\{(4,0)\}$, $\{(4,1)\}$, $\{(4,2)\}$;
- $\{(4,0),(4,1)\}$, $\{(4,0),(4,2)\}$, $\{(4,0),(4,3)\}$, $\{(4,1),(4,2)\}$, $\{(4,1),(4,3)\}$; and
- $\{(4,0),(4,1),(4,2)\}$, $\{(4,0),(4,1),(4,3)\}$, $\{(4,0),(4,2),(4,3)\}$, $\{(4,1),(4,2),(4,3)\}$.

We address $h(n,Q)$ for each of these choices of Q.

We quickly obtain bounds for the first case using results in Ramsey theory (note again that $h(n,4,f) = h(n,4,4-f)$). Recall that the Ramsey number $R_k(s,t)$ is the minimum n such that every red/blue edge-colouring of the complete n-vertex k-graph yields either a monochromatic red s-clique or a monochromatic blue t-clique. It is known [47] that $2^{ct\log t} \leq R_3(4,t) \leq 2^{c't^2\log t}$. This yields positive constants c and c', such that

$$c'\left(\frac{\log n}{\log\log n}\right)^{1/2} < h_3(n,4,0) < c\frac{\log n}{\log\log n}. \tag{5.3}$$

A more recent result of Fox and He [71] constructs n-vertex 3-graphs with every four vertices spanning at most two edges and independence number at most $c\log n/\log\log n$. Together with (5.1) this yields positive a constant c, such that

$$h_3(n,4,1) \leq h_3(n,\{(4,0),(4,1)\}) < c\frac{\log n}{\log\log n}. \tag{5.4}$$

For the remaining cases when $|Q| = 1$ we obtain bounds using recent results by Fox and He [71] and by Gishboliner and Tomon [80]. Recall that $f(n) = (1+o(1))g(n)$ means that there is a function $e(n)$ such that $\lim_{n\to\infty} e(n) \to 0$ and $f(n) = g(n) + e(n)g(n)$.

Proposition 5.3. *There are positive constants c_1, c_2 such that*

$$h_3(n,4,1) > c_1\left(\frac{\log n}{\log\log n}\right)^{1/2} \tag{5.5}$$

and

$$n^{c_2} < h_3(n,4,2) < (1+o(1))n^{1/2}. \tag{5.6}$$

It is unclear if either bound for $h(n,4,1)$ above represents the correct order of magnitude, but the lower bound certainly seems far off; we leave this as an open problem, see Section 5.4.

Our next results address the case when $|Q| = 2$. For the first case we have constants $c, c' > 0$ such that

$$c\,\frac{\log n}{\log\log n} < h_3(n, \{(4,0),(4,1)\} < c'\,\frac{\log n}{\log\log n}.$$

The lower bound follows (after applying (5.2)) from an old result of Erdős and Hajnal [58]. This is the first instance of a (different) conjecture of Erdős and Hajnal [58] about the growth rate of generalized hypergraph Ramsey numbers that correspond to our setting of $h(n,Q)$, where $Q = \{(m,f),(m,f+1),\ldots,(m,\binom{m}{r})\}$. Recent results of Mubayi and Razborov [109] on this problem determine, for each $m > r \geq 4$, the minimum f such that $h_r(n,Q) < c\log^a n$ for some a and $Q = \{(m,f),\ldots,(m,\binom{m}{r})\}$. When $r = 3$, the minimum f was determined by Conlon, Fox and Sudakov [47] for m being a power of 3 and for growing m, as well as some other values.

For the second case when $|Q| = 2$, we have $h_3(n,\{(4,0),(4,2)\}) > n^c$ as follows immediately from (5.1) and (5.6). However, the value of c obtained from [80] is very small (less than 0.005). We improve this below to $1/5$ and also obtain bounds for the other cases.

Theorem 5.4. *There is a positive constant c_1 such that for $n > 5$*

$$h_3(n,\{(4,0),(4,2)\}) > c_1\, n^{1/5}, \tag{5.7}$$

$$h_3(n,\{(4,0),(4,3)\}) > c_1\, n^{1/3}, \tag{5.8}$$

$$c_1(n\log n)^{1/3} < h_3(n,\{(4,1),(4,2)\}) < (1+o(1))n^{1/2},\ \textit{and} \tag{5.9}$$

$$\frac{1}{2}\log n \leq h_3(n,\{(4,1),(4,3)\}) < 4(\log n)^2.$$

We note the upper bound

$$h_3(n,\{(4,0),(4,2)\}) \leq h_3(n,\{(4,0),(4,1),(4,2)\}) < c\sqrt{n\log n}$$

that we will see below. Apart from this we were not able to obtain non-trivial upper bounds in (5.7) or (5.8). Improving the bounds in (5.7), (5.8) and (5.9) seems to be an interesting open problem; see Section 5.4.

Finally, we consider the case when $|Q| = 3$. If $Q = \{(4,0),(4,1),(4,2)\}$, then a $\overline{Q}$-free 3-graph is a partial Steiner triple system (STS), and it is well known [28, 57, 112] that the minimum independence number of an n-vertex partial STS has order of magnitude $\sqrt{n \log n}$. Thus, $h_3(n,Q)$ has order of magnitude $\sqrt{n \log n}$. If $Q = \{(4,1),(4,2),(4,3)\}$, and $n \geq 4$, then it is a simple exercise to show that any Q-free 4-graph on at least four vertices is a clique or co-clique and therefore $h_3(n,Q) = n$ for $n \geq 4$. The two remaining cases are covered below.

Theorem 5.5. *Let $n \geq 4$. Then $h_3(n,\{(4,0),(4,2),(4,3)\}) = n-1$ and*

$$h_3(n,\{(4,0),(4,1),(4,3)\}) = \begin{cases} \frac{n}{2} & \text{if } n \equiv 0 \text{ (mod 6)}, \\ \lceil \frac{n+1}{2} \rceil & \text{if } n \not\equiv 0 \text{ (mod 6)}. \end{cases}$$

The main results of this chapter are joint work with Axenovich and Mubayi [11].

This chapter is structured as follows. In Section 5.2 we prove Proposition 5.2 and in Section 5.3 we prove our results for triple systems. Section 5.4 provides final remarks and open questions.

5.2 Graphs

In this section we prove Proposition 5.2. For a graph G, let $\omega(G)$ and $\alpha(G)$ denote the size of a largest clique and co-clique, respectively.

Proof of Proposition 5.2. We shall use induction on m with basis $m = 2$. In this case $f \in \{0,1\}$. Note that $h(n,2,0) = h(n,2,1) = n = n^1 = n^{1/(m-1)}$, since forbidden graphs are either a non-edge or an edge. Consider an (m,f)-free graph G on n vertices, $m \geq 3$, and assume that the statement of the proposition holds for smaller values of m. We can also assume that G is not a complete graph, an empty graph, a cycle, or the complement of a cycle, since we are done in these cases. Consider Δ and $\overline{\Delta}$, the maximum degree of G and of the complement $\overline{G}$ of G, respectively. Using Brooks' theorem [29], the chromatic number of G and of $\overline{G}$ is at most Δ and $\overline{\Delta}$, respectively. Thus, $\alpha(G) \geq n/\Delta$ and $\omega(G) \geq n/\overline{\Delta}$. Therefore, we can assume that $\Delta \geq n^{(m-2)/(m-1)}$ and $\overline{\Delta} \geq n^{(m-2)/(m-1)}$, otherwise we are done. Thus, there is a vertex with at least

$n^{(m-2)/(m-1)}$ edges incident to it and there is a vertex with at least $n^{(m-2)/(m-1)}$ non-edges incident to it.

Assume first that $f \leq m-1$. Consider a vertex v with at least $n^{(m-2)/(m-1)}$ non-edges incident to it, i.e. with a set X of vertices each non-adjacent to v, $|X| \geq n^{(m-2)/(m-1)}$. Since G is (m, f)-free, $G[X]$ is $(m-1, f)$-free. Thus, by induction $h(G) \geq h(G[X]) \geq |X|^{1/(m-2)} \geq n^{1/(m-1)}$.

Now assume that $f \geq m$. Consider a vertex v with at least $n^{(m-2)/(m-1)}$ edges incident to it, i.e. with a set X of vertices each adjacent to v, $|X| \geq n^{(m-2)/(m-1)}$. Since G is (m, f)-free, $G[X]$ is $(m-1, f-(m-1))$-free. Thus, by induction $h(G) \geq h(G[X]) \geq |X|^{1/(m-2)} \geq n^{1/(m-1)}$. □

5.3 Triple systems

In this section we prove Proposition 5.3, Theorem 5.4 and Theorem 5.5. We will need the following notions and result for our proofs. For an r-graph H and one of its vertices v, we define the *link graph* of v to be the $(r-1)$-graph $L(v)$ whose vertex set is $V(H) \setminus \{v\}$ and edge set is $\{e \subseteq V(H) \setminus \{v\} : e \cup \{v\} \in E(H)\}$. When denoting edges in 3-graphs, we often shall omit parentheses and commas, for example instead of writing $\{x, y, z\}$ we simply shall write xyz. For a 2-graph G, let $L(G)$ be the 3-graph with vertex set $V(G) \cup \{v\}$, $v \notin V(G)$ and edge set $\{uvw : uw \in E(G)\}$. Finally, when we consider a 3-graph H, the link graph of a vertex $u \in V(H)$ *restricted to a vertex set* S, denoted $L_S(u)$ is a graph on vertex set S and edge set $\{vw : v, w \in S, uvw \in E(H)\}$. A clique on s vertices is denoted K_s.

We shall use the following theorem.

Theorem 5.6 (Fox, He [71], Thm. 1.4)**.** *For all $t, s \geq 3$, any 3-graph on more than $(2t)^{st}$ vertices contains either a co-clique on t vertices or $L(K_s)$.*

5.3.1 Forbidden sets of size 1

Proof of Proposition 5.3.

Case 1: $Q = \{(4, 1)\}$.

To prove the lower bound on $h(n, 4, 1)$, we shall consider the complementary setting and an arbitrary n-vertex $(4, 3)$-free 3-graph H. We shall apply Theorem 5.6 with largest possible $t = s$ such that $(2t)^{st} < n$. In this case $t = s \geq c(\log n / \log\log n)^{1/2}$. If H has a co-clique of size t, then $h(H) \geq t$ and we are done.

Otherwise H contains a subgraph isomorphic to $L = L(K_s)$. Let $V(L) = \{v\} \cup V$, where all edges are incident to v, $v \notin V$. Note that V induces a clique in H, otherwise v and three vertices of V not inducing an edge give a $(4,3)$-subgraph. Thus, $h(H) \geq s - 1$. In each case $h(H) \geq c(\log n/\log\log n)^{1/2}$.

Case 2: $Q = \{(4,2)\}$.

The lower bound on $h(n,4,2)$ follows from a result of Gishboliner and Tomon [80]. The upper bound is obtained by taking an affine plane of order q. More precisely, given a sufficiently large n, choose a prime q such that $n^{1/2} < q \leq n^{1/2} + n^{0.29}$; such q exists by density results about primes (see, e.g., [17]). Let $A(2,q)$ be the affine plane of order q. Let H be the 3-graph whose vertex set is some n-element subset of the point set of $A(2,q)$, and whose edge set is the set of triples that are contained in some line in $A(2,q)$. Let S be a set of four vertices in H. If two lines each contain at least three points in S, then they have two points in common, which is impossible, hence at most one line contains at least three points in S. This means that S induces 0, 1 or 4 edges, and consequently, H is $(4,2)$-free. The largest clique in H is the vertex set of a line, and has size at most q. The largest co-clique in H is a cap set in $A(2,q)$ which is well known to have size at most $q+2$. Hence $h(H) \leq q + 2 < n^{1/2} + n^{0.3}$ for sufficiently large n. □

5.3.2 Forbidden sets of size 2

We will need the following special cases of results of de Caen [51] on the hypergraph Turán problem and of Kostochka, Mubayi, and Verstraëte [97] on independent sets in sparse hypergraphs.

Theorem 5.7 (de Caen [51])**.** *Suppose that $n > k \geq 3$ and H is an n-vertex 3-graph with more than $(1 - \binom{k-1}{2}^{-1})(n^3/6)$ edges. Then H contains a clique of size k.*

Theorem 5.8 (Kostochka, Mubayi, Verstraëte [97])**.** *Suppose that H is an n-vertex 3-graph in which every pair of vertices lies in at most d edges, where $0 < d < n/(\log n)^{27}$. Then H has an independent set of size at least $c\sqrt{(n/d)\log(n/d)}$ where c is an absolute constant.*

Proof of Theorem 5.4.

Case 1: $Q = \{(4,0),(4,2)\}$.

Using complementation, we consider a $\{(4,2),(4,4)\}$-free 3-graph H on n vertices. Assume n is sufficiently large. We shall show that $h(H) \geq Cn^{1/5}$, for some constant $C > 0$. For a vertex v in H, let K be a clique in the link graph $L(v)$ of v. Then K is a co-clique in H, for an edge within K in H together with v

yields a 4-clique in H. We will use this observation repeatedly. Suppose that the complement of H has $(1-\gamma)(n^3/6)$ edges for some $0<\gamma<1$ and $k\geq 3$ is defined via

$$\frac{1}{\binom{k}{2}} \leq \gamma < \frac{1}{\binom{k-1}{2}}.$$

Then by Theorem 5.7, H has a co-clique of size at least k. If $k>n^{1/5}$, then we are done so assume from now that $k<n^{1/5}$. As H has at least $\gamma n^3/6 - n^2/2$ edges, by averaging, H has two vertices v,w whose common neighbourhood S has size at least $\gamma n - 4$. If $L_S(v)$ has an induced C_4, then it induces a 4-clique in H, for otherwise we obtain a $(4,2)$-subgraph in H. Hence $L_S(v)$ has no induced C_4 and by known results (see, e.g. [85]) it has a homogeneous set T of size at least $c|S|^{1/3}$. If T is a clique in $L_S(v)$, then by our observation, T is a co-clique in H. If T is a co-clique in $L_S(v)$, then T is a clique in $L_S(w)$ for otherwise we obtain a $(4,2)$-subgraph in H with v,w and two vertices in T. Again the observation implies that T is a co-clique in H. Hence in both cases T is a co-clique in H and $h(H)\geq |T| \geq c|S|^{1/3} \geq (c/2)(\gamma n)^{1/3}$. Since $k<n^{1/5}$, we have $\gamma > n^{-2/5}$ and $h(H) > (c/2)n^{1/5}$ completing the proof.

Case 2: $Q=\{(4,0),(4,3)\}$.

We shall again consider the complementary case. Suppose that H is a 3-graph on n vertices that is $\{(4,1),(4,4)\}$-free. We will prove that $h(H)\geq n^{1/3}$. Let y be an arbitrary vertex of H and consider the link graph $L(y)$ of y.

Assume that there is an induced $2K_2$ in $L(y)$, i.e. that there is a set X of four vertices inducing exactly two disjoint edges in $L(y)$. Any three vertices in X form an edge in H, otherwise these three vertices and y span exactly one edge in H, a contradiction. Thus, X spans exactly 4 edges in H, a contradiction. Thus, $L(y)$ is $2K_2$-free. In the graph case it is known, that $2K_2$ has the Erdős-Hajnal property, and in particular that any n-vertex graph with no induced $2K_2$ contains a homogeneous set of size $cn^{1/3}$ (see e.g. [85]). Thus, $h(L(y))\geq cn^{1/3}$.

Note that a 3-vertex clique in $L(y)$ is not an edge in H, since otherwise there is a 4-clique in H. Similarly, a 3-vertex co-clique in $L(y)$ is not an edge in H, since otherwise together with y it induces a $(4,1)$-subgraph of H. Thus, any set of vertices that is a clique in $L(y)$ or an independent set in $L(y)$ is an independent set in H. Thus, $h(H)\geq h(L(y))\geq cn^{1/3}$ completing the proof.

Case 3: $Q=\{(4,1),(4,2)\}$.

We now prove $c_1(n\log n)^{1/3} \leq h(n,\{(4,1),(4,2)\}) = h(n,\{(4,3),(4,2)\}) \leq n^{1/2} + cn^{0.3}$. The upper bound follows immediately from the construction used in the upper bound in (5.6) so we turn to the lower bound. Using (5.2), consider an

n-vertex 3-graph H that is $\{(4,2),(4,3)\}$-free where n is sufficiently large. Let u, v be a pair of vertices in H whose common neighbourhood S has maximum size $d > 0$. Given vertices $x, y \in S$, the edges xyu and xyv are both in H else $\{u, v, x, y\}$ induces a $(4,2)$ or $(4,3)$-graph. Next, any three vertices $x, y, z \in S$, must form an edge of H otherwise $\{u, x, y, z\}$ induces a $(4,3)$-graph. Therefore S induces a clique in H of size d. If $d > n^{0.4}$, say, then we are done as $h(H) \geq d$. Recalling that n is large enough, we may assume that $d \leq n^{0.4} < n/(\log n)^{27}$. Now Theorem 5.8 yields a co-clique in H of size at least $c\sqrt{(n/d)\log n}$ for some positive constant c. Consequently, there is a constant c' such that

$$h(H) \geq \max\{d,\, c\sqrt{(n/d)\log n}\} > c'\,(n\log n)^{1/3}.$$

Replacing c' by a possibly smaller constant c_1 yields the result for all $n > 4$.

Note that the set of maximal cliques in any $\{(4,2),(4,3)\}$-free 3-graph H forms a linear (maybe non-uniform) hypergraph $\mathcal{H}$. Thus, determining $h(H)$ amounts to finding $\max\{t, |X|\}$, where t is the size of a largest hyperedge and X is a largest set of vertices in $\mathcal{H}$ with no three in the same hyperedge.

Case 4: $Q = \{(4,1),(4,3)\}$.

Finally, we prove $\frac{1}{2}\log n \leq h(n, \{(4,1),(4,3)\}) \leq 4(\log n)^2$. For the lower bound let H be an n-vertex Q-free 3-graph. Pick a vertex v in H and consider its link graph $L(v)$. Since $R_2(t,t) < 4^{t-1}$ (see Erdős and Szekeres [65]), we see that $L(v)$ has a clique or co-clique K of size at least $\frac{1}{2}\log n$. In the first case, K is a clique in H, else we find a $(4,3)$-subgraph in H, and in the second case, K is a co-clique in H, else we find a $(4,1)$-subgraph in H.

We now turn to the upper bound. Let χ be a red/blue colouring of an n-vertex complete graph on vertex set V in which every monochromatic clique has size at most $2\log n$. Such a colouring exists by the classical result of Erdős [53]. Let H be the 3-graph on vertex set V whose edge set consists of all triples of vertices that induce a triangle with one or three red edges under χ.

Consider four vertices x, y, z, and w of H and assume that xyz is an edge in H. Then the triangle xyz has one or three red edges under χ. Assume that xy is red. We need to treat two cases when xz and yz are blue and when xz and yz are red. In each of these cases, consider the fourth vertex w and possible colours on the edges from w to x, y and z. In each of these cases $\{x, y, z, w\}$ induces exactly two or exactly four edges. Thus, any four vertices of H induce none, two, or four edges. So, H is Q-free. Consider a homogeneous set S in H. If it is a clique, all triangles with vertices in S have exactly one or three red edges under χ. Thus, the graph induced by S is a pairwise vertex-disjoint union of red cliques. Either one of these red cliques has size at least $\sqrt{|S|}$ or, by taking a single vertex from each of these red cliques we see that there is a blue clique of size at least $\sqrt{|S|}$ under χ. Since each monochromatic clique in the colouring χ has size at most $2\log n$, we have that $|S| \leq 4(\log n)^2$. Similarly, if S is an independent set in H, all triangles with vertices in S have exactly one or three blue edges under χ and again we get that $|S| \leq 4(\log n)^2$. $\square$

5.3.3 Forbidden sets of size 3

We will need the following structural characterization of Q-free 3-graphs for $Q = \{(4,1),(4,3),(4,4)\}$.

Theorem 5.9 (Frankl, Füredi [74]). *Let H be an $\{(4,1),(4,3),(4,4)\}$)-free 3-graph. Then H is isomorphic to one of the following 3-graphs:*

1. *A blow-up of the 6 vertex 3-graph H' with vertex set $V(H') = [6]$ and edge set $E(H') = \{123, 124, 345, 346, 561, 562, 135, 146, 236, 245\}$. Here for the blow-up we replace every vertex of H' by an independent set, and whenever we have 3 vertices from three distinct of those sets, they induce an edge if and only if the corresponding vertices in H' do.*
2. *The 3-graph whose vertices are the points of a regular n-gon where 3 vertices span an edge if and only if the corresponding points span a triangle whose interior contains the centre of the n-gon.*

Proof of Theorem 5.5.

Case 1: $Q = \{(4,1),(4,3),(4,4)\}$.

We are to prove that

$$h(n,\{(4,0),(4,1),(4,3)\}) = h(n,Q) = \begin{cases} \frac{n}{2} & \text{if } n \equiv 0 \text{ (mod 6)} \\ \lceil \frac{n+1}{2} \rceil & \text{if } n \not\equiv 0 \text{ (mod 6)}. \end{cases}$$

First, let us prove that the second 3-graph H in Theorem 5.9 has independence number exactly $\lceil (n+1)/2) \rceil$. Assume the vertex set is $[n]$ and the vertices are labelled by consecutive integers in clockwise orientation. The lower bound is by taking $\lceil (n+1)/2) \rceil$ consecutive vertices on the n-gon and noting that no three of them contain the centre in their interior. For the upper bound, let us see how many vertices can lie in an independent set containing 1. When n is odd, the triangle formed by $\{1, i, (n-1)/2 + i\}$ contains the centre and hence is an edge. Therefore we may pair the elements of $[n] \setminus \{1\}$ as $(2,(n+3)/2), (3,(n+5)/2), \ldots, ((n+1)/2, n)$ and note that each pair can have at most one vertex in an independent set containing 1. Hence the maximum size of an independent set containing 1 is at most $(n+1)/2$ and by vertex transitivity of H, the independence number of H is at most $(n+1)/2$. For n even we consider the $n/2-1$ pairs $(2, n/2+1), (3, n/2+2), \ldots, (n/2, n-1)$ and add the vertex n to get an upper bound $n/2+1 = \lceil (n+1)/2) \rceil$.

Next we observe that the 6-vertex 3-graph H' in Theorem 5.9 has independence number exactly 3 (we omit the short case analysis needed for the proof). Hence if we blow-up each vertex of H' into sets of the same size, then we obtain n-vertex 3-graphs with independence number exactly $n/2$ whenever $n \equiv 0 \pmod 6$. This concludes the proof of the upper bound.

For the lower bound, let H be Q-free. Then by Theorem 5.9, H is isomorphic to one of the two graphs described in Theorem 5.9. If H is isomorphic to the second graph, then we have already shown that its independence number is at least $(n+1)/2$, so assume that H is isomorphic to the blow-up of the 6-vertex 10-edge 3-graph H'. There are 10 non-edges in H'. Let $V_1, \ldots, V_6$ be the blown-up vertex sets. Since every vertex $i \in [6]$ in H' is contained in exactly 5 non-edges, we obtain

$$5n = 5\sum_{i \in [6]} |V_i| = \sum_{j_1 j_2 j_3 \notin E(H)} |V_{j_1}| + |V_{j_2}| + |V_{j_3}|.$$

By the pigeonhole principle, there is a non-edge $i_1 i_2 i_3$, such that $|V_{i_1}| + |V_{i_2}| + |V_{i_3}| \geq n/2$. Our bound follows by observing that for any non-edge $i_1 i_2 i_3$ in the original 3-graph H' the set $V_{i_1} \cup V_{i_2} \cup V_{i_3}$ is an independent set. This gives an independent set of size at least $n/2$, and if $n \not\equiv 0 \pmod 6$, then equality cannot hold throughout (a short case analysis, which we omit, is needed to prove this) and we obtain an independent set of size strictly greater than $n/2$ as required.

Case 2: $Q = \{(4,0),(4,2),(4,3)\}$.

We now prove $h(n, \{(4,0),(4,2),(4,3)\}) = n-1$, for $n \geq 4$. Let H be a 3-graph that is a clique on $n-1$ vertices and a single isolated vertex, then H is Q-free, giving us the upper bound.

For the lower bound, let H be a Q-free 3-graph on n vertices, $n \geq 4$. Assume that H is not a clique and not a co-clique. We shall show that H is a clique and a single isolated vertex. Consider a maximal clique S in H. Since $|S| < n$, there is a vertex $v \in V(H) \setminus S$. From the maximality of S, $L_S(v)$ is not a clique. If $L_S(v)$ contains an edge, then we have that for some vertices x, y, y', $xy \in E(L_S(v))$ and $xy' \notin E(L_S(v))$. But then $\{v, x, y, y'\}$ induces a $(4,2)$ or a $(4,3)$-graph, a contradiction.

Thus, $L_S(v)$ is an empty graph, i.e. there is no edge in H containing v and two vertices of S. Now assume there exists a second vertex $v' \in V(H) \setminus (S \cup \{v\})$. Then by the same argument as above, v' is also not contained in any edge with two vertices from S. Consider triples $vv'x$, $x \in S$. Since $|S| \geq 3$, by the pigeonhole principle there are two vertices $x, x' \in S$ such that either $vv'x, vv'x' \in E(H)$ or $vv'x, vv'x' \notin E(H)$. Then $\{v, v', x, x'\}$ induces 2 or 0 edges respectively, a contradiction. Thus, $|S| = n-1$ and v is an isolated vertex. □

5.4 Concluding remarks

For $r=3$ we have determined for each family Q of order-size pairs with order 4 whether it has the EH-property. However, there remain some gaps between the upper and lower bounds on $h_3(n, Q)$ for some families Q with $|Q| \in \{1, 2\}$:

Open Problem 5.10. *Improve the exponent* $1/2$ *in the lower bound on* $h_3(n, 4, 1)$.

Open Problem 5.11. *Prove or disprove that*

- $h_3(n, \{(4,0), (4,2)\}) = n^{1/2+o(1)}$,
- $h_3(n, \{(4,0), (4,3)\}) = n^{1+o(1)}$,
- $h_3(n, \{(4,1), (4,2)\}) = n^{1/2+o(1)}$.

Fix integers $m > r$. Say that a set Q of order size pairs $\{(m, f_1), \dots, (m, f_t)\}$ is Erdős-Hajnal (EH) if there exists $\epsilon = \epsilon_Q$ such that $h_r(n, Q) > n^{\epsilon}$. As $|Q|$ grows, the collection of Q-free r-graphs is more restrictive, and hence $h_r(n, Q)$ grows (assuming that large Q-free r-graphs are not forbidden to exist by Ramsey's theorem). The case when $h_r(n, Q) = \Omega(n)$ was treated by Axenovich and Balogh [9] when $r = 2$. A natural question then is to ask what is the smallest t such that every Q of size t is EH. Call this minimum value $EH_r(m)$. Our results for $r = 3$ show that for $m = 4$, all Q of size 3 are EH, but there are Q of size 2 which are not EH. Consequently, $EH_3(4) = 3$.

In order to further study $EH_r(m)$, we need another definition. Given integers $m \geq r \geq 3$, let $g_r(m)$ be the number of edges in an r-graph on m vertices obtained by first taking a partition of the m vertices into almost equal parts, then taking all edges that intersect each part, and then recursing this construction within each part. For example, $g_3(7) = 13$ since we start with a complete 3-partite 3-graph with part sizes $2, 2, 3$ and then add one edge within the part of size 3. It is known (see, e.g. [109]) that

as r grows we have
$$g_r(m) = (1+o(1))\frac{r!}{r^r - r}\binom{m}{r}.$$
Note that $\frac{r!}{r^r-r}$ approaches 0 as r grows. Mubayi and Razborov [109] proved that for all fixed $m > r > 3$, there are n-vertex r-graphs which are Q-free, $Q = \{(m,i) : g_r(m) < i \leq \binom{m}{r}\}$, with $h(G) = O(\log n)$. In other words, there exists Q of size $\binom{m}{r} - g_r(m)$ which is not EH. This proves that $EH_r(m) \geq \binom{m}{r} - g_r(m) + 1$.

Erdős and Hajnal [58] proved that for all $m > r \geq 3$, the set $Q = \{(m,i) : g_r(m) \leq i \leq \binom{m}{r}\}$ is EH. In other words, they proved that every n-vertex r-graph in which every set of m vertices spans less then $g_r(m)$ edges has an independent set of size at least n^ϵ, where ϵ depends only on r and m. This is a particular set Q of size $\binom{m}{r} - g_r(m) + 1$ that is EH and we speculate that every other set Q of this size is also EH.

Open Problem 5.12. *Prove or disprove that for all $m > r > 2$,*
$$EH_r(m) = \binom{m}{r} - g_r(m) + 1.$$

We end by noting that $EH_3(4) = 3 = \binom{4}{3} - g_3(4) + 1$.

Part II

Order-size pairs: absolute avoidability and forcing densities

Introduction and basic notions

One of the central topics of graph theory deals with properties of classes of graphs that contain no subgraph isomorphic to some given fixed graph, see for example Bollobás [26]. Similarly, graphs with forbidden induced subgraphs have been investigated from several different angles – enumerative, structural, algorithmic, and more. One famous example are Erdős-Hajnal-type problems, like the ones discussed in Part I of this thesis.

Erdős, Füredi, Rothschild and Sós [56] initiated a study of a seemingly simpler class of graphs that do not forbid a specific induced subgraph, but rather forbid any induced subgraph on a given number m of vertices and number f of edges. Following their notation we say a graph G *arrows* a pair of non-negative integers, an order-size pair (m, f), and write $G \to (m, f)$ if G has an induced subgraph on m vertices and f edges. We say that a pair (n, e) of non-negative integers *arrows* the pair (m, f), and write $(n, e) \to (m, f)$, if for any graph G on n vertices and e edges, $G \to (m, f)$.

As an example, if $t(n, m-1)$ denotes the number of edges in the complete balanced $(m-1)$-partite graph on n vertices, $T_2(n, m-1)$, then by Turán's Theorem [122] we know that any graph on n vertices with more than $t(n, m-1)$ edges contains a copy of K_m. On the other hand, for any $e \leq t(n, m-1)$ there exists a subgraph of $T_2(n, m-1)$ with e edges, which does not contain a copy of K_m. Equivalently stated, we have $(n, e) \to (m, \binom{m}{2})$ if and only if $e > t(n, m-1)$.

For a fixed pair (m, f) let $S_n(m, f) = \{e : (n, e) \to (m, f)\}$ and define the *forcing density*

$$\sigma(m, f) = \limsup_{n\to\infty} |S_n(m, f)| / \binom{n}{2} .$$

In [56] the authors considered $\sigma(m, f)$ for different choices of (m, f). One of their main results is

Theorem 0.4 (Erdős, Füredi, Rothschild, Sós [56]).
If $(m, f) \notin \{(2,0),(2,1),(4,3),(5,4),(5,6)\}$*, then* $\sigma(m, f) \leq \frac{2}{3}$*; otherwise* $\sigma(m, f) = 1$.

On the other hand, they showed that there are infinitely many pairs of positive forcing density, in particular there are infinitely many pairs (m, f) with $\sigma(m, f) \geq \frac{1}{8}$. He, Ma and Zhao [88] improved this result, by showing that there are infinitely many pairs (m, f) with $\sigma(m, f) \geq \frac{1}{2}$. They also improved the upper bound $2/3$ to $1/2$ and showed that there are infinitely many pairs for which the equality $\sigma(m, f) = \frac{1}{2}$ holds.

Erdős, Füredi, Rothschild and Sós [56] also gave a construction demonstrating that "most of the" forcing densities $\sigma(m, f)$ are 0, by showing that for large n almost all pairs (n, e) can be realised as the vertex-disjoint union of a clique and a high-girth graph, and that for fixed m most pairs (m, f) cannot be realised as the vertex-disjoint union of a clique and a forest. For some other results concerning sizes of induced subgraphs, see for example Alon and Kostochka [4], Alon, Balogh, Kostochka and Samotij [3], Alon, Krivelevich and Sudakov [5], Axenovich and Balogh [9], Bukh and Sudakov [32], Kwan and Sudakov [100,101], Baksys and Chen [18] for a similar result for bipartite graphs, and Narayanan, Sahasrabudhe and Tomon [111]. A similar question on avoidable order-size pairs was considered by Caro, Lauri and Zarb [38] for the class of line graphs.

In Chapter 6 we investigate the existence of pairs (m, f) for which we not only have $\sigma(m, f) = 0$, but the stronger property $S_n(m, f) = \emptyset$ for all sufficiently large n; we call such pairs *absolutely avoidable*. We show that there exist infinitely many absolutely avoidable pairs. Moreover, we give an infinite set of unavoidable pairs of the form $(m, \binom{m}{2}/2)$. We also show that for any m sufficiently large, there exists some f for which the pair (m, f) is absolutely avoidable.

In Chapter 7 we will consider a variant of the problem in the bipartite setting. We only consider a balanced version here, i.e. a bipartite order-size pair (m, f) is the class of all bipartite graphs with m vertices in each part and f edges. We use analogous definitions of avoidability and the bipartite forcing density $\sigma_{bip}(m, f)$. It would be interesting to show whether there are also absolutely avoidable pairs in the bipartite setting. Unfortunately we cannot use our method from Chapter 6 to find such pairs. However, we can show that there exist infinitely many bipartite pairs (m, f) for which $\sigma_{bip}(m, f) = 0$ holds. On the other hand, there also exist infinitely many pairs (m, f) with $\sigma_{bip}(m, f) = 1$.

Finally we focus our attention on forbidden order-size pairs in hypergraphs. We extend our definitions to r-uniform hypergraphs as follows:

We say an r-graph, G *arrows* a pair of non-negative integers (m, f) and write $G \to_r (m, f)$ if G has an induced sub-hypergraph on m vertices and f hyperedges. We say that a pair (n, e) of non-negative integers *arrow* (or simply induces) the pair (m, f), and write

$$(n, e) \to_r (m, f)$$

if for any r-graph G on n vertices and e edges, $G \to_r (m, f)$. We say a pair (n, e) is *realised* by an r-graph G if G has n vertices and e edges. If r is clear from the context, we might omit the index and simply write $(n, e) \to (m, f)$. A pair (m, f) is *absolutely r-avoidable* (or just absolutely avoidable, if the uniformity is clear from the context) if for all n sufficiently large, we have $\{e : (n, e) \to_r (m, f)\} = \emptyset$. The *forcing density* of a pair (m, f) is

$$\sigma_r(m, f) = \limsup_{n\to\infty} \frac{|\{e : (n, e) \to_r (m, f)\}|}{\binom{n}{r}}.$$

In Chapter 8 we show that for any $r \geq 3$ there exists m_0 such that for every $m \geq m_0$ either $(m, \lfloor \binom{m}{r}/2 \rfloor)$ or $(m, \lfloor \binom{m}{r}/2 \rfloor - m - 1)$ is absolutely avoidable. We further show, for any $r, m \in \mathbb{N}$, $r, m \geq 3$, all but at most $m^{\frac{r}{r-1}}$ of all possible $\binom{m}{r}$ pairs (m, f) have forcing density $\sigma_r(m, f) = 0$. We also give some general upper bounds on $\sigma_r(m, f)$ and show that there exists no non-trivial pair with $\sigma_r(m, f) = 1$.

Chapter 8 raises the question whether for $r \geq 3$ there exists any non-trivial pair (m, f) with $\sigma_r(m, f) > 0$ at all and identifies some candidate pairs, the smallest being $(6, 10)$ for $r = 3$. In Chapter 9 we answer the question in the affirmative and prove that $\sigma_3(6, 10) > 0$ indeed. We also give more precise upper and lower bounds on $\sigma_3(6, 10)$ and prove some conditions any other pair (m, f) must satisfy to have forcing density $\sigma_r(m, f) > 0$.

CHAPTER 6 ORDER-SIZE PAIRS IN GRAPHS: ABSOLUTELY AVOIDABLE PAIRS AND FORCING DENSITIES

6.1 INTRODUCTION

In this chapter we investigate the existence of pairs (m, f) for which we not only have $\sigma(m, f) = 0$, but the stronger property $S_n(m, f) = \emptyset$ for large n.

Definition 6.1. *A pair (m, f) is **absolutely avoidable** if there is n_0 such that for each $n > n_0$ and for any $e \in \{0, \dots, \binom{n}{2}\}$, $(n, e) \not\rightarrow (m, f)$.*

Our results show that there are infinitely many absolutely avoidable pairs. Our first result gives an explicit construction of infinitely many absolutely avoidable pairs $(m, \binom{m}{2}/2)$. The second one provides an existence result of infinitely many absolutely avoidable pairs (m, f), where f is "close" to $\binom{m}{2}/2$. Finally, the last result shows that for every sufficiently large m at least one of the pairs $(m, \lfloor \binom{m}{2}/2 \rfloor)$ and $(m, \lfloor \binom{m}{2}/2 \rfloor - 6m)$ is absolutely avoidable.

For the first result we need to define the following set M of integers. Let

$$M = \left\{ \frac{1}{2} \left(\begin{pmatrix} 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 3 & 4 \\ 2 & 3 \end{pmatrix}^s \cdot \begin{pmatrix} 3 \\ 1 \end{pmatrix} + 5 \right) : s \in \mathbb{N}, s \geq 2 \right\}.$$

In particular, we have $M = \{40, 221, 1276 \dots\}$.

Theorem 6.2. *For any $m \in M$, $f = \binom{m}{2}/2$ is an integer and the pair (m, f) is absolutely avoidable.*

Theorem 6.3. *For any monotone integer valued function $q(m)$ such that $|q(m)| = O(m)$, there are infinitely many values of m, such that the pair $(m, \binom{m}{2}/2 - q(m))$ is absolutely avoidable.*

Moreover, there are infinitely many values of m, such that for any integer $f' \in (\binom{m}{2}/2 - 0.175m, \binom{m}{2}/2 + 0.175m)$ the pair (m, f') is absolutely avoidable.

Theorem 6.4. *For any $m \geq 754$ either $(m, \lfloor \binom{m}{2}/2 \rfloor)$ or $(m, \lfloor \binom{m}{2}/2 \rfloor - 6m)$ is absolutely avoidable.*

Theorems 6.2 and 6.3 are joint work with Axenovich and appear in [16], together with Theorem 6.4 for $m \equiv 0, 1 \pmod 4$:

Proposition 6.5. *For any $m \geq 740$ with $m \equiv 0, 1 \pmod 4$ either $(m, \binom{m}{2}/2)$ or $(m, \binom{m}{2}/2 - 6m)$ is absolutely avoidable.*

The result is obtained for $m \equiv 2, 3 \pmod 4$ in a very similar way by carefully changing some of the constants involved. The proof does not appear in [16], but we include it here for completeness.

Proposition 6.6. *For any $m \geq 754$ with $m \equiv 2, 3 \pmod 4$ either $(m, \lfloor\binom{m}{2}/2\rfloor)$ or $(m, \lfloor\binom{m}{2}/2\rfloor - 6m)$ is absolutely avoidable.*

The main idea of the proofs is that for certain pairs (m, f), there is no graph on m vertices and f edges which is the vertex-disjoint union of a clique and a forest or a complement of a vertex-disjoint union of a clique and a forest. In order to do so, we need several number theoretic statements that we prove in several lemmata. After that, we use the observation Erdős, Füredi, Rothschild and Sós [56], that for any sufficiently large n, and any $e \leq c\binom{n}{2}$, for any $0 \leq c < 1$, there is a graph on n vertices and e edges that is the vertex-disjoint union of a clique and a graph of girth greater than m. In particular, any m-vertex induced subgraph of such a graph is a disjoint union of a clique and a forest. Considering the complements, we deduce that (m, f) is absolutely avoidable.

This chapter is structured as follows. We state and prove the lemmata in Section 6.2, prove Theorems 6.2 and 6.3, Proposition 6.5 and Proposition 6.6 in Section in Section 6.3. Section 6.4 provides final remarks and open questions.

6.2 LEMMATA AND NUMBER THEORETIC RESULTS

We say that a pair (m, f) is *realisable* by a graph $H = (V, E)$ if $|V(H)| = m$ and $|E(H)| = f$. Recall that for $x \in \mathbb{R}$ the fractional part of x is denoted by $\{x\}$. A real-valued sequence $(x_n)_{n\in\mathbb{N}}$ is called *uniformly distributed modulo 1* (we write u.d. mod 1) if for any pair of real numbers s, t with $0 \leq s < t \leq 1$ we have

$$\lim_{N\to\infty} \frac{|\{n : 1 \leq n \leq N, \{x_n\} \in [s, t)\}|}{N} = t - s.$$

We will use the following facts:

Lemma 6.7. *(a) The sequence $(x_n) = \alpha n$ is u.d. mod 1 for any $\alpha \in \mathbb{R} \setminus \mathbb{Q}$.*

(b) If a real-valued sequence (x_n) is u.d. mod 1 and a real-valued sequence (y_n) has the property $\lim_{n\to\infty}(x_n - y_n) = \beta$, a real constant, then (y_n) is also u.d. mod 1.

For proofs of these facts see for example Theorem 1.2 and Example 2.1 in [99]. The following lemma is given in [56], we include it here for completeness.

Lemma 6.8. *Let $p \in \mathbb{N}$ and c be a constant $0 \leq c < 1$. Then for $n \in \mathbb{N}$ sufficiently large and any $e \in [c\binom{n}{2}]$, there exists a non-negative integer k and a graph on n vertices and e edges which is the vertex-disjoint union of a clique of size k and a graph on $n-k$ vertices of girth at least p.*

Proof. Let $p > 0$ be given. We use the fact that for any v large enough there exists a graph of girth p on v vertices with $v^{1+\frac{1}{2p}}$ edges. For a probabilistic proof of this fact see for example Bollobás [26] and for an explicit construction see Lazebnik et al. [102]. Let n be a given sufficiently large integer. Let $e \in [c\binom{n}{2}]$. Let k be a non-negative integer such that $\binom{k}{2} \leq e \leq \binom{k+1}{2} - 1$. Note that since $e \leq c\binom{n}{2}$, $\binom{k}{2} \leq c\binom{n}{2}$, thus, $k \leq \sqrt{c}n + 1 \leq c'n$, where c' is a constant, $c' < 1$. We claim that (n, e) could be represented as a vertex-disjoint union of a clique on k vertices and a graph of girth at least p. For that, consider a graph G' on $n-k$ vertices and girth at least p such that $|E(G')| \geq (n-k)^{1+\frac{1}{2p}}$. Consider G'', the vertex-disjoint union of G' and K_k. Then $|E(G'')| \geq \binom{k}{2} + (n-k)^{1+\frac{1}{2p}} \geq \binom{k+1}{2} \geq e$. Here, the second inequality holds since $(n-k)^{1+\frac{1}{2p}} \geq k$ for $k \leq c'n$ and n large enough. Finally, let G be a subgraph of G'' on e edges, obtained from G'' by removing some edges of G'. Thus, G is the vertex-disjoint union of a clique on k vertices and a graph of girth at least p. $\square$

We shall need two number theoretic lemmata for the proof of the main result. Below the set M is defined as in the introduction.

Lemma 6.9. *For any $m \in M$, m is a positive integer congruent to 0 or 1 modulo 4, and $\sqrt{2m^2 - 10m + 9}$ is an odd integer for each $m \in M$.*

Proof. Recall that $M = \left\{ \frac{1}{2}\left(\begin{pmatrix} 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 3 & 4 \\ 2 & 3 \end{pmatrix}^s \cdot \begin{pmatrix} 3 \\ 1 \end{pmatrix} + 5 \right) : s \in \mathbb{N}, s \geq 2 \right\}$. We see, that M corresponds to the following recursion: $(x_0, y_0) = (3, 1)$ and for $s \geq 0$

$$x_{s+1} = 3x_s + 4y_s$$
$$y_{s+1} = 2x_s + 3y_s.$$

I.e., for $s \geq 0$,

$$\begin{pmatrix} x_s \\ y_s \end{pmatrix} = \begin{pmatrix} 3 & 4 \\ 2 & 3 \end{pmatrix}^s \cdot \begin{pmatrix} 3 \\ 1 \end{pmatrix}.$$

Indeed, $M = \{(x_s + 5)/2 : s \geq 2\}$.

From the recursion we see that $x_{2s} \equiv 3 \pmod 8$, $x_{2s+1} \equiv 5 \pmod 8$, $y_{4s} = y_{4s+1} \equiv 1 \pmod 8$, and $y_{4s+2} = y_{4s+3} \equiv 5 \pmod 8$ for $s \in \mathbb{N}_0$. In particular y_s is an odd integer. Let $m_s = (x_s + 5)/2$, i.e. $M = \{m_s : \ s \geq 2\}$. When s is even, $m_s \equiv 0 \pmod 4$, and if s is odd, $m_s \equiv 1 \pmod 4$. This proves the first statement of the lemma.

Next, we observe that $(x, y) = (x_s, y_s)$ gives an integer solution to the generalized Pell's equation

$$x^2 - 2y^2 = 7. \tag{$*$}$$

Indeed, $(x, y) = (x_0, y_0) = (3, 1)$ satisfies $(*)$. Assume that $(x, y) = (x_s, y_s)$ satisfies $(*)$. Let $(x, y) = (x_{s+1}, y_{s+1})$ and insert it into the left hand side of $(*)$. Then we have

$$x_{s+1}^2 - 2y_{s+1}^2 = 9x_s^2 + 24x_sy_s + 16y_s^2 - 8x_s^2 - 24x_sy_s - 18y_s^2 = x_s^2 - 2y_s^2 = 7.$$

Thus, $(x, y) = (x_{s+1}, y_{s+1})$ also satisfies $(*)$.

Since (x_s, y_s) satisfies $(*)$, we have that $y_s = \sqrt{\frac{1}{2}(x_s^2 - 7)}$. Then

$$y_s = \sqrt{\frac{1}{2}((2m_s - 5)^2 - 7)} = \sqrt{\frac{1}{2}(4m_s^2 - 20m_s + 18)} = \sqrt{2m_s^2 - 10m_s + 9}.$$

Since y_s is an odd integer, the second statement of the lemma follows. $\square$

For the next lemmata and theorems we will need the following definitions. Let $m, q \in \mathbb{Z}$, $m \geq 5 + 2\sqrt{|q|}$. Let

$$\begin{aligned} y_q(m) &= \tfrac{\sqrt{2m^2-10m-8q+9}}{2}, & z_q(m) &= \tfrac{\sqrt{2m^2-2m-8q+1}}{2}, \\ t_q(m) &= z_q(m) - y_q(m), & d_q(m) &= \tfrac{3}{2} - t_q(m), \\ L_q(m) &= \left\lfloor \tfrac{5}{2} + y_q(m) \right\rfloor, & R_q(m) &= \left\lfloor \tfrac{1}{2} + z_q(m) \right\rfloor. \end{aligned}$$

Note that since $m \geq 5 + 2\sqrt{|q|}$, we always have $y_q(m), z_q(m) \in \mathbb{R}$.

Lemma 6.10. *Let $q = q(m), m \in \mathbb{Z}, m \equiv 0, 1 \pmod 4, m \geq 5+2\sqrt{|q|}$, and $|q(m)| = O(m)$.*

(a) We have $t_q(m) = \dfrac{2\sqrt{2}(1-\frac{1}{m})}{\sqrt{1-\frac{1}{m}+\frac{1-8q}{2m^2}}+\sqrt{1-\frac{5}{m}+\frac{9-8q}{2m^2}}}$. In particular, $\lim_{m\to\infty} d_q(m) = \frac{3}{2} - \sqrt{2}$.

(b) We have $L_q(m) > R_q(m)$ if and only if $\{y_q(m)\} \in [0, d_q(m)) \cup \left[\frac{1}{2}, 1\right)$. In particular, $L_0(m) > R_0(m)$ if $m \in M$.

Proof. We start by proving (a). By definition of $t_q(m)$ we have

$$\begin{aligned}
t_q(m) &= z_q(m) - y_q(m) \\
&= \frac{1}{2}\sqrt{2m^2-2m-8q+1} - \frac{1}{2}\sqrt{2m^2-10m-8q+9} \\
&= \frac{1}{2}\frac{2m^2-2m-8q+1-2m^2+10m+8q-9}{\sqrt{2m^2-2m-8q+1}+\sqrt{2m^2-10m-8q+9}} \\
&= \frac{2\sqrt{2}(1-\frac{1}{m})}{\sqrt{1-\frac{1}{m}+\frac{1-8q}{2m^2}}+\sqrt{1-\frac{5}{m}+\frac{9-8q}{2m^2}}}.
\end{aligned}$$

This also shows that for $|q| = |q(m)| \in O(m)$, $\lim\limits_{m\to\infty} d_q(m) = \frac{3}{2} - \lim\limits_{m\to\infty} t_q(m) = \frac{3}{2} - \sqrt{2}$, which concludes the proof of (a).

Now we can prove part (b). From part (a) we have in particular that $t_q(m) = \sqrt{2} + \epsilon_q(m)$, where for m sufficiently large $|\epsilon_q(m)| < 0.05$, and thus, $t_q(m) \in (1, \frac{3}{2})$. Thus, $d_q(m) = \frac{3}{2} - t_q(m) \in (0, \frac{1}{2})$ for sufficiently large m. We compare $L_q(m)$ and $R_q(m)$ using the expression $x = \lfloor x \rfloor + \{x\}$:

$$\begin{aligned}
L_q(m) &= \left\lfloor \frac{5}{2} + y_q(m) \right\rfloor \\
&= 2 + \lfloor y_q(m) \rfloor + \left\lfloor \frac{1}{2} + \{y_q(m)\} \right\rfloor \\
&= 2 + \lfloor y_q(m) \rfloor + \begin{cases} 0, & \{y_q(m)\} \in [0, \frac{1}{2}) \\ 1, & \{y_q(m)\} \in [\frac{1}{2}, 1) \end{cases},
\end{aligned}$$

$$\begin{aligned}
R_q(m) &= \left\lfloor \frac{1}{2} + z_q(m) \right\rfloor \\
&= \left\lfloor \frac{1}{2} + y_q(m) + t_q(m) \right\rfloor \\
&= \lfloor y_q(m) \rfloor + \left\lfloor \frac{1}{2} + t_q(m) + \{y_q(m)\} \right\rfloor \\
&= \lfloor y_q(m) \rfloor + \begin{cases} 1, & t_q(m) + \{y_q(m)\} \in [1, \frac{3}{2}) \\ 2, & t_q(m) + \{y_q(m)\} \in [\frac{3}{2}, \frac{5}{2}) \end{cases}.
\end{aligned}$$

Thus,

$$L_q(m) - R_q(m) = 2 + \begin{cases} 0-1, & \{y_q(m)\} \in [0, \frac{1}{2}) \text{ and } t_q(m) + \{y_q(m)\} \in [1, \frac{3}{2}) \\ 0-2, & \{y_q(m)\} \in [0, \frac{1}{2}) \text{ and } t_q(m) + \{y_q(m)\} \in [\frac{3}{2}, \frac{5}{2}) \\ 1-1, & \{y_q(m)\} \in [\frac{1}{2}, 1) \text{ and } t_q(m) + \{y_q(m)\} \in [1, \frac{3}{2}) \\ 1-2, & \{y_q(m)\} \in [\frac{1}{2}, 1) \text{ and } t_q(m) + \{y_q(m)\} \in [\frac{3}{2}, \frac{5}{2}) \end{cases}.$$

So, $L_q(m) - R_q(m) > 0$ in all cases except for the second one, i.e. if and only if

$$\begin{aligned} \{y_q(m)\} &\in [0,1) \setminus \big([0, \tfrac{1}{2}) \cap \big[\tfrac{3}{2} - t_q(m), \tfrac{5}{2} - t_q(m)\big)\big) \\ &= \big[\tfrac{1}{2}, 1\big) \cup ([0,1) \setminus [d_q(m), 1 + d_q(m))) \\ &= \big[\tfrac{1}{2}, 1\big) \cup [0, d_q(m)). \end{aligned}$$

Now let $m \in M$ and consider $y_0(m) = \frac{\sqrt{2m^2-10m+9}}{2}$. Then by Lemma 6.9, $2y_0(m)$ is an odd integer for all $m \in M$, i.e. $\{y_0(m)\} = \frac{1}{2}$. Thus, we have $L_0(m) > R_0(m)$ for all $m \in M$, which concludes the proof of (b). □

Lemma 6.11. *If $q = q(m) \in \mathbb{Z}$, $m \in \mathbb{N}$, $m \equiv 0, 1 \pmod 4$, $m \geq 2\sqrt{|q|} + 5$, and $L_q(m) > R_q(m)$, then the pair $(m, \binom{m}{2}/2 - q)$ cannot be realised as the vertex-disjoint union of a clique and a forest.*

Proof. Let $f = \binom{m}{2}/2 - q$. Suppose that (m, f) can be realised as the vertex-disjoint union of a clique K on x vertices and a forest F on $m - x$ vertices. We shall show that $L_q(m) \leq R_q(m)$.

Claim 1: $x \geq L_q(m)$.

Proof. The forest F has $f - \binom{x}{2} = \binom{m}{2}/2 - q - \binom{x}{2}$ edges. Since F has $m - x$ vertices, it contains strictly less than $m - x$ edges. Thus, we have $\binom{m}{2}/2 - q - \binom{x}{2} < m - x$. Solving for x gives

$$x > \frac{3}{2} + \frac{1}{2}\sqrt{2m^2 - 10m - 8q + 9} \quad \text{or} \quad x < \frac{3}{2} - \frac{1}{2}\sqrt{2m^2 - 10m - 8q + 9}.$$

Since $m \geq 2\sqrt{|q|} + 5$, we have $2m^2 - 10m - 8q + 9 \geq 9$. The second inequality gives $x < \frac{3}{2} - \frac{1}{2}\sqrt{2m^2 - 10m - 8q + 9}$, and thus, $x < 0$, a contradiction. So only the first inequality for x holds and implies that

$$x \geq \left\lfloor \frac{3+\sqrt{2m^2-10m-8q+9}}{2} \right\rfloor + 1 = L_q(m),$$

which proves Claim 1. ■

Claim 2: $x \leq R_q(m)$.

Proof. The number of edges in the clique K is at most f and exactly $\binom{x}{2}$. Thus, $\binom{x}{2} \leq f = \binom{m}{2}/2 - q$, which implies that $2x(x-1) \leq m(m-1) - 4q$. This in turn gives

$$x \leq \left\lfloor \frac{1+\sqrt{2m^2-2m-8q+1}}{2} \right\rfloor = R_q(m),$$

and proves Claim 2. ■

Claims 1 and 2 imply that $L_q(m) \leq R_q(m)$. □

Lemma 6.12. *Let $q = q(m) \in \mathbb{Z}$, $m \in \mathbb{N}$, $m \equiv 0, 1 \pmod 4$, $m \geq 2\sqrt{|q|} + 5$. If both $L_q(m) > R_q(m)$ and $L_{-q}(m) > R_{-q}(m)$, then the pair $(m, f) = (m, \binom{m}{2}/2 - q)$ is absolutely avoidable.*

Proof. Let m satisfy the condition of the lemma and let $f_- = \binom{m}{2}/2 - q$ and $f_+ = \binom{m}{2}/2 + q$. Then by Lemma 6.11, neither (m, f_+) nor (m, f_-) can be represented as the vertex-disjoint union of a clique and a forest.

By Lemma 6.8, for every sufficiently large n, and all $e \leq \lceil \binom{n}{2}/2 \rceil$ we can realise (n, e) as the vertex-disjoint union of a clique and a graph of girth greater than m. Thus, for each $e \in \{0, 1, \ldots, \binom{n}{2}\}$ there is a graph G on n vertices and e edges such that either G or the complement $\overline{G}$ of G is a vertex-disjoint union of a clique and a graph of girth greater than m.

If G is the vertex-disjoint union of a clique and a graph of girth greater than m, then any m-vertex induced subgraph of G is a vertex-disjoint union of a clique and a forest. Since (m, f_-) cannot be represented as a clique and a forest, we have $G \not\to (m, f_-)$. If $\overline{G}$ is the vertex-disjoint union of a clique and a graph of girth greater than m, then as above $\overline{G} \not\to (m, f_+)$. Since $f_- = \binom{m}{2} - f_+$, we have that $G \not\to (m, f_-)$. Thus, (m, f_-) is absolutely avoidable. □

6.3 PROOFS OF THE MAIN THEOREMS

Proof of Theorem 6.2. Let $m \in M$. By Lemma 6.9 we have $m \equiv 0, 1 \pmod 4$, so $f = \binom{m}{2}/2$ is an integer. By Lemma 6.10(b) we have $L_0(m) > R_0(m)$. Now we can apply Lemma 6.12 with $q = 0$. Thus, the pair (m, f) is absolutely avoidable. □

Proof of Theorem 6.3. Let $q = q(m) \in \mathbb{Z}$, $|q(m)| \in O(m)$, be a monotone function. Recall that $y_q(m) = \frac{1}{2}\sqrt{2m^2 - 10m + 9 - 8q}$. Let $a = \lim\limits_{m\to\infty} \frac{q(m)}{m}$.

Claim 1: $\lim\limits_{m\to\infty}\left(\frac{m}{\sqrt{2}} - y_q(m)\right) = \frac{5}{2\sqrt{2}} + \sqrt{2}a$ and $\lim\limits_{m\to\infty}\left(\frac{m}{\sqrt{2}} - y_{-q}(m)\right) = \frac{5}{2\sqrt{2}} - \sqrt{2}a$.

Observe that

$$
\begin{aligned}
\lim_{m\to\infty}\left(\frac{m}{\sqrt{2}} - y_q(m)\right) &= \lim_{m\to\infty}\frac{m}{\sqrt{2}}\left(1 - \sqrt{1 - \frac{5}{m} + \frac{9-8q}{2m^2}}\right) \\
&= \lim_{m\to\infty}\frac{m}{\sqrt{2}}\frac{\frac{5}{m} - \frac{9-8q}{2m^2}}{1 + \sqrt{1 + \frac{5}{m} + \frac{9-8q}{2m^2}}} \\
&= \frac{5}{2\sqrt{2}} + \lim_{m\to\infty}\frac{\sqrt{2}q}{m} \\
&= \frac{5}{2\sqrt{2}} + \sqrt{2}a.
\end{aligned}
$$

Doing a similar calculation for $y_{-q}(m)$ proves Claim 1. ■

Claim 2: $y_q(4m)$ and $y_{-q}(4m)$ are u.d. mod 1; in particular, $y_0(4m)$ is u.d. mod 1.

Since $\frac{1}{\sqrt{2}} \in \mathbb{R} \setminus \mathbb{Q}$, by Lemma 6.7(a) the sequence $(x_{4m}) = (4m)/\sqrt{2}$ is u.d. mod 1. Since we have $\lim\limits_{m\to\infty}(x_{4m} - y_q(4m)) = \frac{5+2\sqrt{2}a}{2\sqrt{2}} \in \mathbb{R}$ and $\lim\limits_{m\to\infty}(x_{4m} - y_{-q}(4m)) = \frac{5-2\sqrt{2}a}{2\sqrt{2}} \in \mathbb{R}$, by Lemma 6.7(b) $(y_q(4m))$ and $(y_{-q}(4m))$ are also u.d. mod 1. This proves Claim 2. ■

Now, to prove the first part of the theorem, from Lemma 6.12 it suffices to find infinitely many integers m such that for $q = q(m)$, $L_q(m) > R_q(m)$ and $L_{-q}(m) > R_{-q}(m)$.

By Lemma 6.10(a), we have that $\lim\limits_{m\to\infty} d_q(m) = \lim\limits_{m\to\infty} d_{-q}(m) = 3/2 - \sqrt{2}$. Let m_0 be large enough so that for any $m \geq m_0$, $d_q(m)$ and $d_{-q}(m)$ are close to these limits, i.e. $|d_q(m) - (3/2 - \sqrt{2})| < (3/2 - \sqrt{2})/3$ and $|d_{-q}(m) - (3/2 - \sqrt{2})| < (3/2 - \sqrt{2})/3$.

Let $\delta > 0$ be a small constant such that $\delta < (3/2 - \sqrt{2})/2$, $2\delta < 1 - \{\sqrt{2}a\}$ and if $\{\sqrt{2}a\} < 1/2$, then $\delta < 1/2 - \{\sqrt{2}a\}$. In addition assume that δ is sufficiently small that for any $m \geq m_0$, $\delta < d_q(m)/3$, and $\delta < d_{-q}(m)/3$. Using Claim 1, define m_δ to be sufficiently large, so that $m_\delta > m_0$ and for any $m \geq m_\delta$, $y_q(m) - \frac{m}{\sqrt{2}}$ and $y_{-q}(m) - \frac{m}{\sqrt{2}}$ are δ-close to the limiting values:

$$y_q(m) \in \left(\left(\frac{m}{\sqrt{2}} - \frac{5}{2\sqrt{2}} - \sqrt{2}a\right) - \delta, \left(\frac{m}{\sqrt{2}} - \frac{5}{2\sqrt{2}} - \sqrt{2}a\right) + \delta\right) \quad \text{and}$$
$$y_{-q}(m) \in \left(\left(\frac{m}{\sqrt{2}} - \frac{5}{2\sqrt{2}} + \sqrt{2}a\right) - \delta, \left(\frac{m}{\sqrt{2}} - \frac{5}{2\sqrt{2}} + \sqrt{2}a\right) + \delta\right).$$

We distinguish two cases based on the values of a:

Case 1: $\{\sqrt{2}a\} \in [0, \frac{1}{4}) \cup [\frac{1}{2}, \frac{3}{4})$, i.e. $\{2\sqrt{2}a\} \in [0, \frac{1}{2})$.

Since $\frac{4m}{\sqrt{2}} - \frac{5}{2\sqrt{2}}$ is a sequence u.d. mod 1, there is an infinite set M_1 of integers at least m_δ, such that for any $m \in M_1$

$$\frac{4m}{\sqrt{2}} - \frac{5}{2\sqrt{2}} \in (k_m + 1/2 + \{\sqrt{2}a\} + \delta, k_m + 1/2 + \{\sqrt{2}a\} + 2\delta),$$

for some integer k_m. Then we have

$$y_q(4m) \in \left((1/2 + k_m + \{\sqrt{2}a\} + \delta) - \sqrt{2}a - \delta, (1/2 + k_m + \{\sqrt{2}a\} + \delta) - \sqrt{2}a + \delta\right),$$

$$y_{-q}(4m) \in \left((1/2 + k_m + \{\sqrt{2}a\} + \delta) + \sqrt{2}a - \delta, (1/2 + k_m + \{\sqrt{2}a\} + \delta) + \sqrt{2}a + \delta\right).$$

This implies that

$$\{y_q(4m)\}, \{y_{-q}(4m)\} \in [1/2, 1).$$

From Lemma 6.10(b), $L_q(4m) > R_q(4m)$ and $L_{-q}(4m) > R_{-q}(4m)$. Note that $f = \binom{4m}{2}/2 - q(4m)$ is an integer. Thus, by Lemma 6.12 the pair $(4m, \binom{4m}{2}/2 - q(4m))$ is absolutely avoidable for any $m \in M_1$.

Case 2: $\{\sqrt{2}a\} \in [\frac{1}{4}, \frac{1}{2}) \cup [\frac{3}{4}, 1)$, i.e. $\{2\sqrt{2}a\} \in [\frac{1}{2}, 1)$.

Since $\frac{4m}{\sqrt{2}} - \frac{5}{2\sqrt{2}}$ is a sequence is u.d. mod 1, there is an infinite set M_2 of integers at least m_δ, such that for any $m \in M_2$

$$\frac{4m}{\sqrt{2}} - \frac{5}{2\sqrt{2}} \in (k_m + \{\sqrt{2}a\} + \delta, k_m + \{\sqrt{2}a\} + 2\delta),$$

for some integer k_m. Then we have

$$y_q(4m) \in \Big((k_m + \{\sqrt{2}a\} + \delta) - \sqrt{2}a - \delta, (k_m + \{\sqrt{2}a\} + \delta) - \sqrt{2}a + \delta\Big) \quad \text{and}$$
$$y_{-q}(4m) \in \Big((k_m + \{\sqrt{2}a\} + \delta) + \sqrt{2}a - \delta, (k_m + \{\sqrt{2}a\} + \delta) + \sqrt{2}a + \delta\Big).$$

This implies that

$$\{y_q(4m)\} \in [0, 2\delta)\,, \{y_{-q}(4m)\} \in [1/2, 1)\,.$$

Recall that for any $m > m_\delta$, $\delta < d_q(m)/3$. Thus, $\{y_{-q}(4m)\} \in [1/2, 1)$ and $\{y_q(4m)\} \in [1/2, 1) \cup [0, d_q(4m))$. From Lemma 6.10(b), $L_q(4m) > R_q(4m)$ and $L_{-q}(4m) > R_{-q}(4m)$. Note that $f = \binom{4m}{2}/2 - q(4m)$ is an integer. Thus, by Lemma 6.12 the pair
$\left(4m, \binom{4m}{2}/2 - q(4m)\right)$ is absolutely avoidable for any $m \in M_2$.

This proves the first part of the theorem.

For the second part, let $c = 0.175 < \frac{1}{4\sqrt{2}}$. We shall show that there is an infinite set M_0 of integers such that for any $m \in M_0$ and for all integers $q \in (-cm, cm)$, the pair $(m, \binom{m}{2}/2 - q)$ is absolutely avoidable. In order to do that, we shall show that $y_0(m)$ does not differ much from $y_q(m)$, for chosen values of m.

Recall that $\lim_{m\to\infty} d_q(m) = 3/2 - \sqrt{2} > 0$ for any $q \in (-cm, cm)$. Thus, the interval $\left[\frac{3}{4}, \frac{3}{4} + d_q(m)\right)$ has positive length for any such q and sufficiently large m. By Claim 2 the sequence $y_0(4m)$ is u.d. mod 1, thus, there are infinitely many values of m that $m \equiv 0 \pmod 4$ and $\{y_0(m)\} \in \left[\frac{3}{4}, \frac{3}{4} + d_q(m)\right)$. Now our choice for m will allow us to use Lemmata 6.10, 6.11 and 6.12.

Let $q \in (-cm, cm)$. It will be easier for us to deal with $y_q(m) - y_0(m)$ instead of $y_q(m)$. Let $s_q(m) = y_q(m) - y_0(m)$. We have

$$\begin{aligned}
\lim_{m\to\infty} s_q(m) &= \lim_{m\to\infty} (y_q(m) - y_0(m)) \\
&= \lim_{m\to\infty} \frac{1}{2}\left(\sqrt{2m^2 - 10m + 9 - 8q} - \sqrt{2m^2 - 10m + 9}\right) \\
&= -\sqrt{2} \lim_{m\to\infty} \frac{q}{m}.
\end{aligned}$$

Thus, since $q \in (-cm, cm)$, $c = 0.175 < \frac{1}{4\sqrt{2}}$, for m sufficiently large we have $s_q(m) \in \left(-\frac{1}{4}, \frac{1}{4}\right)$. Since $y_q = s_q(m) + y_0(m)$, and $\{y_0(m)\} \in \left[\frac{3}{4}, \frac{3}{4} + d_q(m)\right)$, we have that $\{y_q\} = \{s_q(m) + y_0(m)\} \in [0, d_q(m)) \cup [\frac{1}{2}, 1)$. Lemma 6.10(b) implies that $L_q(m) > R_q(m)$

and $L_{-q}(m) > R_{-q}(m)$. Lemmata 6.11 and 6.12 then imply that $(m, \binom{m}{2}/2 - q)$ is absolutely avoidable. □

Proof of Proposition 6.5. Let $m \geq 740$, $m \equiv 0, 1 \pmod 4$. If $L_0(m) > R_0(m)$, by Lemma 6.12 $(m, \binom{m}{2}/2)$ is absolutely avoidable, so we, assume using Lemma 6.10(b), that $\{y_0(m)\} \in [d_0(m), \frac{1}{2})$.

We shall first make some observations about $y_{6m}(m)$ and $y_{-6m}(m)$ by comparing them to $y_0(m)$. From the definition we have

$$y_0(m) = \frac{1}{2}\sqrt{2m^2 - 10m + 9},$$
$$y_{6m}(m) = \frac{1}{2}\sqrt{2m^2 - 58m + 9}, \quad y_{-6m}(m) = \frac{1}{2}\sqrt{2m^2 + 38m + 9}.$$

Thus,

$$\lim_{m\to\infty} y_0(m) - y_{6m}(m) = 6\sqrt{2} \quad \text{and} \quad \lim_{m\to\infty} y_0(m) - y_{-6m}(m) = -6\sqrt{2}.$$

By Lemma 6.10(a),

$$\lim_{m\to\infty} t_0(m) = \lim_{m\to\infty} t_{6m}(m) = \lim_{m\to\infty} t_{-6m}(m) = \sqrt{2}.$$

This implies that

$$\begin{aligned}
\lim_{m\to\infty} y_0(m) - y_{6m}(m) - t_{6m}(m) &= 5\sqrt{2} > 7\\
\lim_{m\to\infty} y_0(m) - y_{6m}(m) + t_0(m) &= 7\sqrt{2} < 10\\
\lim_{m\to\infty} -(y_0(m) - y_{-6m}(m)) + t_{-6m}(m) &= 7\sqrt{2} < 10\\
\lim_{m\to\infty} -(y_0(m) - y_{-6m}(m)) - t_0(m) &= 5\sqrt{2} > 7.
\end{aligned}$$

Thus, for sufficiently large m we have

$$\begin{aligned}
y_{6m}(m) &< y_0(m) - t_{6m}(m) - 7\\
y_{6m}(m) &> y_0(m) + t_0(m) - 10\\
y_{-6m}(m) &< 10 + y_0(m) - t_{-6m}(m)\\
y_{-6m}(m) &> 7 + y_0(m) + t_0(m).
\end{aligned}$$

Thus, combining these inequalities and recalling that $d_q(m) + t_q(m) = 3/2$, for any q, we have

$$y_0(m) - 8 - \frac{1}{2} - d_0(m) < y_{6m}(m) \leq y_0(m) - 8 - \frac{1}{2} + d_{6m}(m),$$

$$y_0(m) + 8 + \frac{1}{2} - d_0(m) < y_{-6m}(m) \leq y_0(m) + 8 + \frac{1}{2} + d_{-6m}(m).$$

Recall that $\{y_0(m)\} \in [d_0(m), \frac{1}{2})$. Recall also that by Lemma 6.10(a), $\lim_{m\to\infty} d_q(m) = \frac{3}{2} - \sqrt{2} \approx 0.086$, for $q \in \{0, 6m, -6m\}$. Then $\{y_{6m}(m)\} \in [0, d_{6m}(m)) \cup [\frac{1}{2}, 1)$ and $\{y_{-6m}(m)\} \in [0, d_{-6m}(m)) \cup [\frac{1}{2}, 1)$.

This implies by Lemma 6.10(b) that $L_{6m}(m) > R_{6m}(m)$ and $L_{-6m}(m) > R_{-6m}(m)$. Therefore by Lemma 6.12, the pair $(m, \binom{m}{2}/2 - 6m)$ is absolutely avoidable.

In particular, one can check that all the above inequalities hold for each $m \geq 740$. □

Proof of Proposition 6.6

Here we prove Proposition 6.6, i.e. Theorem 6.4 for $m \equiv 2, 3 \pmod 4$. It is very similar to the proof of Proposition 6.5 for $m \equiv 0, 1 \pmod 4$, so it is not included in the paper containing the other results [16], but we include it here for completeness.

We will need the following definitions. Let $m, q \in \mathbb{Z}$, $m \geq 6 + 2\sqrt{|q|}$. Let

$$\begin{aligned} y'_q(m) &= \tfrac{\sqrt{2m^2-10m-8q+5}}{2}, & z'_q(m) &= \tfrac{\sqrt{2m^2-2m-8q-3}}{2}, \\ t'_q(m) &= z'_q(m) - y'_q(m), & d'_q(m) &= \tfrac{3}{2} - t'_q(m), \\ L'_q(m) &= \left\lfloor \tfrac{5}{2} + y'_q(m) \right\rfloor, & R_q(m) &= \left\lfloor \tfrac{1}{2} + z'_q(m) \right\rfloor. \end{aligned}$$

Note that since $m \geq 6 + 2\sqrt{|q|}$, we always have $y'_q(m), z'_q(m) \in \mathbb{R}$.

Lemma 6.13. *Let $q = q(m)$, $m \in \mathbb{Z}$, $m \equiv 2, 3 \pmod 4$, $m \geq 6+2\sqrt{|q|}$, and $|q(m)| = O(m)$.*

(a) We have $t'_q(m) = \frac{2\sqrt{2}(1-\frac{1}{m})}{\sqrt{1-\frac{1}{m}+\frac{-3-8q}{2m^2}}+\sqrt{1-\frac{5}{m}+\frac{5-8q}{2m^2}}}$. *In particular,* $\lim_{m\to\infty} d'_q(m) = \frac{3}{2} - \sqrt{2}$.

(b) We have $L'_q(m) > R'_q(m)$ if and only if $\{y'_q(m)\} \in \left[0, d'_q(m)\right) \cup \left[\frac{1}{2}, 1\right)$.

Proof. We start by proving (a). By definition of $t'_q(m)$ we have

$$\begin{aligned}
t'_q(m) &= z'_q(m) - y'_q(m)\\
&= \frac{1}{2}\sqrt{2m^2-2m-8q-3} - \frac{1}{2}\sqrt{2m^2-10m-8q+5}\\
&= \frac{1}{2}\frac{2m^2-2m-8q-3-2m^2+10m+8q-5}{\sqrt{2m^2-2m-8q-3}+\sqrt{2m^2-10m-8q+5}}\\
&= \frac{2\sqrt{2}(1-\frac{1}{m})}{\sqrt{1-\frac{1}{m}+\frac{-3-8q}{2m^2}}+\sqrt{1-\frac{5}{m}+\frac{5-8q}{2m^2}}}.
\end{aligned}$$

This also shows that for $|q| = |q(m)| \in O(m)$, $\lim_{m\to\infty} d'_q(m) = \frac{3}{2} - \lim_{m\to\infty} t'_q(m) = \frac{3}{2} - \sqrt{2}$, which concludes the proof of (a).

Now we can prove part (b). From part (a) we have in particular that $t'_q(m) = \sqrt{2} + \epsilon_q(m)$, where for m sufficiently large $|\epsilon_q(m)| < 0.05$, and thus, $t'_q(m) \in (1, \frac{3}{2})$. Thus, $d'_q(m) = \frac{3}{2} - t'_q(m) \in (0, \frac{1}{2})$ for sufficiently large m. We compare $L'_q(m)$ and $R'_q(m)$ using the expression $x = \lfloor x \rfloor + \{x\}$, similar as in the proof of Lemma 6.10(b):

$$\begin{aligned}
L'_q(m) &= \left\lfloor \frac{5}{2} + y'_q(m) \right\rfloor\\
&= 2 + \lfloor y'_q(m) \rfloor + \begin{cases} 0, & \{y'_q(m)\} \in [0, \frac{1}{2})\\ 1, & \{y'_q(m)\} \in [\frac{1}{2}, 1) \end{cases},
\end{aligned}$$

$$\begin{aligned}
R'_q(m) &= \left\lfloor \frac{1}{2} + z'_q(m) \right\rfloor\\
&= \lfloor y'_q(m) \rfloor + \begin{cases} 1, & t'_q(m) + \{y'_q(m)\} \in [1, \frac{3}{2})\\ 2, & t'_q(m) + \{y'_q(m)\} \in [\frac{3}{2}, \frac{5}{2}) \end{cases}.
\end{aligned}$$

Thus,

$$L'_q(m) - R'_q(m) = 2 + \begin{cases} 0-1, & \{y'_q(m)\} \in [0, \frac{1}{2}) \text{ and } t'_q(m) + \{y'_q(m)\} \in [1, \frac{3}{2})\\ 0-2, & \{y'_q(m)\} \in [0, \frac{1}{2}) \text{ and } t'_q(m) + \{y'_q(m)\} \in [\frac{3}{2}, \frac{5}{2})\\ 1-1, & \{y'_q(m)\} \in [\frac{1}{2}, 1) \text{ and } t'_q(m) + \{y'_q(m)\} \in [1, \frac{3}{2})\\ 1-2, & \{y'_q(m)\} \in [\frac{1}{2}, 1) \text{ and } t'_q(m) + \{y'_q(m)\} \in [\frac{3}{2}, \frac{5}{2}) \end{cases}.$$

So, $L'_q(m) - R'_q(m) > 0$ in all cases except for the second one, i.e. if and only if

$$\begin{aligned}\{y'_q(m)\} &\in [0,1) \setminus \left([0,\tfrac{1}{2}) \cap \left[\tfrac{3}{2} - t'_q(m), \tfrac{5}{2} - t'_q(m)\right)\right) \\ &= \left[\tfrac{1}{2}, 1\right) \cup \left([0,1) \setminus \left[d'_q(m), 1 + d'_q(m)\right)\right) \\ &= \left[\tfrac{1}{2}, 1\right) \cup [0, d'_q(m)).\end{aligned}$$

□

Lemma 6.14. *If $q = q(m) \in \mathbb{Z}$, $m \in \mathbb{N}$, $m \equiv 2, 3 \pmod 4$, $m \geq 2\sqrt{|q|} + 6$, and $L'_q(m) > R'_q(m)$, then the pair $(m, \lfloor \binom{m}{2}/2 \rfloor - q)$ cannot be realised as the vertex-disjoint union of a clique and a forest.*

Proof. Let $f = \lfloor \binom{m}{2}/2 \rfloor - \frac{1}{2} - q$. Suppose that (m, f) can be realised as the vertex-disjoint union of a clique K on x vertices and a forest F on $m - x$ vertices. We shall show that $L'_q(m) \leq R'_q(m)$.

Claim 1: $x \geq L'_q(m)$.

The forest F has $f - \binom{x}{2} = \binom{m}{2}/2 - \frac{1}{2} - q - \binom{x}{2}$ edges. Since F has $m - x$ vertices, it contains strictly less than $m - x$ edges. Thus, we have $\binom{m}{2}/2 - \frac{1}{2} - q - \binom{x}{2} < m - x$. Solving for x gives

$$x > \frac{3}{2} + \frac{1}{2}\sqrt{2m^2 - 10m - 8q + 5} \quad \text{or} \quad x < \frac{3}{2} - \frac{1}{2}\sqrt{2m^2 - 10m - 8q + 5}.$$

Since $m \geq 2\sqrt{|q|} + 6$, we have $2m^2 - 10m - 8q + 5 \geq 9$. The second inequality gives $x < \frac{3}{2} - \frac{1}{2}\sqrt{2m^2 - 10m - 8q + 5}$, and thus, $x < 0$, a contradiction. So only the first inequality for x holds and implies that

$$x \geq \left\lfloor \frac{3 + \sqrt{2m^2 - 10m - 8q + 5}}{2} \right\rfloor + 1 = L'_q(m),$$

which proves Claim 1.

Claim 2: $x \leq R'_q(m)$.

The number of edges in the clique K is at most f and exactly $\binom{x}{2}$. Thus, $\binom{x}{2} \leq f = \binom{m}{2}/2 - \frac{1}{2} - q$, which implies that $2x(x-1) \leq m(m-1) - 2 - 4q$. This in turn gives

$$x \leq \left\lfloor \frac{1 + \sqrt{2m^2 - 2m - 8q - 3}}{2} \right\rfloor = R'_q(m),$$

and proves Claim 2.

Claims 1 and 2 imply that $L'_q(m) \leq R'_q(m)$. □

Lemma 6.15. *Let* $q = q(m) \in \mathbb{Z}$, $m \in \mathbb{N}$, $m \equiv 2, 3 \pmod 4$, $m \geq 2\sqrt{|q|} + 6$. *If both* $L'_q(m) > R'_q(m)$ *and* $L'_{-(q+1)}(m) > R'_{-(q+1)}(m)$, *then the pair* $(m, f) = (m, \lfloor \binom{m}{2}/2 \rfloor - q)$ *is absolutely avoidable.*

Proof. Let m satisfy the condition of the lemma and let $f_- = \lfloor \binom{m}{2}/2 \rfloor - q$ and $f_+ = \lfloor \binom{m}{2}/2 \rfloor + q + 1 = \lceil \binom{m}{2}/2 \rceil + q$. Then by Lemma 6.14, neither (m, f_+) nor (m, f_-) can be represented as the vertex-disjoint union of a clique and a forest.

The rest of the proof is identical to the proof of Lemma 6.12. □

Proof of Proposition 6.6. Let $m \geq 754$, $m \equiv 2, 3 \pmod 4$.

If $L'_0(m) > R'_0(m)$ and $L'_{-1}(m) > R'_{-1}(m)$, by Lemma 6.15 $(m, \lfloor \binom{m}{2}/2 \rfloor)$ is absolutely avoidable, so we assume, using Lemma 6.13(b), that $\{y'_0(m)\} \in [d'_0(m), \frac{1}{2})$ or $\{y'_{-1}(m)\} \in [d'_0(m), \frac{1}{2})$.

We shall first make some observations about $y'_{6m}(m)$ and $y'_{-(6m+1)}(m)$ by comparing them to $y_0(m)$. From the definition we have

$$y'_0(m) = \frac{1}{2}\sqrt{2m^2 - 10m + 5}, \quad y'_{-1}(m) = \frac{1}{2}\sqrt{2m^2 - 10m + 13}$$
$$y'_{6m}(m) = \frac{1}{2}\sqrt{2m^2 - 58m + 5}, \quad y'_{-6m}(m) = \frac{1}{2}\sqrt{2m^2 + 38m + 13}.$$

Thus,

$$\lim_{m\to\infty} y'_0(m) - y'_{6m}(m) = y'_{-1}(m) - y'_{6m}(m) = 6\sqrt{2} \qquad \text{and}$$
$$\lim_{m\to\infty} y'_0(m) - y'_{-(6m+1)}(m) = y'_{-1}(m) - y'_{-(6m+1)}(m) = -6\sqrt{2}.$$

By Lemma 6.13(a),

$$\lim_{m\to\infty} t'_0(m) = \lim_{m\to\infty} t'_{-1}(m) = \lim_{m\to\infty} t'_{6m}(m) = \lim_{m\to\infty} t'_{-(6m+1)}(m) = \sqrt{2}.$$

This implies that

$$\begin{aligned}
\lim_{m\to\infty} y'_0(m) - y'_{6m}(m) - t'_{6m}(m) &= 5\sqrt{2} > 7\\
\lim_{m\to\infty} y'_0(m) - y'_{6m}(m) + t'_0(m) &= 7\sqrt{2} < 10\\
\lim_{m\to\infty} -(y'_0(m) - y'_{-(6m+1)}(m)) + t'_{-(6m+1)}(m) &= 7\sqrt{2} < 10\\
\lim_{m\to\infty} -(y'_0(m) - y'_{-(6m+1)}(m)) - t'_0(m) &= 5\sqrt{2} > 7,
\end{aligned}$$

and the same holds if we replace $y'_0(m)$ by $y'_{-1}(m)$ and $t'_0(m)$ by $t'_{-1}(m)$.

Thus, for sufficiently large m we have

$$\begin{aligned}
y'_{6m}(m) &< y'_0(m) - t'_{6m}(m) - 7\\
y'_{6m}(m) &> y'_0(m) + t'_0(m) - 10\\
y'_{-(6m+1)}(m) &< 10 + y'_0(m) - t'_{-(6m+1)}(m)\\
y'_{-(6m+1)}(m) &> 7 + y'_0(m) + t'_0(m),
\end{aligned}$$

and the same holds if we replace $y'_0(m)$ by $y'_{-1}(m)$ and $t'_0(m)$ by $t'_{-1}(m)$.

Thus, combining these inequalities and recalling that $d'_q(m) + t'_q(m) = 3/2$, for any q, we have

$$y'_0(m) - 8 - \frac{1}{2} - d'_0(m) < y'_{6m}(m) \leq y'_0(m) - 8 - \frac{1}{2} + d'_{6m}(m),$$

$$y'_0(m) + 8 + \frac{1}{2} - d'_0(m) < y'_{-(6m+1)}(m) \leq y'_0(m) + 8 + \frac{1}{2} + d'_{-6m}(m),$$

and the same holds if we replace $y'_0(m)$ by $y'_{-1}(m)$ and $d'_0(m)$ by $d'_{-1}(m)$.

Now assume that $\{y'_0(m)\} \in [d'_0(m), \frac{1}{2})$ or $\{y'_{-1}(m)\} \in [d'_0(m), \frac{1}{2})$. Recall that by Lemma 6.13(a), $\lim_{m\to\infty} d'_q(m) = \frac{3}{2} - \sqrt{2} \approx 0.086$, for $q \in \{0, -1, 6m, -6m-1\}$. Then $\{y'_{6m}(m)\} \in [0, d'_{6m}(m)) \cup [\frac{1}{2}, 1)$ and $\{y'_{-(6m+1)}(m)\} \in [0, d_{-(6m+1)}(m)) \cup [\frac{1}{2}, 1)$.

This implies by Lemma 6.13(b) that $L'_{6m}(m) > R'_{6m}(m)$ and $L'_{-(6m+1)}(m) > R'_{-(6m+1)}(m)$. Therefore by Lemma 6.15, the pair $(m, \lfloor \binom{m}{2}/2 \rfloor - 6m)$ is absolutely avoidable.

In particular, one can check that all the above inequalities hold for each $m \geq 754$. $\square$

6.4 Concluding remarks

We showed that there are infinite sets of absolutely avoidable pairs (m, f). One could further extend our results and provide more absolutely avoidable pairs.

The arguments in the proof of Proposition 6.5 should still hold if we deviate from $f_0 = \binom{m}{2}/2$ by a small term, as in Theorem 6.3. The reason here is that this change does not affect the limit computations for $d_q(m)$ and $y_q(m)$. Thus, for each large enough m, one should be able to obtain a small interval for f so that each (m, f) is absolutely avoidable. We cannot hope to do much better though: In infinitely many cases, if (m, f_0) is absolutely avoidable, then already for $(m, f_0 - m)$ or $(m, f_0 + m)$ our method does not give a contradiction. The constant 6 is the smallest integer for which the argument in the proof of Proposition 6.5 works (since $\{6\sqrt{2}\}$ is *close* to $\frac{1}{2}$ while $\{c\sqrt{2}\}$, $c \in [5]$ is not). We believe that one could show by an argument very similar to that used in the proof, that for sufficiently large m, for any constants a, b which satisfy that $\{a\sqrt{2} - b\sqrt{2}\}$ is *close enough* to $\frac{1}{2}$, we have that either $(m, f_0 - am)$ or $(m, f_0 - bm)$ is absolutely avoidable.

CHAPTER 7 BIPARTITE ORDER-SIZE PAIRS

7.1 INTRODUCTION

In Chapter 6 we were mainly concerned with the existence of absolutely avoidable pairs in graphs. A similar question on avoidable pairs can be asked in the bipartite setting. Recall that a *biclique* is a complete bipartite graph and a *bihole* is an empty bipartite graph, i.e. the bipartite complement of a biclique.

We say a bipartite graph G *bipartite arrows* the pair (m, f), and write $G \overset{bip}{\to} (m, f)$ if G has an induced subgraph with parts of size m each, contained in the respective parts of G, with exactly f edges. We say that a pair (n, e) of non-negative integers *bipartite arrows* the pair (m, f), written $(n, e) \overset{bip}{\to} (m, f)$ if for any bipartite graph G with parts of size n each and with e edges, $G \overset{bip}{\to} (m, f)$.

We say that a bipartite graph H *bipartite realises* a pair (m, f) if H has m vertices in each part and f edges. We also call H a bipartite (m, f)-graph. We call a pair (m, f) *absolutely bipartite avoidable* if there exists n_0, such that for each $n \geq n_0$ and for any $e \in \{0, \dots, n^2\}$, $(n, e) \overset{bip}{\not\to} (m, f)$. We define the *bipartite forcing density* of a bipartite order-size pair as $\sigma_{bip}(m, f) = \limsup\limits_{n \to \infty} \frac{\{e : (n,e) \overset{bip}{\to} (m,f)\}}{n^2}$.

In Section 7.2 we will show that the methods for showing the existence of absolutely avoidable pairs in the graph case from Chapter 6 are not extendable to the bipartite setting.

Proposition 7.1. *Let (m, f) be a bipartite order size pair. Then either (m, f) or its bipartite complementary pair $(m, m^2 - f)$ can be bipartite realised as the vertex-disjoint union of a biclique and a forest.*

Proposition 7.1 also appears in [16] together with most of the results from Chapter 6.

In Section 7.3 we will show that there is a family of three *unavoidable* bipartite graphs, one of which appears as an induced subgraph of any bipartite graph with sufficiently large order and sufficiently many edges and non-edges. We will connect this to avoidability of order-size pairs.

Proposition 7.2. *Let $m, f \in \mathbb{N}$ with $0 \leq f \leq m^2$. Then for all n sufficiently large there exists a positive $q = q_m$ and a number $\varphi_{n,t} \in O(n^{2-1/q})$ such that for all $e \in [\varphi_{n,m}, n^2 - \varphi_{n,m}]$, $(n, e) \overset{bip}{\to} (m, a \cdot m)$ for all $a \in [m]$.*

In Section 7.4 we will look at a specific pair, namely $(3, 4)$ and completely characterise which graphs (and hence pairs) bipartite arrow $(3, 4)$. In Section 7.5 we will use the results from the previous two sections to derive some results on the bipartite forcing density.

Proposition 7.3. *(a) There are infinitely many pairs (m, f) with $\sigma_{bip}(m, f) = 1$. In particular, for any $m \in \mathbb{N}, a \in [m]$, we have $\sigma_{bip}(m, am) = 1$ and $\sigma_{bip}(3, 4) = 1$.*

(b) There are infinitely many pairs (m, f) with $\sigma_{bip}(m, f) = 0$. In particular, for any $m, f \in \mathbb{N}$ with $m \geq 2$, $(m-1)m < f < m^2$ we have $\sigma_{bip}(m, f) = 0$.

In particular, we did not find any bipartite pair for which $\sigma_{bip}(m, f) \in (0, 1)$; we leave the existence of such a pair as an open question, see Section 7.6.

7.2 Realising bipartite order-size pairs as the vertex-disjoint unions of a biclique and a forest or its complement

Our entire argument for the existence of absolutely avoidable pairs in the graph setting (see Chapter 6) built on the fact that certain pairs (m, f) cannot be realised as the disjoint union of a clique and a forest. The following lemma shows that our argument for the existence of such absolutely avoidable pairs in the non-bipartite setting cannot be extended to the bipartite setting.

Note that a *biclique* is an induced subgraph of a complete bipartite graph, i.e. could be in particular an empty set or a single vertex.

Lemma 7.4. *For any positive integer m and any non-negative integer f, $f \leq \left\lfloor \frac{m^2}{2} \right\rfloor$, there is a bipartite graph H with m vertices in each part and f edges, which is the vertex-disjoint union of a biclique and a forest.*

Proof. Fix a pair (m, f) with $f \leq \left\lfloor \frac{m^2}{2} \right\rfloor$. Let $x = \left\lfloor \frac{m}{2} \right\rfloor$ and let y be the largest integer such that $xy \leq f$. In particular

$$xy > f - x \quad \text{and} \quad y \leq \left\lfloor \frac{m^2}{2} \right\rfloor / \left\lfloor \frac{m}{2} \right\rfloor.$$

We shall use the fact that for any non-negative integers v' and e', with $e' < v'$ and for any partition $v' = v'' + v'''$, with v'', v''' positive integers, there is a forest with partite sets of sizes v'' and v''' and e' edges. To see this, let $F = K_{1,v'-1}$ or $K_{v'-1,1}$ if $\min\{v'', v'''\} = 1$, or otherwise, let F be the vertex-disjoint union of $K_{1,v'''-1}$ and $K_{v''-1,1}$, connected by a single edge between two leaves. Then F is a tree with parts of sizes v'' and v''' with $v''+v'''-1=v'-1$ edges. Then any subgraph of this with $e<v'$ edges is a forest with desired part sizes.

Case 1: $y<m$.

If $y=0$ then $f<\lfloor\frac{m}{2}\rfloor$. In this case (m,f) is bipartite realisable as a forest. So, assume that $y>0$. We shall show that (m,f) is bipartite realisable as the vertex disjoint union of $K_{x,y}$ and a forest. Let $e'=f-xy$, $v'=2m-x-y$. We have that $e'\leq x-1=\lfloor\frac{m}{2}\rfloor-1$. On the other hand, using the upper bound on y, we have that $v'\geq 2m-\lfloor\frac{m}{2}\rfloor-\left(\left\lfloor\frac{m^2}{2}\right\rfloor/\lfloor\frac{m}{2}\rfloor\right)$. Considering the cases when m is even or odd, one can immediately verify that $e'<v'$. Since $x+y+v'=2m$ and $xy+e'=f$, we have that (m,f) is bipartite realisable as the vertex-disjoint union of $K_{x,y}$ and a forest on v' vertices and e' edges. Note that in this case we needed $y<m$ so that $K_{x,y}$ does not span one of the parts completely.

Case 2: $y=m$.

In particular, we have that $f\geq\lfloor\frac{m}{2}\rfloor m$. If m is even, we have that $f\geq m^2/2$ and from our original upper bound $f\leq m^2/2$ it follows that $f=m^2/2$. Thus (m,f) is bipartite representable as $K_{m/2,m}$ and isolated vertices. If m is odd, let $m=2k+1$, $k\geq 1$. Then $f\leq\left\lfloor\frac{m^2}{2}\right\rfloor=2k^2+2k$ and $f\geq y\lfloor\frac{m}{2}\rfloor=2k^2+k$. Consider $K_{k+1,2k-1}$ and let $e'=f-(k+1)(2k-1)$ and $v'=2m-3k$. Then $e'\leq 2k^2+2k-(2k^2+k-1)=k+1$ and $v'=4k+2-3k=k+2$. Thus, $v'>e'$. Therefore (m,f) is bipartite realisable as a vertex-disjoint union of $K_{k+1,2k-1}$ and a forest on v' vertices and e' edges.

Case 3: $y=m+1$.

This case could happen only if m is odd. Let $m=2k+1$. Then we have $x=k$ and $y=2k+2$ and $f=2k^2+2k$. We see that (m,f) is bipartite representable by $K_{2k,k+1}$ and isolated vertices. □

Proof of Proposition 7.1. If $f\leq\left\lfloor\frac{m^2}{2}\right\rfloor$, then by Lemma 7.4 we can bipartite realise (m,f) as the vertex-disjoint union of a biclique and a forest. Otherwise, we have $m^2-f\leq\left\lfloor\frac{m^2}{2}\right\rfloor$, so we can apply Lemma 7.4 to the complementary pair (m,m^2-f). In particular, we can bipartite realise (m,f) as the bipartite complement of the vertex-disjoint union of a biclique and a forest. □

7.3 Unavoidable bipartite patterns

In this section we will show that any $(n \times n)$ bipartite graph with positive edge-density in both G and G^c contains at least one of three induced *unavoidable* subgraphs. This result is inspired by the following result for the general graph case:

Theorem 7.5 (Cutler, Montagh [50])**.** *Let $\mathcal{F}_k$ be the family of 2-graphs on $2k$ vertices, which are isomorphic to either K_k and k isolated vertices, $K_{2k} - E(K_k)$, $2K_k$ or $K_{2k} - E(2K_k)$.*

Then for any $\epsilon > 0$ and positive integer k, there exists $n(k, \epsilon)$, such that any complete graph of order $n \geq n(k, \epsilon)$ and size e with $\epsilon\binom{n}{2} \leq e \leq (1-\epsilon)\binom{n}{2}$ contains a member of $\mathcal{F}_k$.

Theorem 7.5 was first proven by Cutler and Montagh [50] who showed that $n(k, \epsilon) < 4^{\epsilon/k}$. This was further improved by Fox and Sudakov [73] to $n(k, \epsilon) \leq \left(\frac{16}{\epsilon}\right)^{2k+1}$ for $\epsilon < \frac{1}{7}$, using probabilistic arguments. Caro, Hansberg and Montejano [37] reproved the theorem using only classical Ramsey and Turán numbers for bipartite graphs and the results on Zarankiewicz numbers.

This result on unavoidable patterns is also related to the notion of *balanceability* of graphs which was introduced by Caro, Hansberg and Montejano [37]. A graph H is called *balanceable* if in any 2-colouring of the edges of a large complete graph with "enough" edges in both colours, there exists a copy of H having exactly $\left\lfloor \frac{|E(H)|}{2} \right\rfloor$ edges of one colour. One might use Theorem 7.6 below to obtain bipartite balanceability results. For more on Zero-Sum problems and balanceability in the general graph case, see also Caro, Hansberg and Montejano [35, 36] and the survey by Caro [34].

Recall that for the symmetric Zarankiewicz number $z(n;t)$ we have $z(n;t) < (t-1)^{1/t}n^{2-1/t} + \frac{1}{2}(t-1)n$. We define the following bipartite graphs with $2t$ vertices in each part: A_t is isomorphic to $K_{2t,2t} - E(K_{t,2t})$, and B_t is isomorphic to $K_{2t,2t} - E(K_{t,t})$.

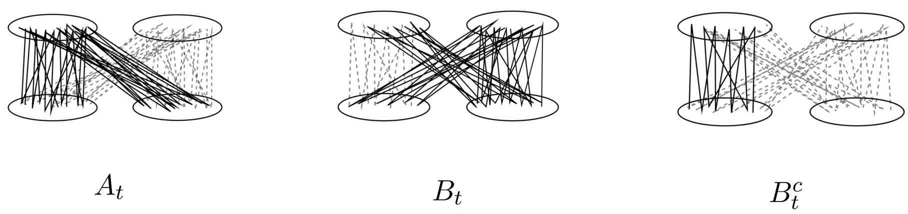


Figure 7.1: The bipartite graphs A_t, B_t and B_t^c with $2t$ vertices in each part

Theorem 7.6. *Let t be a positive integer. For all sufficiently large n there exists a positive $q = q_t$ and a number $\varphi = \varphi_{n,t} \in O(n^{2-\frac{1}{q}})$, such that any $(n \times n)$ bipartite graph G with*

$|E(G)| \in [\varphi_{n,t}, n^2 - \varphi_{n,t}]$ contains an induced copy of either A_t, B_t or B_t^c, the bipartite complement of B_t.

Proof. Let $q \geq t$ be an integer satisfying

$$\frac{1}{2}q^2 > (t-1)^{1/t}q^{2-1/t} + \frac{1}{2}(t-1)q,$$

and set

$$\varphi_{n,t} := (q-1)^{1/q}n^{2-1/q} + \frac{1}{2}(q-1)n + 2(n-q)q + 1,$$

which is clearly in $O(n^{2-1/q})$.

Let G be an $(n \times n)$ bipartite graph with $|E(G)| \in [\varphi_{n,t}, n^2 - \varphi_{n,t}]$ edges. Then by definition of $z(n;q)$, there is a copy of $K_{q,q}$ in G, let its parts be V_1, W_1.

Consider the $((n-q) \times (n-q))$ bipartite graph G_1, obtained by removing V_1 and W_1 from G. Then the bipartite complement G_1^c of G_1 has at least $E(G^c) - |V_1|(n - |W_1|) - |W_1|(n - |V_1|) \geq \varphi_{n,t} - 2(n-q)q > z(n;q) > z(n-q;q)$ edges, so we find $K_{q,q}$ in G_1^c. Denote the parts of this $K_{q,q}$ by V_2 and W_2.

Now consider the two $(q \times q)$ bipartite graphs $G_3 = G[V_1, W_2]$ and $G_4 = G[V_2, W_1]$. Each of them has either at least $\frac{1}{2}q^2$ edges or non-edges, so there is $K_{t,t}$ in G_i or G_i^c for $i = 3, 4$, i.e. either a biclique or a bihole. If we have two bicliques, we have a copy of B_t, if we have two biholes, we have a copy of B_t^c, and if we have a bihole and a biclique, we have a copy of A_t. □

Lemma 7.7. *If $G \overset{bip}{\to} (m,f)$ for some $G \in \{A_t, B_t, B_t^c\}$, we have $(n,e) \overset{bip}{\to} (m,f)$ for all n sufficiently large and all $e \in [\varphi_{n,m}, n^2 - \varphi_{n,m}]$.*

Proof. By Theorem 7.6, for pairs (n,e) as given, we have $(n,e) \overset{bip}{\to} G$ for some $G \in \{A_t, B_t, B_t^c\}$. Then by assumption, we have $(n,e) \overset{bip}{\to} G \overset{bip}{\to} (m,f)$. □

Clearly if $t \geq m$, the graph A_t contains all pairs (m,f) for which $f = a \cdot m$ and $a \in [m]$. These pairs are also induced by B_t and B_t^c, so we obtain Proposition 7.2 as a Corollary of Theorem 7.6 and Lemma 7.7.

7.4 A characterisation of graphs that bipartite arrow $(3,4)$

The goal of this section is to find all bipartite graphs G that bipartite arrow the pair $(3,4)$.

For $n > 7$ consider the following graph classes $\mathcal{G}_i = \mathcal{G}_i(n)$ of $(n \times n)$ bipartite graphs. We shall claim that these classes of graphs contain all graphs for which we have $G \overset{bip}{\nrightarrow} (3,4)$:

1. $\mathcal{G}_1 = \mathcal{G}_1(n) = \{P_4\text{+isolated vertices}\} \cup \{G \subseteq K_{n,n} : G \text{ is a union of pairwise vertex disjoint stars (with centres in the same part)}\}$.
   Then $G \in \mathcal{G}_1$ does not contain $(3,4)$: If $G = P_4$, then G has only 3 edges. Otherwise, all vertices in one part of G have degree at most 1, i.e. there cannot be $(3,4)$.
   For each $e \in \{0, \ldots, n\}$ there exists $G \in \mathcal{G}_1(n)$ with $|E(G)| = e$.

2. $\mathcal{G}_2 = \mathcal{G}_2(n) = \{G = K_{a,n} - \{e\} \subseteq K_{n,n} : 2 \leq a \leq n-1\}$, where e is an arbitrary edge of the $K_{a,n}$.
   Then $G \in \mathcal{G}_2$ does not contain $(3,4)$: In any (3×3) induced subgraph of G, there is at most one vertex of degree 2, any other vertex has degree 0 or 3. Thus, any (3×3) induced subgraph H of G has $|E(H)| \in \{0,2,3,5,6,8,9\}$ edges.
   For each $a \in \{2, \ldots, n-1\}$ there is $G \in \mathcal{G}_2(n)$ with $|E(G)| = an - 1$.

3. $\mathcal{G}_3 = \mathcal{G}_3(n) = \{K_{n,n} - E(K_{a,b}) : a, b \in [n]\}$.
   For $G \in \mathcal{G}_3$ we have $G^c = K_{a,b}+$ isolated vertices for some $a, b \in [n]$. This clearly does not contain $(3,5)$, and thus, G does not contain $(3,4)$.
   For any integers a, b with $0 \leq a \leq b \leq n$ there is $G \in \mathcal{G}_3(n)$ with $|E(G)| = n^2 - ab$.

4. $\mathcal{G}_4 = \mathcal{G}_4(n) = \{K_{n,n} - E(\bigcup_i H_i) : H_i \in \mathcal{H}_n, \bigcup_i H_i \subseteq K_{n,n}\}$, where $\mathcal{H}_n = \{C_6\} \cup \{H \subseteq K_{n,n} : H \text{ is a tree with at most 2 vertices in one part}\}$. Note that we can have one tree component in G^c with a part of size at least 3 in U and another tree with a part of size at least 3 in V.
   Then $G \in \mathcal{G}_4$ does not contain $(3,4)$: No component of G^c contains $(3,5)$, and since any $(3,5)$-graph is connected, G^c cannot contain $(3,5)$. In particular, the complement of any $G \in \mathcal{G}_4$ is the vertex-disjoint union of C_6's and a forest, so $|E(G)| \geq n^2 - 2n$.

We also define the following set:

$$\mathcal{E}(n) := \{e : e = |E(G)|,\ G \in \mathcal{G}_i,\ i = 1, \ldots, 4\}.$$

Now we can state our characterisation lemma:

Lemma 7.8. *Let G be a bipartite graph with n vertices in each part, $n > 7$. Then $G \overset{bip}{\nrightarrow} (3,4)$ if and only if $G \in \bigcup_{i=1}^{4} \mathcal{G}_i$. In particular, for $n > 7$ we have $(n,e) \overset{bip}{\to} (3,4)$ if and only if $e \notin \mathcal{E}(n)$.*

Before we prove this lemma, we need an auxiliary result:

Lemma 7.9. *Let G be a C_4-free $(n \times n)$ bipartite graph. Then either $G \overset{bip}{\to} (3,5)$ or $G^c \in \mathcal{G}_4(n)$.*

Proof. Let $G = (U \dot\cup V, E)$ be a C_4-free bipartite graph with $|U| = |V| = n$ and assume that $G \overset{bip}{\not\to} (3,5)$. We want to show that each connected component of G is either C_6 or a tree with at most 2 vertices in one part.

Let G' be a connected component of G.

Assume that G' is a tree. If G' has at most 2 vertices in one part, we are done, so assume G' has at least 3 vertices in each part. Then there must be a P_4 in G', otherwise G' would be a star, a contradiction to G' having at least 3 vertices in each part. Then this P_4 and two vertices adjacent to it, one from each part, induce a $(3,5)$-graph, a contradiction.

Consider a longest induced cycle in G'. If it has length at least 8, we find induced P_6, which is a $(3,5)$-graph, a contradiction. Since G' is C_4-free, it must contain induced C_6. If $G' = C_6$, we are done, so there must be a vertex u' incident to the $C_6 = u_1v_1u_2v_2u_3v_3$. Since G' is C_4-free, u' is incident to exactly one vertex of C_6, w.l.o.g. v_3. But then $\{u_1, u_2, u', v_1, v_2, v_3\}$ induces $(3,5)$, a contradiction. □

Now we can prove Lemma 7.8.

Proof of Lemma 7.8. Let $G = (U \dot\cup V, E)$ be a bipartite (n,e)-graph, and assume that $G \overset{bip}{\not\to} (3,4)$. We will show that $G \in G_i(n)$ for some $i \in [4]$.

Let $u_1 \in V(G)$ be a vertex with $d(u_1) = \Delta(G)$, w.l.o.g. $u_1 \in U$.

If $\Delta(G) = 1$, then G is a matching and isolated vertices, in particular, $G \in \mathcal{G}_1(n)$. Thus, we can assume that $d(u_1) = \Delta(G) \geq 2$.

Assume $N(u_1) \cap N(u) = \emptyset$ for all $u \in U \setminus \{u_1\}$. Assume there is a vertex $v \in V \setminus N(u_1)$ of degree $d(v) \geq 2$. But then we find induced $K_{1,2} \cup K_{2,1}$, which is a $(3,4)$-graph, a contradiction, see Figure 7.2 for an illustration. Thus, we have $d(v) \in \{0,1\}$ for all $v \in V \setminus N(u_1)$, so G is a vertex-disjoint union of stars with centres in U, and in particular, $G \in \mathcal{G}_1(n)$

Thus, there exists some $u \in U \setminus \{u_1\}$ such that $N(u_1) \cap N(u) \neq \emptyset$. Let u_2 be such that $|N(u_1) \cap N(u_2)|$ is maximal and with maximal degree among all such vertices.

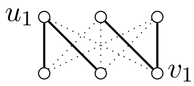
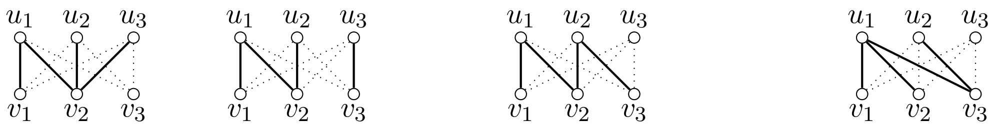


Figure 7.2: $K_{1,2} \cup K_{2,1}$

Assume that $|N(u_1) \cap N(u_2)| = 1$, i.e. there is no vertex in U that shares more than one neighbour with u_1. Note that since $d(u_1) \geq 2$, we have $|N(u_1) \setminus N(u_2)| \geq 1$.

- If $\Delta(G) = d(u_1) = 2$, there are $v_1, v_2 \in V$ such that v_1, u_1, v_2, u_2 induce a P_4. In particular, $d(u_2) \in \{1, 2\}$.
  - Assume $d(u_2) = 1$. If $d(u) = 0$ for all $u \in U \setminus \{u_1, u_2\}$, then G is P_4 and isolated vertices, in particular, $G \in \mathcal{G}_1(n)$. So let $u_3 \in U \setminus \{u_1, u_2\}$ with $d(u_3) > 0$.
    Then u_3 either has (exactly) one neighbour in $\{v_1, v_2\}$, w.l.o.g. v_2, then let $v_3 \in V \setminus \{v_1, v_2\}$, or it has no neighbour in $\{v_1, v_2\}$, then let $v_3 \in N(u_3)$. In either case, there is induced $(3, 4)$, see Figure 7.3a for an illustration.
  - If $d(u_2) = 2$, let $v_3 \in N(u_2) \setminus N(u_1)$. Then for $n \geq 5$ there is $u_3 \in U \setminus \{u_1, u_2\}$ with $N(u_3) \cap \{v_1, v_2, v_3\} = \emptyset$: since $\Delta(G) = 2$, the set $\{v_1, v_2, v_3\}$ has at most 2 more neighbours in $U \setminus \{u_1, u_2\}$, i.e. for $n \geq 5$, such a vertex u_3 exists, so we find $(3, 4)$, a contradiction. See Figure 7.3a for an illustration.
- If $\Delta(G) = d(u_1) = 3$, we find vertices $v_1, v_2 \in N(u_1) \setminus N(u_2)$ and $v_3 \in N(u_1) \cap N(u_2)$. Claim: For $n > 7$ there is a vertex $u_3 \in U \setminus \{u_1, u_2\}$ with $N(u_3) \cap \{v_1, v_2, v_3\} = \emptyset$. Since $\Delta(G) = 3$, we have $d(v_1) + d(v_2) + d(v_3) \leq 9$, so there are at most 5 vertices in $U \setminus \{u_1, u_2\}$ which are incident to $\{v_1, v_2, v_3\}$. Thus, if $|U| \geq 8$, u_3 exists. Then $\{u_1, u_2, u_3, v_1, v_2, v_3\}$ induces $(3, 4)$. See Figure 7.3b for an illustration.
- If $\Delta(G) = d(u_1) \geq 4$, i.e. we have $|N(u_1) \setminus N(u_2)| \geq 3$.
  Recall that each vertex $u \in U \setminus \{u_1\}$ has at most 1 neighbour in $N(u_1)$. Let $v_3 \in N(u_1) \cap N(u_2)$. Assume there is a vertex $u_3 \in U \setminus \{u_1, u_2\}$ with $v_3 \notin N(u_3)$. Since $|N(u_3) \cap N(u_1)| \leq 1$, there exist two vertices $v_1, v_2 \in (N(u_1) \setminus N(u_2)) \setminus N(u_3)$, i.e. $\{u_1, u_2, u_3, v_1, v_2, v_3\}$ induces $(3, 4)$, a contradiction. See Figure 7.3 for an illustration.
  Thus, we have $v_3 \in N(u)$ for all $u \in U$, i.e. $d(v_3) = n = d(u_1)$ and have $d(u) = d(v) = 1$ for $u, v \in U \cup V \setminus \{u_1, v_3\}$, i.e. $e = 2n - 1$ and $G^c = K_{n-1,n-1}$, i.e. $G \in \mathcal{G}_3(n)$.

Thus, we can assume that $|N(u_1) \cap N(u_2)| \geq 2$. Recall that $d(u_1) = \Delta(G)$ and that



$d(u_2) = 1$ $d(u_2) = 2$

(a) $d(u_1) = 2$ (b) $d(u_1) > 2$

Figure 7.3: $|N(u_1) \cap N(u_2)| = 1$

$u_2 \in U \setminus \{u_1\}$ is such that $|N(u_1) \cap N(u_2)|$ is maximal and with maximal degree among all such vertices.

Let $V_1 = N(u_1)$, $V_2 = N(u_2)$, $V_1' = V_1 \setminus V_2$ and $V_2' = V_2 \setminus V_1$. We split the remaining proof into the following 4 cases:

Case 1: $\exists u \in U \setminus \{u_1, u_2\} : |(V_1' \cup V_2') \setminus N(u)| \geq 2$ and $|(V_1 \cap V_2) \setminus N(u)| \geq 1$,

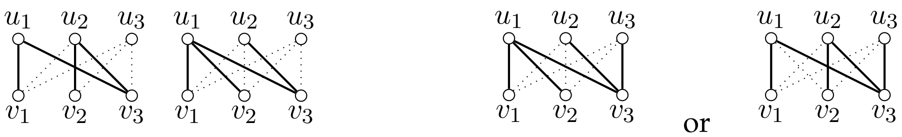
Case 2: $\exists u \in U \setminus \{u_1, u_2\} : |(V_1' \cup V_2') \setminus N(u)| \geq 3$,

Case 3: $\exists u \in U \setminus \{u_1, u_2\} : |(V_1 \cap V_2) \setminus N(u)| \geq 3$,

Case 4: $\forall u \in U \setminus \{u_1, u_2\} : |(V_1 \cup V_2) \setminus N(u)| \leq 2$.

In particular we will show that Case 1 is impossible, in Case 2 we have $G \in \mathcal{G}_3$, in Case 3.1 we have $G \in \mathcal{G}_2$, and in Cases 3.2 and 4 we have $G \in \mathcal{G}_4$.

Case 1: There is $u_3 \in U \setminus \{u_1, u_2\}$ with $|(V_1' \cup V_2') \setminus N(u_3)| \geq 2$ and $|(V_1 \cap V_2) \setminus N(u_3)| \geq 1$. Then let $v_1, v_2 \in (V_1' \cup V_2') \setminus N(u_3)$ and $v_3 \in (V_1 \cap V_2) \setminus N(u_3)$. Then $\{u_1, u_2, u_3, v_1, v_2, v_3\}$ induces $(3, 4)$. See Figure 7.4a for an illustration.

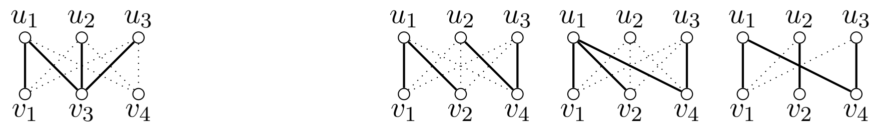


(a) Case 1 (b) Case 2

Figure 7.4: Cases 1 and 2

Case 2: There is $u_3 \in U \setminus \{u_1, u_2\}$ with $|(V_1' \cup V_2') \setminus N(u_3)| \geq 3$. Since we are not in Case 1, $(V_1 \cap V_2) \setminus N(u_3) = \emptyset$. Let $v_1, v_2 \in (V_1' \cup V_2') \setminus N(u_3)$, $v_3 \in V_1 \cap V_2$. See Figure 7.4b for an illustration.

Assume there is $v_4 \in V \setminus (V_1 \cup V_2)$ with $v_4 u_3 \notin E(G)$. Then $\{u_1, u_2, u_3, v_1, v_3, v_4\}$ induces $(3, 4)$. Assume there is a vertex $v_4 \in V_1' \cup V_2'$ with $v_4 u_3 \in E(G)$. Then $\{u_1, u_2, u_3, v_1, v_2, v_4\}$ induces $(3, 4)$. See Figure 7.5 for an illustration.



$\exists v_4 \in V \setminus (V_1 \cup V_2) : u_3v_4 \notin E(G)$ $\quad$ $\exists v_4 \in V_1' \cup V_2' : u_3v_4 \in E(G)$

Figure 7.5: Case 2 continued

Thus, $N(u_3) = V \setminus (V_1' \cup V_2')$, i.e. $V = V_1' \dot\cup V_2' \dot\cup N(u_3) \dot\cup V_1 \cap V_2$ and any vertex in V is either in exactly one or in all three neighbourhoods $N(u_1), N(u_2), N(u_3)$.

Let $u_4 \in U \setminus \{u_1, u_2, u_3\}$. Assume u_4 has a non-neighbour $v_3 \in V_1 \cap V_2$. Since we are not in Case 1, u_4 has at most 1 non-neighbour in $V_1' \cup V_2'$. Recall that $|V_1'| + |V_2'| \geq 3$, i.e. $|N(u_4) \cap (V_1' \cup V_2')| \geq 2$. If $|N(u_4) \cap V_1'| \geq 2$ or $|N(u_4) \cap V_2'| \geq 2$, w.l.o.g. the former, let $v_1, v_2 \in N(u_4) \cap V_1'$. Then $\{u_2, u_3, u_4, v_1, v_2, v_3\}$ induces $(3, 4)$. Thus, we have $|N(u_4) \cap V_1'| = |N(u_4) \cap V_2'| = 1$. Since $|V_1' \cup V_2'| \geq 3$, w.l.o.g. there is $v_1 \in V_1'$ with $v_1u_4 \notin E(G)$. Let $v_2 \in N(u_4) \cap V_2'$. Then $\{u_2, u_3, u_4, v_1, v_2, v_3\}$ induces $(3, 4)$. See Figure 7.6

$|N(u_4) \cap V_1'| \geq 2$ $\quad$ $|N(u_4) \cap V_1'| = |N(u_4) \cap V_2'| = 1$

Figure 7.6: Case 2 continued

Thus, $V_1 \cap V_2 \subseteq N(u)$ for all $u \in U$, and thus, since $V_1 \cap V_2 \neq \emptyset$, we have $n = \Delta(G) = d(u_1)$. In particular, we have $V = (V_1 \cap V_2) \dot\cup V_1'$, and $G = K_{n,n} - E(K_{n-1,|V_1'|})$, i.e $G \in \mathcal{G}_3(n)$.

Case 3: There is $u_3 \in U \setminus \{u_1, u_2\}$ with $|(V_1 \cap V_2) \setminus N(u_3)| \geq 3$.

Since we are not in Case 1, u_3 is incident to all vertices in $V_1' \cup V_2'$.

We partition V into four sets $V = V_1' \dot\cup V_2' \dot\cup (V_1 \cap V_2) \dot\cup V'$. Note that $v'u_3 \in E(G)$ for all $v' \in V'$: assume not. Then there exists $v' \in V'$ with $vu_i \notin E(G)$ for $i = 1, 2, 3$, and since u_3 has two non-neighbours $v_1, v_2 \in V_1 \cap V_2$, $\{u_1, u_2, u_3, v_1, v_2, v'\}$ induces $(3, 4)$.

Claim 0: $|V'| \leq 1$.
Assume there exist $v', v'' \in V'$. Then $u_3v', u_3v'' \in E(G)$, i.e. $\{u_1, u_2, u_3, v_1, v', v''\}$ induces $(3, 4)$ (with $v_1 \in (V_1 \cap V_2) \setminus N(u_3)$), a contradiction. ■

Consider the bipartite complement G^c of G. We will find a $(3, 5)$-graph in G^c which corresponds to a $(3, 4)$-graph in G.

Case 3.1: There is a C_4 $u_3v_3u_4v_4u_3$ in G^c with $u_3, u_4 \in U \setminus \{u_1, u_2\}$ and $v_3, v_4 \in V_1 \cap V_2$.

Claim 1: $|V'| = \emptyset$.

Assume there is $v' \in V'$. If $u_3v' \in E(G^c)$ (or u_4v'), then $\{u_1, u_2, u_3, v', v_3, v_4\}$ induces $(3,5)$ in G^c. If both $u_3v', u_4v' \notin E(G^c)$, then $\{u_1, u_3, u_4, v', u_3, u_4\}$ induces $(3,5)$ in G^c, a contradiction. ■

Claim 2: $|V_1' \cup V_2'| \leq 1$. We have $vu_3, vu_4 \in E(G^c)$ for $v \in V_1' \cup V_2'$.

Let $v_1 \in V_1' \cup V_2'$ (w.l.o.g. $v_1 \in V_1'$). Then v_1 is incident to at least one of $\{u_3, u_4\}$ in G^c (else $\{u_2, u_3, u_4, v_1, v_3, v_4\}$ induces $K_2 \cup C_4$, a $(3,5)$-graph, a contradiction), not incident to only one of $\{u_3, u_4\}$ (else $\{u_1, u_3, u_4, v_1, v_3, v_4\}$ induces C_4 with a pendant edge, a $(3,5)$-graph, a contradiction). Thus, v_1 is incident to both $\{u_3, u_4\}$. Assume there is a second vertex $v_2 \in V_1' \cup V_2'$. Then by the same argument it is also incident to both $\{u_2, u_3\}$. Then $\{u_1, u_2, u_3, v_1, v_2, v_3\}$ induces $(3,5)$ in G^c, a contradiction. ■

Thus, we have $|V_1 \cap V_2| \geq n-1$ and $0 = d_{G^c}(u_1) \leq d_{G^c}(u_2) \leq 1$, i.e. $0 = |V_2'| \leq |V_1'| \leq 1$.

Claim 3: $N(u_3) = N(u_4)$.

Assume not. Then w.l.o.g. there is $v_5 \in V = (V_1 \cap V_2) \cup V_1'$, such that $u_3v_5 \in E(G^c)$ and $u_4v_5 \notin E(G^c)$. Then $\{u_1, u_3, u_4, v_3, v_4, v_5\}$ induces $(3,5)$ in G^c, a contradiction. ■

Let $u_5 \in U \setminus \{u_1, u_2, u_3, u_4\}$.

Claim 4: If $|N_{G^c}(u_5) \cap N_{G^c}(u_3)| \geq 2$, then $N(u_3) = N(u_5)$:

Assume not, i.e. w.l.o.g. we have $v, v', v'' \in N_{G^c}(u_3)$ such that $u_5v, u_5v' \in E(G^c)$ and $u_5v'' \notin E(G^c)$. Then $\{u_1, u_3, u_5, v, v', v''\}$ induces $(3,5)$ in G^c, a contradiction. ■

If $|N_{G^c}(u_5) \cap N_{G^c}(u_3)| \geq 2$, by Claim 4 we have $N(u_3) = N(u_5)$.

If $N_{G^c}(u_5) \cap N_{G^c}(u_3) = \emptyset$, then $d_{G^c}(u_5) = 0$:

Otherwise there exists $v_5 \in N(u_5) \setminus N(u_3)$, i.e. $\{u_3, u_4, u_5, v_3, v_4, v_5\}$ induces $K_2 \cup C_4$, i.e. $(3,5)$ in G^c, a contradiction.

If $|N(u_5) \cap N(u_3)| = 1$, (and hence $|N(u_3) \setminus N(u_5)| \geq 1$), then $V \setminus N(u_3) \subseteq N(u_5)$:

Otherwise there is $v_5 \in V \setminus (N(u_3) \cup N(u_5))$ and thus, $\{u_3, u_4, u_5, v_3, v_4, v_5\}$ induces $(3,5)$ in G^c (where w.l.o.g. $u_5v_3 \in E(G^c)$ and $u_5v_4 \notin E(G^c)$).

Note that we must have $d_{G^c}(u_5) \leq 2$, since otherwise there are two vertices $v_5, v_6 \in N(u_5) \setminus N(u_3)$, i.e. $\{u_3, u_4, u_5, v_3, v_5, v_6\}$ induces $(3,5)$. Thus, we have $|N(u_3)| \geq |V| - 1$. Now assume $d(u_5) = 2$. Then since $n \geq 4$, we have $|N(u_3)| \geq 3$, and thus, there exist $v_1, v_2 \in N(u_3) \setminus N(u_5)$ and $v_3 \in N(u_5) \setminus N(u_3)$, i.e. $\{u_3, u_4, u_5, v_1, v_2, v_3\}$ induces $C_4 \cup K_2$, i.e. $(3,5)$. Thus, if

$|N(u_5) \cap N(u_3)| = 1$ for some $u_5 \in U \setminus \{u_1, u_2, u_3, u_4\}$, then $N(u_3) = V$.

Assume there is a second vertex $u_6 \in U \setminus \{u_1, u_2, u_3, u_4, u_5\}$ with $|N(u_6) \cap N(u_3)| = 1$. Then we find a $(3, 5)$ with $\{u_3, u_5, u_6\}$, the neighbour(s) of u_5, u_6 and an arbitrary other vertex in $V = N(u_3)$. Thus, there exists at most one vertex with this property.

In conclusion, if no vertex in U has degree 1 in $N_{G^c}(u_3)$, for all vertices $u \in U$ we have $N_{G^c}(u) \in \{\emptyset, N(u_3)\}$. Thus, $G^c = K_{a,n}$ for some $a \geq 2$ is a complete bipartite graph, i.e $G \in \mathcal{G}_3(n)$. Otherwise, there is exactly one vertex in U with degree 1 in $N_{G^c}(u_3)$ we have that $G^c = K_{a,n}$ with a pendant edge, i.e. $G = K_{b,n} - \{e\}$ for some a with $2 \leq a \leq n-1$, so $G \in \mathcal{G}_2(n)$.

Case 3.2: There is no C_4 in $G^c[U \setminus \{u_1, u_2\}, V_1 \cap V_2]$.

We will show that then there is no C_4 in G^c.

There cannot be a C_4 in G^c containing both u_1 and u_2, since u_1 and u_2 share at most one neighbour (By Claim 0, $|V'| \leq 1$).

Assume there is a $C_4 = u_3v_1u_4v_2$ in G^c with $u_3, u_4 \in U \setminus \{u_1, u_2\}$.

Assume $v_2 \in V_1 \cap V_2$. Then we must have $v_1 \in V' \cup V_1' \cup V_2'$. Note that then no vertex $v_3 \in V_1 \cap V_2 \setminus \{v_2\}$ is incident to both u_3, u_4, since otherwise $\{u_3, u_4, v_2, v_3\}$ is a C_4, a contradiction.

If $v_1 \in V'$, let $v_3 \in (V_1 \cap V_2) \setminus \{v_2\}$. If v_3 is incident to none of u_3, u_4, then $\{u_1, u_3, u_4, v_1, v_2, v_3\}$ induces $(3, 5)$ in G^c, a contradiction. Thus, v_3 is incident to exactly one of u_3, u_4, say $v_3u_3 \in E(G^c)$. Then $\{u_1, u_2, u_3, v_1, v_2, v_3\}$ induces $(3, 5)$ in G^c, a contradiction.

If $v_1 \in V_1' \cup V_2'$, w.l.o.g. $v_1 \in V_1'$: If there is $v_3 \in V_1 \cap V_2$ incident to none of u_2, u_3, then $\{u_2, u_3, u_4, v_1, v_2, v_3\}$ induces $(3, 5)$. Otherwise, there is $v_3 \in V_1 \cap V_2$ incident to exactly one of $\{u_3, u_4\}$, say $u_3v_3 \in E(G^c)$, then $\{u_1, u_3, u_4, v_1, v_2, v_3\}$ induces $(3, 5)$ in G^c, a contradiction.

So assume $v_1, v_2 \in V_1' \cup V_2' \cup V'$, and there is no $v \in V_1 \cap V_2$ which is incident to both u_3 and u_4. If there is $v_3 \in V_1 \cap V_2$ incident to none of u_3, u_4, then w.l.o.g. u_1 has exactly one neighbour in v_1, v_2, so $\{u_1, u_2, u_3, v_1, v_2, v_3\}$ induces $(3, 5)$. Otherwise all $v \in V_1 \cap V_2$ are incident to exactly one of u_3, u_4, pick $v_3, v_4 \in V_1 \cap V_2$. Then w.l.o.g. u_1 is incident to v_1, so $\{u_1, u_2, u_3, v_1, v_3, v_4\}$ induces $(3, 5)$, a contradiction.

Thus, if there is a C_4 in G^c, it contains exactly one of u_1, u_2, so it consists w.l.o.g. of $u_1, u_3 \in U$ and $v_1, v_2 \in V_2' \cup V'$.

Assume that $v_1 \in V'$ and $v_2 \in V_2'$. Then u_3 is incident to every vertex in $V_1 \cap V_2$ in G^c (otherwise there is $v_3 \in V_1 \cap V_2 \setminus N_{G^c}(u_3)$, i.e. $\{u_1, u_2, u_3, v_1, v_2, v_3\}$ induces $(3, 5)$in G^c). Since $|V_1 \cap V_2| \geq 2$, there are $v_3, v_4 \in V_1 \cap V_2$. Then $\{u_1, u_2, u_3, v_1, v_3, v_4\}$ induces $(3, 5)$.

Thus, both $v_1, v_2 \in V_2'$. Then $N_{G^c}(u_3) \cap (V_1 \cap V_2) = \emptyset$: otherwise there exists $v_3 \in N_{G^c}(u_3) \cap (V_1 \cap V_2)$, i.e. $\{u_1, u_2, u_3, v_1, v_2, v_3\}$ induces $(3, 5)$.

Let $u_4 \in U \setminus \{u_1, u_2, u_3\}$. Then u_4 has at most 1 neighbour in $\{v_1, v_2\}$ (else we have a C_4 on $\{u_3, u_4, v_1, v_2\}$, a contradiction). Assume u_4 has exactly one neighbour in $\{v_1, v_2\}$, say v_1. If it has a non-neighbour v_3 in $V_1 \cap V_2$, then $\{u_1, u_3, u_4, v_1, v_2, v_3\}$ induces $(3, 5)$, i.e. v_3 is incident to all vertices in $V_1 \cap V_2$. Let $v_3, v_4 \in V_1 \cap V_2$. Then $\{u_1, u_3, u_4, v_1, v_3, v_4\}$ induces $(3, 5)$. So assume u_4 has no neighbour in $\{v_1, v_2\}$. If it has at least one neighbour in $v_3 \in V_1 \cap V_2$, then $\{u_1, u_3, u_4, v_1, v_2, v_3\}$ induces $(3, 5)$ in G^c. Thus, no vertex $u \in U$ has any neighbours in $V_1 \cap V_2$, so $\delta(G^c) = 0$. This is a contradiction to $\delta(G^c) = d(u_1) \geq 2$.

Thus, there is no C_4 in G^c, so by Lemma 7.9, we have $G \in \mathcal{G}_4(n)$.

Case 4: For all $u \in U \setminus \{u_1, u_2\}$ we have $|(V_1 \cup V_2) \setminus N(u)| \leq 2$.

That means in particular, that $d(u_1) \geq d(u) \geq d(u_1) - 2$ for all $u \in U$.

Thus, there are at least $d(u_1)n - 2(n-1)$ edges between U and V_1, i.e. the maximum degree in V_1 is at least $n - \frac{2(n-1)}{d(u_1)}$. Since $d(u_1) = \Delta(G)$, we must have $n - \frac{2(n-1)}{d(u_1)} \leq d(u_1)$, i.e.

$$d(u_1)^2 - nd(u_1) \geq -2n + 2 \quad \Leftrightarrow \quad \left(d(u_1) - \frac{n}{2}\right)^2 \geq \left(\frac{n}{2}\right)^2 - 2n + 2 = \left(\frac{n}{2} - 2\right)^2 - 2$$

and in particular, we either have

$$d(u_1) \geq \frac{n}{2} + \sqrt{((n/2) - 2)^2 - 2} > \frac{n}{2} + \sqrt{((n/2) - 3)^2} = n - 3 \text{ (for } n > 7\text{)},$$

or

$$d(u_1) \leq \frac{n}{2} - \sqrt{((n/2) - 2)^2 - 2} < \frac{n}{2} - \frac{n}{2} + 3 = 3.$$

In the second case, we have $\Delta(G) = d(u_1) = 2$, and u_1, u_2, v_1, v_2 induce a C_4 disjoint from the remaining graph. Since $|U \setminus \{u_1, u_2\}| \geq 3$, there must be $u_3 \in U \setminus \{u_1, u_2\}$ and $v_3 \in V \setminus \{v_1, v_2\}$ with $uv \notin E(G)$. But then $\{u_1, u_2, u_3, v_1, v_2, v_3\}$ induces $(3, 4)$, a contradiction. Thus, we have $d(u_1) \in \{n-2, n-1, n\}$.

Case 4.1: $d(u_1) = n$.

Then G has at least $n + (n-1)(n-2) = n^2 - 2n + 2$ edges. In particular we have $0 = d_{G^c}(u_1) \leq d_{G^c}(u_2) \leq d_{G^c}(u) \leq 2$. Assume there is a C_4 $u_3v_3u_4v_4$ in G^c. Then any vertex $u \in U$ has $N(u) \in \{\emptyset, \{v_3, v_4\}\}$, since otherwise there is $(3, 5)$, i.e. $G^c = K_{2,a}$ for some $a \in [n-1]$, i.e. $G \in \mathcal{G}_3(n)$.

If there is no C_4 in G^c, by Lemma 7.9 we have $G \in \mathcal{G}_4(n)$.

Case 4.2: $d(u_1) = n - 1$.

We consider the complement G^c. Then $d(u_1) = 1$ and $d(u) \in \{1, 2, 3\}$ for $u \in U$. Assume there is a $C_4 = u_2v_2u_3v_3$ in G^c.

If $N(u_1) \cap \{v_2, v_3\} \neq \emptyset$, then $\{u_1, u_2, u_3, v_2, v_3\}$ and any vertex in $V \setminus (N(u_2) \cup N(u_3))$ (which exists, since $|V| \geq 5$) induce C_4 with a pendant edge, i.e. a $(3, 5)$-graph, a contradiction.

If $(N(u_2) \cup N(u_3)) \cap N(u_1) = \emptyset$, then $\{u_1, u_2, u_3, v_2, v_3\} \cup N(u_1)$ induces $K_2 \cup C_4$, i.e. a $(3, 5)$-graph, a contradiction.

Let v_1 be the neighbour of u_1. Then w.l.o.g. $u_2v_1 \in E$ and $u_3v_1 \notin E$.

Now let $u_4 \in U \setminus \{u_1, u_2, u_3\}$.

If u_4 has no neighbour in $\{v_1, v_2, v_3\}$, then $\{u_2, u_3, u_4, v_1, v_2, v_3\}$ induced C_4 and a pendant edge, i.e. a $(3, 5)$-graph, a contradiction.

So u_4 has at least one neighbour in $\{v_1, v_2, v_3\}$. Note that since $d_{G^c}(u) \leq 3$ for all $u \in U$ and $n > 7$, there must exist a vertex $v_4 \in V \setminus (\bigcup\limits_{i \in \{1,2,3,4\}} N(u_i))$

Assume u_4 has exactly one neighbour in $\{v_1, v_2, v_3\}$. If $u_4v_1 \in E(G^c)$, then $\{u_1, u_2, u_4, v_1, v_2, v_3\}$ induces $(3, 5)$, and if w.l.o.g. $u_4v_2 \in E(G^c)$, then $\{u_2, u_3, u_4, v_2, v_3, v_4\}$ induces $(3, 5)$, a contradiction.

Thus, u_4 has at least 2 neighbours in $\{v_1, v_2, v_3\}$.

If $\{v_1, v_2\} \subseteq N(u_4)$, then $\{u_1, u_2, u_4, v_1, v_2, v_4\}$ induces $(3, 5)$. If $\{v_1, v_3\} \subseteq N(u_4)$, then $\{u_1, u_2, u_3, v_1, v_3, v_4\}$ induces $(3, 5)$. So $N(u_3) \cap \{v_1, v_2, v_3\} = \{v_2, v_3\}$. Then $\{u_1, u_3, u_4, v_1, v_2, v_3\}$ induces $K_2 \cup C_4$, a $(3, 5)$-graph. Thus, in either case we have a contradiction.

Thus, G^c does not contain C_4, so by Lemma 7.9 we have $G \in \mathcal{G}_4(n)$.

Case 4.3: $d(u_1) = n - 2$.

Consider the bipartite complement G^c. Let v_1, v_2 be the neighbours of u_1 in G^c. Note that in G^c every vertex in U has degree $\in \{2, 3, 4\}$, while in V every vertex has degree ≥ 2. Then since $\delta(G^c) = 2$, both v_1 and v_2 have at least one more neighbour. Assume v_1, v_2 have a common neighbour, say $u_2 \in U \setminus \{u_1\}$, choose u_2 of maximum degree with that property.

- $\underline{d_{G^c}(u_2) = 2:}$ Let $u_3 \in U \setminus \{u_1, u_2\}$.
  Assume $N(u_3) \cap \{v_1, v_2\} = \emptyset$. Since $d(u_3) \geq 2$, there exists $v_3 \in N(u_3)$, i.e. $\{u_1, u_2, u_3, v_1, v_2, v_3\}$ induces $(3, 5)$. Assume $|N(u_3) \cap \{v_1, v_2\}| = 1$, say $v_2 \in N(u_3)$. Then there exists $v_3 \in V \setminus \{v_1, v_2\}$ with $u_3v_3 \notin E(G^c)$, since $d(u_3) \leq 4$ and $|V| \geq 6$), so $\{u_1, u_2, u_3, v_1, v_2, v_4\}$ induces $(3, 5)$. Thus, $\{v_1, v_2\} \subseteq N(u)$ for all $u \in U$ and $d(u_2) \geq d(u)$, so we have $N(u) = \{v_1, v_2\}$ for all $u \in U$, and hence $d(v) = 0$ for all $v \in V \setminus \{v_1, v_2\}$, a contradiction to $\delta(G^c) \geq 2$.

- $\underline{d_{G^c}(u_2) = 3\text{:}}$ Let $N(u_2) = \{v_1, v_2, v_3\}$. Assume there exists $u_3 \in U \setminus \{u_1, u_2\}$ with $v_3 \in N(u_3)$.
If $|N(u_3) \cap \{v_1, v_2\}| = 0$, then let $v_4 \in N(u_3) \setminus \{v_3\}$. Then $\{u_1, u_2, u_3, v_1, v_2, v_4\}$ induces $(3, 5)$. If $|N(u_3) \cap \{v_1, v_2\}| \geq 1$, w.l.o.g. $v_2 \in N(u_3)$, then let $v_4 \in V \setminus (N(u_2) \cup N(u_3))$, which exists for $n \geq 6$. Then $\{u_1, u_2, u_3, v_2, v_3, v_4\}$ induces $(3, 5)$.
Thus, for all vertices $u \in U \setminus \{u_2\}$ we have $uv_3 \notin E(G^c)$, i.e. $d_{G^c}(v_3) = 1$, a contradiction to $\delta(G^c) \geq 2$.
- $\underline{d_{G^c}(u_2) = 4\text{:}}$ Let $N(u_2) = \{v_1, v_2, v_3, v_4\}$ and let $u_3 \in U \setminus \{u_1, u_2\}$.
If $|N(u_3) \cap N(u_2)| \leq 1$, w.l.o.g. $v_2, v_3 \notin N(u_3)$. If $v_1 \in N(u_3)$, then $\{u_1, u_2, u_3, v_1, v_3, v_4\}$ induces $(3, 5)$, otherwise $\{u_1, u_2, u_3, v_1, v_2, v_3\}$ induces $(3, 5)$. If $|N(u_3) \cap N(u_2)| \geq 3$, w.l.o.g. we have $v_2u_3, v_3u_3 \in E(G^c)$, so for $v_4 \in V \setminus (N(u_3) \cup N(u_2))$, the set $\{u_1, u_2, u_3, v_2, v_3, v_4\}$ induces $(3, 5)$.
Thus, we have $|N(u_3) \cap N(u_2)| = 2$. If $N(u_3) \cap N(u_2) = \{v_1, v_2\}$, then $\{u_1, u_2, u_3, v_1, v_3, v_4\}$ induces $(3, 5)$. If $|N(u_3) \cap \{v_1, v_2\}| = 1 = |N(u_3) \cap \{v_3, v_4\}$, say $\{v_1, v_4\} \in N(u_3)$, then let $v_5 \in V \setminus (N(u_3) \cup N(u_2))$, so $\{u_1, u_2, u_3, v_1, v_2, v_5\}$ induces $(3, 5)$.
Thus, all vertices $u \in U \setminus \{u_1, u_2\}$ satisfy $N(u) \cap N(u_2) = \{v_3, v_4\}$. But then for $u_3, u_4 \in U \setminus \{u_1, u_2\}$, the set $\{u_1, v_1, u_3, u_4, v_3, v_4\}$ induces $(3, 5)$.
Thus, there is no vertex $u \in U \setminus \{u_1\}$ with $\{v_1, v_2\} \in N(u)$. Pick vertices $u_2, u_3 \in U$ with $v_1u_2, v_2u_3 \in E(G^c)$.
Assume there is a C_4 in G^c with vertex set $C = \{u_2, u_3, v_3, v_4\}$. Then $u_1 \notin C$ and $|\{v_1, v_2\} \cap \{v_3, v_4\}| \leq 1$. If $\{v_1, v_2\} \cap \{v_3, v_4\} = \emptyset$, then $\{u_1, u_2, u_3, v_1, v_3, v_4\}$ induces $(3, 5)$. If $\{v_1, v_2\} \cap \{v_3, v_4\} \neq \emptyset$, say $v_2 = v_3$, then let $v_5 \in V \setminus \{N(u_1) \cup N(u_2) \cup N(u_3)\}$ ($\neq \emptyset$ for $n \geq 8$), then $\{u_1, u_2, u_3, v_3, v_4, v_5\}$ induces $(3, 5)$. Thus, G^c is C_4-free, so we can apply Lemma 7.9, and in particular, $G \in \mathcal{G}_4(n)$. $\square$

7.5 DENSITY OBSERVATIONS IN THE BIPARTITE SETTING

Consider the class of $(n \times n)$ bipartite graphs which are the vertex-disjoint union of bicliques and isolated vertices, i.e. $\mathcal{M}_n = \{\bigcup K_{a_i, b_i} : a_i, b_i \in [n], \sum a_i \leq n, \sum b_i \leq n\}$. Then the set of sizes of graphs in $\mathcal{M}_n$ is given by

$$T(n) = \{|E(G)| : G \in \mathcal{M}_n\} = \left\{\sum a_i b_i : a_i, b_i \in [n], \sum a_i \leq n, \sum b_i \leq n\right\}.$$

Lemma 7.10. *For any $c \in [0, 1)$, there exists an $n \geq n_0$ such that $T(n) \geq cn^2$.*

Proof. We will show that for any e with $0 \leq e \leq cn^2$ there exists a graph in $\mathcal{M}_n$ with exactly e edges.

Let $e \leq cn^2$ and let x be an integer such that $x^2 \leq e < (x+1)^2$.

We want to realise the pair (n, e) as a union $K \cup K' \cup K_{1,\ell}$ and isolated vertices, with $K \in \{K_{x,x}, K_{x,(x+1)}\}$ and $K' = K_{y,y}$ for some suitable $y \leq \lfloor \sqrt{x} \rfloor$ and $\ell \leq 2y + 1$.

If $e < x(x+1)$, let $K = K_{x,x}$. Otherwise, let $K = K_{x,x+1}$. Then $e - |E(K)| \leq x$, so we need to show that for any f with $0 \leq f \leq x$ we can find a graph in $\mathcal{M}_{n-x-1}$ with exactly f edges. Fix some f with $0 \leq f \leq x$.

Let $y = \lfloor \sqrt{f} \rfloor$, let $K' = K_{y,y}$ and let $\ell := e - E(K) - E(K') = f - y^2$. Recall that $\{y\}$ denotes the fractional part of y. Thus, we have

$$\ell = f - (\sqrt{f} - \{\sqrt{f}\})^2 = 2\sqrt{f}\underbrace{\{\sqrt{f}\}}_{\in(0,1)} - \underbrace{\{\sqrt{f}\}}_{\in(0,1)}^2 < 2\sqrt{f} \leq 2\sqrt{x}.$$

Thus, if we can show that $2\sqrt{x} \leq n - x - 1 - y$, we can realise the remaining $\ell \leq 2\sqrt{x}$ edges as $K_{1,\ell}$. Equivalently, we need to show that $n \geq x + 2\sqrt{x} + y + 1$. We have

$$x + 2\sqrt{x} + y + 1 \leq x + 3\sqrt{x} + 1 \leq e^{\frac{1}{2}} + 3e^{\frac{1}{4}} + 1 \leq c^{\frac{1}{2}}n + 3c^{\frac{1}{4}}\sqrt{n} + 1.$$

Note that depending on c, there exists n_0 such that $c^{\frac{1}{2}}n + 3c^{\frac{1}{4}}\sqrt{n} + 1 \leq n$ for all $n \geq n_0$.

Thus, for $n \geq n_0$, we can realise all pairs (n, e) with $0 \leq e \leq cn^2$ as $K \cup K' \cup K_{1,\ell} \in \mathcal{M}_n$ with K, K', ℓ defined as above. □

Lemma 7.11. *If for some bipartite pair (m, f) we have $G \overset{bip}{\nrightarrow} (m, f)$ for all $G \in \mathcal{M}_n$, then $\sigma_{bip}(m, f) = 0$.*

Proof. Assume we have $G \overset{bip}{\nrightarrow} (m, f)$ for all $G \in \mathcal{M}_n$. Then by the definition of the bipartite forcing density, we obtain for all $c \in [0, 1)$:

$$\begin{aligned}
\sigma_{bip}(m, f) &= \limsup_{n\to\infty} \frac{|\{e : (n, e) \overset{bip}{\rightarrow} (m, f)\}|}{n^2} = \limsup_{n\to\infty} \frac{n^2 - |\{e : (n, e) \overset{bip}{\nrightarrow} (m, f)\}|}{n^2} \\
&\leq 1 - \limsup_{n\to\infty} \frac{|\mathcal{M}_n|}{n^2} = 1 - \limsup_{n\to\infty} \frac{T(n)}{n^2} \leq 1 - c,
\end{aligned}$$

where the last inequality follows from Lemma 7.10. Thus, we obtain $\sigma_{bip}(m, f) = 0$. □

Observation 7.12. *Any bipartite pair (m, f) which cannot be realised as H with $H \in \mathcal{M}_m$ satisfies $\sigma_{bip}(m, f) = 0$.*

Proof. Any induced subgraph of a component of a graph in $\mathcal{M}_n$ is a biclique. Thus, any induced (m, f)-subgraph H of a graph in $\mathcal{M}_n$ is the vertex-disjoint union of bicliques and isolated vertices, and thus, $H \in \mathcal{M}_m$. □

Lemma 7.13. *There are infinitely many bipartite pairs (m, f) which satisfy $\sigma_{bip}(m, f) = 0$. In particular, for any $m \geq 4$ and f with $(m-1)^2 + 1 < f < m(m-1)$ we have $\sigma_{bip}(m, f) = 0$.*

Proof. Let $m \geq 4$ and f with $(m-1)^2 + 1 < f < m(m-1)$. Note that for $m \geq 4$, $[(m-1)^2 + 2, m(m-1) - 1] \neq \emptyset$. Let H be a bipartite (m, f) graph.

Assume H is disconnected and has two components H' and H'', both containing at least one edge. But then $f \leq |E(H')| + |E(H'')| \leq |E(K_{m-1,m-1})| + |E(K_2)| = (m-1)^2 + 1$, a contradiction. Thus, H is either connected or $H = H' + \{v\}$ and H' is connected.

Now assume H is connected. Then it cannot be a complete bipartite graph, since $f < m^2 = E(|K_{m,m}|)$. So assume $H = H' \cup \{v\}$. Then H' is not complete bipartite since $f \leq m(m-1) = |K_{m,m-1}|$.

Thus, $H \notin \mathcal{M}_m$, and thus, by Observation 7.12, $\sigma_{bip}(m, f) = 0$. □

Lemma 7.14. *For any $m \geq 2$ and f with $m(m-1) < f < m^2$ we have $\sigma_{bip}(m, f) = 0$.*

Proof. Let $m \geq 2$ and f with $m(m-1) < f < m^2$. Note that for $m \geq 2$, $[m(m-1) + 1, m^2 - 1] \neq \emptyset$. Let H be a bipartite (m, f) graph.

Assume that H is disconnected. Then $|E(H)| \leq |E(K_{m,m-1})| = m(m-1)$, a contradiction. So H is connected, and H is also not complete bipartite, since $E(H) < m^2 = |E(K_{m,m})|$.

Thus, $H \notin \mathcal{M}_m$, and thus, by Observation 7.12, $\sigma_{bip}(m, f) = 0$. □

This shows that there are infinitely many pairs which have forcing density 0. On the other hand, we can show, using the results from the previous sections, that there also exist pairs with positive forcing density:

Lemma 7.15. *Let $m \in \mathbb{N}$. We have $\sigma_{bip}(m, am) = 1$, for any $a \in [m]$.*

Proof. By Proposition 7.2, we have $(n, e) \overset{bip}{\to} (m, am)$ for any n sufficiently large and $e \in [n^{2-\frac{1}{q}}, n^2 - n^{2-\frac{1}{q}}]$ for some q depending only on m. Thus, by definition of the bipartite forcing density, we obtain

$$\sigma_{bip}(m, am) = \limsup_{n\to\infty} \frac{|\{e : (n,e) \overset{bip}{\to} (m, am)\}|}{n^2} \geq \limsup_{n\to\infty} \frac{n^2 - 2n^{2-\frac{1}{q}}}{n^2}$$
$$= 1 - \limsup_{n\to\infty} \frac{2}{n^{\frac{1}{q}}} = 1. \qquad \square$$

In the previous section, we have characterised the bipartite graphs that bipartite arrow $(3,4)$. Note that $4 \neq 3 \cdot a$ for any $a \in [3]$, so we cannot apply Lemma 7.15 to $(3,4)$, so the following lemma shows that not all pairs of bipartite forcing density 0 are of the form (m, am).

Lemma 7.16. *We have* $\sigma_{bip}(3,4) = 1$.

Proof. By Lemma 7.8 we know that $G \overset{bip}{\to} (3,4)$ if and only if $G \notin \bigcup_{i=1}^{4} \mathcal{G}_i(n)$, for the families $\mathcal{G}_i(n)$ given in Section 7.4. Let $E_i(n) = \{|E(G)| : G \in \mathcal{G}_i(n)\}$. In particular we have

1. Any graph in $\mathcal{G}_1(n)$ is a vertex-disjoint union of stars, i.e. a graph with at most n edges. In particular, $|E_1(n)| \in O(n)$.
2. Any graph in $\mathcal{G}_2(n)$ is of the form $K_{a,n} \setminus \{e\}$. Since there are at most n choices for a, we obtain $|E_2(n)| \in O(n)$.
3. Any graph in $\mathcal{G}_3(n)$ is $K_{n,n} - E(K_{a,b})$ for some $0 \leq a, b \leq n$. Thus, $|E_3(n)| = \{ab : 0 \leq a, b \leq n\}$. This set was considered by Erdős [54], who proved that $|E_3(n)| \in o(n^2)$. For the correct asymptotics of the cardinality of this set see also Ford [67, 68].
4. There at most $2n$ edges in the complement of any graph in $\mathcal{G}_4(n)$, and thus, $|E_4(n)| \in O(n)$.

By the definition of the forcing density, we obtain

$$\sigma_{bip}(3,4) = \limsup_{n\to\infty} \frac{|\{e : (n,e) \overset{bip}{\to} (3,4)\}|}{n^2} = \limsup_{n\to\infty} \frac{n^2 - \left|\bigcup_{i=1}^{4} E_i(n)\right|}{n^2}$$
$$\geq 1 - \limsup_{n\to\infty} \frac{1}{n^2}\left(\sum_{i=1}^{4} |E_i(n)|\right) \geq 1 - \limsup_{n\to\infty} \frac{o(n^2)}{n^2} = 1. \qquad \square$$

Now Proposition 7.3 follows immediately from Lemma 7.14, Lemma 7.15 and Lemma 7.16.

7.6 CONCLUDING REMARKS

We have shown that there exist infinitely many bipartite pairs with forcing density 0, but we were not able to extend our results on absolutely avoidable pairs to the bipartite setting. This leaves the following open problem:

Open Problem 7.17. *Are there any absolutely avoidable pairs (m, f) in the bipartite setting?*

Clearly good candidates for absolutely avoidable pairs are those for which we already know that the bipartite forcing density is 0. Lemma 7.11 and Observation 7.12 tells us, that we might want to look at pairs which are not realisable as the vertex disjoint union of bicliques and isolated vertices, i.e not as graphs in $\mathcal{M}_m$. Pairs (m, f) with $m(m-1) < f < m^2$ are not absolutely bipartite avoidable, since $(n, n^2-(m^2-f)) \overset{bip}{\to} (m, f)$, but it might be interesting to look at the pairs identified in Lemma 7.14.

While finding infinitely many pairs (m, am) and the additional pair $(3, 4)$ with bipartite forcing density 1, we also did not find any pair of non-trivial density, i.e. not 0 or 1. This leaves the second open question for this chapter:

Open Problem 7.18. *Are there any bipartite pairs (m, f) with $\sigma_{bip}(m, f) \in (0, 1)$?*

Chapter 8 Order-size pairs in hypergraphs: absolute avoidability and forcing densities

8.1 Introduction

This chapter is concerned with order-size pairs in r-graphs with $r \geq 3$. We will look at questions about avoidability and forcing densities.

Recall that in Chapter 6 we investigated the existence of so-called absolutely avoidable pairs (m, f) for which we not only have $\sigma_2(m, f) = 0$, but the stronger property $\{e : (n, e) \to (m, f)\} = \emptyset$ for all sufficiently large n. We showed that for $r = 2$ there are infinitely many absolutely avoidable pairs and amongst others constructed an infinite family of absolutely avoidable pairs of the form $(m, \binom{m}{2}/2)$ and showed that for any sufficiently large m, there exists an f such that (m, f) is absolutely avoidable. Here, we extend this result to higher uniformities:

Theorem 8.1. *Let $r \geq 3$. Then there exists m_0 such that that for any $m \geq m_0$ either $(m, \lfloor\binom{m}{r}/2\rfloor)$ or $(m, \lfloor\binom{m}{r}/2\rfloor - m - 1)$ is absolutely r-avoidable.*

Recall that the *forcing density* of a pair (m, f) is defined as

$$\sigma_r(m, f) = \limsup_{n\to\infty} \frac{|\{e : (n, e) \to_r (m, f)\}|}{\binom{n}{r}}.$$

Erdős, Füredi, Rothschild and Sós [56] claimed for $r = 2$ that "almost all pairs" have forcing density $\sigma_2(m, f) = 0$. Here we prove the following for higher uniformities:

Proposition 8.2. *For $r, m \in \mathbb{N}$, $r, m \geq 3$, all but $O\left(m^{\frac{r}{r-1}}\right)$ of all possible $\binom{m}{r}$ pairs (m, f) satisfy $\sigma_r(m, f) = 0$.*

As seen in Theorem 0.4, for $r = 2$ there exist pairs with $\sigma_2(m, f) = 1$. This changes for $r \geq 3$, as seen in Proposition 8.4 below, for which we need some additional definitions and notation. Recall that $T_r(n, l)$ denotes the complete l-partite r-graph on n vertices and part sizes $n_1, \ldots, n_l \in \{\lfloor\frac{n}{l}\rfloor, \lceil\frac{n}{l}\rceil\}$. Note that for $l < r$, $T_r(l, n)$ is empty, and for $r = 2$ this is just the Turán graph. The number of edges in $T_r(n, l)$ is denoted by $t_r(n, l)$ and for $l \geq r$ we have

$$t_r(n, l) = \sum_{S\in\binom{[l]}{r}} \prod_{i\in S} n_i = \frac{(l)_r}{l^r}\binom{n}{r} + o(n^r),$$

where $(l)_r = \prod_{i=0}^{r-1}(l-i) = l(l-1)\cdots(l-r+1)$. Note that for $r \geq 2$, $\frac{(l)_r}{l^r} = \frac{(l-1)_r}{(l-1)^r}(1-\frac{1}{l})^r\frac{l}{l-r}$, and by Bernoulli's inequality we have $(1-\frac{1}{l})^r > 1-\frac{r}{l}$, i.e. $\frac{(l)_r}{l^r} > \frac{(l-1)_r}{(l-1)^r}$, so $\frac{(l)_r}{l^r}$ is strictly increasing in l, and $\lim\limits_{l\to\infty} \frac{(l)_r}{l^r} = 1$. Also note that we have $\frac{(r)_r}{r^r} = \frac{r!}{r^r} \leq \frac{1}{r}$.

Let $l_{m,r}$ be the largest $l \in \mathbb{N}$ for which $t_r(m,l) < \frac{1}{2}\binom{m}{r}$. Note that this is well-defined by the previous observation, in particular, $l_{m,r} \geq r$ and $l_{m,r}$ is increasing in m. In particular, for fixed r there exists some m_0 such that $l_{m,r} = l_{m_0,r}$ for all $m \geq m_0$. One can verify that the following holds for $r=3$:

Observation 8.3. *We have* $l_{m,3} = \begin{cases} 3, & 4 \leq m \leq 11, \\ 4, & 12 \leq m \leq 72, \\ 5, & m \geq 73. \end{cases}$

Proposition 8.4. *Let $m > r \geq 3$, $0 \leq f \leq \binom{m}{r}$. Then $\sigma_r(m,f) < 1$.*
In particular, we can give the following upper bounds on σ_r for any $l \in \mathbb{N}$ satisfying $l \leq l_{m,r}$:

(a) We have $\sigma_r(m,f) \leq 1 - \frac{(l)_r}{l^r}$.

(b) If $t_r(m,l) < f < \binom{m}{r} - t_r(m,l)$ for some l, then $\sigma_r(m,f) < 1 - 2\frac{(l)_r}{l^r}$.

Note that He, Ma and Zhao [88] mentioned in their conclusion without proof, that for pairs (m,f) with $m > r \geq 3$, $0 \leq f \leq \binom{m}{r}$, the bound $\sigma_r(m,f) \leq 1 - \frac{r!}{r^r}$ holds.

This chapter is organised as follows. In Section 8.2 of this chapter we will build on one of the proof ideas used in [16] for 2-graphs, and extend these methods to higher uniformities in order to prove Theorem 8.1. In Section 8.3 we will make some observations on the forcing density σ_r and prove Proposition 8.2 and Proposition 8.4.

8.2 Existence of absolutely avoidable pairs

We will call an r-graph G *m-sparse* if every subset of m vertices in G induces at most m edges. We call an r-graph with at most m edges an *$\leq m$-edge ("at most m-edge") r-graph.*

In order to prove results in the 2-uniform case, the following fact is used in [16, 56]: Let $m > 0$ be given. Then for any v large enough there exists a graph of girth at least m on v vertices with $v^{1+\frac{1}{2m}}$ edges. For a probabilistic proof of this fact see for example Bollobás [26] and for an explicit construction see Lazebnik et al. [102].

Our first lemma is inspired by this: A 2-graph of girth $\geq m$ is an $(m-1)$-sparse graph. The proof of the lemma follows a standard probabilistic argument:

Lemma 8.5. *Let $m > 0$, $r \geq 2$ be given. Then for any n large enough there exists an n-vertex r-graph with $\Omega(n^{r-1+\frac{1}{m+1}})$ edges which is m-sparse.*

Proof. Let $c_{m,f,r} = \left(\frac{m}{e}\right)^{\frac{m}{f}} \frac{r!f}{m^r e 2^{1/f}}$. Consider a random r-graph $G \in G_r(n,p)$, for $p < c_{m,r,f} n^{-m/f}$. Then the probability that some m-subset contains at least f edges is less than

$$\binom{n}{m}\binom{\binom{m}{r}}{f}p^f \leq \left(\frac{ne}{m}\right)^m \left(\frac{m^r ep}{r!f}\right)^f < \frac{1}{2}. \tag{$\star$}$$

Now let X be the number of edges in G. Using the the standard bound $\binom{n}{r} \geq \left(\frac{n}{r}\right)^r$, we have

$$\mathbb{E}[X] = \binom{n}{r}p \geq c_{m,f,r}\frac{n^{r-m/f}}{r^r}.$$

Using Chernoff's bound for $\mathrm{Bin}(n,p)$ distributed random variables ([90], see Lemma 2.9), we obtain that for $\delta \in (0,1)$ the probability that G has fewer than a $(1-\delta)$-fraction of the expected number of edges is

$$\mathbb{P}(X \leq (1-\delta)\mathbb{E}(X)) \leq \exp\left(-\tfrac{\delta^2}{2}\mathbb{E}[X]\right) \leq \exp\left(-c_{m,f,r}\frac{\delta^2 n^{r-\frac{m}{f}}}{2r^r}\right) < \frac{1}{2},$$

where the last inequality holds for $\frac{m}{f} \leq r$ and n sufficiently large. Together with $(\star)$ that gives there exists an r-graph on n vertices with at least $(1-\delta)\binom{n}{r}p$ edges in which each m-element subset spans at most $f-1$ edges.

Note that, by choosing $f = m+1$, we obtain the existence of an r-graph with $c'_{r,f,m} n^{r-1+\frac{1}{m+1}}$ hyperedges and no m-subset which spans more than m hyperedges, i.e. an m-sparse graph. $\square$

The next two lemmata show that for many possible order-size pairs (n,e) we can find an r-graph which realises this pair and has a "nice", i.e. easy to analyse, structure. We will use them repeatedly throughout this chapter.

Lemma 8.6. *Let $m, r \in \mathbb{N}$, $m, r \geq 2$, and c be a constant, $0 \leq c < 1$. Then for $n \in \mathbb{N}$ sufficiently large and any $e \in [c\binom{n}{r}]$, there exists a non-negative integer k and an r-graph on n vertices and e edges which is the vertex-disjoint union of a $K_k^{(r)}$ and an m-sparse r-graph on $n-k$ vertices.*

Proof. Let $m, r > 0$ be given and let n be a given sufficiently large integer. Let $e \in [c\binom{n}{r}]$. Let k be the non-negative integer such that $\binom{k}{r} \leq e \leq \binom{k+1}{r} - 1$. Note that since $e \leq c\binom{n}{r}$, $\binom{k}{r} \leq c\binom{n}{r}$, and thus, $k \leq \sqrt[r]{c}n + 1 \leq c'n$, where c' is a constant with $c' < 1$. We claim that the pair (n, e) could be represented as the vertex-disjoint union of a $K_k^{(r)}$ and an m-sparse r-graph.

Let G' be a m-sparse graph on $n - k$ vertices with exactly $e - \binom{k}{r}$ edges. Lemma 9.13 guarantees the existence of such a graph, since

$$e - \binom{k}{r} < \binom{k+1}{r} - \binom{k}{r} = \binom{k}{r-1} \overset{k \leq c'n}{\leq} (n-k)^{r-1+\frac{1}{m+1}}.$$

Now let G be the vertex-disjoint union of $K_k^{(r)}$ and G. □

Lemma 8.7. *Let $m, r \in \mathbb{N}$, $m, r \geq 2$, and c be a constant, $0 < c \leq 1$. Then for $n \in \mathbb{N}$ sufficiently large and any integer e with $c\binom{n}{r} \leq e \leq \binom{n}{r}$, there exists a non-negative integer $k \leq n$ and an r-graph on k vertices and e edges which is the complement of an m-sparse r-graph.*

Proof. Note that adding isolated vertices to the complement of an m-sparse graph results in the complement of the vertex-disjoint union of a clique and an m-sparse graph. Thus, the statement immediately follows from Lemma 8.6 by taking complements. □

Lemma 8.8. *If for some integers m, r, f with $m \geq r \geq 2$ and $0 \leq f \leq \binom{m}{r}$ neither (m, f) nor $(m, \binom{m}{r} - f)$ can be realised as an r-graph which is the vertex-disjoint union of a complete r-graph and an $\leq$$m$-edge r-graph, then the pair (m, f) is absolutely avoidable.*

Proof. Assume we can realise neither (m, f) nor $(m, \binom{m}{r} - f)$ as the vertex-disjoint union of a complete r-graph and an $\leq$$m$-edge r-graph.

By the previous lemma, for n sufficiently large and any $e \leq \lceil \binom{n}{r}/2 \rceil$, there exists an r-graph G with e hyperedges which is the vertex-disjoint union of a clique and an r-graph which is m-sparse. In particular, for every $e \in \{0, 1, \ldots, \binom{n}{r}\}$, there is an r-graph G on n vertices with e edges, such that either G or its complement is the vertex-disjoint union of a clique and an m-sparse r-graph.

If G is the union of a clique and an m-sparse r-graph, then clearly $G \not\to_r (m, f)$, since (m, f) cannot be realised as the union of a clique and an $\leq$$m$-edge r-graph.

If $\bar{G}$ is the union of a clique and an m-sparse r-graph, then any induced r-graph on m vertices is the complement of the vertex-disjoint union of a clique and an $\leq$$m$-edge r-graph. Since $(m, \binom{m}{r} - f)$ cannot be realised as the union of a clique and an $\leq$$m$-edge r-graph, the pair $(m, f) = (m, \binom{m}{r} - (\binom{m}{r} - f))$ cannot be realised by a graph whose complement is the union of a clique and an $\leq$$m$-edge r-graph. Thus, $G \not\to_r (m, f)$. □

In the 2-uniform case we used a slightly stronger statement (i.e. no m-subset spans more than $m-1$ edges), to find absolutely avoidable pairs. For $r > 2$, it suffices to find pairs (m, f), which cannot be realised as the vertex-disjoint union of a clique $K_x^{(r)}$ and an $\leq m$-edge r-graph.

Good candidates for such pairs (m, f) are again, as in the 2-uniform case, pairs which look roughly like $(m, \binom{m}{r}/2 + o(1))$.

We will use the following lemmata several times:

Lemma 8.9. *Let $r \geq 2, m, f$ be integers with $m \geq r$, $0 \leq f \leq \binom{m}{r}$. If for some $k \in \mathbb{N}$, $\binom{k}{r} + m < f < \binom{k+1}{r}$, then the pair (m, f) cannot be realised as an r-graph which is the vertex disjoint union of a clique and an $\leq m$-edge r-graph.*

Proof. Assume (m, f) is realised as $K_l^{(r)} + H$, where $l \geq 0$ and H is an $\leq m$-edge r-graph. Then from the lower bound on f, we have that $l > k$, and from the upper bound on f, we see that $l < k + 1$. Thus, no such l exists. □

Lemma 8.10. *Let $r \geq 2, m, f$ be integers with $m \geq r$, $0 \leq f \leq \binom{m}{r}$. If for some $k \in \mathbb{N}$, $\binom{k-1}{r} < f < \binom{k}{r} - m$, then the pair (m, f) cannot be realised as an r-graph which is the union of the complement of an $\leq m$-edge hypergraph and some isolated vertices.*

Proof. Assume (m, f) is realised as $K_l^{(r)} - H$, where $l \geq 0$ and H is an $\leq m$-edge r-graph, and some isolated vertices. Then from the upper bound on f, we have that $l < k$, and from the lower bound on f, we see that $l > k - 1$. Thus, no such l exists. □

Proof of Theorem 8.1. Let $r \geq 3$, $m \geq m_0$ and let $f_0 = \left\lfloor \binom{m}{r}/2 \right\rfloor$.
Using Lemma 8.8, we need to show that either (m, f_0) or both $(m, f_0 - (m+1))$ and $(m, \binom{m}{r} - f_0 + (m+1))$ are not realisable as the vertex-disjoint union of a clique and an $\leq m$-edge r-graph. To this end we will show that the condition of Lemma 8.9 is satisfied.
Let x be an integer such that $\binom{x}{r} \leq \left\lfloor \binom{m}{r}/2 \right\rfloor < \binom{x+1}{r}$. By standard bounds we observe the following:

$$\frac{1}{2}\left(\frac{m}{r}\right)^r \leq \left\lfloor \frac{1}{2}\binom{m}{r} \right\rfloor < \binom{x+1}{r} < \left(\frac{(x+1)e}{r}\right)^r,$$

and thus,

$$x + 1 > \frac{1}{2^{1/r}e}m.$$

Thus, by choosing m_0 sufficiently large, we have for $m \geq m_0$ that $x - 1 > \frac{m}{4}$, since for $r \geq 3$ $\sqrt[r]{2}e \geq \sqrt[3]{2}e > 4$, and

$$\binom{x-1}{r-1} \geq \left(\frac{x-1}{r-1}\right)^{r-1} \geq \left(\frac{m}{4(r-1)}\right)^{r-1} > 2m + 2. \tag{$*$}$$

Let $f_- = \lfloor \binom{m}{r}/2 \rfloor - (m+1)$ and $f_+ = \lceil \binom{m}{r}/2 \rceil + (m+1)$.

Case 1: $\binom{x}{r} + m < f_0$. Then by Lemma 8.9, (m, f_0) cannot be realised by $K_k^{(r)} + H$, where $k \in \mathbb{N}$ and H has at most m edges. If $\lceil \binom{m}{r}/2 \rceil < \binom{x+1}{r}$, then again by Lemma 8.9, $(m, \binom{m}{r} - f_0)$ cannot be realised as the disjoint union of a clique and an $\leq m$-edge r-graph, i.e. by Lemma 8.8, (m, f_0) is absolutely avoidable.
Otherwise, we have $\lceil \binom{m}{r}/2 \rceil = \binom{x+1}{r} = \lfloor \binom{m}{r}/2 \rfloor + 1$. We clearly have $f_- < \binom{x+1}{r}$ and $f_+ > \binom{x+1}{r} + m$, so it remains to show that $f_- > \binom{x}{r} + m$ and $f_+ < \binom{x+2}{r}$. Indeed, we have

$$f_- - \binom{x}{r} = \binom{x+1}{r} - m - 1 - \binom{x}{r} = \binom{x}{r-1} - (m+1) \overset{(*)}{>} 2m + 1 - m - 1 = m,$$

i.e. $f_- > \binom{x}{r} + m$, and also

$$f_+ = \binom{x+1}{r} + m + 1 < \binom{x+1}{r} + 2m + 1 \overset{(*)}{<} \binom{x+1}{r} + \binom{x-1}{r-1} < \binom{x+2}{r},$$

and thus, $f_+ < \binom{x+2}{r}$.

Case 2: $\binom{x}{r} \leq \lfloor \binom{m}{r}/2 \rfloor \leq \binom{x}{r} + m$.
It remains to check that neither (m, f_-) nor (m, f_+) can be realised as the vertex-disjoint union of a clique and an $\leq m$-edge r-graph. On the one hand we have $f_- \leq \binom{x}{r} + m - (m+1) < \binom{x}{r}$. Thus, in order to use Lemma 8.9, it remains to verify that we also have $f_- > \binom{x-1}{r} + m$. Indeed, we have

$$\begin{aligned} f_- - \binom{x-1}{r} &\geq \binom{x}{r} - (m+1) - \left(\binom{x}{r} - \binom{x-1}{r-1}\right) = \binom{x-1}{r-1} - (m+1) \\ &\overset{(*)}{>} 2m + 1 - m - 1 = m \end{aligned}$$

for $m \geq m_0$, i.e. we have $\binom{x-1}{r} + m < f_- < \binom{x}{r}$, so by Lemma 8.9, (m, f_-) cannot be realised as the vertex-disjoint union of a clique and an $\leq m$-edge r-graph.

On the other hand, we clearly have $f_+ > \binom{x}{r} + m$, and also,

$$f_+ \leq \binom{x}{r} + m + 1 + (m+1) \overset{(*)}{<} \binom{x}{r} + \binom{x-1}{r-1} < \binom{x}{r} + \binom{x}{r-1} = \binom{x+1}{r},$$

so by Lemma 8.9, (m, f_+) cannot be realised as $K_k^{(r)} + H$, where $k \in \mathbb{N}$ and H is an $\leq m$-edge r-graph.

Thus, by Lemma 8.8 the pair (m, f_-) is absolutely avoidable. □

8.3 DENSITY OBSERVATIONS

Let $r \geq 3$, m, $f \leq \binom{m}{r}$. Recall from the Introduction to Part II that the forcing density is defined as

$$\sigma_r(m, f) = \limsup_{n \to \infty} \frac{|\{e : (n, e) \to (m, f)\}|}{\binom{n}{r}}.$$

It immediately follows that $\sigma_r(m, f) = \sigma_r(m, \binom{m}{r} - f)$ by considering complementary pairs. Recall that for a family of r-graphs $\mathcal{G}$, $\mathrm{ex}_r(n, \mathcal{G})$ denotes the extremal number, and for an r-graph H, $\pi_r(H)$ denotes the Turán density.

Note that for $f = 0$, σ_r can be expressed in terms of the Turán density, i.e. $\sigma_r(m, 0) = \sigma_r(m, \binom{m}{r}) = 1 - \pi_r(K_m^{(r)})$, where the currently best known general bounds on the Turán density are

$$1 - \left(\frac{r-1}{m-1}\right)^{r-1} \leq \pi(K_m^r) \leq 1 - \binom{m-1}{r-1}^{-1},$$

due to Sidorenko [118] and de Caen [51]. Also note that $\sigma_r(r, 1) = \sigma_r(r, 0) = 1$. Thus, the only non-trivial cases are $m > r$, which are dealt with in Proposition 8.2 and Proposition 8.4.

Before we prove Proposition 8.2, we show the following auxiliary lemma:

Lemma 8.11. *Let $m, r, f \in \mathbb{N}$ with $m \geq r \geq 3$ and $0 \leq f \leq \binom{m}{r}$.*

(a) If (m, f) cannot be realised as the disjoint union of a clique and an $\leq m$-edge r-graph, then $\sigma_r(m, f) = 0$. In particular, if there is no $x \in [m]$, such that $0 \leq f - \binom{x}{r} < m$, then $\sigma_r(m, f) = 0$.

(b) If (m, f) cannot be realised as the complement of an $\leq m$-edge r-graph and some isolated vertices, then $\sigma_r(m, f) = 0$. In particular, if there is no $x \in [m]$, such that $0 \leq \binom{x}{r} - f < m$, then $\sigma_r(m, f) = 0$.

(c) If $\sigma_r(m, f) > 0$, then there exist $x, \bar{x} \in [m]$ such that $0 \leq f - \binom{x}{r} < m$ and $0 \leq (\binom{m}{r} - f) - \binom{\bar{x}}{r} < m$.

(d) If for some $l \in \mathbb{N}$ we have $\binom{l}{r-1} > 2m$, then for $f > \binom{l}{r}$ and $f \neq \binom{x}{r}$ for $x \in [m]$ we have $\sigma_r(m, f) = 0$.

Proof. (a) By Lemma 8.6, for any $0 < c' < 1$, n sufficiently large, and $e \in \mathcal{E}_n := \left[c'\binom{n}{r}\right]$ there exists an r-graph G on n vertices with e edges which is the vertex-disjoint union of a clique and an m-sparse r-graph.

Note that any induced subgraph on m vertices of G is the union of a clique and an r-graph with at most m edges. Thus, by definition of σ_r, if a pair (m, f) cannot be realised by a clique and an $\leq m$-edge r-graph, we have

$$\sigma_r(m, f) = \limsup_{n\to\infty} \frac{|\{e : (n, e) \to (m, f)\}|}{\binom{n}{r}} \leq \limsup_{n\to\infty} \frac{|\binom{[n]}{r} - \mathcal{E}_n|}{\binom{n}{r}} = 1 - c'.$$

Since this holds for any $c' \in (0, 1)$, $\sigma_r(m, f) = 0$.

If there is no $x \in [m]$, such that $0 \leq f - \binom{x}{r} \leq \binom{m}{r}$, then the pair (m, f) cannot be realised as the union of a clique and an $\leq m$-edge r-graph. Thus, in this case we have $\sigma_r(m, f) = 0$.

(b) By Lemma 8.7, for any $0 < c' < 1$, n sufficiently large, and $e \in \mathcal{E}_n := \left[\binom{n}{r}\right] - \left[c'\binom{n}{r}\right]$ there exists an r-graph G on n vertices with e edges which is the complement of an m-sparse r-graph and some isolated vertices.

Note that any induced subgraph on m vertices of G is the union of a clique with at most m edges removed and an empty graph. Thus, by definition of σ_r, if a pair (m, f) cannot be realised as the complement of an $\leq m$-edge r-graph and some isolated vertices, we have

$$\sigma_r(m, f) = \limsup_{n\to\infty} \frac{|\{e : (n, e) \to (m, f)\}|}{\binom{n}{r}} \leq \limsup_{n\to\infty} \frac{|\binom{[n]}{r} - \mathcal{E}_n|}{\binom{n}{r}} = c'.$$

Since this holds for all $c' \in (0, 1)$, we have $\sigma_r(m, f) = 0$.

The "in particular" part follows similarly as in part (a).

(c) The first part is the contrapositive of the "in particular" statement in part (a). The second statement follows trivially using $\sigma_r(m, f) = \sigma_r(m, \binom{m}{r} - f)$.

(d) Let $f > \binom{l}{r}$, $f \neq \binom{x}{r}$ for $x \in [m]$, and let t be the unique integer satisfying $\binom{t}{r} < f < \binom{t+1}{r}$. Since $\binom{l}{r-1} > 2m$, it implies that $\binom{t+1}{r} - \binom{t}{r} = \binom{t}{r-1} \geq \binom{l}{r-1} > 2m$. Thus, we either have $f > \binom{t}{r} + m$ or $\binom{t+1}{r} - m > f$. In particular, by part (a) or (b), we have $\sigma_r(m, f) = 0$.

Note: The results by Axenovich, Balogh, Clemen and the author in [10] imply that the condition $f > \binom{l}{r}$ might not be needed. It is shown there for $r = 3$. □

Proof of Proposition 8.2. Let m be fixed and $f \leq \binom{m}{r}$; write f uniquely as $f = \binom{l}{r} + l'$, where $l \in [m]$ and $0 \le l' < \binom{l}{r-1}$. By Lemma 8.11(d) it follows that we have $\sigma_r(m,f) = 0$ if $\binom{l}{r} > 2m$ and $l' > 0$. In particular, any pair (m,f) with $\sigma_r(m,f) > 0$ must satisfy either $l' = 0$ or $\binom{l}{r} \le 2m$. In the first case, there are exactly $m+1$ possible choices for f (i.e. $f = \binom{x}{r}$ for some $x \in \{0,\ldots,m\}$). In the second case, we obtain that $\left(\frac{l}{r}\right)^{r-1} \le \binom{l}{r-1} \le 2m$, i.e. $l \le (2m)^{1/r-1}r$, i.e. $f \le \binom{l}{r} \le \left(\frac{el}{r}\right)^r \le e^r(2m)^{r/r-1}$. Thus, at most $(m+1)+e^r(2m)^{r/r-1} \in O\left(m^{\frac{r}{r-1}}\right)$ of all possible pairs (m,f) satisfy $\sigma_r(m,f) > 0$. Note that for $r \ge 3$, we have $m^{\frac{r}{r-1}} \in o(\binom{m}{r})$. $\square$

Proof of Proposition 8.4. Note that $t_r(n,l)$ is the maximum number of edges an l-partite r-graph can have.

Let $m > r \ge 2$, $0 \le f \le \binom{m}{r}$.

Let $l \in \mathbb{N}$, such that $t_r(m,l) < \frac{1}{2}\binom{m}{r}$. Note that for $r \ge 3$, such an l always exists, since we have $t_r(m,r) < \frac{1}{2}\binom{m}{r}$, so we can always choose $l = r$.

Thus, in particular, we are in one of two cases: Either we have $f \ge \frac{1}{2}\binom{m}{r} > t_r(m,l)$, or we have $f \le \frac{1}{2}\binom{m}{r}$, i.e. $f - \binom{m}{r} \ge \frac{1}{2}\binom{m}{r} > t_r(m,l)$.

Case 1: $f > t_r(m,l)$. Then any r-graph that realises the pair (m,f) is not l-partite. If $e \le t_r(n,l)$, we have $(n,e) \not\to (m,f)$, since taking any subgraph of $T_r(n,l)$ with e edges yields an l-partite (n,e)-graph, which cannot contain any (m,f)-graph. In particular, this implies that

$$\sigma_r(m,f) \le \lim_{n\to\infty} \frac{\binom{n}{r} - t_r(n,l)}{\binom{n}{r}} < 1.$$

Case 2: $\binom{m}{r} - f > t_r(m,l)$. Then any graph that realises $(m, \binom{m}{r} - f)$ is not l-partite. Then any r-graph G with $G \to_r (m, \binom{m}{r} - f)$ cannot be l-partite, i.e. $|E(G)| > t_r(n,l)$. Thus, for each $e \le t_r(n,l)$, $(n,e) \not\to_r (m, \binom{m}{r} - f)$, and thus, by using complementation, for each $e \ge \binom{n}{r} - t_r(n,l)$, $(n,e) \not\to_r (m,f)$. In particular, we have

$$\sigma_r(m,f) \le \lim_{n\to\infty} \frac{\binom{n}{r} - t_r(n,l)}{\binom{n}{r}} < 1.$$

Thus, in either case we have

$$\sigma_r(m,f) \le 1 - \limsup_{n\to\infty} \frac{t_r(n,l)}{\binom{n}{r}} = 1 - \frac{(l)_r}{l^r} < 1.$$

This proves part (a).

To obtain part (b), assume that for some $l \in \mathbb{N}$ we have $t_r(m,l) < f < \binom{m}{r} - t_r(m,l)$. Then by Cases 1 and 2, we see that $(n,e) \to_r (m,f)$ requires $t_r(n,l) < e < \binom{n}{r} - t_r(n,l)$.

Thus, we obtain that

$$\sigma_r(m,f) \leq 1 - 2\limsup_{n\to\infty} \frac{t_r(n,l)}{\binom{n}{r}} = 1 - 2\frac{(l)_r}{l^r},$$

which completes the proof. □

Remark 8.12. *For the proof of Proposition 8.4(a) one can also use the following extension of Turán's theorem to hypergraphs by Mubayi [108]. This is also the proof that appeared in the paper [126] containing the result. For fixed $l, r \geq 2$ let $\mathcal{F}_l^{(r)}$ be the family of r-graphs with at most $\binom{l}{2}$ edges, that contain a* core *S of l vertices, such that every pair of vertices in S is contained in an edge.*

Theorem 8.13 (Mubayi [108])**.** *Let $r, l, n \geq 2$. Then*

$$\mathrm{ex}(n, \mathcal{F}_{l+1}^{(r)}) = t_r(n,l)$$

and the unique r-graph on n vertices containing no copy of any member of $\mathcal{F}_{l+1}^{(r)}$ for which equality holds is $T_r(n,l)$, the complete balanced l-partite r-graph on n vertices.

Alternative proof of Proposition 8.4. Now let $m > r \geq 2$, $0 \leq f \leq \binom{m}{r}$.
Let $l \in \mathbb{N}$, such that $t_r(m,l) < \frac{1}{2}\binom{m}{r}$. Note that for $r \geq 3$, such an l always exists, since we have $t_r(m,r) < \frac{1}{2}\binom{m}{r}$, so we can always choose $l = r$.
Thus, in particular, we are in one of two cases: Either we have $f \geq \frac{1}{2}\binom{m}{r} > t_r(m,l)$, or we have $f \leq \frac{1}{2}\binom{m}{r}$, i.e. $f - \binom{m}{r} \geq \frac{1}{2}\binom{m}{r} > t_r(m,l)$.
Case 1: $f > t_r(m,l)$. Then by Theorem 8.13, any r-graph that realises the pair (m,f) contains a member of $\mathcal{F}_{l+1}^{(r)}$. If $e \leq t_r(n,l)$, we have $(n,e) \not\to (m,f)$, since taking any subgraph of $T_r(n,l)$ with e edges yields an (n,e) graph not containing any member of $\mathcal{F}_{l+1}^{(r)}$, and thus, a graph not containing induced (m,f). In particular, this implies that

$$\sigma_r(m,f) \leq \lim_{n\to\infty} \frac{\binom{n}{r} - t_r(n,l)}{\binom{n}{r}} < 1.$$

Case 2: $\binom{m}{r} - f > t_r(m,l)$. Then by Theorem 8.13 any graph that realises $(m, \binom{m}{r} - f)$ contains a member of $K_{l+1}^{(r)}$. Then any graph G with $G \to_r (m, \binom{m}{r} - f)$ must contain a member of $\mathcal{F}_{l+1}^{(r)}$, i.e. $|E(G)| > t_r(n,l)$. Thus, for each $e \leq t_r(n,l)$, $(n,e) \not\to_r (m, \binom{m}{r} - f)$, and thus, by considering the complement, for each $e \geq \binom{n}{r} - t_r(n,l)$, $(n,e) \not\to_r (m,f)$. In particular, we have

$$\sigma_r(m,f) \leq \lim_{n\to\infty} \frac{\binom{n}{r} - t_r(n,l)}{\binom{n}{r}} < 1.$$

Thus, in either case we have

$$\sigma_r(m, f) \leq 1 - \limsup_{n\to\infty} \frac{t_r(n,l)}{\binom{n}{r}} = 1 - \frac{(l)_r}{l^r} < 1. \qquad \square$$

Corollary 8.14. *For an integer m with $m > 3$ and any integer f with $0 < f < \binom{m}{3}$, we have:*

1. $\sigma_3(m, f) \leq \frac{7}{9}$.
2. *If $t_3(m, 3) < f < \binom{m}{r} - t_3(m, 3)$, then $\sigma_3(m, f) \leq \frac{5}{9}$.*
3. *If $m \geq 12$, then $\sigma_3(m, f) \leq \frac{5}{8}$.*
4. *If $m \geq 73$, then $\sigma_3(m, f) \leq \frac{13}{25}$.*

For an integer m with $m > 4$ and any integer f with $0 < f < \binom{m}{4}$, we have:

1. $\sigma_4(m, f) \leq \frac{29}{32}$,
2. *There is m_0, such that for all $m \geq m_0$ we have $\sigma_4(m, f) \leq \frac{131}{243} \approx 0.54$.*

Proof. Recall that Proposition 8.4(a) says that for a pair (m, f) and any l satisfying $l \leq l_{m,r}$, we have $\sigma_r(m, f) \leq 1 - \frac{(l)_r}{l^r}$. Recall that $l_{m,r}$ is the largest integer l for which $t_r(m, l) < \frac{1}{2}\binom{m}{r}$. Note again that $l_{m,r}$ is increasing in m and that for fixed r there exists some m_0 such that $l_{m,r} = l_{m_0,r}$ for all $m \geq m_0$.

We start with $r = 3$. In order to obtain our bounds, we can compute the fraction $\frac{(l)_3}{l^3}$ for different $l \geq 3$. We have that

$$\frac{(3)_3}{3^3} = \frac{2}{9}, \qquad \frac{(4)_3}{4^3} = \frac{3}{8}, \qquad \frac{(5)_3}{5^3} = \frac{12}{25}, \qquad \frac{(6)_3}{6^3} = \frac{5}{9}.$$

Note that $\frac{(6)_3}{6^3} > \frac{1}{2}$, so for $r = 3$, the best possible upper bound on $\sigma_3(m, f)$ one can achieve for any pair (m, f) using Proposition 8.4 will use $l = 5$.
Now (1) and (2) immediately follow from Proposition 8.4, by setting $l = r = 3$ and observing that for $r \geq 3$, we always have $t_r(m, r) < \frac{1}{2}\binom{m}{r}$. Then by Proposition 8.4(a) we have

$$\sigma_3(m, f) \leq 1 - \frac{3!}{3^3} = \frac{7}{9},$$

and for all f with $t_3(m, 3) < f < \binom{m}{r} - t_3(m, 3)$, by Proposition 8.4(b) we have

$$\sigma_3(m, f) < 1 - 2\frac{(3)_3}{3^3} = \frac{5}{9}.$$

Now for (3), by Observation 8.3 we know that for $m \geq 12$ we have $l_{m,3} \geq 4$, i.e. $t_3(m,4) < \frac{1}{2}\binom{m}{3}$ for all $m \geq 12$, and thus, by Proposition 8.4(a), for all pairs (m, f) with $m \geq 12$ we have

$$\sigma_3(m,f) \leq 1 - \left(\frac{(4)_3}{4^3}\right) = 1 - \frac{6}{16} = \frac{5}{8}.$$

For (4), by Observation 8.3 we know that for $m \geq 73$ we have $l_{m,3} = 5$, i.e. $t_3(m,5) < \frac{1}{2}\binom{m}{3}$ for all $m \geq 73$, so by Proposition 8.4(a), for all pairs (m, f) with $m \geq 73$ we have

$$\sigma_3(m,f) \leq 1 - \frac{(5)_3}{5^3} = 1 - \frac{12}{25} = \frac{13}{25}.$$

For the case $r = 4$, we obtain the first part by computing $\frac{(4)_4}{4^4} = \frac{3}{32}$. As before, since $t_4(m,4) < \frac{1}{2}\binom{m}{4}$ for all m, by Proposition 8.4(a) we have

$$\sigma_4(m,f) \leq 1 - \frac{4!}{4^4} = 1 - \frac{3}{32} = \frac{29}{32}.$$

The second part is obtained by computing $\frac{(9)_4}{9^4} = \frac{112}{243} < \frac{1}{2}$ and $\frac{(10)_4}{10^4} = \frac{63}{125} > \frac{1}{2}$. This implies, that for $r = 4$ there exists some m_0, such that $l_{m,4} = 9$ for all $m \geq m_0$. In particular, we then have $t_4(m,9) < \frac{1}{2}\binom{m}{4}$ for all $m \geq m_0$ and thus, by Proposition 8.4(a), we have

$$\sigma_4(m,f) \leq 1 - \frac{(9_4)}{9^4} = 1 - \frac{112}{243} = \frac{131}{243}$$

for $m \geq m_0$. □

8.4 CONCLUDING REMARKS

We have shown that for $r \geq 3$ and m sufficiently large there always exists an f such that (m, f) is absolutely r-avoidable; however, in the cases considered f is always roughly $\binom{m}{r}/2$. This inspires the following interesting question:

Open Problem 8.15. *Are there absolutely avoidable pairs where $f/\binom{m}{r}$ is bounded away from $\frac{1}{2}$ in the limit?*

We have proven that for $m > r \geq 3$ and f with $0 \leq f \leq \binom{m}{r}$, we always have $\sigma_r(m,f) < 1$. We have also shown that for fixed r, most pairs (m, f) satisfy $\sigma_r(m,f) = 0$. On the other hand, for $r \geq 3$ we have not identified any pair (m, f) with $0 < f < \binom{m}{r}$ for which we could show that $\sigma_r(m,f) > 0$. We will now use Lemma 8.11 to identify candidate pairs (m, f) for $r = 3, m \leq 15$ which might satisfy $\sigma_3(m,f) > 0$:

Lemma 8.16. *Let $r = 3$, $4 \leq m \leq 15$ and $0 < f < \binom{m}{r}$. If $(m,f) \neq (6,10)$, then $\sigma_3(m,f) = 0$.*

Proof. Let (m, f) be a pair with $4 \le m \le 15$, $0 < f < \binom{m}{r}$ with $\sigma_3(m, f) > 0$. One can verify that for $4 \le m \le 15$, we have $2\binom{m-3}{3} + 4 \le \binom{m}{3}$, i.e. $\binom{m-3}{3} + 2 \le \binom{m}{3} - \binom{m-3}{3} - 2$. Thus, for any $f \le \binom{m}{3}$, we either have $f \ge \binom{m-3}{3} + 2$ or $\binom{m}{3} - f \ge \binom{m-3}{3} + 2$, assume w.l.o.g. that we have $f \ge \binom{m-3}{3} + 2$. Let x be the unique value with $\binom{x}{3} \le f < \binom{x+1}{3}$ and write $f = \binom{x}{3} + x'$, $0 \le x' < \binom{x}{2}$. Then by assumption we have $x \in \{m-3, m-2, m-1\}$.

By Lemma 8.11(a) we know that if we cannot realise the pair (m, f) as the vertex-disjoint union of a clique and an $\le m$-edge 3-graph, then we have $\sigma_3(m, f) = 0$. Thus, we can assume that the pair (m, f) can be realised as the vertex-disjoint union of a clique and an $\le m$-edge 3-graph. By Lemma 8.9, it follows that (m, f) can be realised as the disjoint union of $K_x^{(3)}$ and an $\le m$-edge 3-graph.

Assume we have $x = m - 3$. Since by assumption $f \ge \binom{x}{3} + 2$, clearly the pair (m, f) cannot be realised as the vertex-disjoint union of $K_x^{(3)}$ and a graph with at most 1 edge. Since $m - x = 3$, it also cannot be realised as the vertex-disjoint union of $K_x^{(3)}$ and a graph with 2 edges, a contradiction. Thus, for $x = m - 3$ we have $\sigma_3(m, f) = 0$.

So assume $x \in \{m-2, m-1\}$. Note that since $m - x \le 2$, the pair (m, f) cannot be realised as the vertex-disjoint union of a clique on x vertices and a graph with at least 1 edge. Thus, the only pairs (m, f) which might satisfy $\sigma_3(m, f) > 0$ have $x' = 0$, i.e. $f \in \{\binom{m-2}{3}, \binom{m-1}{3}\}$. Then $(m, f) \in \mathcal{A}_1 \cup \mathcal{A}_2 \cup \{(6, 10), (10, 84), (13, 165), (15, 169), (15, 91)\}$ with

$$\mathcal{A}_1 = \{(5,4), (7,20), (8,35), (9,56), (11,120), (12,165), (13,220), (14,286)\},$$
$$\mathcal{A}_2 = \{(5,1), (6,4), (7,10), (8,20), (9,35), (10,56), (11,84), (12,120), (14,220)\}.$$

Now let $(m, f) \in \mathcal{A}_1 \cup \mathcal{A}_2$. Let $\bar{f} = \binom{m}{3} - f$ and let $y \in [m]$ such that $\binom{y}{3} \le \bar{f} < \binom{y+1}{3}$, i.e. $\bar{f} = \binom{y}{3} + y'$ for some $y' \le \binom{y}{2}$. Then it is easy to verify that we are in one of three cases:

- $y \in \{m-1, m-2\}$ and $y' > 0$,
- $y = m - 3$ and $y' > 1$,
- $y \le m - 4$ and $y' > m$.

In each case, by Lemma 8.9 $(m, \bar{f})$ cannot be realised as the disjoint union of a clique and an $\le m$-edge 3-graph, and thus, by Lemma 8.11(a), $\sigma_3(m, f) = \sigma_3(m, \bar{f}) = 0$.

The pair $(6, 10)$ is self-complementary with $10 = \binom{5}{3} = \binom{6}{3}/2$.

For the pair $(10, 84)$ we have $\binom{10}{3} - 84 = 36 = \binom{7}{3} + 1 = \binom{10}{3} - \binom{9}{3}$. Note that for $m = 10$, we have $2m < \binom{7}{2}$, i.e. by Proposition 8.2, $\sigma_3(10, 36) = 0$.

For the pair $(13, 165)$ we have $\binom{13}{3} - 165 = 121 = \binom{10}{3} + 1 = \binom{13}{3} - \binom{11}{3}$. Note that for $m = 13$, we have $2m < \binom{10}{2}$, i.e. by Proposition 8.2, $\sigma_3(13, 121) = 0$.

For the pair $(15, 91)$ we have $91 = \binom{15}{3} - \binom{14}{3} = \binom{9}{3} + 7$, and for the pair $(15, 169)$ we have $169 = \binom{15}{3} - \binom{13}{3} = \binom{11}{3} + 4$. Note that for $m = 15$, we have $2m < \binom{9}{2} < \binom{11}{2}$, i.e. by Proposition 8.2, $\sigma_3(15, 91) = \sigma_3(15, 169) = 0$. $\square$

Lemma 8.16 implies that for $r = 3$, the smallest value of m for which we might have $\sigma_3(m, f) > 0$ for some f is $m = 6$. In this case, $f = 10$ is the only possible value for which we might have $\sigma_3(6, f) > 0$. We will show in Chapter 9 that indeed $\sigma_3(6, 10) > 0$. This is based on the work by Axenovich, Balogh, Clemen and the author [10], where both upper and lower bounds on $\sigma_3(6, 10)$ are provided.

It would be interesting to further investigate this problem, as currently there is no other known non-trivial pair (m, f) for $r \geq 3$ which satisfies $\sigma_r(m, f) > 0$. Here the results from Chapter 9 also provide some further insight.

CHAPTER 9 ORDER-SIZE PAIR IN HYPERGRAPHS: POSITIVE FORCING DENSITY

9.1 INTRODUCTION

Recall that the *Turán function* or *extremal number* $\mathrm{ex}_r(n, H)$ is the maximum number of edges in an H-free n-vertex r-graph, and the *Turán density* of H is defined as $\pi(H) = \lim\limits_{n\to\infty} \frac{\mathrm{ex}_r(n,H)}{\binom{n}{r}}$.

Determining the Turán function for graphs and hypergraphs is a central topic in extremal graph theory with many challenging open problems, trying to identify what graph density forces the occurrence of a specific subgraph. Here, we are concerned with conditions on the graph density that forces the occurrence of an induced subgraph on a given number of vertices and a given number of edges, i.e. a given order-size pair.

As seen in Chapter 8, there are many pairs (m, f) for which $\sigma_r(m, f) = 0$, but not a single (non-trivial) pair with positive forcing density was known for r-graphs when $r \geq 3$. Note that $\sigma_r(r, 1) = \sigma_r(r, 0) = 1$ and for $f = 0$, σ_r corresponds to the Turán density, i.e. $\sigma_r(m, 0) = \sigma_r(m, \binom{m}{r}) = \pi(K_m^{(r)})$, where the best currently known general bounds on the Turán density are

$$1 - \left(\frac{r-1}{m-1}\right)^{r-1} \leq \pi(K_m^{(r)}) \leq 1 - \binom{m-1}{r-1}^{-1},$$

due to Sidorenko [118] and de Caen [51]. In the previous chapter we asked whether for $m > r \geq 3$, there is any f with $0 < f < \binom{n}{r}$ such that $\sigma_r(m, f) > 0$ and suggested the pair $(6, 10)$ as a candidate. We answer this question in the affirmative and prove $\sigma_3(6, 10) > 0$.

Given families of r-graphs $\mathcal{F}, \mathcal{G}$, we denote by $\mathrm{ex}(n, {}_{\mathrm{ind}}\mathcal{F}, \mathcal{G})$ the maximum number of edges in an n-vertex r-graph not containing any $F \in \mathcal{F}$ as an induced copy and also not any $G \in \mathcal{G}$ as a copy. Further, denote by $\pi({}_{\mathrm{ind}}\mathcal{F}, \mathcal{G})$ the limit

$$\pi({}_{\mathrm{ind}}\mathcal{F}, \mathcal{G}) = \limsup_{n\to\infty} \frac{\mathrm{ex}(n, {}_{\mathrm{ind}}\mathcal{F}, \mathcal{G})}{\binom{n}{r}}.$$

We mostly consider 3-graphs in this chapter. When clear from context, we shall write abc for the set $\{a, b, c\}$ corresponding to an edge in a 3-graph. The 3-graph on vertex set $[4]$ with edge set $\{123, 124, 124\}$ is denoted by K_4^{3-}. Let $\mathcal{F}_6^{10}$ be the family of 6-vertex 3-graphs containing exactly 10 edges.

Theorem 9.1. *We have that* $\sigma_3(6,10) = 1 - 2\pi({}_{\mathrm{ind}}\mathcal{F}_6^{10}, \{K_4^{3-}\})$. *Moreover,* $0.42622 \le \sigma_3(6,10) \le 0.47106$.

We do not know whether other pairs (m,f) with $m > 3$, $0 < f < \binom{m}{3}$ exist, such that $\sigma_3(m,f) > 0$. We conjecture that for $r = 3$ there are indeed no other pairs with positive forcing density. The following result provides evidence for this conjecture to be true.

Theorem 9.2. *Let m and f be positive integers, $0 < f < \binom{m}{3}$. If $\sigma_3(m,f) > 0$, then there exist $x_1, x_2, x_3 \in [m-1]$ such that*

$$f = \binom{x_1}{3} = \binom{m}{3} - \binom{x_2}{3} = \binom{x_3}{3} + \binom{x_3}{2}(m - x_3). \tag{9.1}$$

Thus, in particular if there are no other non-trivial solutions except for $m = 6$, $x_1 = 5, x_2 = 5, x_3 = 3$, to the above Diophantine equation, then Conjecture 9.19 is true. A computer search for suitable solutions of (9.1) did not give a result for $m \le 10^6$.

The main results of this chapter are joint work with Axenovich, Balogh and Clemen [10].

This chapter is organised as follows. In Section 9.2 we prove Theorem 9.1. In Section 9.3 we prove Theorem 9.2. Finally, in Section 9.4 we make concluding remarks and state open problems.

9.2 Proof of Theorem 9.1

We say a 3-graph G *induces* $(6,10)$ if G contains an induced copy of some $F \in \mathcal{F}_6^{10}$. If G does not contain any $F \in \mathcal{F}_6^{10}$ as an induced copy, we say G is $(6,10)$-free, i.e. a 3-graph is $(6,10)$-free if no 6-vertex set induces exactly 10 edges.

9.2.1 Proof idea

Before proving Theorem 9.1, we give a short sketch of the proof. We shall show that for every $\epsilon > 0$ there is n_0 such that for every $n > n_0$ if G is an n-vertex 3-graph satisfying

$$\frac{e(G)}{\binom{n}{3}} \in \left[\pi({}_{\mathrm{ind}}\mathcal{F}_6^{10}, \{K_4^{3-}\}) + \varepsilon, 1 - \pi({}_{\mathrm{ind}}\mathcal{F}_6^{10}, \{K_4^{3-}\}) - \varepsilon\right], \tag{9.2}$$

then G induces $(6,10)$. Then we first use a standard Ramsey type argument to partition most of the vertices of G into many large homogeneous sets. First, we rule out the case that there is a large clique and a large independent set that are disjoint. Thus, most of the vertex set of G or its complement $\overline{G}$ can be partitioned into large independent sets. Due to the symmetry of the problem, if we find a $(6,10)$-set in $\overline{G}$, we also find a $(6,10)$-set in G. Thus, without loss of generality, we can assume that most of the vertices of G can be partitioned into many large independent sets. Using a classical supersaturation result and the density assumption on G, we find many copies of K_4^{3-} in G and thus, in particular, four large independent sets spanning many transversal copies of K_4^{3-}. Using a final cleaning argument, we find a $(6,10)$-set in this substructure.

On the other hand, we fix an arbitrary 3-graph G on $\mathrm{ex}(n, {}_{\mathrm{ind}}\mathcal{F}_6^{10}, \{K_4^{3-}\})$ edges that is $(6,10)$-free and K_4^{3-}-free. Then every set of 6 vertices spans at most 9 edges, so there is a graph on n vertices and e edges, for any $e \leq \mathrm{ex}(n, {}_{\mathrm{ind}}\mathcal{F}_6^{10}, \{K_4^{3-}\})$, that is $(6,10)$-free. By taking complements, there also is a graph on n vertices and e edges for every $e \geq \binom{n}{3} - \mathrm{ex}(n, {}_{\mathrm{ind}}\mathcal{F}_6^{10}, \{K_4^{3-}\})$, that is $(6,10)$-free.

9.2.2 Definitions, notations, and construction

Let G be a 3-graph and let $X, Y, Z \subseteq V(G)$, not necessarily disjoint from each other. Then, let $E_G(X,Y,Z) = \{(x,y,z) \in E(G) : x \in X, y \in Y, z \in Z, x, y, z$ pairwise distinct$\}$. We say $E_G(X,Y,Z)$ is *complete* if $E_G(X,Y,Z) = \{(x,y,z) : x \in X, y \in Y, z \in Z, x, y, z$ pairwise distinct$\}$, and $E_G(X,Y,Z)$ is *empty* if $E_G(X,Y,Z) = \emptyset$. If the 3-graph G is clear from the context, we might omit the index and simply write $E(X,Y,Z)$.

Let H be an r-graph and $t \in \mathbb{N}$. The *t-blow-up* of H, denoted by $H(t)$, is the r-graph with its vertex set partitioned in $|V(H)|$ sets $V_1, V_2, \dots, V_{|V(H)|}$, each of size t and edge set $\{\{a_1, \dots, a_r\} : a_j \in V_{i_j}, j = 1, \dots, r, \{i_1, \dots, i_r\} \in E(H)\}$. Informally, $H(t)$ is obtained from H by replacing each vertex i with an independent set V_i and each hyperedge e of H with a complete r-partite hypergraph with parts corresponding to the vertices of e.

We say that a 3-graph G is a *weak t-blow-up* of H, which we also call *weak $H(t)$*, if the vertex set of G can be partitioned into $|V(H)|$ sets $V_1, V_2, \dots, V_{|V(H)|}$ each of size t such that if $ijk \in E(H)$ then for every $a \in V_i, b \in V_j, c \in V_k$ we have $abc \in E(G)$, and if $ijk \notin E(H)$ then for every $a \in V_i, b \in V_j, c \in V_k$ we have $abc \notin E(G)$. Moreover, V_i is an independent set for $i = 1, \dots, |V(H)|$. Note that we do not impose any condition on 3-tuples of vertices with exactly two vertices in some part V_i.

Recall that $R_r(t,t)$ denotes the *Ramsey number* of $K_t^{(r)}$ versus $K_t^{(r)}$. Erdős, Hajnal and Rado [61] showed that there exists a constant $c > 0$ such that $R_3(t,t) < 2^{2^{ct}}$.

Next, we shall provide a construction of a $(6, 10)$-free graph that we shall use to provide an upper bound in Theorem 9.1.

CONSTRUCTION OF THE 3-GRAPH H_n^{it}

Let H be the 3-graph with vertex set $[6]$ and edges $123, 124, 345, 346, 561, 562, 135, 146,$ and 236. Note that adding the edge 245 to H results in a 5-regular 3-graph on 6 vertices, which is K_4^{3-}-free and the basis for the construction for the lower bound on $\pi(K_4^{3-})$ by Frankl and Füredi [74].

We define the following iterated unbalanced blow-up of this graph. Denote by H_n the 3-graph on n vertices where the vertex set is partitioned into six sets A_1, A_2, A_3, A_4, A_5 and A_6, where

$$\begin{aligned}
|A_2| = |A_4| = |A_5| &= \left\lceil \frac{n}{3\sqrt{3}} \right\rceil, \\
|A_1| = |A_3| &= \left\lceil n\left(\frac{1}{3} - \frac{1}{3\sqrt{3}}\right) \right\rceil \quad \text{and} \\
|A_6| &= n\left(\frac{1}{3} - \frac{1}{3\sqrt{3}}\right) + O(1).
\end{aligned}$$

The 3-graph H_n consists of all triples xyz, where $x \in A_i, y \in A_j$ and $z \in A_k$ and $ijk \in E(H)$. Now, let H_n^{it} be the 3-graph constructed from H_n by iteratively adding a copy of $H_{|A_i|}$ with vertex set A_i for all $i \in [6]$ if $|A_i|$ is sufficiently large.

Lemma 9.3. *The graph H_n^{it} is an n-vertex 3-graph with $\frac{4}{3+7\sqrt{3}}\binom{n}{3} + o(n^3)$ edges such that every 6 vertices in H_n^{it} induce at most 9 edges. In particular, H_n^{it} is $(6, 10)$-free.*

Proof. We have

$$\begin{aligned}
|E(H_n)| &= 3\left(\frac{n}{3\sqrt{3}}\right)^2\left(\frac{1}{3} - \frac{1}{3\sqrt{3}}\right)n + 6\left(\frac{n}{3\sqrt{3}}\right)\left(\frac{1}{3} - \frac{1}{3\sqrt{3}}\right)^2 n^2 + o(n^3) \\
&= \frac{2\sqrt{3}}{81}n^3 + o(n^3).
\end{aligned}$$

Since H_n^{it} is an n-vertex 3-graph, it has at most $\binom{n}{3} \leq n^3/6$ edges. Let $|E(H_n^{\mathrm{it}})| = dn^3 + o(n^3)$ for some $d \in [0, \frac{1}{6}]$. We have

$$\begin{aligned}
|E(H_n^{\mathrm{it}})| &= \frac{2\sqrt{3}}{81}n^3 + 3d\left(\frac{n}{3\sqrt{3}}\right)^3 + 3d\left(\frac{1}{3} - \frac{1}{3\sqrt{3}}\right)^3 n^3 + o(n^3) \\
&= \left(\frac{2\sqrt{3}}{81} + \frac{d}{9}(2 - \sqrt{3})\right)n^3 + o(n^3).
\end{aligned}$$

Comparing the two expressions for $|E(H_n^{\text{it}})|$, we get $d = 2/(9+21\sqrt{3})$. In particular,

$$\frac{|E(H_n^{\text{it}})|}{\binom{n}{3}} = \frac{4}{3+7\sqrt{3}} + o(1) \approx 0.26447 + o(1).$$

Next we show that every set of six vertices in H_n^{it} spans at most 9 edges. Recall that H_n^{it} is obtained as an iterated blow-up construction with a "seed" graph H, where H is the 3-graph with vertex set [6] and edges 123, 124, 345, 346, 561, 562, 135, 146, and 236. At the first iteration, the vertices $1,\dots,6$ or H correspond to parts $A_1,\dots,A_6$. We have that H has three vertices of degree 4 and three vertices of degree 5, and H is K_4^{3-}-free, so every subset of four vertices spans at most two edges. Moreover, the link graph of any vertex of H is a subgraph of a 5-cycle. Here, the link graph of a vertex x is a 2-graph which has contains an edge yz if and only if xyz is an edge of H.

Let X be an arbitrary set of six vertices of H_n^{it}.

Case 1: X contains vertices from six distinct parts $A_1,\dots A_6$.
Then $|X|$ induces a copy of H, i.e. exactly 9 edges.

Case 2: X contains vertices from five distinct parts, say $A_{i_1}, A_{i_2}, A_{i_3}, A_{i_4}$, and A_{i_5}.
Assume we have two vertices in A_{i_1}, and one vertex in each of $A_{i_2}, A_{i_3}, A_{i_4}$, and A_{i_5}. Note that $A_{i_1}, A_{i_2}, A_{i_3}, A_{i_4}$, and A_{i_5} correspond to the vertices $i_1,\dots,i_5 \in V(H)$. Let $H' = H[\{i_1,i_2,i_3,i_4,i_5\}]$. Since the link graph of any vertex in H is a subgraph of C_5, the link graph of any vertex in H' has at most three edges, so the maximum degree of H' is at most three. This implies that the total number of edges in H' is at most $3\cdot 5/3 = 5$. Since the subgraph of H_n^{it} induced by X corresponds to H' with an added copy of i_1 which contributes at most three edges, X induces at most $5+3=8$ edges.

Case 3: X contains vertices from four distinct parts: $A_{i_1}, A_{i_2}, A_{i_3}$, and A_{i_4}.

Case 3.1: X contains 3 vertices from A_{i_1} and one vertex from each of $A_{i_2}, A_{i_3}, A_{i_4}$.
Then $H[\{i_1,i_2,i_3,i_4\}]$ contains at most two edges, so X induces at most $2\cdot 3$ edges between the parts and at most one additional edge inside A_{i_1}, so in total at most 7 edges.

Case 3.2: X contains two vertices from each of A_{i_1}, A_{i_2} and one vertex in each of A_{i_3}, A_{i_4}. Again, since $H[\{i_1,i_2,i_3,i_4\}]$ contains at most two edges, X induces at most $2\cdot 4 = 8$ edges.

Case 4: X contains vertices from three distinct parts: A_{i_1}, A_{i_2}, and A_{i_3}.
If we have two vertices in each of the three parts, they induce at most $2\cdot 2\cdot 2 = 8$

edges. If there are exactly three vertices in one of the parts, then there is a part with two vertices and a part with one vertex, i.e. there are at most $3 \cdot 2 \cdot 1 = 6$ edges between the parts, and at most one additional edge inside the first part, giving at most 7 edges. If there are four vertices in one part, then there are at most $4 \cdot 1 \cdot 1$ edges between the parts, and at most $\binom{4}{3} = 4$ additional edges inside the first part, i.e. at most 8 edges in total.

Case 5: X contains vertices from only one or two distinct parts.
Then there are no edges between these parts, and all possible edges induced by the six vertices are inside the A_i's. Since the construction is iterative, we can use the previous cases to conclude that X induces at most 9 edges. □

9.2.3 LEMMATA

The following lemma shows that every sufficiently large 3-graph can be partitioned into many large homogeneous sets.

Lemma 9.4. *Let $t > 0$. Then there exists $n_0 = n_0(t)$ such that for every $n \geq n_0$, if G is an n-vertex 3-graph, then G or $\overline{G}$ contains at least $n/t - \sqrt{n}$ pairwise disjoint homogeneous sets of size t.*

Proof. Let $t > 0$ be fixed. Set $n_0 = (\lceil 2^{2^{ct}} \rceil)^2$ and let $n \geq n_0$. Let $G = G_0$ be an n-vertex 3-graph. Since $n \geq R_3(t,t)$, there exists a homogeneous set of size t in G. Call it D_0 and define $G_1 = G_0 \setminus D_0$. We iteratively repeat this process. Define $G_{i+1} := G_i \setminus D_i$, where D_i is a homogeneous set of size t in G_i. We can proceed as long as $|V(G_i)| > R_3(t,t)$. Since $R_3(t,t) \leq \left\lceil 2^{2^{ct}} \right\rceil \leq \sqrt{n_0} \leq \sqrt{n}$, we have found at least $(n - \sqrt{n})/t \geq n/t - \sqrt{n}$ pairwise disjoint homogeneous sets of size t each. □

The following lemma analyses the structure "between" two large vertex sets. This is partly motivated by a result by Fox and Sudakov [73] for 2-graphs.

Lemma 9.5. *Let $t \geq 0$. Then there exists n_0 such that for all $n \geq n_0$ the following holds. Let G be a 3-graph with vertex set $V(G) = A \cup B$ with $A \cap B = \emptyset$, $|A| = |B| = n$. Then there exist sets $A' \subseteq A$, $B' \subseteq B$ with $|A'| = |B'| = t$ such that each of the edge sets $E(A', A', B')$ and $E(A', B', B')$ is either empty or complete.*

Proof. Let $m = \underbrace{4^{4^{\cdot^{\cdot^{\cdot^{4^t}}}}}}_{2t}$, let $n_0 = \underbrace{4^{4^{\cdot^{\cdot^{\cdot^{4^{2t-1}}}}}}}_{m}$. Let A and B be sets of size $n \geq n_0$. For $a \in A, X \subseteq B$ we define an auxiliary 2-graph $G_a^X = (X, \binom{X}{2})$ and an edge-colouring

$c_a^X : E(G_a^X) \to \{r, b\}$ with $c_a^X(\{b_1, b_2\}) = \begin{cases} r, & \{a, b_1, b_2\} \in E(G), \\ b, & \text{else.} \end{cases}$

Note that by the standard bound on the diagonal Ramsey number $r_2(s,s) \leq 4^s$, each 2-coloured 2-clique on k vertices contains a monochromatic clique of size $\log_4(k)$.

Let $A = \{a_1, \ldots, a_n\}$, let $B_1 \subseteq B$ be the vertex set of a monochromatic clique in $G_{a_1}^B$ of size $\log_4(|B|)$. Now assume B_i, $i \geq 1$, has been chosen. Let $B_{i+1} \subseteq B_i$ be a monochromatic clique in $G_{a_{i+1}}^{B_i}$ of size $\log_4(|B_i|)$. Thus, after m iterations we obtain a set B_m of size $|B_m| = \underbrace{\log_4 \cdots \log_4}_{m}(n) \geq 2t - 1$, such that for each a_i, $i \in [m]$, the set $E(\{a_i\}, B_m, B_m)$ is either empty or complete. Thus, there exists a subset $A'' \subset A$, $|A''| = \lceil \frac{m}{2} \rceil \geq \underbrace{4^{4^{\cdots^{4^t}}}}_{2t-1}$, such that the set $E(A'', B_m, B_m)$ is either empty or complete.

Now we repeat this process with vertices in $B'' = B_m$, to obtain a subset $A' \subseteq A''$, $|A'| = \underbrace{\log_4 \cdots \log_4}_{|B''|}(|A''|) \geq t$, such that for each vertex $b \in B''$, the set $E(A', A', \{b\})$ is either empty or complete. Thus, there exists a subset $B' \subseteq B''$, $|B'| \geq \left\lceil \frac{|B''|}{2} \right\rceil = t$ such that the set $E(A', A', B')$ is either empty or complete. The sets A', B' satisfy the conditions of the lemma, completing the proof. $\square$

The next lemma shows that in a $(6, 10)$-free 3-graph there cannot be a large independent set and a large clique that are disjoint.

Lemma 9.6. *There exists $t_0 > 0$ such that for all $t \geq t_0$ the following holds. Let G be a $2t$-vertex 3-graph with vertex set $V(G) = A \cup B$ where $A \cap B = \emptyset$, $|A| = |B| = t$, $G[A]$ is a clique and $G[B]$ is an independent set. Then G induces $(6, 10)$.*

Proof. By Lemma 9.5, for sufficiently large t, we can we find subsets $A' \subseteq A$, $B' \subseteq B$ with $|A'| = |B'| = 5$ such that the two sets $E(A', A', B')$ and $E(A', B', B')$ are either empty or complete.

If $E(A', B', B')$ is complete, then any vertex from A' together with the 5 vertices from B' induces $(6, 10)$. If $E(A', A', B')$ is empty, then any vertex from B' together with the five vertices from A' induces $(6, 10)$. Hence, we may assume that $E(A', B', B')$ is empty and $E(A', A', B')$ is complete. But then three arbitrary vertices from A' together with three arbitrary vertices from B' induce $(6, 10)$. $\square$

Lemma 9.4 together with Lemma 9.6 immediately implies the following lemma.

Lemma 9.7. *There exists t_0 such for all $t \geq t_0$ the following holds. There exists $n_0 = n_0(t)$ such that for all $n \geq n_0$, if G is a $(6,10)$-free n-vertex 3-graph, then either G or $\overline{G}$ contains at least $n/t - \sqrt{n}$ pairwise disjoint independent sets of size t.*

Lemma 9.8. *Let $t' > 0$. Then there exists $t_0 > 0$ such that for all $t \geq t_0$ the following holds. Let G be a $(6,10)$-free $2t$-vertex 3-graph with vertex set $V(G) = A \cup B$ where $|A| = |B| = t$, $A \cap B = \emptyset$, $G[A]$ and $G[B]$ are independent sets. Then there exists $A' \subseteq A$, $B' \subseteq B$ of sizes $|A'| = |B'| = t'$ such that the two sets $E(A', B', B')$ and $E(A', A', B')$ are empty.*

Proof. We apply Lemma 9.5 for t'. Then there exists t_0 such that for $t \geq t_0$, we find $A' \subseteq A, B' \subseteq B$, such that the two sets $E(A', A', B')$ and $E(A', B', B')$ are either empty or complete. Assume the set $E(A', A', B')$ is complete. Then we find induced $(6,10)$ by taking any 5 vertices from A' and 1 vertex from B. By symmetry the same holds for the set $E(A', B', B')$, so in particular, $G[A' \cup B']$ is the empty graph. $\square$

Lemma 9.9. *There exists $t_0 > 0$ such that for all $t \geq t_0$ a weak $K_4^{(3)}(t)$ and also a weak $K_4^{3-}(t)$ induces $(6,10)$.*

Proof. Let G be a weak $K_4^{3-}(t)$ with independent sets V_1, V_2, V_3, V_4. By iteratively applying Lemma 9.8 to all of the tuples (V_i, V_j), $1 \leq i < j \leq 4$, we obtain an induced copy $H \subseteq G$ of $K_4^{3-}(2)$ with sets X_1, X_2, X_3, X_4, $X_i \subset V_i$, $i \in [4]$, i.e. $H[X_i \cup X_j]$ is empty for all $i \neq j$, the sets $E(X_i, X_j, X_k)$ are complete for $\{i,j,k\} \in \binom{[4]}{3}$ except for $E(X_2, X_3, X_4)$, which is empty. Let $x_1, x_1' \in X_1$, $x_2, x_2' \in X_2$, $x_3 \in X_3$ and $x_4 \in X_4$. Then $\{x_1, x_1', x_2, x_2', x_3, x_4\}$ induces $(6,10)$.

Now assume there is a weak $K_4^3(t)$ called G with independent sets V_1, V_2, V_3, V_4. By iteratively applying Lemma 9.8 to all of the tuples (V_i, V_j), $1 \leq i < j \leq 4$, we obtain an induced copy $H \subseteq G$ of $K_4^3(3)$ with sets X_1, X_2, X_3, X_4, $X_i \subset V_i$, $i \in [4]$, i.e. $H[X_i \cup X_j]$ is empty for all $i \neq j$ and the sets $E(X_i, X_j, X_k)$ are complete for all $\{i,j,k\} \in \binom{[4]}{3}$. Let $x_2 \in X_2, x_3 \in X_3, x_4 \in X_4$. Then $H[X_1 \cup \{x_2, x_3, x_4\}]$ is a 6-vertex 3-graph spanning exactly 10 edges. $\square$

Lemma 9.10. *Let $t > 0$ be an integer and $\delta > 0$. Then there exists $m_0 = m_0(t, \delta)$ such that for all $m \geq m_0$ the following holds. Let G be a 3-graph on $4m$ vertices such that the vertex set of G can be partitioned into four independent sets V_1, V_2, V_3, V_4 of size m each and the number of copies of K_4^{3-} with one endpoint from each of the V_i's is at least δm^4. Then G contains an induced copy of a weak $K_4^3(t)$ or a weak $K_4^{3-}(t)$.*

Proof. Define the auxiliary 4-graph H on $4m$ vertices where a 4-set spans an edge iff the corresponding four vertices in G form a copy of K_4^{3-}. We 5-colour the edges of H

in the following way: An edge $\{v_1, v_2, v_3, v_4\}$ of H with $v_i \in V_i$ for $i \in [4]$ is coloured with $j \in [4]$ if $\{v_1, v_2, v_3, v_4\} \setminus \{v_j\}$ is not an edge in G, and it is coloured with colour 5 if $\{v_1, v_2, v_3, v_4\}$ induces a K_4^3 in G.

By the pigeonhole principle, there exists $(\delta/5)m^4$ edges of the same colour. Erdős [55] proved that $\pi(K_4^4(t)) = 0$ and thus, there exists a monochromatic copy of $K_4^4(t)$ in H. Denote by T the vertex set of this monochromatic copy. The 3-graph $G[T]$ is a weak $K_4^3(t)$ or weak $K_4^{3-}(t)$. □

We will use a supersaturation result discovered by Erdős and Simonovits [64]. The proof presented below follows a proof given by Keevash (Lemma 2.1. in [92]).

Lemma 9.11. *For $\varepsilon > 0$ and families $\mathcal{F}, \mathcal{G}$ of r-graphs, there exists constants $\delta > 0$ and $n_0 > 0$ so that if G is an r-graph on $n > n_0$ vertices with $e(G) > (\pi(_{\text{ind}}\mathcal{F}, \mathcal{G}) + \varepsilon)\binom{n}{r}$, then G contains at least $\delta\binom{n}{|V(H)|}$ copies of H for some $H \in \mathcal{G}$, or at least $\delta\binom{n}{|V(H)|}$ induced copies of H for some $H \in \mathcal{F}$.*

Proof. Let G be an r-graph on sufficiently many vertices n with $e(G) > (\pi(_{\text{ind}}\mathcal{F}, \mathcal{G}) + \varepsilon)\binom{n}{r}$. Fix an integer $k \geq r$, $k \geq |V(H)|$ for all $H \in \mathcal{F} \cup \mathcal{G}$ so that $\text{ex}(k, _{\text{ind}}\mathcal{F}, \mathcal{G}) \leq \left(\pi(_{\text{ind}}\mathcal{F}, \mathcal{G}) + \frac{\varepsilon}{2}\right)\binom{k}{r}$. There are at least $\frac{\varepsilon}{2}\binom{n}{k}$ k-sets $K \subseteq V(G)$ with $e(G[K]) > (\pi(_{\text{ind}}\mathcal{F}, \mathcal{G}) + \frac{\varepsilon}{2})\binom{k}{r}$. Otherwise, we would have

$$\sum_{\substack{K \subseteq V(G) \\ |K| = k}} e(G[K]) \leq \binom{n}{k}\left(\pi(_{\text{ind}}\mathcal{F}, \mathcal{G}) + \frac{\varepsilon}{2}\right)\binom{k}{r} + \frac{\varepsilon}{2}\binom{n}{k}\binom{k}{r}$$
$$= (\pi(_{\text{ind}}\mathcal{F}, \mathcal{G}) + \varepsilon)\binom{n}{k}\binom{k}{r},$$

but we also have

$$\sum_{\substack{K \subseteq V(G) \\ |K| = k}} e(G[K]) = \binom{n-r}{k-r}e(G) > \binom{n-r}{k-r}(\pi(_{\text{ind}}\mathcal{F}, \mathcal{G}) + \varepsilon)\binom{n}{r}$$
$$= (\pi(_{\text{ind}}\mathcal{F}, \mathcal{G}) + \varepsilon)\binom{n}{k}\binom{k}{r},$$

a contradiction. By the choice of k, each of these k-sets K contains an induced copy of some $H \in \mathcal{F}$ or a copy of some $H \in \mathcal{G}$. By the pigeonhole principle, there exists $H_1 \in \mathcal{F}$ such that at least $\frac{\varepsilon}{2(|\mathcal{F}|+|\mathcal{G}|)}\binom{n}{k}$ of these k-sets K contain an induced copy of H_1, or there exists $H_2 \in \mathcal{G}$ such that at least $\frac{\varepsilon}{2(|\mathcal{F}|+|\mathcal{G}|)}\binom{n}{k}$ of these k-sets K contain a copy of

H_2. Thus, in the first case, the number of induced copies of H_1 is at least

$$\frac{\frac{\varepsilon}{2(|\mathcal{F}|+|\mathcal{G}|)}\binom{n}{k}}{\binom{n-|V(H_1)|}{k-|V(H_1)|}} = \delta\binom{n}{|V(H_1)|}, \quad \text{for} \quad \delta = \frac{\varepsilon}{2(|\mathcal{F}|+|\mathcal{G}|)\binom{k}{|V(H_1)|}}.$$

Similarly, in the second case, the number of copies of H_2 is at least

$$\delta\binom{n}{|V(H_2)|} \quad \text{for} \quad \delta = \frac{\varepsilon}{2(|\mathcal{F}|+|\mathcal{G}|)\binom{k}{|V(H_2)|}}. \qquad \square$$

9.2.4 Proof of Theorem 9.1.

Proof of Theorem 9.1. Let $\varepsilon > 0$. Fix an integer t whose existence is guaranteed by Lemma 9.9, such that every weak $K_4^{(3)}(t)$ and also every weak $K_4^{3-}(t)$ induces $(6,10)$, see the paragraph before Lemma 9.9 for the definition of a weak blow-up. Fix $\delta > 0$ and $n_1 \in \mathbb{N}$, given by Lemma 9.10, such that every $(6,10)$-free 3-graph G on $n \geq n_1$ vertices satisfying $e(G) \geq (\pi(_{\text{ind}}\mathcal{F}_6^{10}, \{K_4^{3-}\}) + \varepsilon)\binom{n}{3}$ contains at least $2\delta\binom{n}{4}$ copies of K_4^{3-}. Let $m_0 = m_0(t,\delta)$ be given by Lemma 9.10. Fix integers m_1 and n_2 whose existence is guaranteed by Lemma 9.7, such that $m_1 \geq m_0$ and for all $n \geq n_2$, if G is $(6,10)$-free n-vertex 3-graph, then either G or $\overline{G}$ contains at least $n/m_1 - \sqrt{n}$ pairwise disjoint independent sets of size m_1. Choose $n_0 := \max\{n_1, n_2, m_1^2, \lceil 40000\delta^{-2}\rceil\}$ and let $n \geq n_0$.

Let G be a $(6,10)$-free n-vertex 3-graph satisfying the density assumption (9.2):

$$\frac{e(G)}{\binom{n}{3}} \in \left[\pi(_{\text{ind}}\mathcal{F}_6^{10}, \{K_4^{3-}\}) + \varepsilon, 1 - \pi(_{\text{ind}}\mathcal{F}_6^{10}, \{K_4^{3-}\}) - \varepsilon\right].$$

By Lemma 9.7 either G or $\overline{G}$ contains at least $n' := n/m_1 - \sqrt{n}$ pairwise disjoint independent sets, each of size m_1. Since the density assumption is symmetric, and since G induces $(6,10)$ if and only if $\overline{G}$ induces $(6,10)$, we can assume, without loss of generality, that G contains at least n' pairwise disjoint independent sets $V_1, V_2, \ldots, V_{n'}$ of size m_1 each.

By Lemma 9.11, G contains at least $2\delta\binom{n}{4}$ (not necessarily induced) copies of K_4^{3-}. We call a 4-set *transversal* in G if each of the four vertices is in a different V_i. A copy of K_4^{3-} in G is called *transversal* if the vertex set of the copy is transversal in G. The number of 4-sets which are not transversal in G is at most

$$\sqrt{n}n^3 + n'\binom{m_1}{2}n^2 \leq n^{\frac{7}{2}} + m_1 n^3 \leq 2n^{\frac{7}{2}},$$

for $n \geq m_1^2$. The number of transversal copies of K_4^{3-} in G is at least $\frac{3}{2}\delta\binom{n}{4}$, since

$$2\delta\binom{n}{4} - \frac{3}{2}\delta\binom{n}{4} = \frac{\delta}{2}\binom{n}{4} \geq \frac{\delta}{2}\frac{n^4}{2\cdot 4!} = \frac{\delta}{96}n^4 > 2n^{7/2},$$

where the last inequality holds for $n \geq 40000\delta^{-2}$. By pigeonhole principle there exist $1 \leq i_1 < i_2 < i_3 < i_4 \leq n'$, such that the number of copies of K_4^{3-} with one endpoint in each of $V_{i_1}, V_{i_2}, V_{i_3}, V_{i_4}$ is at least

$$\frac{\frac{3}{2}\delta\binom{n}{4}}{\binom{n'}{4}} \geq \frac{\delta\frac{n^4}{4!}}{\frac{\left(\frac{n}{m_1}\right)^4}{4!}} = \delta m_1^4.$$

By Lemma 9.10, the 3-graph $G[V_{i_1} \cup V_{i_2} \cup V_{i_3} \cup V_{i_4}]$ contains a weak $K_4^{3-}(t)$ or a weak $K_4^{(3)}(t)$ as an induced subhypergraph. This contradicts Lemma 9.9.

We conclude $\sigma_3(6,10) \geq 1 - 2\pi({}_{\text{ind}}\mathcal{F}_6^{10}, \{K_4^{3-}\})$. In fact, we have $\sigma_3(6,10) = 1 - 2\pi({}_{\text{ind}}\mathcal{F}_6^{10}, \{K_4^{3-}\})$ holds by the following argument: Let G be an n-vertex K_4^{3-}-free and $(6,10)$-free 3-graph with exactly $\text{ex}(n, {}_{\text{ind}}\mathcal{F}_6^{10}, \{K_4^{3-}\})$ many edges. Since G is K_4^{3-}-free, every four vertices span at most 2 edges, so using double counting, we see that every 6 vertices span at most $\binom{6}{4}\cdot 2/3 = 10$ edges. Since G is also $(6,10)$-free, every 6 vertices span only at most 9 edges. We conclude that every subgraph $G' \subseteq G$ is $(6,10)$-free. Further, by symmetry, also the complement 3-graph of any $G' \subseteq G$ is $(6,10)$-free. This proves the first part of the theorem.

To get specific numerical bounds on the forcing density, recall again that if

$$\frac{e(G)}{\binom{n}{3}} \in \left[\pi({}_{\text{ind}}\mathcal{F}_6^{10}, \{K_4^{3-}\}) + \varepsilon, 1 - \pi({}_{\text{ind}}\mathcal{F}_6^{10}, \{K_4^{3-}\}) - \varepsilon\right],$$

then G induces $(6,10)$. In particular, if $\frac{e(G)}{\binom{n}{3}} \in \left[\pi(K_4^{3-}) + \varepsilon, 1 - \pi(K_4^{3-}) - \varepsilon\right]$, then G induces $(6,10)$. The Turán density of K_4^{3-} is not known precisely. The best currently known bounds on the Turán density of K_4^{3-} are $0.28571 \approx \frac{2}{7} \leq \pi(K_4^{3-}) \leq 0.28689$, where the lower bound construction was given by Frankl and Füredi [74]. The upper bound was proved by Vaughan [123] who applied the flag algebra method, see also the webpage of Lidický [103]. Thus, $\sigma_3(6,10) \geq 1 - 2\cdot 0.28689 = 0.42622$. However, from Lemma 9.3, we have that there is a 3-graph on n vertices and $\frac{4}{3+7\sqrt{3}}\binom{n}{3}(1+o(1))$ hyperedges, such that each of its subgraphs is $(6,10)$-free. Moreover, the complement of this 3-graph has $\left(1 - \frac{4}{3+7\sqrt{3}}\binom{n}{3}\right)(1+o(1))$ hyperedges and each of its supergraphs is $(6,10)$-free. Thus, $\sigma_3(6,10) \leq 1 - 2\frac{4}{3+7\sqrt{3}} = 0.47105$. $\square$

9.3 Proof of Theorem 9.2

9.3.1 Constructions and notations

We shall first construct a special class of 3-graphs.

Let $n, k \in \mathbb{N}$, $k \le n$ and $S \subseteq [2]$. Let $G(S, n, k)$ be the 3-graph with vertex set $A \cup B$, $|A| = k$, $|B| = n - k$, where A and B are disjoint such that A induces a clique, B induces an independent set, called *base set*, and we have the additional edges $\bigcup_{i \in S} E_i$, where $E_i = \{A' \cup B' : A' \in \binom{A}{i}, B' \in \binom{B}{3-i}\}$. Thus, $G_\emptyset(n, k)$ is just a clique on k vertices and $n - k$ isolated vertices, and $G_{[2]}(n, k)$ is the complete graph on n vertices with a clique of size $n - k$ removed. For an illustration of $G(\{2\}, n, k)$ see Figure 9.1.

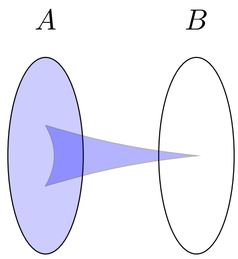


Figure 9.1: Illustration of $G(\{2\}, n, k)$.

Note that the complement of $G(S, n, k)$ is $G([2] - S, n, n - k)$. Let $f(S, n, k) = |E(G(S, n, k))|$. We call a 3-graph G *m-sparse* if every subset of m vertices in G induces at most m edges. We say that a 3-graph G is *canonical plus with parameters* (S, n, k), or simply *canonical plus* if G is a 3-graph obtained as a union of $G(S, n, k)$ and an m-sparse graph whose vertex set is the base independent set of $G(S, n, k)$. A 3-graph G is *canonical minus with parameters* (S, n, k), or simply *canonical minus*, if G is the complement of a canonical plus graph with parameters $([2] - S, n, n - k)$. Note that a canonical minus graph with parameters (S, n, k) is obtained from the graph $G(S, n, k)$ by removing edges of a copy of an m-sparse graph from the clique A. We see that (letting $\binom{y}{x} = 0$ for $y < x$), that

$$f(S, n, k) = \binom{k}{3} + \sum_{i \in S} \binom{k}{i} \binom{n-k}{3-i}.$$

Moreover, $|f(S, n, x) - f(S, n, x - 1)| \in O(n^2)$. Note that any induced subgraph of a canonical plus 3-graph with parameters (S, n, k) is a canonical plus 3-graph with parameters (S, n', k'), for some n' and k'. A similar statement holds for canonical minus graphs. Thus, these two classes of graphs are hereditary. We see that if an m-vertex 3-graph is canonical plus with parameters (S, m, x), then the number of edges in such a graph is in the interval $[f(S,m,x), f(S,m,x)+m]$. Similarly, the number of edges in a canonical minus graph with parameters (S,m,x) is in the interval $[f(S,m,x)-m, f(S,m,x)]$. Thus, if f is the number of edges of a graph that could be represented as both a canonical plus and a canonical minus graph with first parameter S and m vertices, then $f \in F(S,m)$, where

$$\begin{aligned} F(S,m) &= \bigcup_{x=0}^{m-1} [f(S,m,x), f(S,m,x)+m] \cap \bigcup_{x=1}^{m} [f(S,m,x)-m, f(S,m,x)] \\ &\subseteq \left\{0,1,\ldots,\binom{m}{3}\right\}. \end{aligned}$$

9.3.2 Proof idea

We are using the following general principle:

Proposition 9.12. *Let $\mathcal{C}_1,\ldots,\mathcal{C}_k$ be hereditary classes of r-graphs such that for any c, $0<c<1/2$, any sufficiently large n, and any e with $c\binom{n}{r} \le e \le (1-c)\binom{n}{r}$, there is a graph $G_i \in \mathcal{C}_i$ on n vertices and e edges for all $i=1,\ldots,k$. If for any sufficiently large n and some $i \in [k]$, each n-vertex graph in $\mathcal{C}_i$ is (m,f)-free, then $\sigma_r(m,f)=0$.*

Here, we use two classes $\mathcal{C}_1$ and $\mathcal{C}_2$ of 3-graphs that are canonical plus and canonical minus with the same first parameter S. Specifically, the main idea of the proof of Theorem 9.2 is that for any sufficiently large n, any $S \subseteq [2]$, and any e in the interval $[c\binom{n}{3}, (1-c)\binom{n}{3}]$ for $0<c<1/2$, there is a canonical plus 3-graph $G^+_{c,S}$ and a canonical minus 3-graph $G^-_{c,S}$ with first parameter S, on n vertices and e edges. If, for a pair (m,f), $f \notin F(S,m)$ for some $S \subseteq [2]$, then the pair (m,f) is not representable as a canonical plus or canonical minus graph with first parameter S. Then in particular, $G^+_{c,S}$ and $G^-_{c,S}$ are (m,f)-free and $(n,e) \not\to (m,f)$. Letting c be arbitrarily small, we conclude that $\sigma_3(m,f)=0$ for such a pair (m,f). Finally, we derive number theoretic conditions for a pair (m,f) not being representable by a canonical plus or a canonical minus graph.

9.3.3 Lemmata

In the following lemmata, n,m,f,e are non-negative integers with $m>3$, $0<f<\binom{m}{3}$. In [126] it was shown that for any $m \le 15$ and for any $0<f<\binom{m}{3}$ such that $(m,f) \ne (6,10)$, $\sigma_3(m,f)=0$. Thus, we can assume that $m \ge 16$. The following folklore result can be obtained by a standard probabilistic argument.

Lemma 9.13. *Let $m > 0$. Then for any sufficiently large n there exists an n-vertex 3-graph with $\Omega(n^{2+\frac{1}{m+1}})$ edges which is m-sparse.*

For a proof of Lemma 9.13 see e.g. [126]. The next lemma is a generalization of a similar statement proven in [56] for graphs.

Lemma 9.14. *Let $S \subseteq [2]$ and c be a constant, $0 < c < 1/2$. For $n \in \mathbb{N}$ sufficiently large and any e where $c < e < (1-c)\binom{n}{3}$, there exist 3-graphs $G_1(n,e)$ and $G_2(n,e)$ on n vertices and e edges that are canonical plus and canonical minus respectively, with first parameter S.*

Proof. Let n be a given sufficiently large integer. Let k be a non-negative integer such that either $f(S,n,k) \leq e \leq f(S,n,k+1)$ or $f(S,n,k) \leq e \leq f(S,n,k-1)$ holds. Without loss of generality assume that $f(S,n,k) \leq e \leq f(S,n,k+1)$. Let $c_1 = 1-c$. Note that since $e \leq c_1\binom{n}{3}$, $\binom{k}{3} \leq c_1\binom{n}{3}$, we have $k \leq \sqrt[3]{c_1}n+1 \leq c'n$, where $c' < 1$ is a constant.

Let G' be an m-sparse 3-graph on $n-k$ vertices with $|E(G')| \geq (n-k)^{2+\frac{1}{m+1}}$. The existence of G' is guaranteed by Lemma 9.13. Define G'' to be the 3-graph obtained as a union of $G(S,n,k)$ and a copy of G' on the vertex set that is the base independent set of $G(S,n,k)$. Then $|E(G'')| \geq f(S,n,k) + (n-k)^{2+\frac{1}{m+1}} \geq f(S,n,k+1) \geq e$. Here, the second inequality holds since $f(S,n,k+1) - f(S,n,k) = O(n^2)$. Finally, let $G_1(n,e)$ be a subgraph of G'' with e edges, obtained from G'' by removing some edges of G'.

For the second part of the lemma, take $G_2(n,e)$ to be the complement of $G_1(n,\binom{n}{3}-e)$ with first parameter $[2]-S$, guaranteed by the first part of the lemma. □

Lemma 9.15. *Let $S \subseteq [2]$. If $f \notin F(S,m)$, then $\sigma_3(m,f) = 0$.*

Proof. Assume we have integers m, f as above, some $S \subseteq [2]$ and $f \notin F(S,m)$. Let c be a constant, $0 < c < 1/10$, $n \geq n_0$, and e be any integer satisfying $c\binom{n}{3} \leq e \leq (1-c)\binom{n}{3}$. Define graphs $G_1 = G_1(n,e)$ and $G_2 = G_2(n,e)$ whose existence is guaranteed by Lemma 9.14. Any induced subgraph of G_1 on m vertices is canonical plus with parameters (S,m,x) for some x and thus, its number of edges is in $\bigcup_{x=0}^{m-1} [f(S,m,x), f(S,m,x)+m]$. Any induced subgraph of G_2 on m vertices is canonical minus with parameters (S,m,x) for some x and thus, its number of edges is in $\bigcup_{x=1}^{m} [f(S,m,x)-m, f(S,m,x)]$. Since $f \notin F(S,m)$, we get that G_1 and G_2 are (m,f)-free. Letting c go to zero, we see that $\sigma_3(m,f) = 0$. □

In the following lemmata we shall use the set $S = \emptyset$, $S = \{1\}$, or $S = \{2\}$, to claim that for many pairs (m, f), $\sigma_3(m, f) = 0$.

Lemma 9.16. *Let $m \geq 7$ and $0 < f < \binom{m-1}{2}$. Then $\sigma_3(m, f) = 0$.*

Proof. Let $S = \{1\}$. By Lemma 9.15, it is sufficient to verify that $f \notin F(\{1\}, m)$. For that it is sufficient to check that $F(\{1\}, m) \cap [1, \binom{m-1}{2} - 1] = \emptyset$. Recall that

$$F(\{1\}, m) = \bigcup_{x=0}^{m-1} [f(\{1\}, m, x), f(\{1\}, m, x) + m] \cap \bigcup_{x=1}^{m} [f(\{1\}, m, x) - m, f(\{1\}, m, x)].$$

Note that $f(\{1\}, m, 0) = 0$, $f(\{1\}, m, 1) = \binom{m-1}{2}$, and $f(\{1\}, m, x) \geq \binom{m-1}{2}$, for $x > 1$. Thus, we have

$$\begin{aligned} F(\{1\}, m) \cap [1, \tbinom{m-1}{2} - 1] &= \bigcup_{x=0}^{m-1} [f(\{1\}, m, x), f(\{1\}, m, x) + m] \cap [1, \tbinom{m-1}{2} - 1] \\ &= [f(\{1\}, m, 0), f(\{1\}, m, 0) + m] \cap [1, \tbinom{m-1}{2} - 1] = [1, m], \end{aligned}$$

and

$$\begin{aligned} \bigcup_{x=1}^{m} [f(\{1\}, m, x) - m, f(\{1\}, m, x)] \cap [1, \tbinom{m-1}{2} - 1] &= [f(\{1\}, m, 1) - m, f(\{1\}, m, 1) - 1] \\ &= [\tbinom{m-1}{2} - m, \tbinom{m-1}{2} - 1]. \end{aligned}$$

In particular, we have

$$F(\{1\}, m) \cap [1, \tbinom{m-1}{2} - 1] = [0, m] \cap [\tbinom{m-1}{2} - m, \tbinom{m-1}{2} - 1] = \emptyset,$$

where in the last step we used that $\binom{m-1}{2} > 2m$. Thus, $\sigma_3(m, f) = 0$. □

Lemma 9.17. *Let f be an integer such that $\binom{m-1}{2} \leq f < \binom{m}{3}$ and for any $x \in [m]$, $f \neq \binom{x}{3}$. Then $\sigma_3(m, f) = 0$.*

Proof. Define f as given in the statement of the lemma and $S = \emptyset$. By Lemma 9.15, it is sufficient to prove that $f \notin F(\emptyset, m)$ and in particular it is sufficient to show that $F(\emptyset, m) \cap [\binom{m-1}{2}, \binom{m}{3} - 1] \subseteq \{\binom{x}{3} : x \in [m]\}$. Since $f(\emptyset, n, x) = \binom{x}{3}$, we have

$$F(\emptyset, m) = \bigcup_{x=0}^{m-1} [\tbinom{x}{3}, \tbinom{x}{3} + m] \cap \bigcup_{x=1}^{m} [\tbinom{x}{3} - m, \tbinom{x}{3}],$$

see Figure 9.2 for an illustration of the set $F(\emptyset, m)$. Note that $\binom{x}{3} \geq \binom{m-1}{2}$ implies

$\binom{x}{2} > 2m$, which is equivalent to $\binom{x+1}{3} - m > \binom{x}{3} + m$. In particular, in this case the interval $[\binom{x}{3}, \binom{x+1}{3}]$ is long enough that we have $[\binom{x}{3}, \binom{x}{3} + m] \cap [\binom{x'}{3} - m, \binom{x'}{3}] = \emptyset$ for $x \neq x'$ and $\binom{x}{3}, \binom{x'}{3} \geq \binom{m-1}{2}$. Thus,

$$F(\emptyset, m) \cap [\binom{m-1}{2}, \binom{m}{3}] \subseteq \{\binom{x}{3} : x \in [m]\}. \qquad \square$$

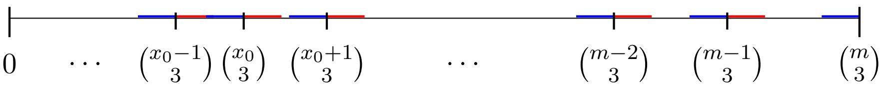


Figure 9.2: This figure displays the set $\bigcup_{x=0}^{m-1} [\binom{x}{3}, \binom{x}{3}+m]$ in red and the set $\bigcup_{x=1}^{m} [\binom{x}{3}-m, \binom{x}{3}]$ in blue on the number line. Here, x_0 is the smallest integer x such that $\binom{x+1}{3} - m > \binom{x}{3} + m$.

Lemma 9.18. *Let $m \geq 13$ and f be an integer, such that $\binom{m-1}{2} \leq f \leq \binom{m}{3} - \binom{m-1}{2}$ and for any $x \in [m]$, $f \neq \binom{x}{3} + \binom{x}{2}(m-x)$. Then $\sigma_3(m, f) = 0$.*

Proof. Consider m and f as given in the statement of the lemma and let $S = \{2\}$. By Lemma 9.15, it is sufficient to prove that $f \notin F(S, m)$ and in particular, it is sufficient to show that

$$F(\{2\}, m) \cap \left[\binom{m-1}{2}, \binom{m}{3} - \binom{m-1}{2}\right] \subseteq \left\{\binom{x}{3} + \binom{x}{2}(m-x) : \; x \in [m]\right\}.$$

Recall that

$$F(\{2\}, m) = \bigcup_{x=0}^{m-1} [f(\{2\}, m, x), f(\{2\}, m, x) + m] \cap \bigcup_{x=1}^{m} [f(\{2\}, m, x) - m, f(\{2\}, m, x)].$$

From the definition of f, we have that $f(\{2\}, m, x) = \binom{x}{3} + \binom{x}{2}(m-x)$. Note that for $x < 4$ we have $f(\{2\}, m, x) + m < \binom{m-1}{2}$ and for $x > m-4$, $f(\{2\}, m, x) - m > \binom{m}{3} - \binom{m-1}{2}$. Therefore it is sufficient to consider only

$$\bigcup_{x=4}^{m-4} [f(\{2\}, m, x), f(\{2\}, m, x) + m] \cap \bigcup_{x=4}^{m-4} [f(\{2\}, m, x) - m, f(\{2\}, m, x)].$$

One can verify, that for $m \geq 13$ and $4 \leq x \leq m-4$, $f(\{2\}, m, x) - f(\{2\}, m, x-1) > 2m$. Thus,

$$\begin{aligned}\bigcup_{x=4}^{m-4} [f(\{2\}, m, x), f(\{2\}, m, x) + m] \cap \bigcup_{x=4}^{m-4} [f(\{2\}, m, x) - m, f(\{2\}, m, x)]& \\ = \{f(\{2\}, m, x) : 4 \leq x \leq m-4\}&.\end{aligned}$$

In particular, we have

$$F(\{2\}, m) \cap [\tbinom{m-1}{2}, \tbinom{m}{3} - \tbinom{m-1}{2}] \subseteq \{\tbinom{x}{3} + \tbinom{x}{2}(m-x) : 4 \leq x \leq m-4\}. \qquad \square$$

9.3.4 Proof of Theorem 9.2

Proof. For $m \leq 15$ it was already shown in [126], that the only possible pair (m, f) with $0 < f < \binom{m}{3}$ and $\sigma_3(m, f) > 0$ is $(6, 10)$, where $10 = \binom{5}{3} = \binom{6}{3} - \binom{5}{3} = \binom{3}{3} + \binom{3}{2}(6-3)$. Now let $m > 15$, and assume that for some f we have $\sigma_3(m, f) > 0$. Then applying Lemma 9.16 to (m, f) and $(m, \binom{m}{3} - f)$, we obtain that $\binom{m-1}{2} \leq f \leq \binom{m}{3} - \binom{m-1}{2}$. Applying Lemma 9.17 to (m, f) gives us that $f = \binom{x_1}{3}$, for some x_1; applying it again to $(m, \binom{m}{3} - f)$ gives us that $f = \binom{m}{3} - \binom{x_2}{3}$, for some x_2. Lemma 9.18 shows the existence of some x_3, for which we have $f = \binom{x_3}{3} + \binom{x_3}{2}(m - x_3)$. This completes the proof. $\square$

9.4 Concluding remarks

In this chapter we have investigated 3-uniform hypergraphs and forcing densities $\sigma_3(m, f)$. We have shown that $\sigma_3(6, 10) > 0$ and provided more specific bounds. Apart from the pairs $(m, 0)$, $(m, \binom{m}{3})$, the pair $(6, 10)$ is the only known non-trivial pair for which the forcing density is positive. We conjecture that $(6, 10)$ is the unique pair (m, f) with $0 < f < \binom{m}{3}$ for which $\sigma_3(m, f) > 0$:

Conjecture 9.19. *Let m and f be positive integers, $0 < f < \binom{m}{3}$. If $\sigma_3(m, f) > 0$, then $(m, f) = (6, 10)$.*

Theorem 9.2 implies that if there is no $m \neq 6$ for which there is a solution (x_1, x_2, x_3), $x_i \in [m-1]$, of the system of Diophantine equations

$$\binom{x_1}{3} = \binom{m}{3} - \binom{x_2}{3} = \binom{x_3}{3} + \binom{x_3}{2}(m - x_3), \tag{9.3}$$

then Conjecture 9.19 is true. However, we do not know much about solutions (x_1, x_2, x_3) to the above system of equations. A computer search for suitable solutions of (9.3) for any given $m \leq 10^6$ did not give a result. Considering only the equation $\binom{x_1}{3} = \binom{m}{3} - \binom{x_2}{3}$, Sierpiński [119] found an infinite class of solutions.

It might be possible to find stronger necessary conditions for a pair to have positive forcing density using different constructions than the ones used in the proof of Theorem 9.2. In particular, the reader might wonder why Lemma 9.15 and the corresponding constructions in Lemma 9.14 were not used when $S = \{1\}$. The reason for this is that the respective function $f(\{1\}, m, x) = \binom{x}{3} + x\binom{m-x}{2}$ is not monotone, making it difficult to capture the structure of the set $F(\{1\}, m)$. However, this construction could very well be used to conclude that certain pairs (m, f) have forcing density zero.

Determining the exact value of $\sigma_3(6, 10)$ remains open. We believe that the upper bound from Theorem 9.1, coming from the iterated construction H_n^{it} in Lemma 9.3, is tight.

Conjecture 9.20. *We have* $\sigma_3(6, 10) = 1 - 2\frac{12}{9+21\sqrt{3}} \approx 0.47105.$

We remark that a standard flag algebra calculation yields that $\pi({}_{\mathrm{ind}}\mathcal{F}_6^{10}, \{K_4^{3-}\}) \leq 0.275 < 2/7$. Using the first part of Theorem 9.1, this gives $\sigma_3(6, 10) \geq 0.45$ which improves the lower bound on $\sigma_3(6, 10)$ given in the second part of Theorem 9.1.

Index

Bibliography

[1] Miklós Ajtai, János Komlós, and Endre Szemerédi. **A note on Ramsey numbers**. *Journal of Combinatorial Theory, Series A*, **29**(3):354–360, nov 1980. Available from: https://doi.org/10.1016%2F0097-3165%2880%2990030-8.

[2] Bogdan Alecu, Aistis Atminas, Vadim Lozin, and Viktor Zamaraev. **Graph classes with linear Ramsey numbers**. *Discrete Mathematics*, **344**(4):112307, apr 2021. Available from: https://doi.org/10.1016%2Fj.disc.2021.112307.

[3] Noga Alon, József Balogh, Alexandr Kostochka, and Wojciech Samotij. **Sizes of Induced Subgraphs of Ramsey Graphs**. *Combinatorics, Probability & Computing*, **18**(4):459–476, jul 2009. Available from: https://doi.org/10.1017%2Fs0963548309009869.

[4] Noga Alon and Alexandr V. Kostochka. **Induced subgraphs with distinct sizes**. *Random Structures and Algorithms*, **34**(1):45–53, jan 2009. Available from: https://doi.org/10.1002%2Frsa.20250.

[5] Noga Alon, Michael Krivelevich, and Benny Sudakov. **Induced subgraphs of prescribed size**. *Journal of Graph Theory*, **43**(4):239–251, jul 2003. Available from: https://doi.org/10.1002%2Fjgt.10117.

[6] Noga Alon, János Pach, and József Solymosi. **Ramsey-type theorems with forbidden subgraphs**. *Combinatorica*, **21**(2):155–170, apr 2001. Available from: https://doi.org/10.1007%2Fs004930100016.

[7] Noga Alon and Vojtěch Rödl. **Sharp bounds for some multicolor Ramsey numbers**. *Combinatorica*, **25**(2):125–141, mar 2005. Available from: https://doi.org/10.1007%2Fs00493-005-0011-9.

[8] Noga Alon, Lajos Rónyai, and Tibor Szabó. **Norm-graphs: variations and applications**. *Journal of Combinatorial Theory, Series B*, **76**(2):280–290, jul 1999. Available from: https://doi.org/10.1006%2Fjctb.1999.1906.

[9] MARIA AXENOVICH AND JÓZSEF BALOGH. **Graphs having small number of sizes on induced k-subgraphs**. *SIAM Journal on Discrete Mathematics*, **21**(1):264–272, jan 2007. Available from: https://doi.org/10.1137%2F05064357x.

[10] MARIA AXENOVICH, JÓZSEF BALOGH, FELIX CHRISTIAN CLEMEN, AND LEA WEBER. **Unavoidable order-size pairs in hypergraphs – positive forcing density**. *arXiv preprint*, 2022. Available from: arXiv.org/abs/2208.06626.

[11] MARIA AXENOVICH, DHRUV MUBAYI, AND LEA WEBER. **Homogeneous sets in hypergraphs with forbidden order-size pairs**. *arXiv preprint*, 2023. Available from: https://arxiv.org/abs/2303.09578.

[12] MARIA AXENOVICH, ALEX RIASANOVSKY, AND LEA WEBER. **Multicolour and size version of the Erdős-Hajnal problem**. *[Unpublished Manuscript]*, 2021.

[13] MARIA AXENOVICH, JEAN-SÉBASTIEN SERENI, RICHARD SNYDER, AND LEA WEBER. **Bipartite independence number in graphs with bounded maximum degree**. *SIAM Journal on Discrete Mathematics*, **35**(2):1136–1148, jan 2021. Available from: https://doi.org/10.1137%2F20m1321760.

[14] MARIA AXENOVICH, RICHARD SNYDER, AND LEA WEBER. **The Erdős-Hajnal conjecture for three colors and triangles**. *Discrete Mathematics*, **345**(5):112791, may 2022. Available from: https://doi.org/10.1016%2Fj.disc.2021.112791.

[15] MARIA AXENOVICH, CASEY TOMPKINS, AND LEA WEBER. **Large homogeneous subgraphs in bipartite graphs with forbidden induced subgraphs**. *Journal of Graph Theory*, **97**(1):34–46, oct 2020. Available from: https://doi.org/10.1002%2Fjgt.22639.

[16] MARIA AXENOVICH AND LEA WEBER. **Absolutely avoidable order-size pairs for induced subgraphs**. *To appear in Journal of Combinatorics. arXiv preprint*, 2021. Available from: arXiv.org/abs/2106.14908.

[17] ROGER C. BAKER, GLYN HARMAN, AND JÁNOS PINTZ. **The difference between consecutive primes, II**. *Proceedings of the London Mathematical Society*, **83**(3):532–562, nov 2001. Available from: https://doi.org/10.1112%2Fplms%2F83.3.532.

[18] MANTAS BAKSYS AND XUANANG CHEN. **On number of different sized induced subgraphs of bipartite-Ramsey graphs**. *arXiv preprint*, 2021. Available from: arXiv.org/abs/2109.08485.

[19] CAMINO BALBUENA, PEDRO GARCÍA-VÁZQUEZ, XAVIER MARCOTE, AND JUAN CARLOS VALENZUELA. **New results on the Zarankiewicz problem**. *Discrete mathematics*,

307(17-18):2322–2327, aug 2007. Available from: https://doi.org/10.1016%2Fj.disc.2006.11.002.

[20] Camino Balbuena, Pedro García-Vázquez, Xavier Marcote, and Juan Carlos Valenzuela. **Extremal $K_{(s,t)}$-free bipartite graphs**. *Discrete Mathematics and Theoretical Computer Science*, **10**(3):35–48, jan 2008. Available from: https://doi.org/10.46298%2Fdmtcs.435.

[21] Lowell W. Beineke and Allen J. Schwenk. **On a bipartite form of the Ramsey problem**. In *Proceedings of the Fifth British Combinatorial Conference*, pages 17–22. Congressus Numerantium, No. XV. Utilitas Math., Winnipeg, Man., 1976.

[22] Claude Berge and Pierre Duchet. **Strongly perfect graphs**. In *Topics on Perfect Graphs*, pages 57–61. Elsevier, 1984. Available from: https://doi.org/10.1016%2Fs0304-0208%2808%2972922-0.

[23] Tom Bohman. **The triangle-free process**. *Advances in Mathematics*, **221**(5):1653–1677, aug 2009. Available from: https://doi.org/10.1016%2Fj.aim.2009.02.018.

[24] Tom Bohman and Peter Keevash. **The early evolution of the H-free process**. *Inventiones mathematicae*, **181**(2):291–336, may 2010. Available from: https://doi.org/10.1007%2Fs00222-010-0247-x.

[25] Béla Bollobás. **The independence ratio of regular graphs**. *Proceedings of the American Mathematical Society*, **83**(2):433–436, oct 1981. Available from: https://doi.org/10.2307%2F2043545.

[26] Béla Bollobás. *Extremal graph theory*. Courier Corporation, 2004.

[27] Nicolas Bousquet, Aurélie Lagoutte, and Stéphan Thomassé. **The Erdős–Hajnal conjecture for paths and antipaths**. *Journal of Combinatorial Theory, Series B*, **113**:261–264, jul 2015. Available from: https://doi.org/10.1016%2Fj.jctb.2015.01.001.

[28] Marcia D.E. Brandes, Kevin T. Phelps, and Vojtech Rödl. **Coloring Steiner Triple Systems**. *SIAM Journal on Algebraic Discrete Methods*, **3**(2):241–249, jun 1982. Available from: https://doi.org/10.1137%2F0603023.

[29] Rowland Leonard Brooks. **On colouring the nodes of a network**. *Mathematical Proceedings of the Cambridge Philosophical Society*, **37**(2):194–197, apr 1941. Available from: https://doi.org/10.1017%2Fs030500410002168x.

[30] William G Brown. **On graphs that do not contain a Thomsen graph**. *Canadian Mathematical Bulletin*, **9**(3):281–285, aug 1966. Available from: https://doi.org/10.4153%2Fcmb-1966-036-2.

[31] Matija Bucić, Tung Nguyen, Alex Scott, and Paul Seymour. **Induced subgraph density. I. A loglog step towards Erdős-Hajnal**. *arXiv preprint*, 2023. Available from: arXiv.org/abs/2301.10147.

[32] Boris Bukh and Benny Sudakov. **Induced subgraphs of Ramsey graphs with many distinct degrees**. *Journal of Combinatorial Theory, Series B*, **97**(4):612–619, jul 2007. Available from: https://doi.org/10.1016%2Fj.jctb.2006.09.006.

[33] Marcelo Campos, Simon Griffiths, Robert Morris, and Julian Sahasrabudhe. **An exponential improvement for diagonal Ramsey**. *arXiv preprint*, 2023. Available from: arXiv.org/abs/2303.09521.

[34] Yair Caro. **Zero-Sum problems—A survey**. *Discrete Mathematics*, **152**(1-3):93–113, may 1996. Available from: https://doi.org/10.1016%2F0012-365x%2894%2900308-6.

[35] Yair Caro, Adriana Hansberg, Josef Lauri, and Christina Zarb. **On Zero-Sum spanning trees and Zero-Sum connectivity**. *The Electronic Journal of Combinatorics*, **29**(1), jan 2022. Available from: https://doi.org/10.37236%2F10289.

[36] Yair Caro, Adriana Hansberg, and Amanda Montejano. **Zero-Sum K_m over $\mathbb{Z}$ and the story of K_4**. *Graphs and Combinatorics*, **35**(4):855–865, apr 2019. Available from: https://doi.org/10.1007%2Fs00373-019-02040-3.

[37] Yair Caro, Adriana Hansberg, and Amanda Montejano. **Unavoidable chromatic patterns in 2-colorings of the complete graph**. *Journal of Graph Theory*, **97**(1):123–147, nov 2020. Available from: https://doi.org/10.1002%2Fjgt.22645.

[38] Yair Caro, Josef Lauri, and Christina Zarb. **The feasibility problem for line graphs**. *Discrete Applied Mathematics*, **324**:167–180, jan 2023. Available from: https://doi.org/10.1016%2Fj.dam.2022.09.019.

[39] Yair Caro, Yusheng Li, Cecil C. Rousseau, and Yuming Zhang. **Asymptotic bounds for some bipartite graph: complete graph Ramsey numbers**. *Discrete Mathematics*, **220**(1-3):51–56, jun 2000. Available from: https://doi.org/10.1016%2Fs0012-365x%2899%2900399-4.

[40] YAIR CARO AND CECIL ROUSSEAU. **Asymptotic bounds for bipartite Ramsey numbers**. *The Electronic Journal of Combinatorics*, **8**(1), feb 2001. Available from: https://doi.org/10.37236%2F1561.

[41] DEBSOUMYA CHAKRABORTI. **Extremal bipartite independence number and balanced coloring**. *European Journal of Combinatorics*, **113**:103750, oct 2023. Available from: https://doi.org/10.1016%2Fj.ejc.2023.103750.

[42] GUANTAO CHEN AND RICHARD H. SCHELP. **Ramsey Problems with Bounded Degree Spread**. *Combinatorics, Probability and Computing*, **2**(3):263–269, sep 1993. Available from: https://doi.org/10.1017%2Fs0963548300000663.

[43] MARIA CHUDNOVSKY. **The Erdös-Hajnal conjecture – A survey**. *Journal of Graph Theory*, **75**(2):178–190, jan 2013. Available from: https://doi.org/10.1002%2Fjgt.21730.

[44] MARIA CHUDNOVSKY AND SHMUEL SAFRA. **The Erdős–Hajnal conjecture for bull-free graphs**. *Journal of Combinatorial Theory, Series B*, **98**(6):1301–1310, nov 2008. Available from: https://doi.org/10.1016%2Fj.jctb.2008.02.005.

[45] MARIA CHUDNOVSKY, ALEX SCOTT, PAUL SEYMOUR, AND SOPHIE SPIRKL. **Erdős–Hajnal for graphs with no 5-hole**. *Proceedings of the London Mathematical Society*, **126**(3):997–1014, jan 2023. Available from: https://doi.org/10.1112%2Fplms.12504.

[46] DAVID CONLON. **A new upper bound for diagonal Ramsey numbers**. *Annals of Mathematics*, **170**(2):941–960, sep 2009. Available from: https://doi.org/10.4007%2Fannals.2009.170.941.

[47] DAVID CONLON, JACOB FOX, AND BENNY SUDAKOV. **Hypergraph Ramsey numbers**. *Journal of the American Mathematical Society*, **23**(1):247–266, aug 2009. Available from: https://doi.org/10.1090%2Fs0894-0347-09-00645-6.

[48] DEREK GORDON CORNEIL, H. LERCHS, AND L. STEWART BURLINGHAM. **Complement reducible graphs**. *Discrete Applied Mathematics*, **3**(3):163–174, jul 1981. Available from: https://doi.org/10.1016%2F0166-218x%2881%2990013-5.

[49] K. ČULIK. **Teilweise Lösung eines verallgemeinerten Problems von K. Zarankiewicz**. In *Annales Polonici Mathematici*, **3(1)**, pages 165–168. Institute of Mathematics, Polish Academy of Sciences, 1956. Available from: https://doi.org/10.4064%2Fap-3-1-165-168.

[50] JONATHAN CUTLER AND BALÁZS MONTÁGH. **Unavoidable subgraphs of colored graphs**. *Discrete Mathematics*, **308**(19):4396–4413, oct 2008. Available from: https://doi.org/10.1016%2Fj.disc.2007.08.102.

[51] D. DE CAEN. **Extension of a theorem of Moon and Moser on complete subgraphs**. *Ars Combin.*, **16**:5–10, 1983.

[52] REINHARD DIESTEL. *Graph Theory*. Springer-Verlag, Heidelberg, 2017.

[53] PAUL ERDŐS. **Some remarks on the theory of graphs**. *Bulletin of the American Mathematical Society*, **53**(4):292–294, 1947. Available from: https://doi.org/10.1090%2Fs0002-9904-1947-08785-1.

[54] PAUL ERDŐS. **An asymptotic inequality in the theory of numbers**. *Vestnik Leningrad Univ*, **15**(13):41–49, 1960.

[55] PAUL ERDŐS. **On extremal problems of graphs and generalized graphs**. *Israel J. Math.*, **2**(3):183–190, sep 1964. Available from: https://doi.org/10.1007%2Fbf02759942.

[56] PAUL ERDŐS, ZOLTÁN FÜREDI, BRUCE L. ROTHSCHILD, AND VERA T. SÓS. **Induced subgraphs of given sizes**. *Discrete Mathematics*, **200**(1-3):61–77, apr 1999. Available from: https://doi.org/10.1016%2Fs0012-365x%2898%2900387-2.

[57] PAUL ERDŐS AND ANDRÁS HAJNAL. **On chromatic number of graphs and set-systems**. *Acta Mathematica Academiae Scientiarum Hungaricae*, **17**(1-2):61–99, mar 1966. Available from: https://doi.org/10.1007%2Fbf02020444.

[58] PAUL ERDŐS AND ANDRÁS HAJNAL. **On Ramsey like theorems. Problems and results**. In *Combinatorics (Proc. Conf. Combinatorial Math., Math. Inst., Oxford, 1972)*, pages 123–140. Citeseer, 1972.

[59] PAUL ERDŐS AND ANDRÁS HAJNAL. **Ramsey-type theorems**. *Discrete Applied Mathematics*, **25**(1-2):37–52, oct 1989. Available from: https://doi.org/10.1016%2F0166-218x%2889%2990045-0.

[60] PAUL ERDŐS, ANDRÁS HAJNAL, AND JÁNOS PACH. **A Ramsey-type theorem for bipartite graphs**. *Geombinatorics*, **10**(2):64–68, 2000.

[61] PAUL ERDŐS, ANDRÁS. HAJNAL, AND RICHARD RADO. **Partition relations for cardinal numbers**. *Acta Math. Acad. Sci. Hungar.*, **16**:93–196, mar 1965. Available from: https://doi.org/10.1007%2Fbf01886396.

[62] Paul Erdős and János Pach. **On a Quasi-Ramsey problem**. *Journal of Graph Theory*, **7**(1):137–147, 1983. Available from: https://doi.org/10.1002%2Fjgt.3190070117.

[63] Paul Erdős, Alfréd Rényi, and Vera T. Sós. **On a problem of graph theory**. *Studia Sci. Math. Hungar.*, **1**:215–235, 1966.

[64] Paul Erdős and Miklós Simonovits. **Supersaturated graphs and hypergraphs**. *Combinatorica*, **3**(2):181–192, jun 1983. Available from: https://doi.org/10.1007%2Fbf02579292.

[65] Paul Erdős and George Szekeres. **A combinatorial problem in geometry**. *Compositio mathematica*, **2**:463–470, 1935.

[66] Uriel Feige and Shimon Kogan. **Balanced coloring of bipartite graphs**. *Journal of Graph Theory*, **64**(4):277–291, 2009. Available from: https://doi.org/10.1002%2Fjgt.20456.

[67] Kevin Ford. **The distribution of integers with a divisor in a given interval**. *Annals of Mathematics*, **168**(2):367–433, sep 2008. Available from: https://doi.org/10.4007%2Fannals.2008.168.367.

[68] Kevin Ford. **Integers with a divisor in (y, 2y]**. In *Anatomy of Integers*, pages 65–80. American Mathematical Society, sep 2008. Available from: https://doi.org/10.4007%2Fannals.2008.168.367.

[69] Jean-Luc Fouquet, Vassilis Giakoumakis, and Jean-Marie Vanherpe. **Bipartite graphs totally decomposable by canonical decomposition**. *International Journal of Foundations of Computer Science*, **10**(04):513–533, dec 1999. Available from: https://doi.org/10.1142%2Fs0129054199000368.

[70] Jacob Fox, Andrey Grinshpun, and János Pach. **The Erdős–Hajnal conjecture for rainbow triangles**. *Journal of Combinatorial Theory, Series B*, **111**:75–125, mar 2015. Available from: https://doi.org/10.1016%2Fj.jctb.2014.09.005.

[71] Jacob Fox and Xiaoyu He. **Independent sets in hypergraphs with a forbidden link**. *Proceedings of the London Mathematical Society*, **123**(4):384–409, mar 2021. Available from: https://doi.org/10.1112%2Fplms.12400.

[72] Jacob Fox, János Pach, and Andrew Suk. **Erdős–Hajnal conjecture for graphs with bounded VC-dimension**. *Discrete & Computational Geometry*, **61**(4):809–829, nov 2018. Available from: https://doi.org/10.1007%2Fs00454-018-0046-5.

[73] JACOB FOX AND BENNY SUDAKOV. **Unavoidable patterns**. *Journal of Combinatorial Theory, Series A*, **115**(8):1561–1569, nov 2008. Available from: https://doi.org/10.1016%2Fj.jcta.2008.04.003.

[74] PÉTER FRANKL AND ZOLTÁN FÜREDI. **An exact result for 3-graphs**. *Discrete Math.*, **50**(2-3):323–328, 1984. Available from: https://doi.org/10.1016%2F0012-365x%2884%2990058-x.

[75] SHINYA FUJITA, COLTON MAGNANT, AND KENTA OZEKI. **Rainbow generalizations of Ramsey theory: A survey**. *Graphs and Combinatorics*, **26**(1):1–30, jan 2010. Available from: https://doi.org/10.1007%2Fs00373-010-0891-3.

[76] ZOLTÁN FÜREDI. **An upper bound on Zarankiewicz' problem**. *Combinatorics, Probability and Computing*, **5**(1):29–33, 1996. Available from: https://doi.org/10.1017/s0963548300001814.

[77] ZOLTÁN FÜREDI AND MIKLÓS SIMONOVITS. **The history of degenerate (bipartite) extremal graph problems**. In *Erdős Centennial*, pages 169–264. Springer, 2013. Available from: https://doi.org/10.1007%2F978-3-642-39286-3_7.

[78] ZOLTÁN FÜREDI, ALEXANDR V. KOSTOCHKA, RISTE ŠKREKOVSKI, MICHAEL STIEBITZ, AND DOUGLAS B. WEST. **Nordhaus-Gaddum-type theorems for decompositions into many parts**. *Journal of Graph Theory*, **50**(4):273–292, 2005. Available from: https://doi.org/10.1002%2Fjgt.20113.

[79] TIBOR GALLAI. **Transitiv orientierbare Graphen**. *Acta Mathematica Academiae Scientiarum Hungaricae*, **18**(1-2):25–66, mar 1967. Available from: https://doi.org/10.1007%2Fbf02020961.

[80] LIOR GISHBOLINER AND ISTVÁN TOMON. **On 3-graphs with no four vertices spanning exactly two edges**. *Bulletin of the London Mathematical Society*, **54**(6):2117–2134, may 2022. Available from: https://doi.org/10.1112%2Fblms.12681.

[81] ROBERT E. GREENWOOD AND ANDREW MATTEI GLEASON. **Combinatorial relations and chromatic graphs**. *Canadian Journal of Mathematics*, **7**:1–7, 1955. Available from: https://doi.org/10.4153%2Fcjm-1955-001-4.

[82] JERROLD R. GRIGGS AND JIANXIN OUYANG. $(0,1)$**-Matrices with no half–half submatrix of ones**. *European Journal of Combinatorics*, **18**(7):751–761, oct 1997. Available from: https://doi.org/10.1006%2Feujc.1996.0133.

[83] JERROLD R GRIGGS, MIKLÓOS SIMONOVITS, AND GEORGE RUBIN THOMAS. **Extremal graphs with bounded densities of small subgraphs**. *Journal of Graph Theory*, **29**(3):185–207, nov 1998. Available from: https:

//doi.org/10.1002%2F%28sici%291097-0118%28199811%2929%3A3%3C185%3A%3Aaid-jgt6%3E3.0.co%3B2-m.

[84] He Guo and Lutz Warnke. **Packing nearly optimal Ramsey $R(3,t)$ graphs**. *Combinatorica*, **40**(1):63–103, feb 2020. Available from: https://doi.org/10.1007%2Fs00493-019-3921-7.

[85] András Gyárfás. **Reflections on a Problem of Erdős and Hajnal**. In *The Mathematics of Paul Erdős II*, pages 135–141. Springer New York, 2013. Available from: https://doi.org/10.1007%2F978-1-4614-7254-4_11.

[86] András Gyárfás and Gábor Simonyi. **Edge colorings of complete graphs without tricolored triangles**. *Journal of Graph Theory*, **46**(3):211–216, apr 2004. Available from: https://doi.org/10.1002%2Fjgt.20001.

[87] Johannes H. Hattingh and Michael A. Henning. **Bipartite Ramsey theory**. *Utilitas Mathematica*, **53**:217–230, 1998.

[88] Jialin He, Jie Ma, and Lilu Zhao. **Improvements on induced subgraphs of given sizes**. *arXiv preprint*, 2021. Available from: arXiv.org/abs/2101.03898.

[89] Robert W. Irving. **A bipartite Ramsey problem and the Zarankiewicz numbers**. *Glasgow Mathematical Journal*, **19**(1):13–26, jan 1978. Available from: https://doi.org/10.1017%2Fs0017089500003323.

[90] Svante Janson, Tomasz Łuczak, and Andrzej Rucinski. *Random Graphs*, **45** of *Wiley-Interscience Series in Discrete Mathematics and Optimization*. John Wiley & Sons, Inc., feb 2011. Available from: https://doi.org/10.1002%2F9781118032718.

[91] Ross J. Kang, János Pach, Viresh Patel, and Guus Regts. **A precise threshold for Quasi-Ramsey numbers**. *SIAM Journal on Discrete Mathematics*, **29**(3):1670–1682, jan 2015. Available from: https://doi.org/10.1137%2F14097313x.

[92] Peter Keevash. **Hypergraph Turán problems**. *Surveys in combinatorics 2011*, **392**:83–139, jun 2011. Available from: https://doi.org/10.1017%2Fcbo9781139004114.004.

[93] Jeong Han Kim. **The Ramsey number $R(3,t)$ has order of magnitude $t^2/\log t$**. *Random Structures & Algorithms*, **7**(3):173–207, oct 1995. Available from: https://doi.org/10.1002%2Frsa.3240070302.

[94] Rivka Klein and Jochanan Schönheim. **Decomposition of K_n into degenerate graphs**. *Combinatorics and Graph Theory*, pages 141–155, 1992.

[95] Jerzy Konarski and Andrzej Żak. **Near packings of two graphs**. *Discrete Mathematics*, **340**(5):963–968, may 2017. Available from: https://doi.org/10.1016%2Fj.disc.2016.12.022.

[96] Dániel Korándi, János Pach, and István Tomon. **Large homogeneous submatrices**. *SIAM Journal on Discrete Mathematics*, **34**(4):2532–2552, jan 2020. Available from: https://doi.org/10.1137%2F19m125786x.

[97] Alexandr Kostochka, Dhruv Mubayi, and Jacques Verstraëte. **On independent sets in hypergraphs**. *Random Structures & Algorithms*, **44**(2):224–239, aug 2012. Available from: https://doi.org/10.1002%2Frsa.20453.

[98] Tamás Kővári, Vera T. Sós, and Pál Turán. **On a problem of K. Zarankiewicz**. *Colloquium Mathematicum*, **3**(1):50–57, 1954. Available from: https://doi.org/10.4064%2Fcm-3-1-50-57.

[99] Lauwerens Kuipers and Harald Niederreiter. *Uniform distribution of sequences*. Courier Corporation, 2012.

[100] Matthew Kwan and Benny Sudakov. **Proof of a conjecture on induced subgraphs of Ramsey graphs**. *Transactions of the American Mathematical Society*, **372**(8):5571–5594, dec 2018. Available from: https://doi.org/10.1090%2Ftran%2F7729.

[101] Matthew Kwan and Benny Sudakov. **Ramsey graphs induce subgraphs of quadratically many sizes**. *International Mathematics Research Notices*, **2020**(6):1621–1638, apr 2018. Available from: https://doi.org/10.1093%2Fimrn%2Frny064.

[102] Felix Lazebnik, Vasiliy Ustimenko, and Andrew Woldar. **A new series of dense graphs of high girth**. *Bulletin of the American mathematical society*, **32**(1):73–79, 1995. Available from: https://doi.org/10.1090%2Fs0273-0979-1995-00569-0.

[103] Bernard Lidický. **Flagmatic webpage**. Available from: http://lidicky.name/flagmatic/.

[104] Vadim V. Lozin. **Bipartite graphs without a skew star**. *Discrete Mathematics*, **257**(1):83–100, nov 2002. Available from: https://doi.org/10.1016%2Fs0012-365x%2801%2900471-x.

[105] Tomasz Łuczak. **Highly connected monochromatic subgraphs of two-colored complete graphs**. *Journal of Combinatorial Theory, Series B*, **117**:88–92, mar 2016. Available from: https://doi.org/10.1016%2Fj.jctb.2015.11.006.

[106] Brendan D. McKay, Nicholas C. Wormald, and Beata Wysocka. **Short cycles in random regular graphs**. *The Electronic Journal of Combinatorics*, **11**(1), sep 2004. Available from: https://doi.org/10.37236%2F1819.

[107] Dhruv Mubayi. **Generalizing the Ramsey problem through diameter**. *The Electronic Journal of Combinatorics*, **9**(1), nov 2001. Available from: https://doi.org/10.37236%2F1658.

[108] Dhruv Mubayi. **A hypergraph extension of Turán's theorem**. *Journal of Combinatorial Theory, Series B*, **96**(1):122–134, 2006. Available from: https://doi.org/10.1016/j.jctb.2005.06.013.

[109] Dhruv Mubayi and Alexander Razborov. **Polynomial to exponential transition in Ramsey theory**. *Proceedings of the London Mathematical Society*, **122**(1):69–92, may 2020. Available from: https://doi.org/10.1112%2Fplms.12320.

[110] Dhruv Mubayi and Andrew Suk. **New lower bounds for hypergraph Ramsey numbers**. *Bulletin of the London Mathematical Society*, **50**(2):189–201, dec 2017. Available from: https://doi.org/10.1112%2Fblms.12133.

[111] Bhargav Narayanan, Julian Sahasrabudhe, and István Tomon. **Ramsey graphs induce subgraphs of many different sizes**. *Combinatorica*, **39**(1):215–237, feb 2018. Available from: https://doi.org/10.1007%2Fs00493-017-3755-0.

[112] Kevin T. Phelps and Vojtech Rödl. **Steiner triple systems with minimum independence number**. *Ars Combin*, **21**:167–172, 1986.

[113] Svatopluk Poljak and Zsolt Tuza. **Bipartite subgraphs of triangle-free graphs**. *SIAM Journal on Discrete Mathematics*, **7**(2):307–313, may 1994. Available from: https://doi.org/10.1137%2Fs0895480191196824.

[114] Frank P. Ramsey. **On a problem of formal logic**. *Proceedings of the London Mathematical Society*, **2**(1):264–286, 1930.

[115] Istvan Reiman. **Über ein Problem von K. Zarankiewicz**. *Acta Mathematica Academiae Scientiarum Hungaricae*, **9**(3-4):269–273, sep 1958. Available from: https://doi.org/10.1007%2Fbf02020254.

[116] M. Rosenfeld. **Independent sets in regular graphs**. *Israel Journal of Mathematics*, **2**(4):262–272, dec 1964. Available from: https://doi.org/10.1007%2Fbf02759743.

[117] Alex Scott, Paul Seymour, and Sophie Spirkl. **Pure pairs. IV. Trees in bipartite graphs**. *Journal of Combinatorial Theory, Series B*, **161**:120–146, jul 2023. Available from: https://doi.org/10.1016%2Fj.jctb.2023.02.005.

[118] ALEXANDER F. SIDORENKO. **Systems of sets with T-property**. *Vestnik Moskovskogo Universiteta Seriya 1 Matematika Mekhanika*, **5**:19–22, 1981.

[119] WACŁAW SIERPIŃSKI. **Sur une propriété des nombres tétraédraux**. *Elemente der Mathematik*, **17**:29–30, 1962.

[120] JOEL SPENCER. **Asymptotic lower bounds for Ramsey functions**. *Discrete Mathematics*, **20**:69–76, 1977. Available from: https://doi.org/10.1016%2F0012-365x%2877%2990044-9.

[121] ANDREW THOMASON. **On finite Ramsey numbers**. *European Journal of Combinatorics*, **3**(3):263–273, sep 1982. Available from: https://doi.org/10.1016%2Fs0195-6698%2882%2980038-3.

[122] PAUL TURÁN. **On an extremal problem in graph theory**. *Matematikai és Fizikai Lapok (in Hungarian)*, **48**:436–452, 1941.

[123] EMIL VAUGHAN. **Flagmatic software package**. Available from: http://jakubsliacan.eu/flagmatic/.

[124] LUTZ WARNKE. **The C_ℓ-free process**. *Random Structures & Algorithms*, **44**(4):490–526, nov 2012. Available from: https://doi.org/10.1002%2Frsa.20468.

[125] LEA WEBER. **On the structure of graphs without forbidden induced subgraphs**. *[Master Thesis, Karlsruhe Institute of Technology]*, 2019. Available from: https://www.math.kit.edu/iag6/~weberl/media/thesis.pdf.

[126] LEA WEBER. **Avoidable order-size pairs in hypergraphs**. *To appear in Journal of Combinatorics. arXiv preprint*, 2022. Available from: arXiv.org/abs/2205.15197.

[127] DOUGLAS BRENT WEST ET AL. *Introduction to graph theory*, **2**. Prentice hall Upper Saddle River, 2001.

[128] ŠTEFAN ZNÁM. **On a combinatorical problem of K. Zarankiewicz**. *Colloquium Mathematicum*, **11**(1):81–84, 1963. Available from: https://doi.org/10.4064%2Fcm-11-1-81-84.